US 20160032396A1

(19) United States

(12) Patent Application Publication

Diehn et al.

(10) Pub. No.: US 2016/0032396 A1

(43) Pub. Date: Feb. 4, 2016

(54) IDENTIFICATION AND USE OF CIRCULATING NUCLEIC ACID TUMOR MARKERS

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, Stanford, CA (US)

(72) Inventors: **Maximilian Diehn**, San Carlos, CA (US); **Arash Ash Alizadeh**, San Mateo, CA (US); **Aaron M. Newman**, Palo Alto, CA (US); **Scott V. Bratman**, Palo Alto, CA (US)

(21) Appl. No.: **14/774,518**

(22) PCT Filed: **Mar. 12, 2014**

(86) PCT No.: **PCT/US14/25020**

§ 371 (c)(1),
(2) Date: **Sep. 10, 2015**

Related U.S. Application Data

(60) Provisional application No. 61/798,925, filed on Mar. 15, 2013.

Publication Classification

(51) **Int. Cl.**
C12Q 1/68 (2006.01)
G06F 19/22 (2006.01)

(52) **U.S. Cl.**
CPC ***C12Q 1/6886*** (2013.01); ***G06F 19/22*** (2013.01); *C12Q 2600/156* (2013.01)

(57) **ABSTRACT**

Methods for creating a selector of mutated genomic regions and for using the selector set to analyze genetic alterations in a cell-free nucleic acid sample are provided. The methods can be used to measure tumor-derived nucleic acids in a blood sample from a subject and thus to monitor the progression of disease in the subject. The methods can also be used for cancer screening, cancer diagnosis, cancer prognosis, and cancer therapy designation.

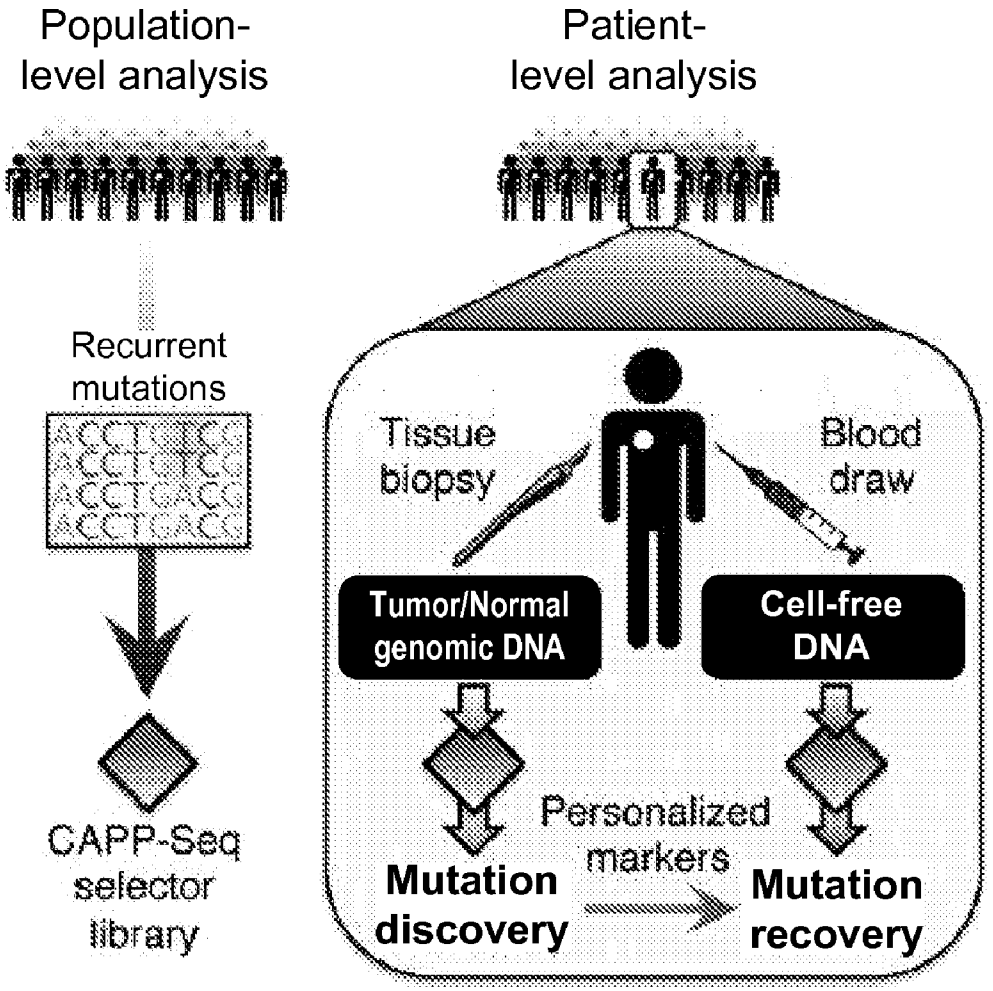


Figure 1A

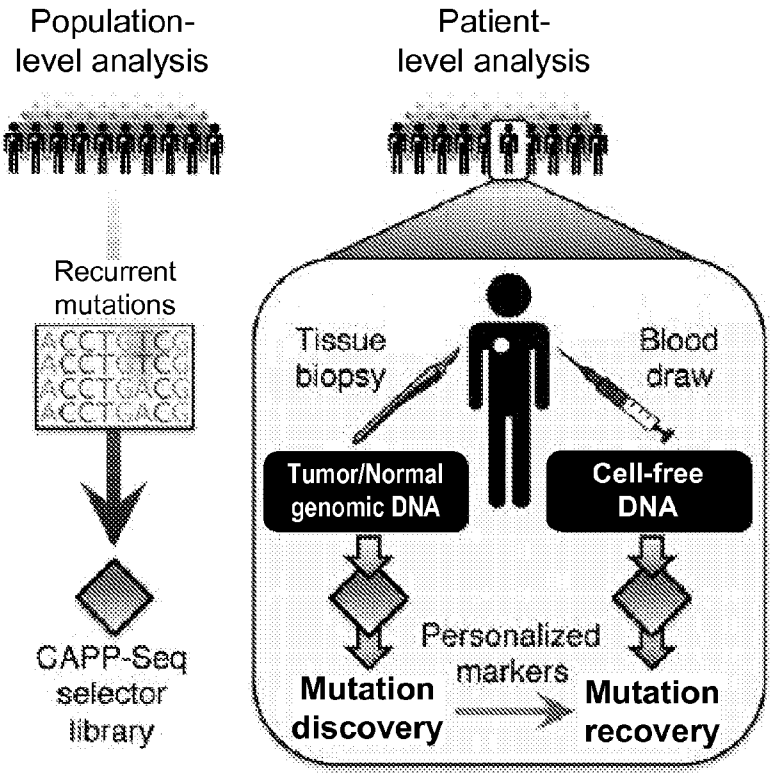


Figure 1B

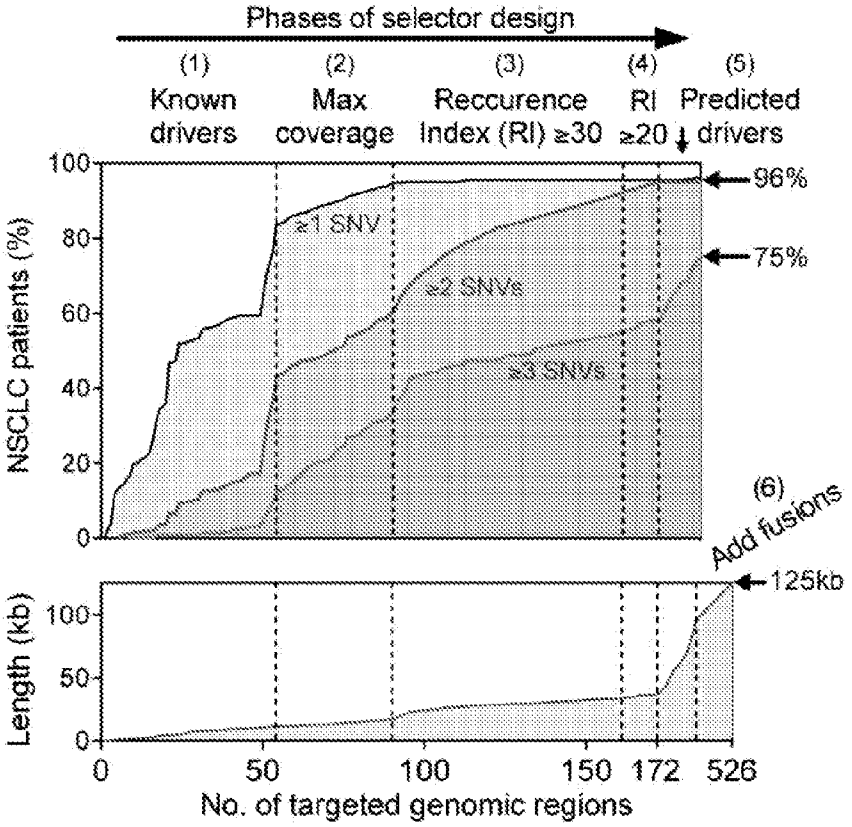


Figure 1C

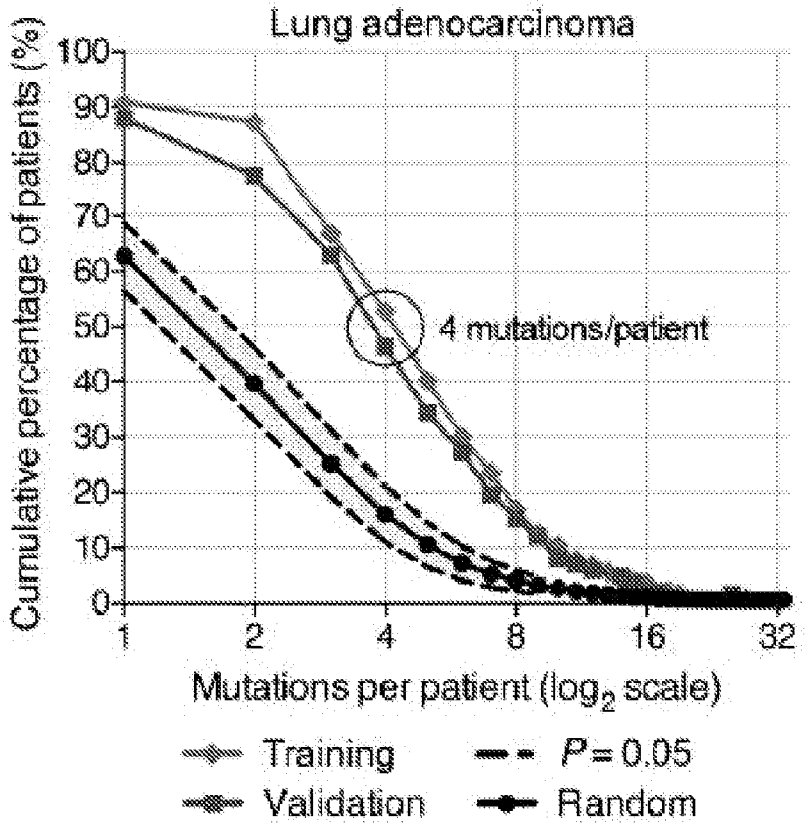


Figure 1D

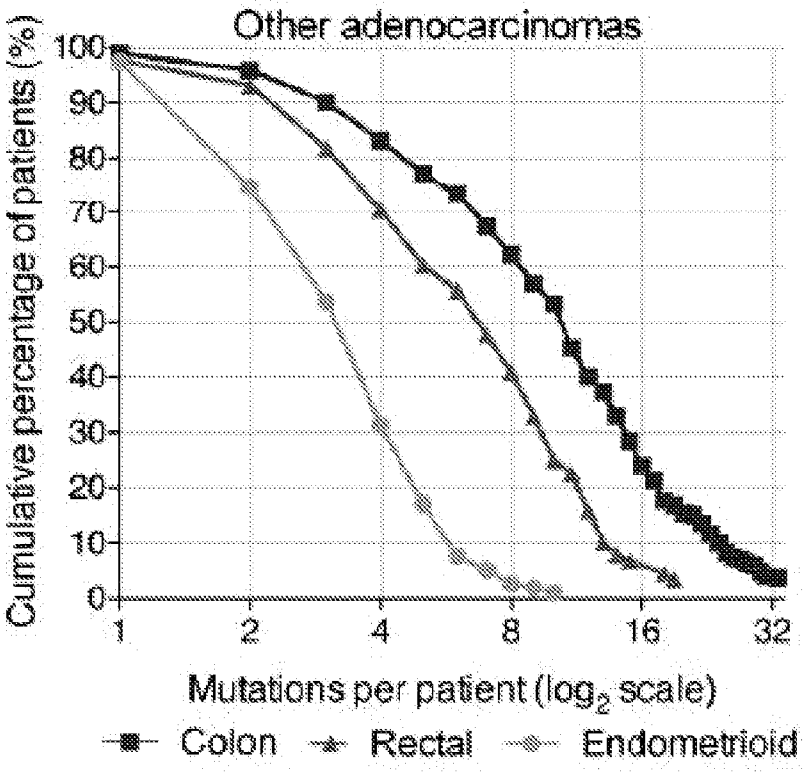


Figure 2A

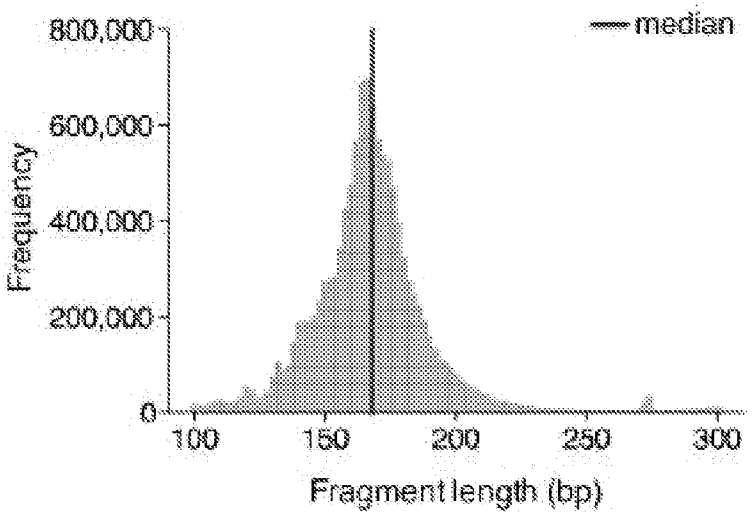


Figure 2B

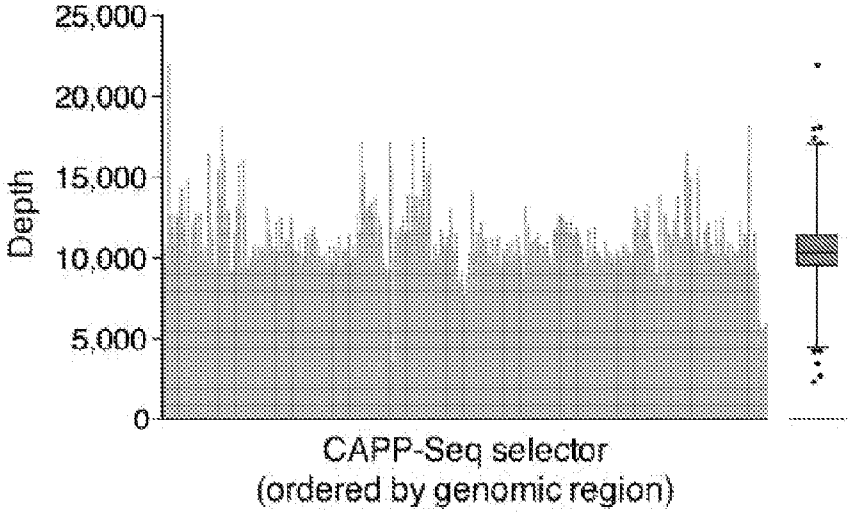


Figure 2C

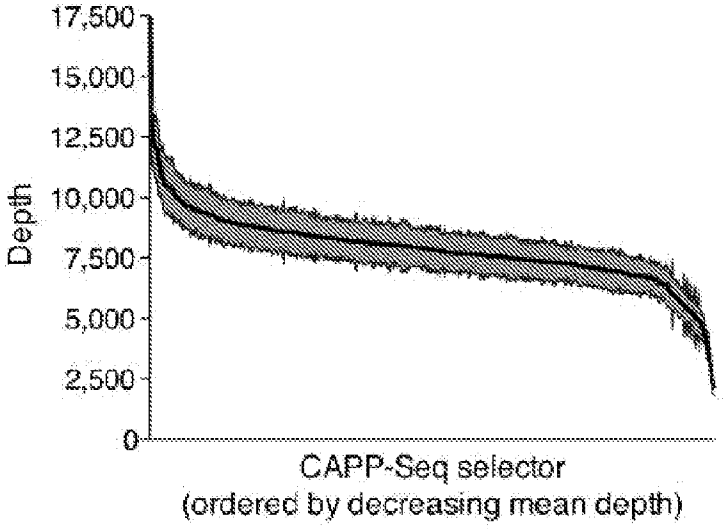


Figure 2D

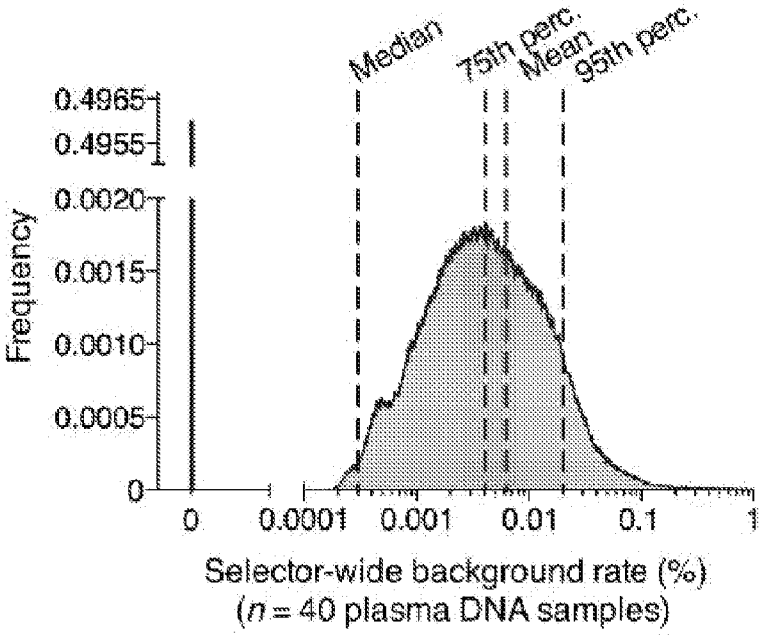
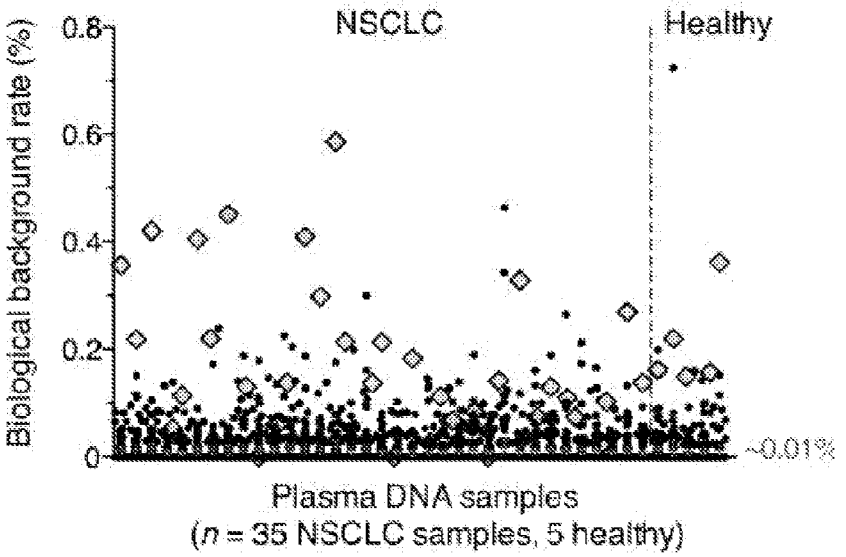


Figure 2E



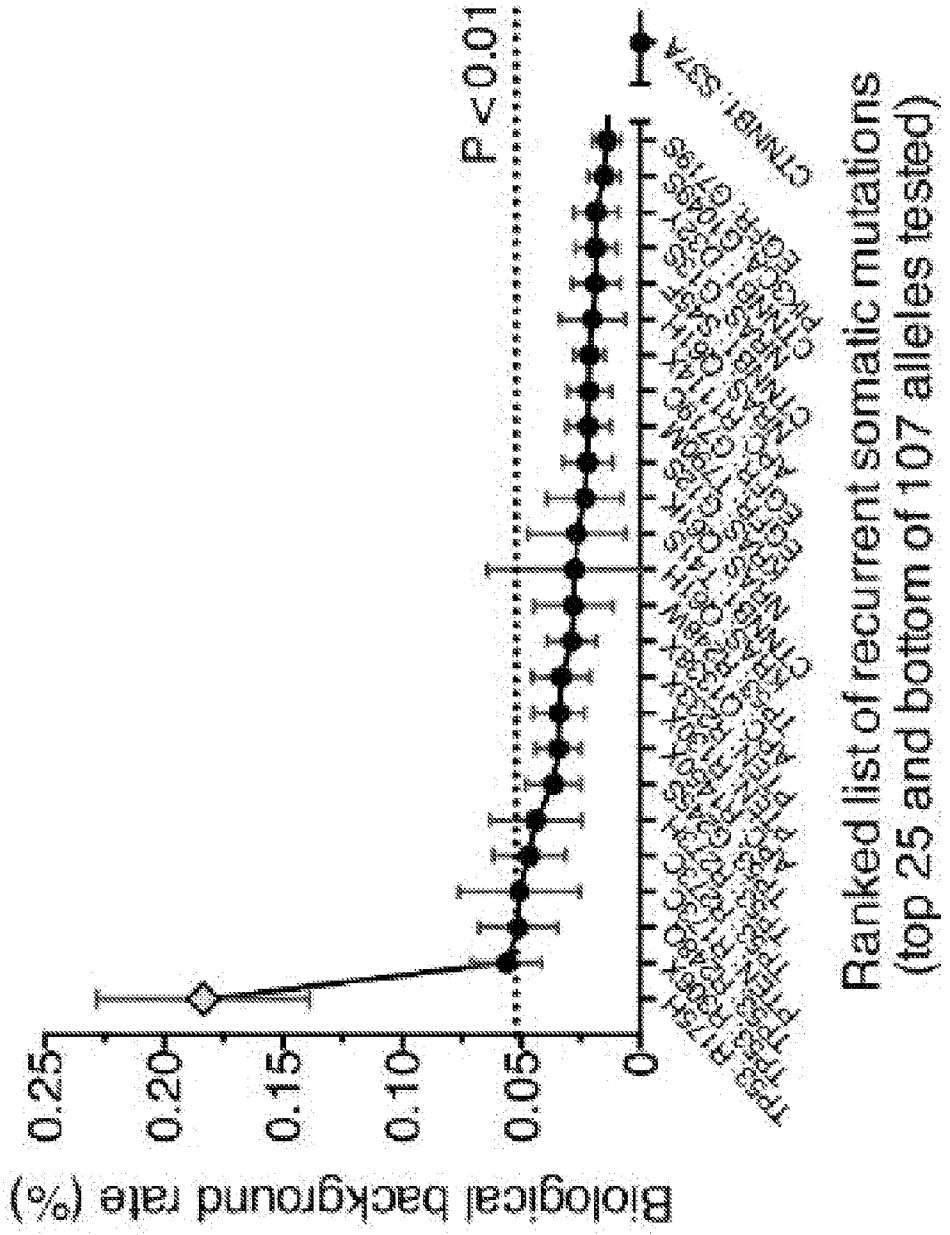

Biological background rate (%)
0.25
0.20
0.15
0.10
0.05
0
P <0.01
CTNNB1 S37A
Ranked list of recurrent somatic mutations
(top 25 and bottom of 107 alleles tested)

Figure 2F

Figure 2G

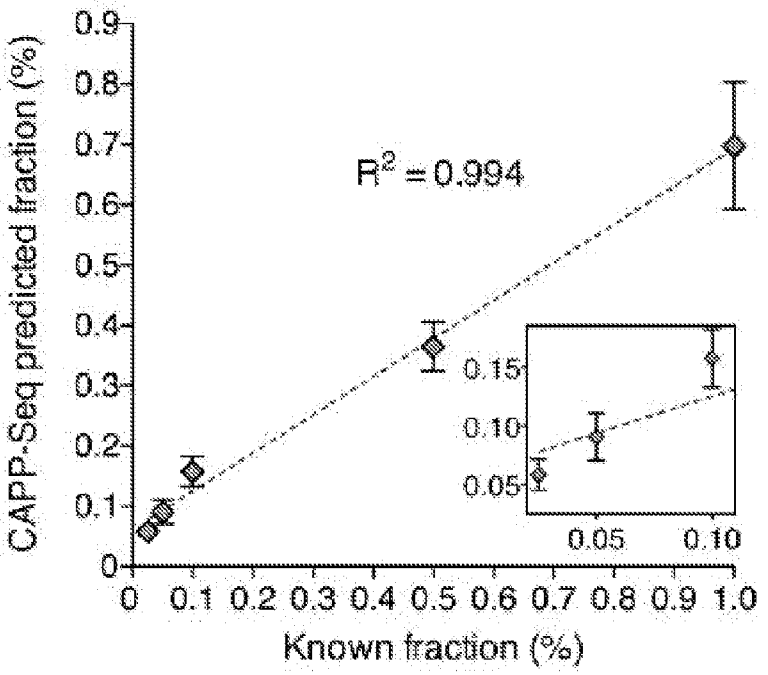


Figure 2H

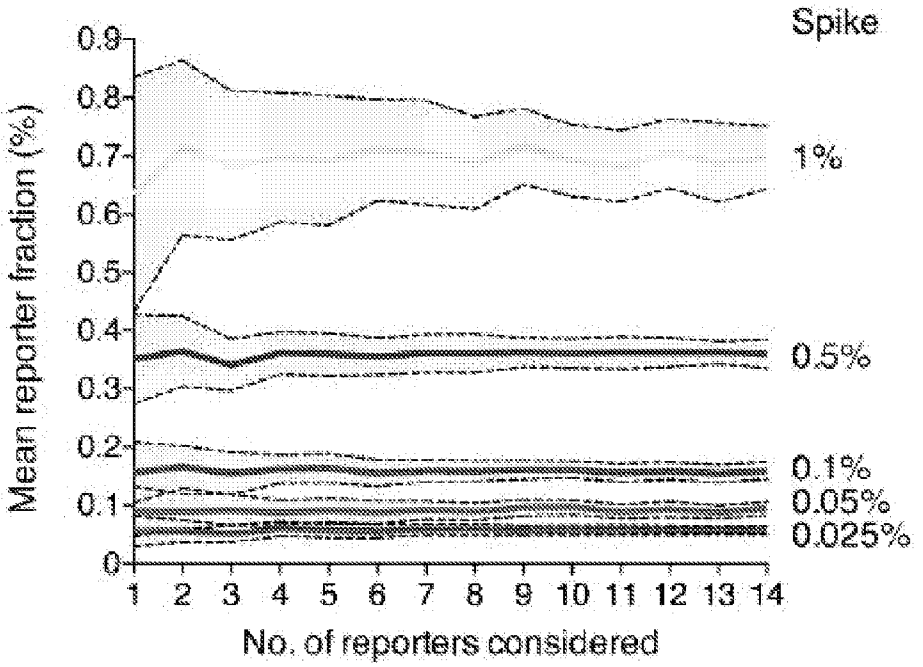


Figure 2I

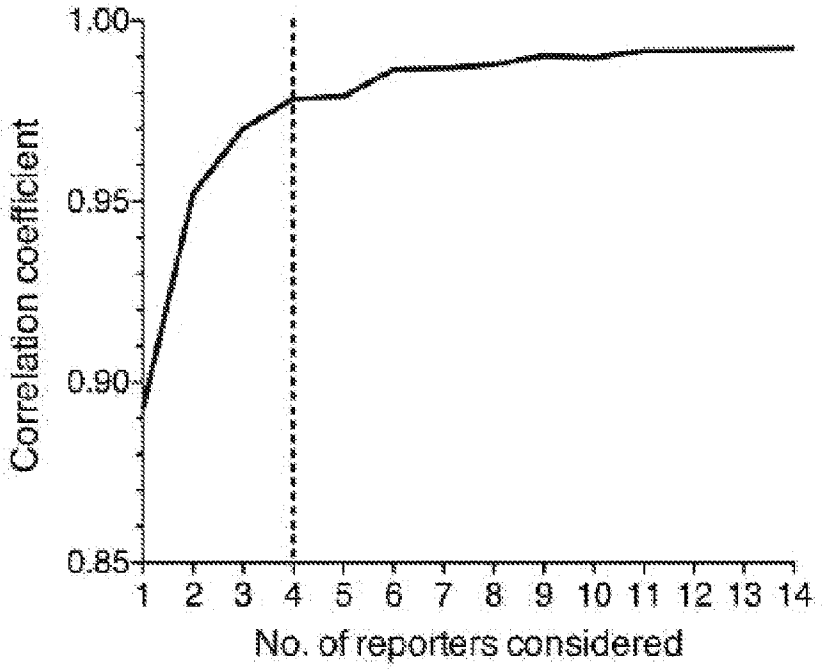


Figure 3A

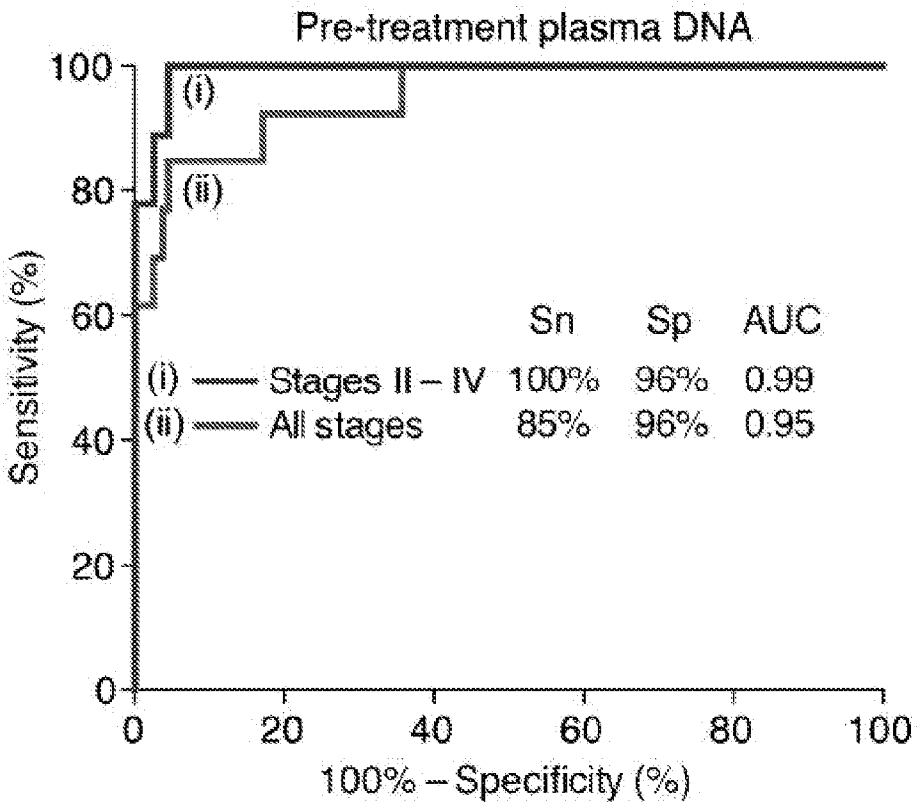


Figure 3B

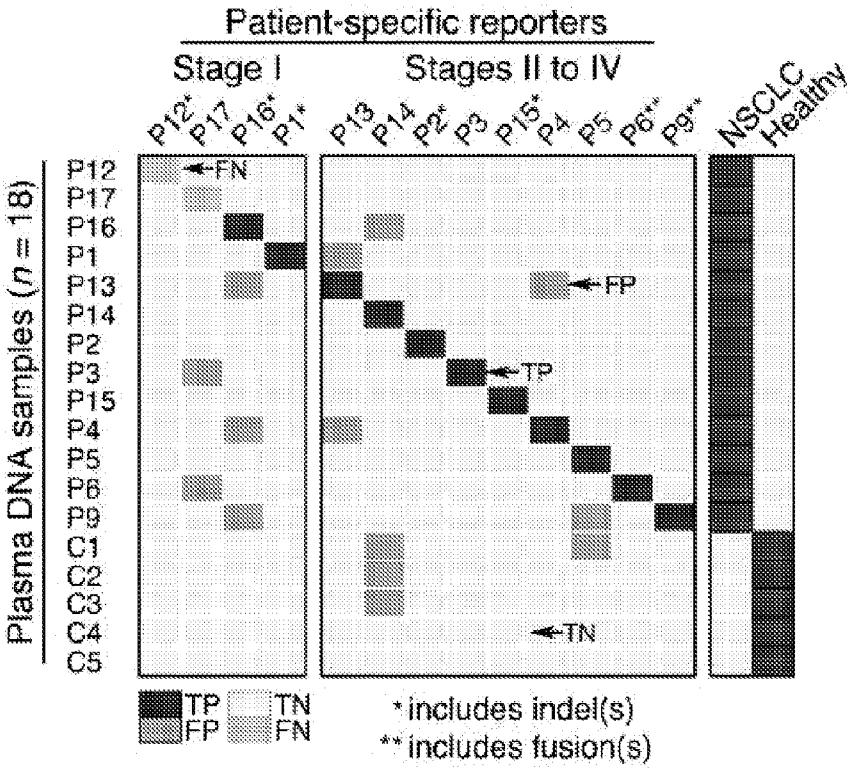


Figure 3C

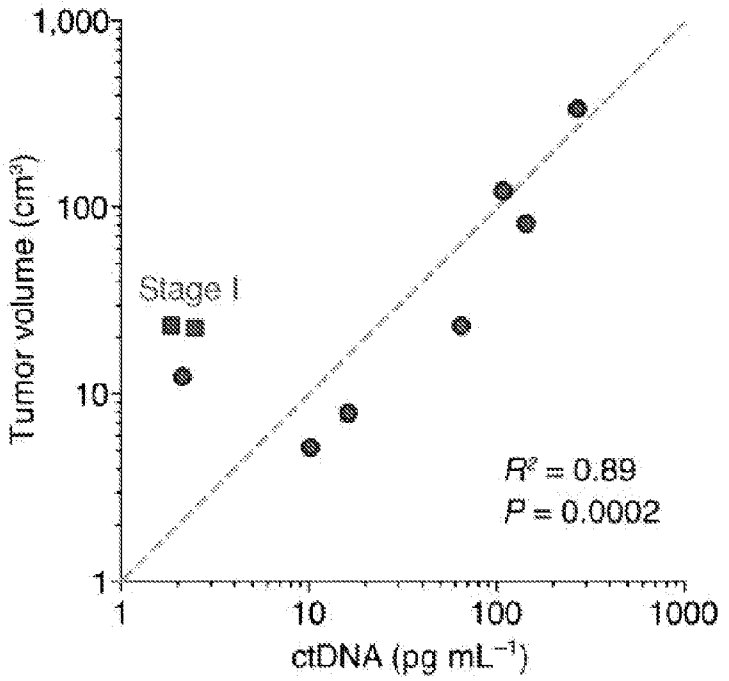


Figure 4A

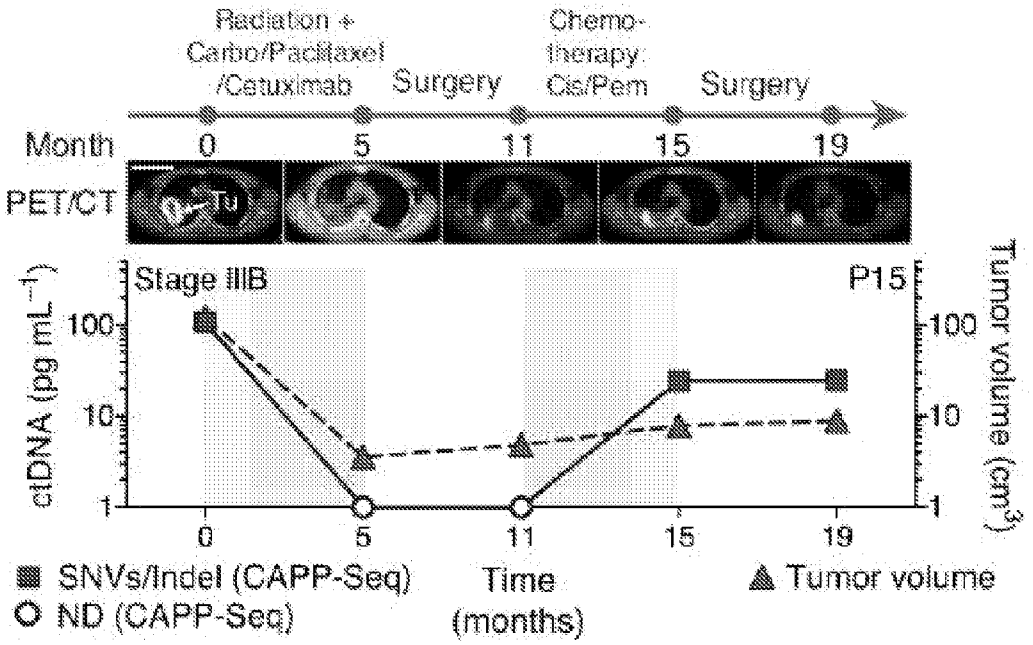


Figure 4B

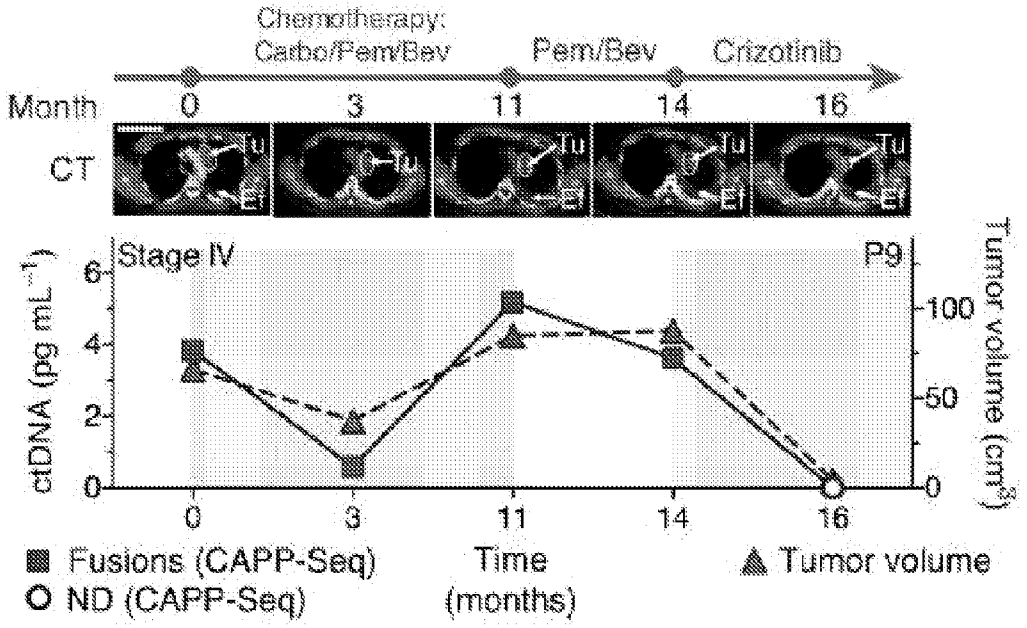


Figure 4C

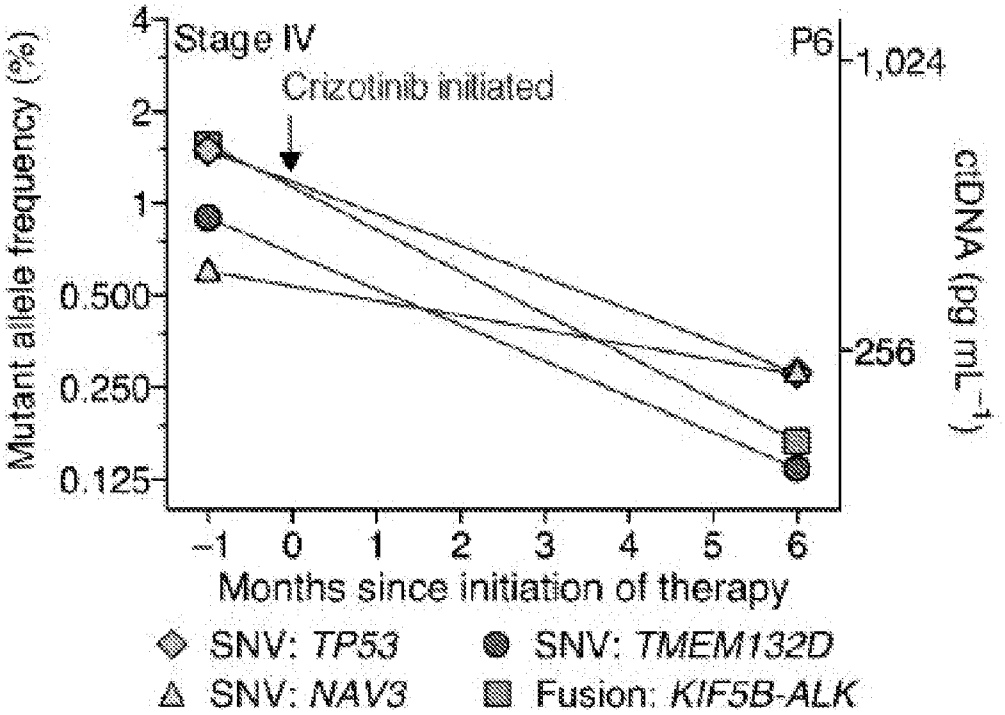


Figure 4D

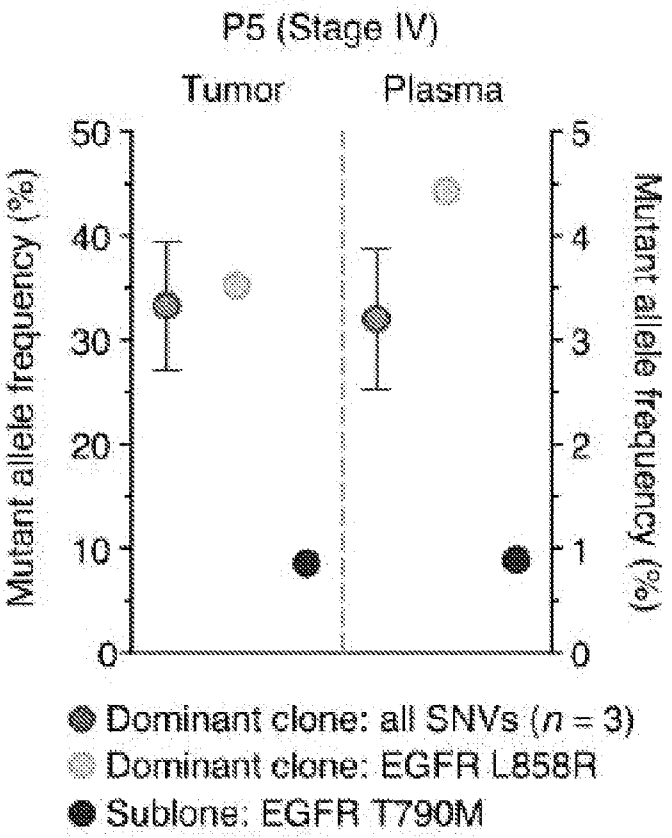


Figure 4E

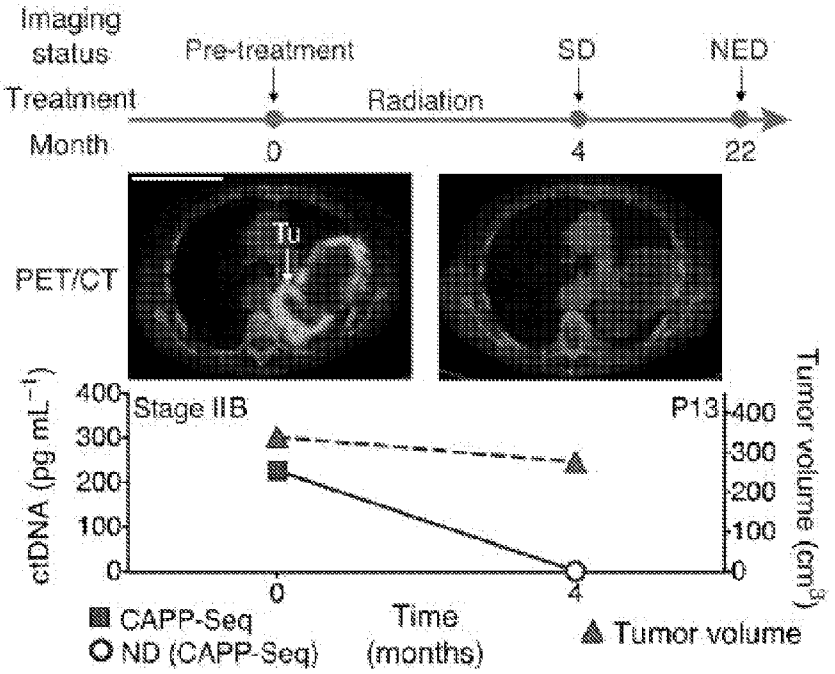


Figure 4F

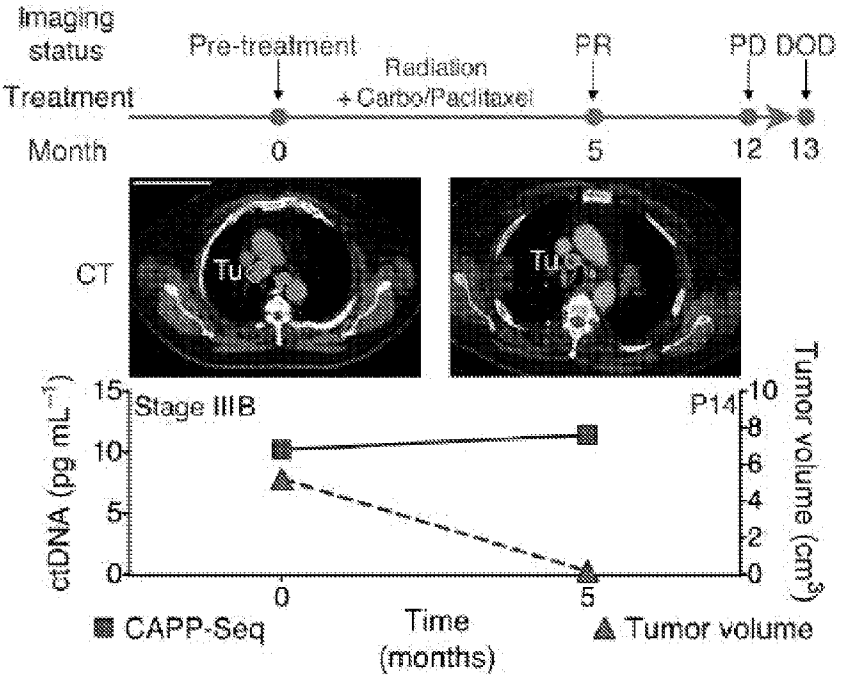


Figure 4G

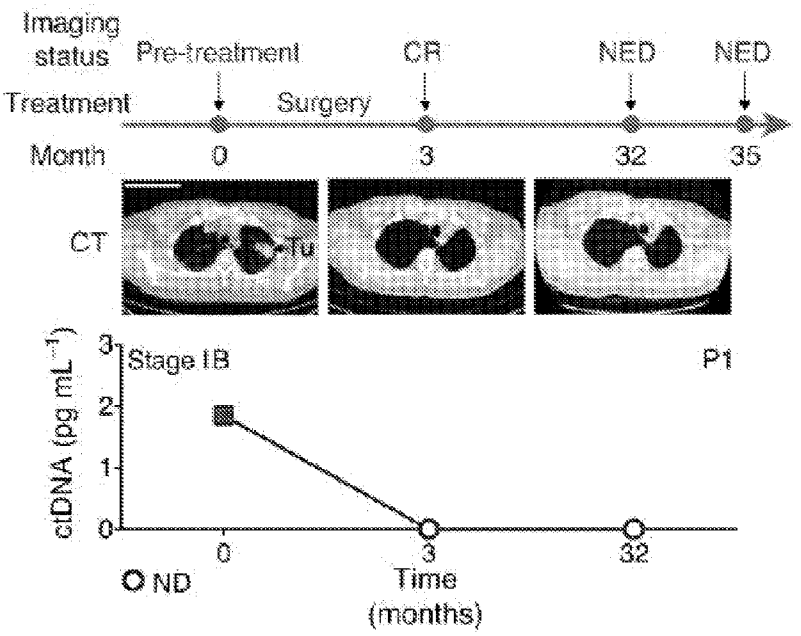


Figure 4H

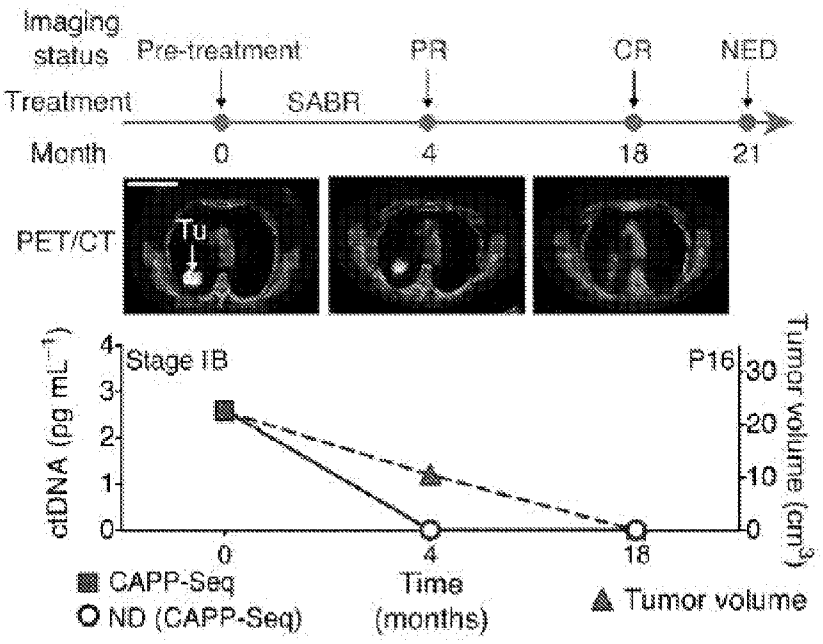


Figure 4I

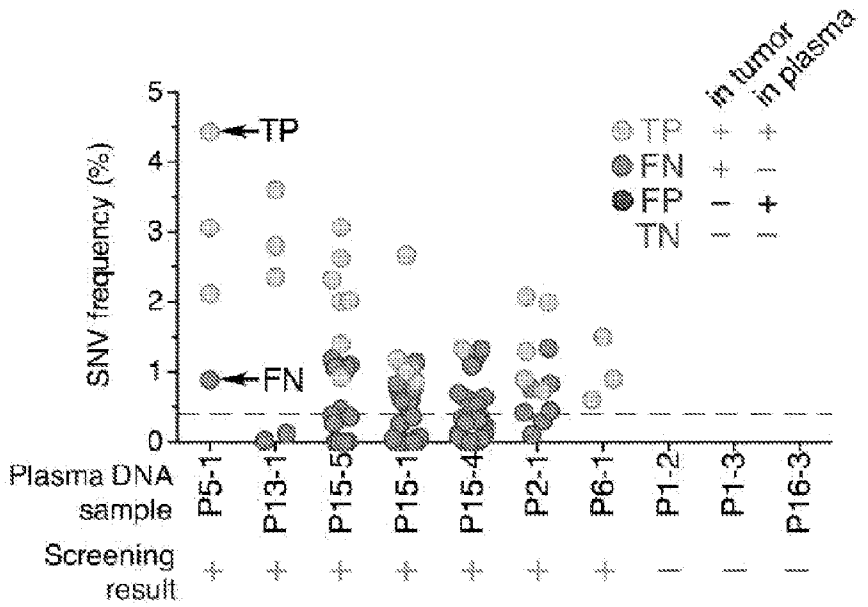

SNV frequency (%)
5
4
3
2
1
0
TP
FN
in tumor
in plasma
TP + +
FN + −
FP − +
TN − −
Plasma DNA sample
P5-1
P13-1
P15-5
P15-1
P15-4
P2-1
P6-1
P1-2
P1-3
P16-3
Screening result
+ + + + + + + − − −

Figure 5A

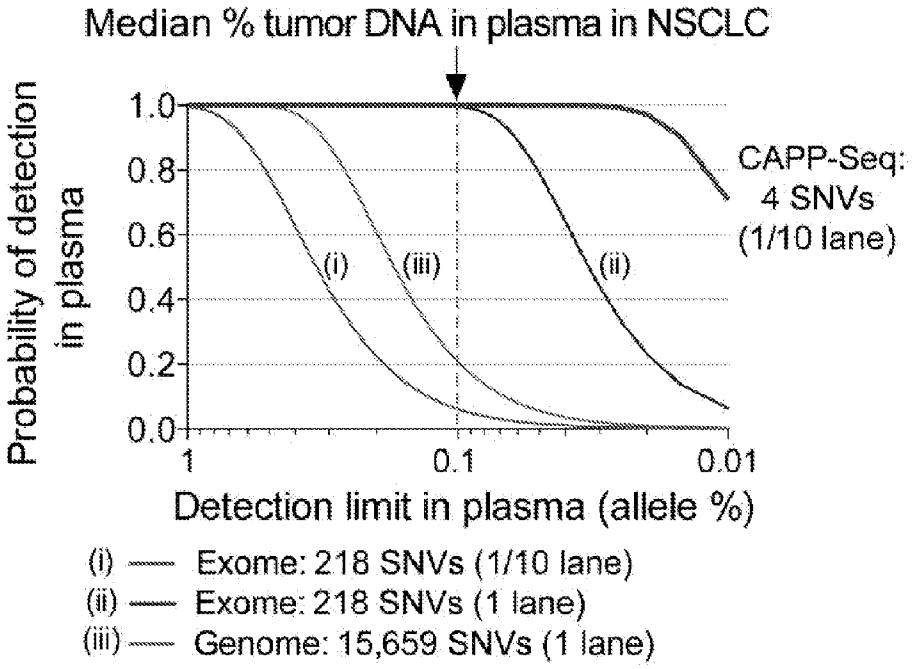


Figure 5B

Cost comparison for equal detection limit in plasma

Feature	CAPP-Seq	Whole exome	Whole genome
Bases covered	0.125 Mb	50 Mb	3,000 Mb
No. mutations	4	218	15,659
Depth	10,000x	1,100x	120x
Gbs sequenced	2.1	92	360
Relative cost	**1x**	**44x**	**171x**

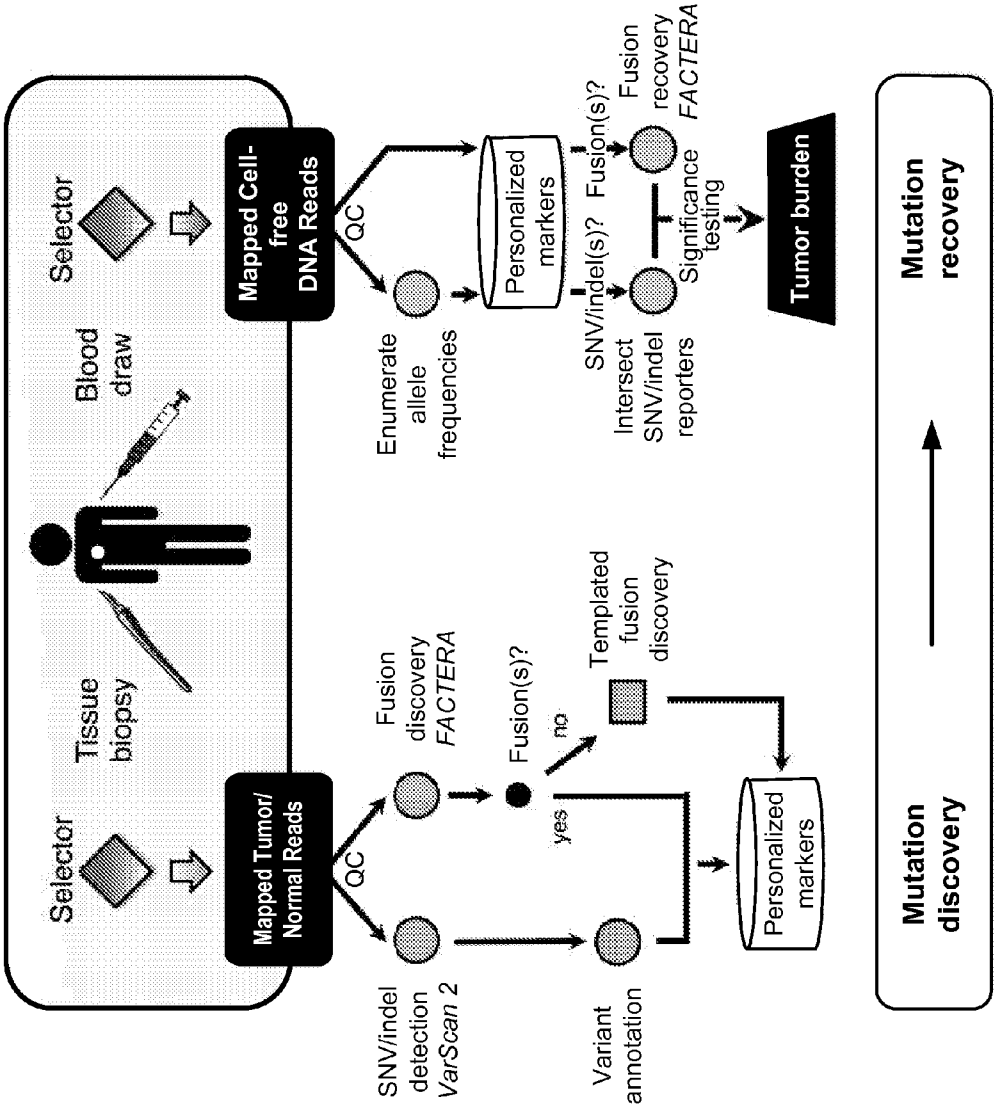

Selector
Blood draw
Tissue biopsy
Selector
Mapped Cell-free DNA Reads
Mapped Tumor/ Normal Reads
QC
QC
Enumerate allele frequencies
Personalized markers
Fusion(s)?
Fusion recovery FACTERA
SNV/indel(s)?
Intersect SNV/indel reporters
Significance testing
Tumor burden
Fusion discovery FACTERA
Fusion(s)?
yes
no
Templated fusion discovery
SNV/indel detection VarScan 2
Variant annotation
Personalized markers
Mutation discovery
Mutation recovery

Figure 6

Figure 7A

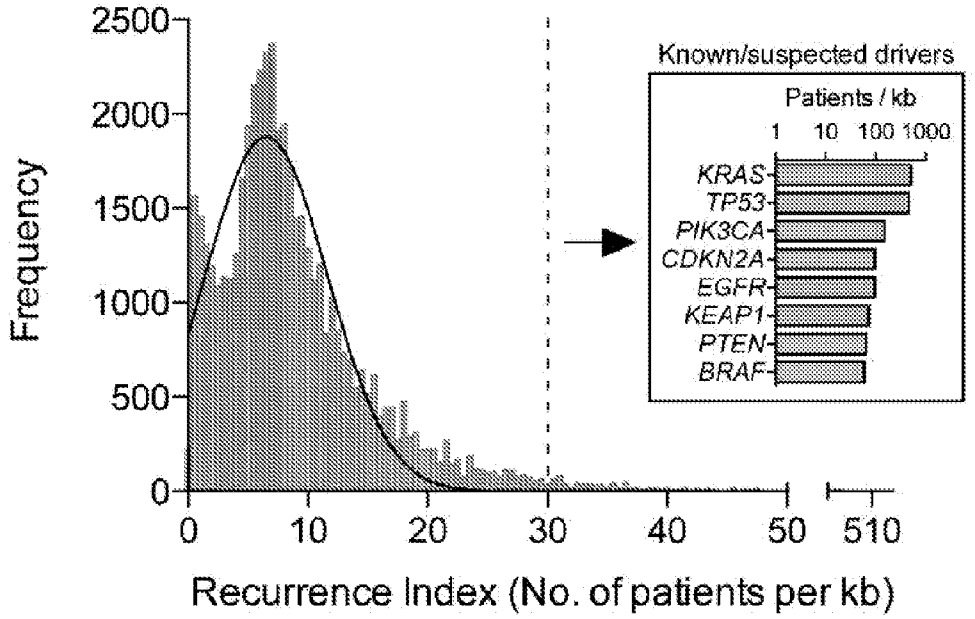
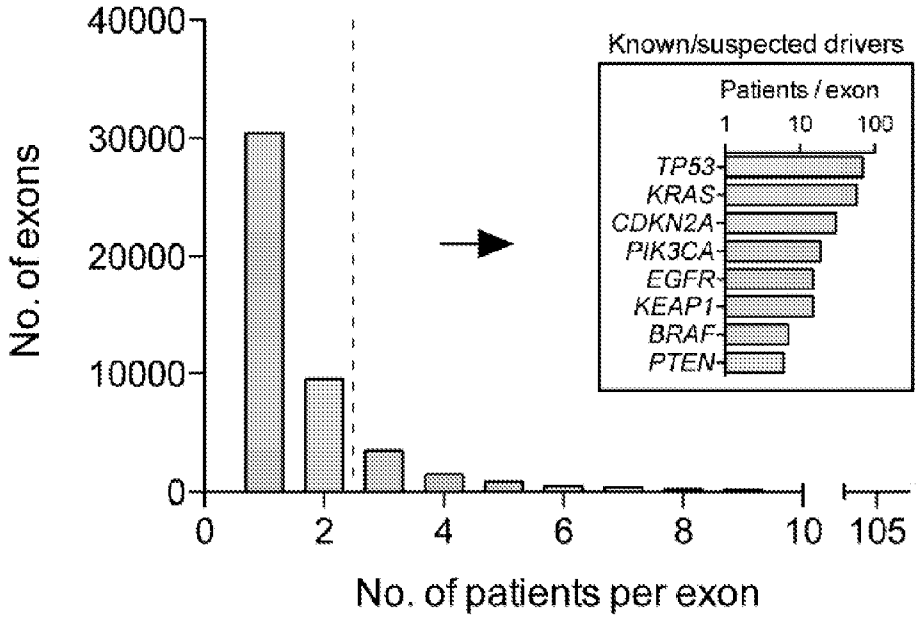


Figure 7B



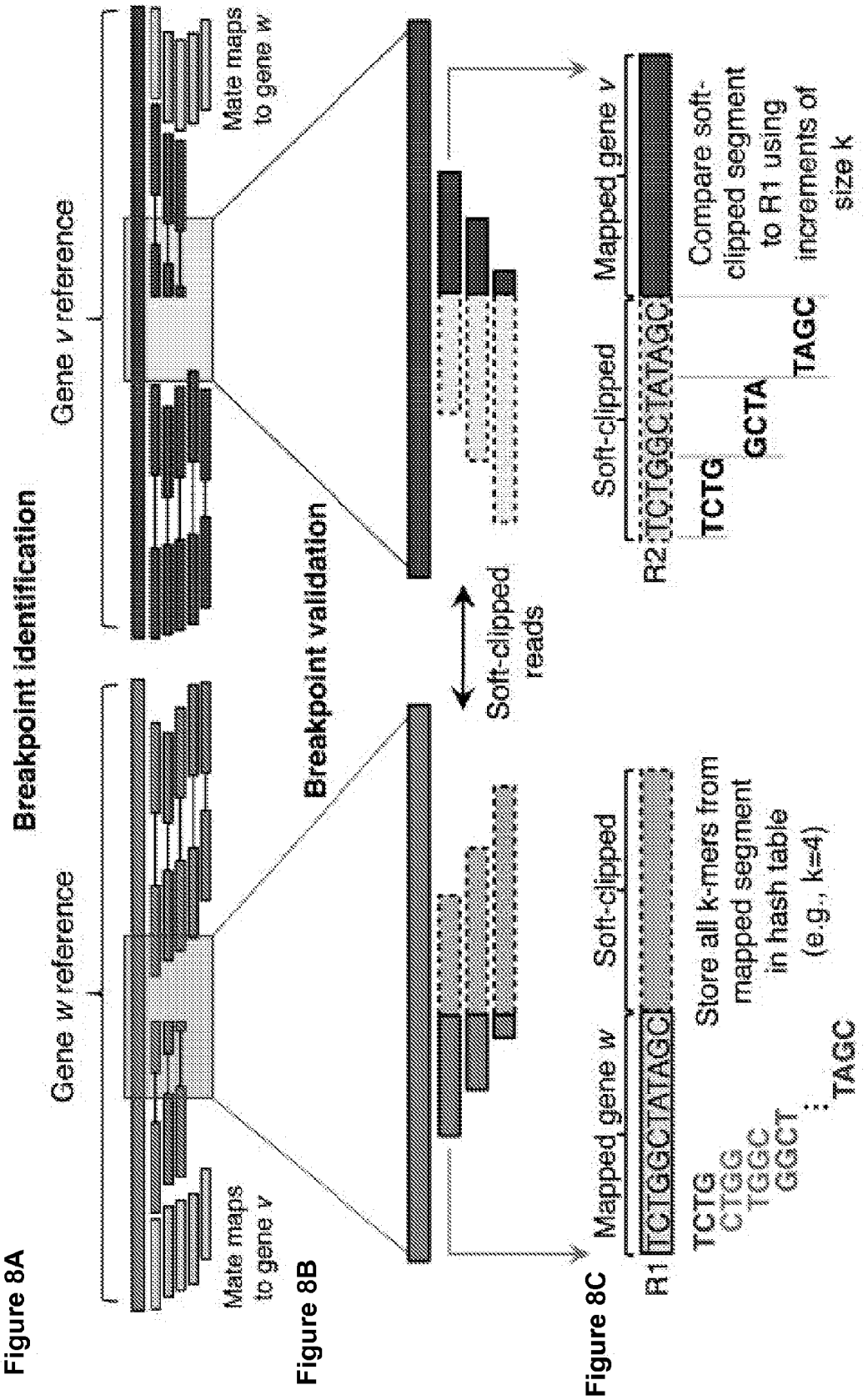


Figure 8A

Figure 8B

Figure 8C

Figure 8D

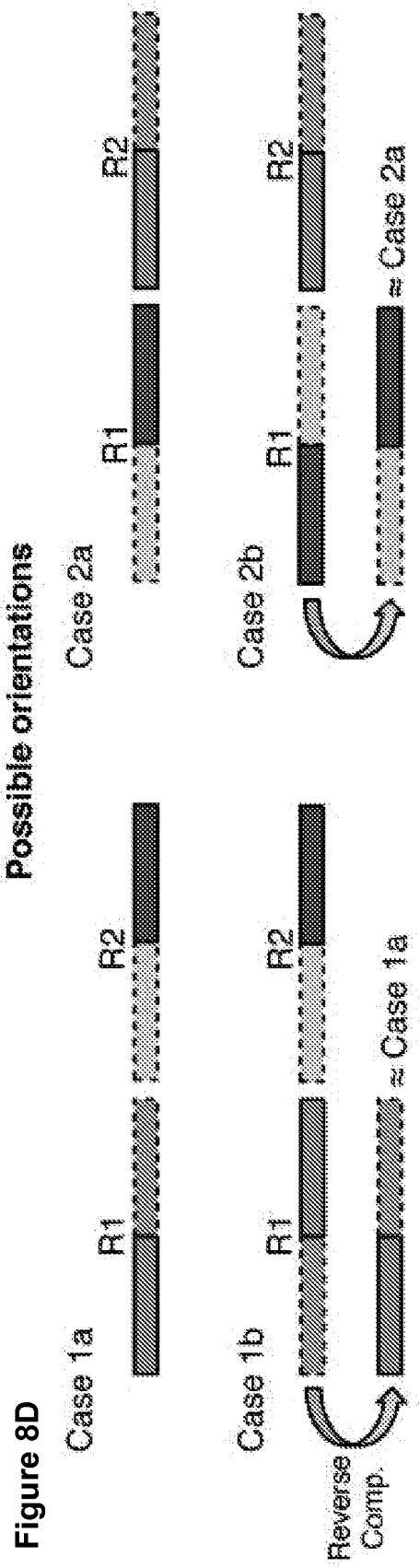


Figure 8E

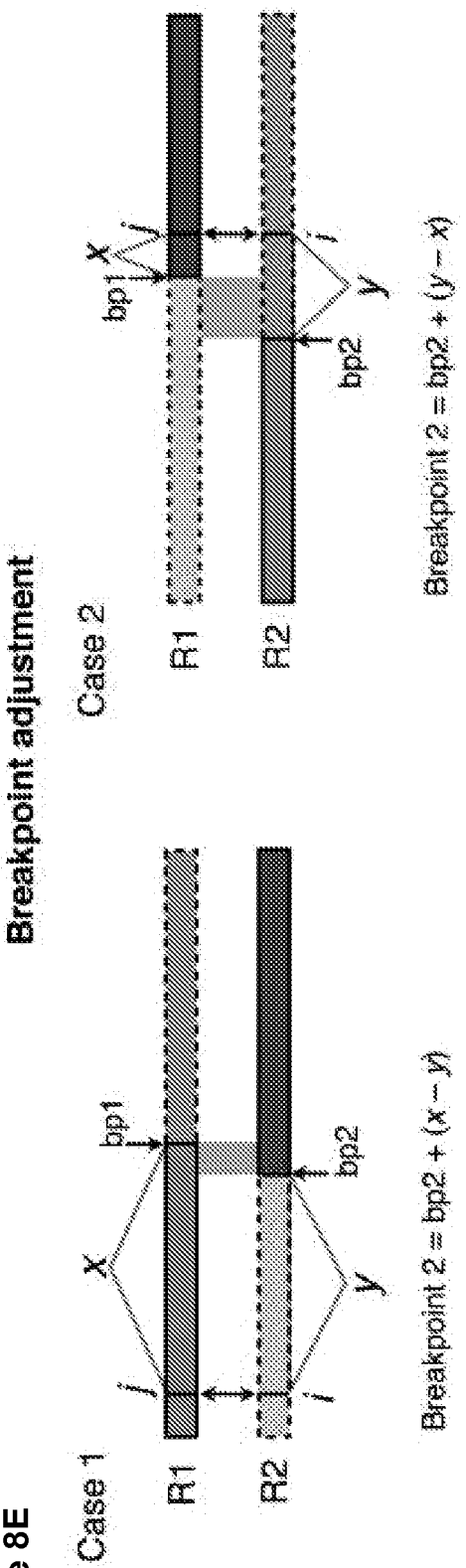


Figure 9A

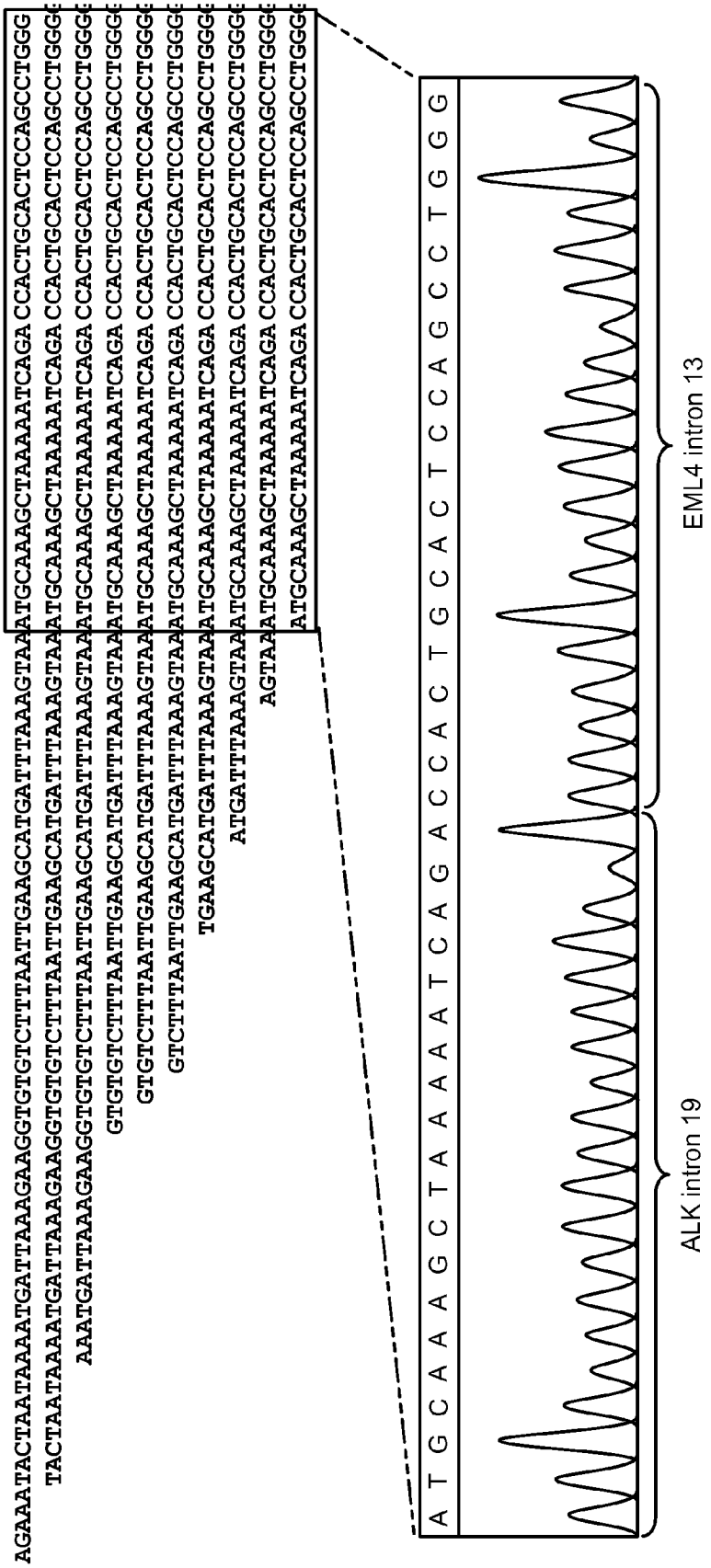

AGAAATACTAATAAAATGATTAAAGAAGGTGTGTCTTTAATTGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
TACTAATAAAATGATTAAAGAAGGTGTGTCTTTAATTGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
AAATGATTAAAGAAGGTGTGTCTTTAATTGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
GTGTGTCTTTAATTGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
GTGTCTTTAATTGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
GTCTTTAATTGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
TGAAGCATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
ATGATTTAAAGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
AGTAATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
ATGCAAAGCTAAAAATCAGA CCACTGCACTCCAGCCTGGG
A T G C A A A G C T A A A A A T C A G A C C A C T G C A C T C C A G C C T G G G
EML4 intron 13
ALK intron 19

Figure 9A (contd.)

```
GAACA
GAACAAGAGTGAA
GAACAAGAGTGAAACCCCATCTCAAAAA
GAACAAGAGTGAAACCCCATCTCAAAAAC
GAACAAGAGTGAAACCCCATCTCAAAAACAA
GAACAAGAGTGAAACCCCATCTCAAAAACAAA
GAACAAGAGTGAAACCCCATCTCAAAAACAAACAAACA
GAACAAGAGTGAAACCCCATCTCAAAAACAAACAAACAAAACAAAAC
GAACAAGAGTGAAACCCCATCTCAAAAACAAACAAACAAAACAAAACAAAAAAAACTAAG
```

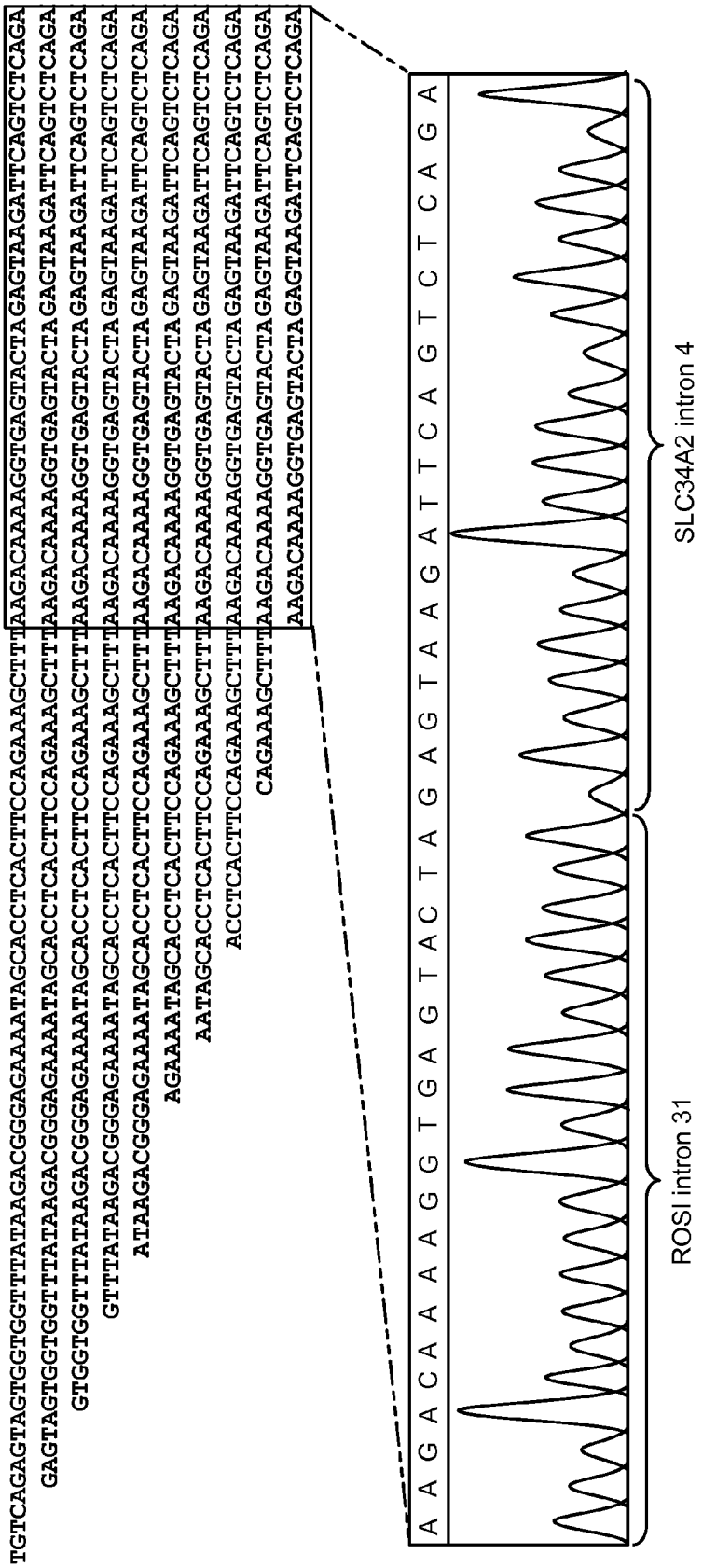

TGTCAGAGTAGTGGTGGTTTATAAGACGGGAGAAAATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
GAGTAGTGGTGGTTTATAAGACGGGAGAAAATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
GTGGTGGTTTATAAGACGGGAGAAAATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
GTTTATAAGACGGGAGAAAATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
ATAAGACGGGAGAAAATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
AGAAAATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
AATAGCACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
ACCTCACTTCCAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
CAGAAAGCTTTAAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
AAGACAAAAGGTGAGTACTAGAGTAAGATTCAGTCTCAGA
A A G A C A A A A G G T G A G T A C T A G A G T A A G A T T C A G T C T C A G A
SLC34A2 intron 4
ROSI intron 31

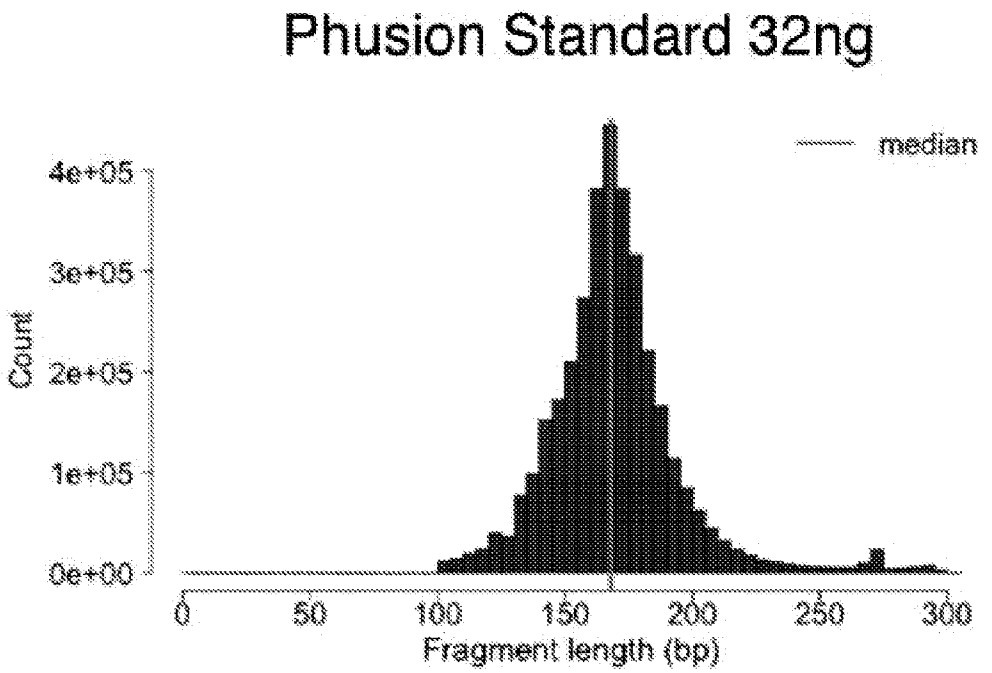
Figure 9B

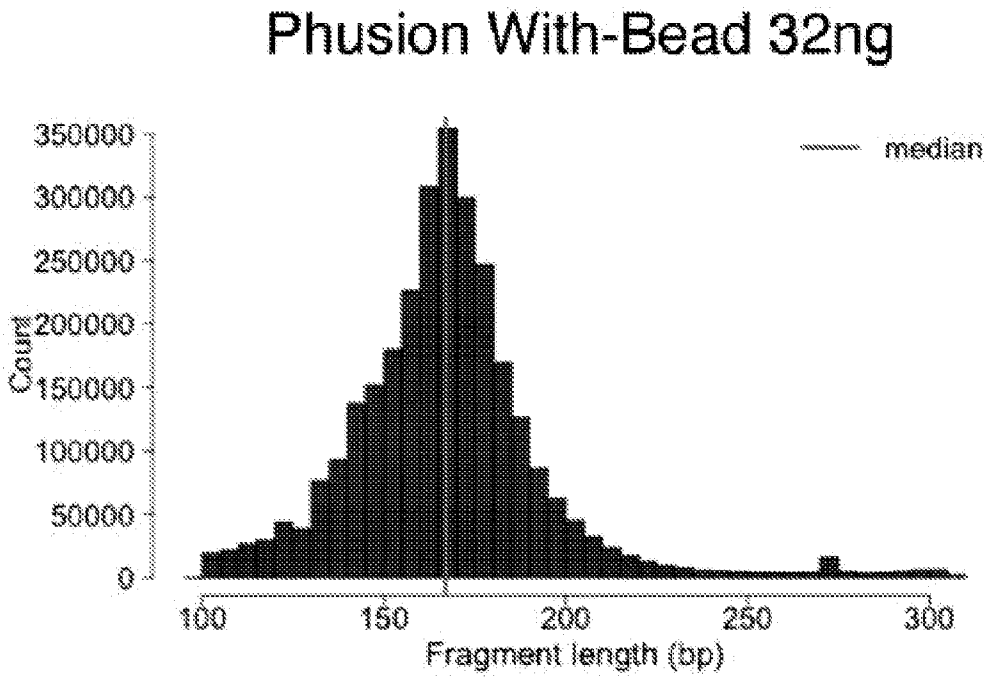
TCTGG
TCTGGGTGACAC
TCTGGGTGACACAAAG
TCTGGGTGACACAAAGGACC
TCTGGGTGACACAAAGGACCATGGATTTCT
TCTGGGTGACACAAAGGACCATGGATTTCTGCAA
TCTGGGTGACACAAAGGACCATGGATTTCTGCAACCCTTGGTG
TCTGGGTGACACAAAGGACCATGGATTTCTGCAACCCTTGGTGCCTTTCT
TCTGGGTGACACAAAGGACCATGGATTTCTGCAACCCTTGGTGCCTTTCTTGGGAACCCAT

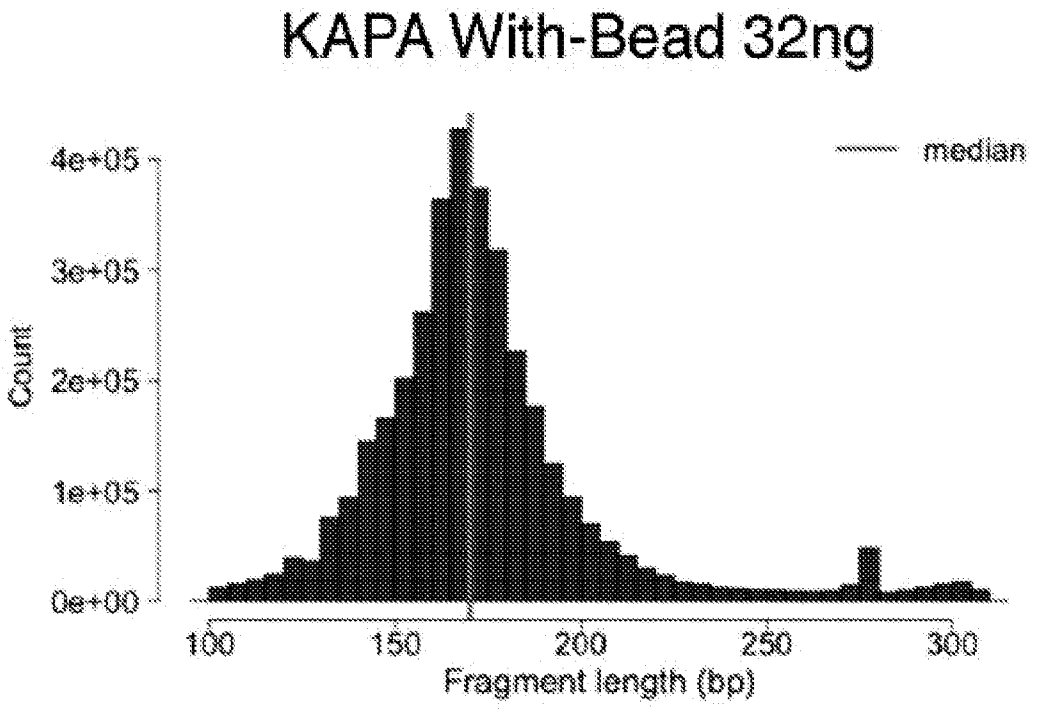
Figure 9B

(Contnued)

Figure 10A


Phusion Standard 32ng
median
4e+05
3e+05
2e+05
1e+05
0e+00
Count
0
50
100
150
200
250
300
Fragment length (bp)


Phusion With-Bead 32ng
median
350000
300000
250000
200000
150000
100000
50000
0
Count
100
150
200
250
300
Fragment length (bp)



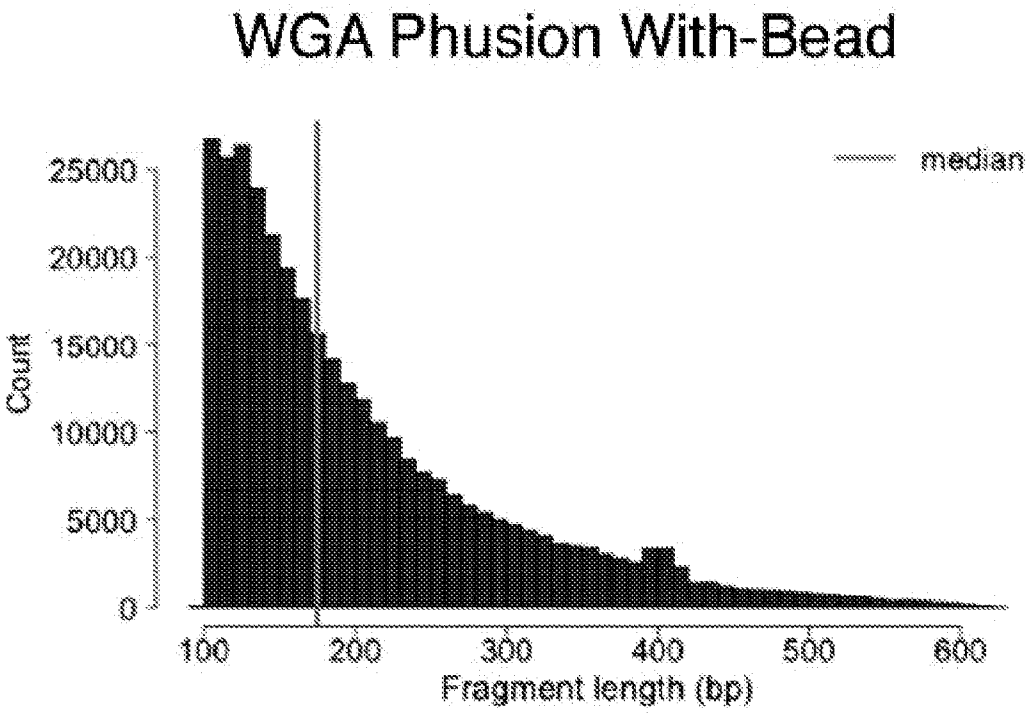
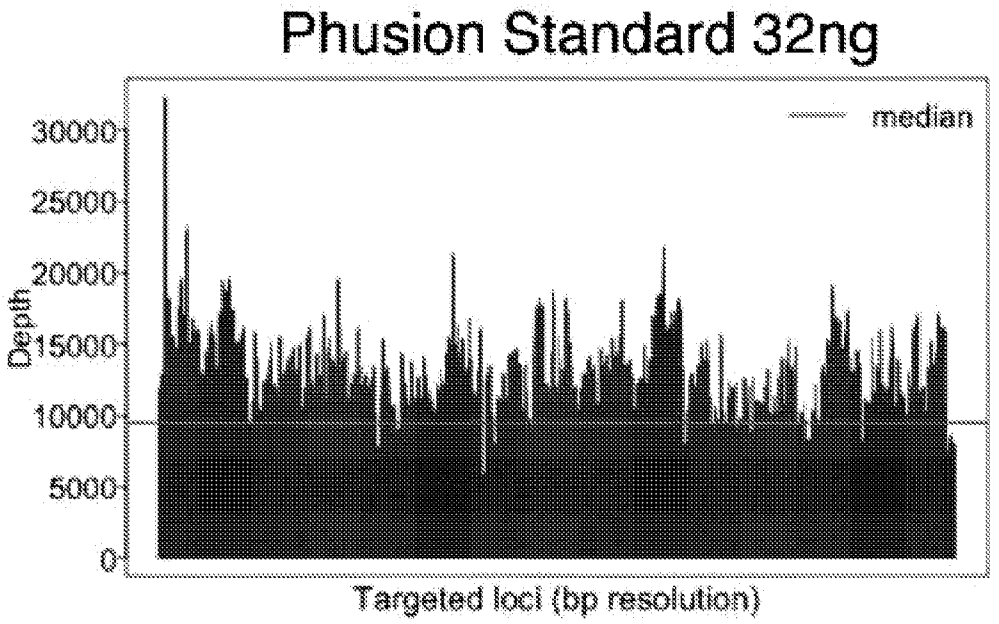
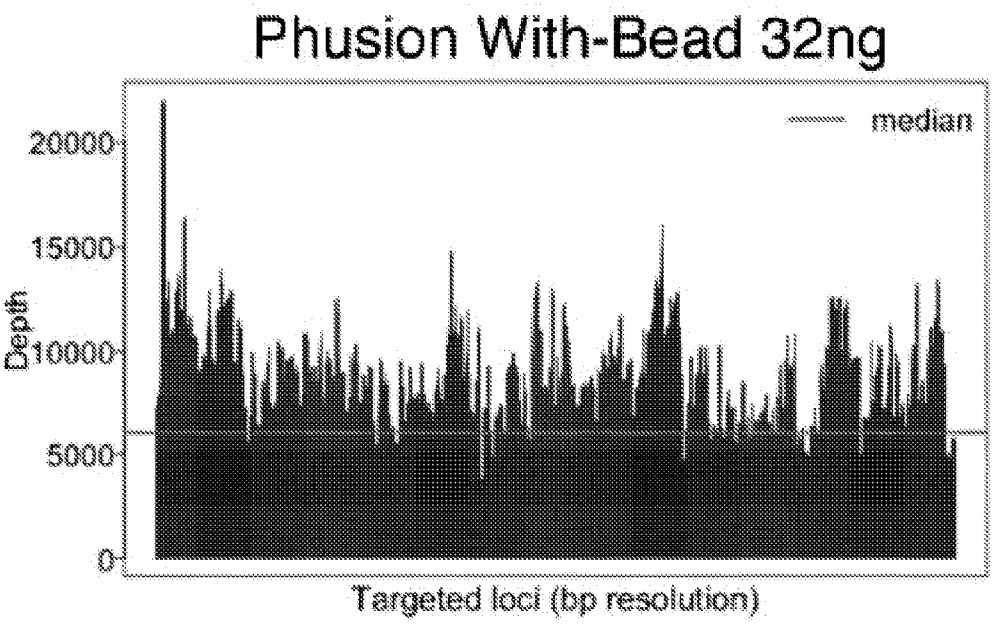
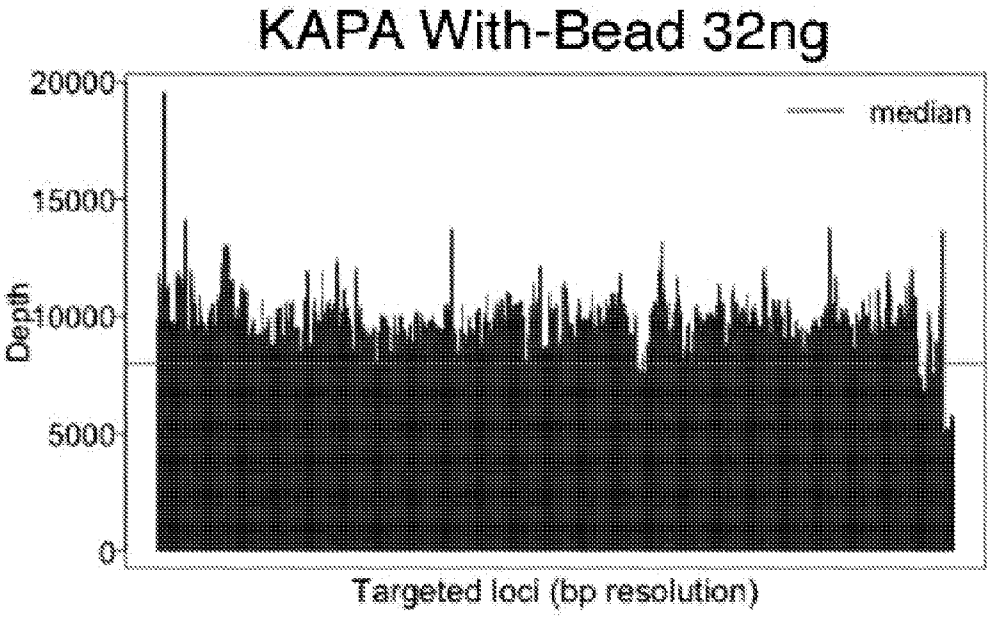


Figure 10A
(Continued)

Figure 10B


Phusion Standard 32ng
median
30000
25000
20000
15000
10000
5000
0
Depth
Targeted loci (bp resolution)


Phusion With-Bead 32ng
median
20000
15000
10000
5000
0
Depth
Targeted loci (bp resolution)



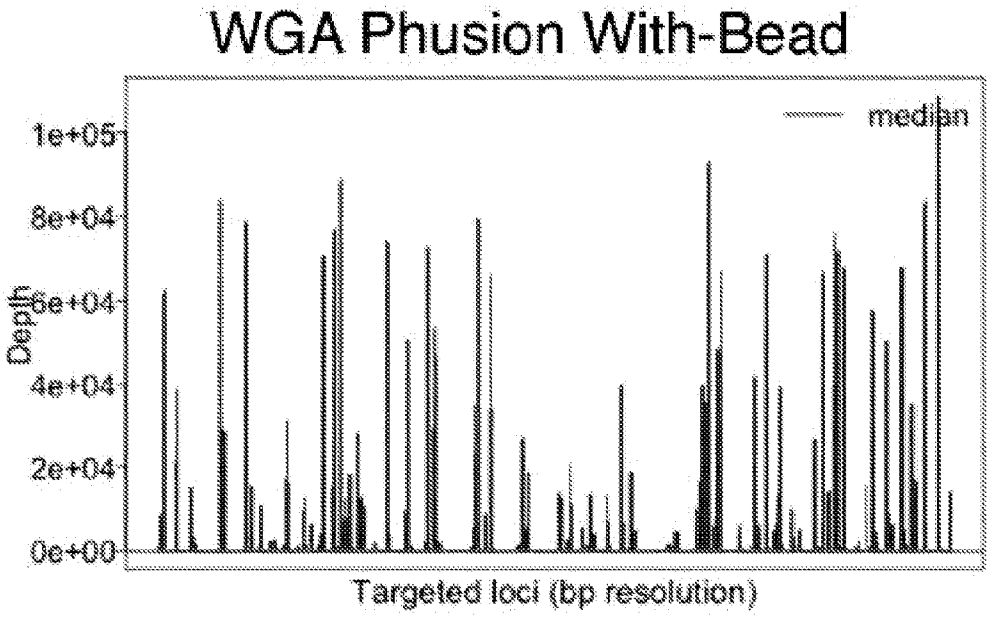


Figure 10B
(Continued)

Figure 10C

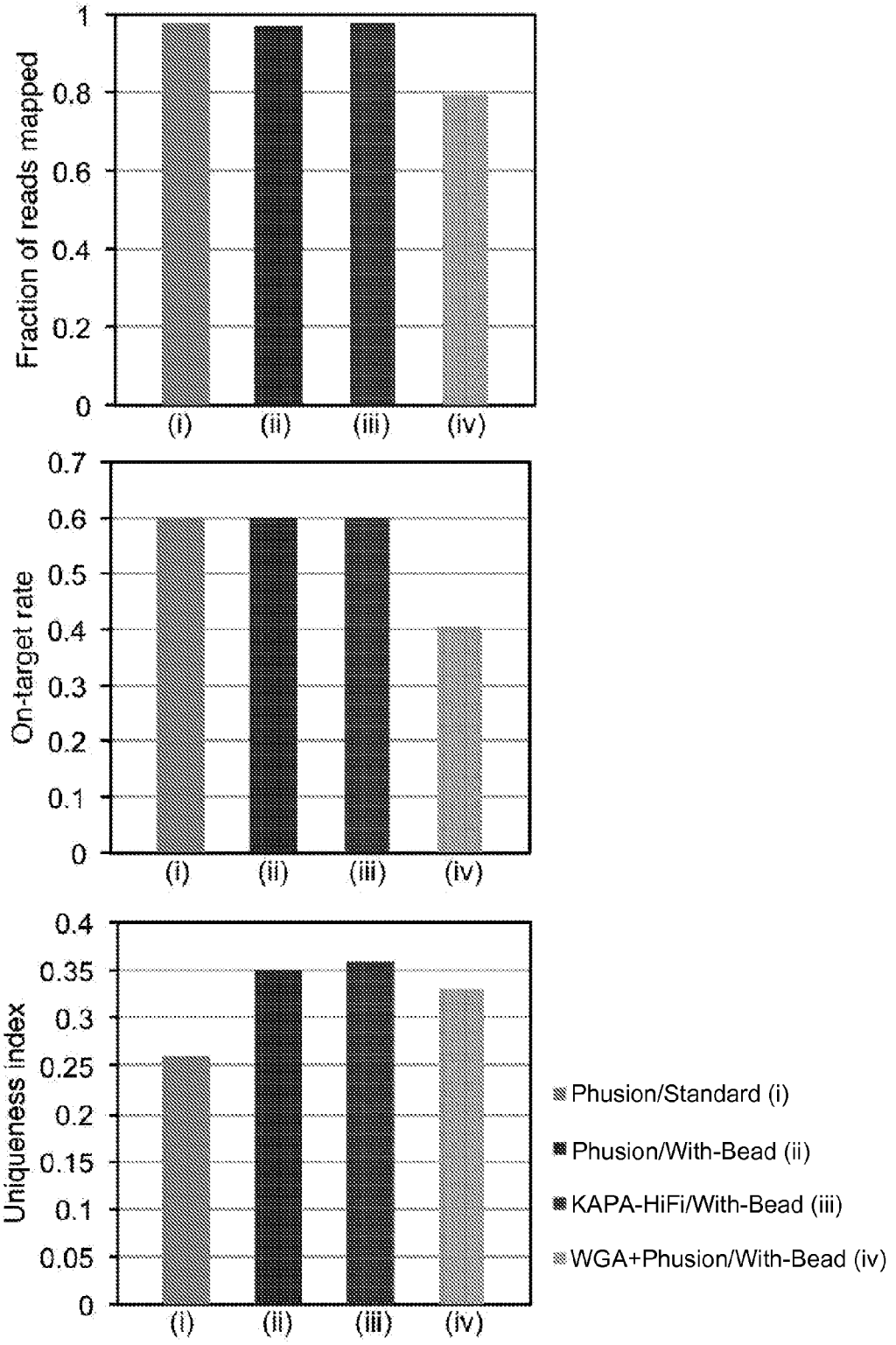

Fraction of reads mapped
1
0.8
0.6
0.4
0.2
0
(i)
(ii)
(iii)
(iv)
On-target rate
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
(i)
(ii)
(iii)
(iv)
Uniqueness index
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
(i)
(ii)
(iii)
(iv)
Phusion/Standard (i)
Phusion/With-Bead (ii)
KAPA-HiFi/With-Bead (iii)
WGA+Phusion/With-Bead (iv)

Figure 11A

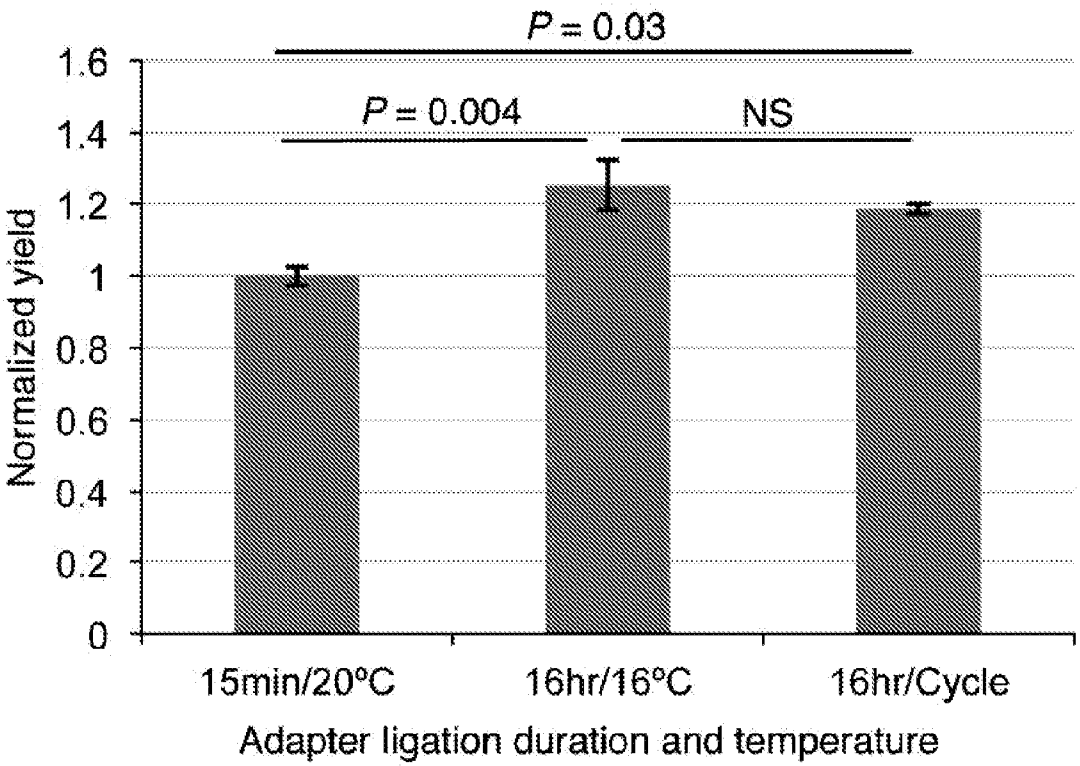


Figure 11B

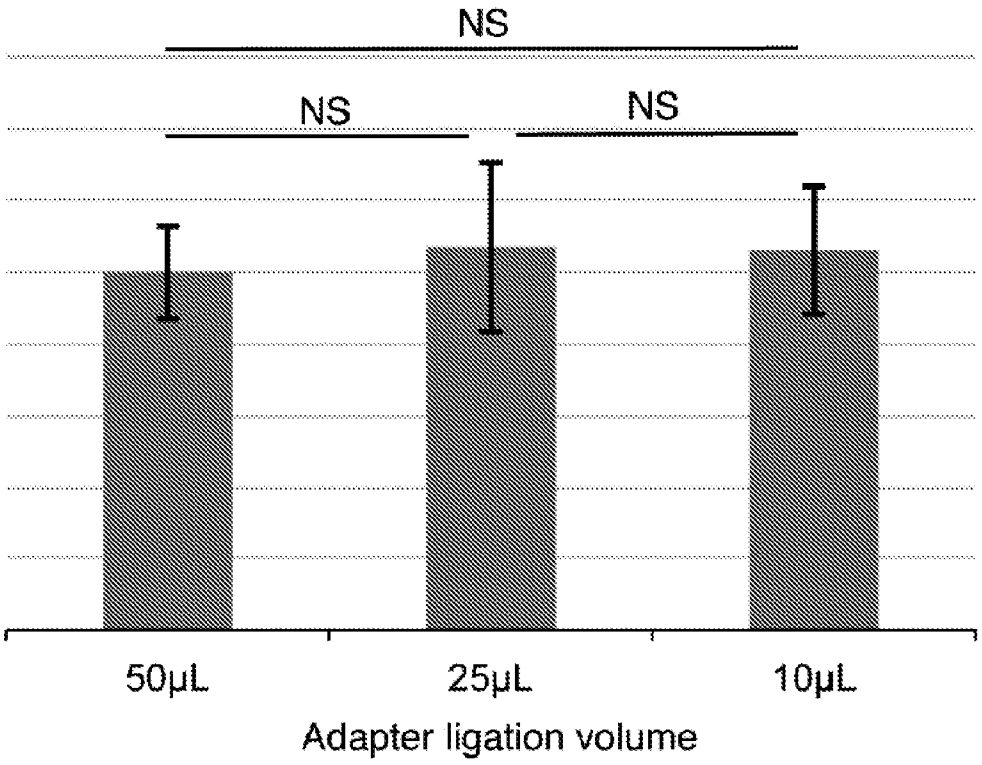


Figure 11C

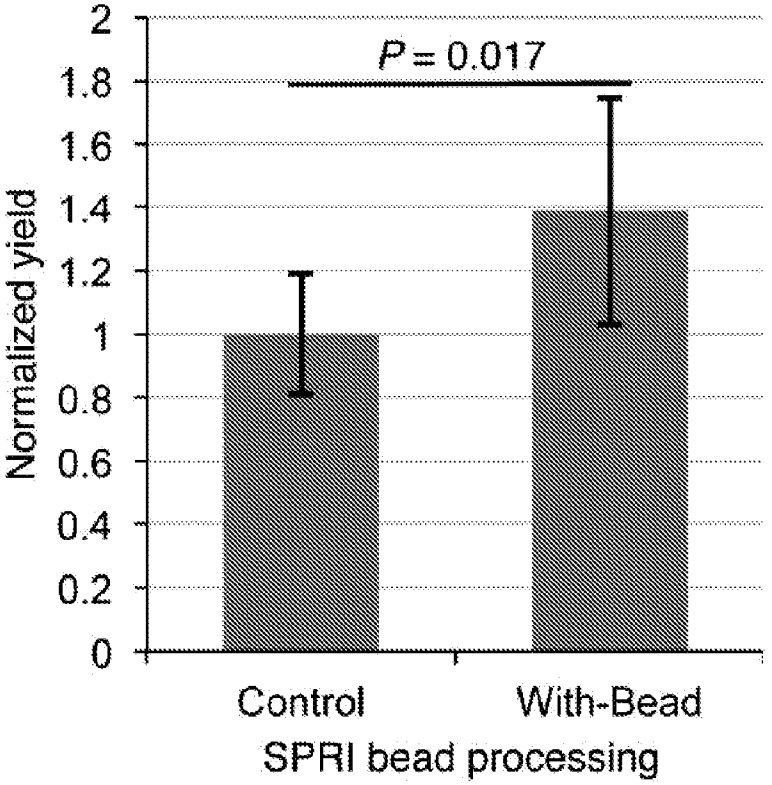


Figure 11D

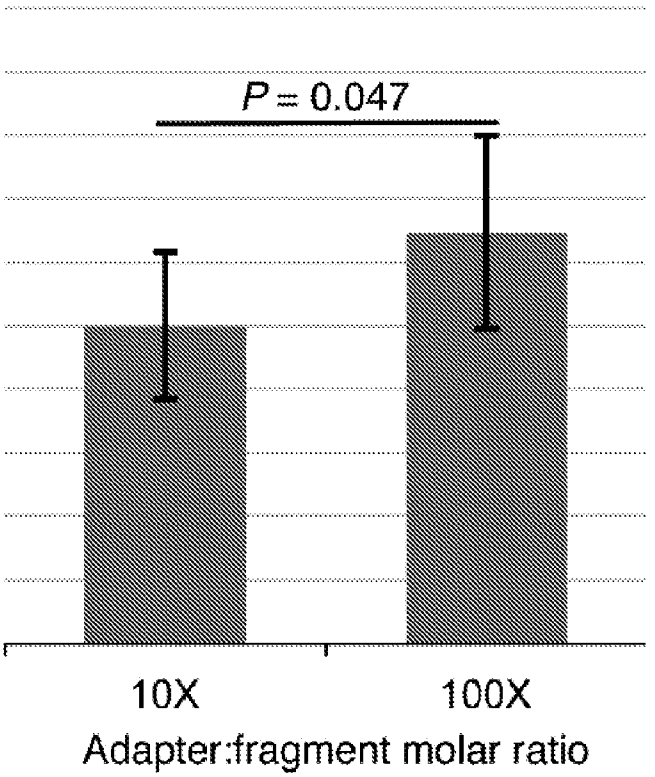


Figure 11E

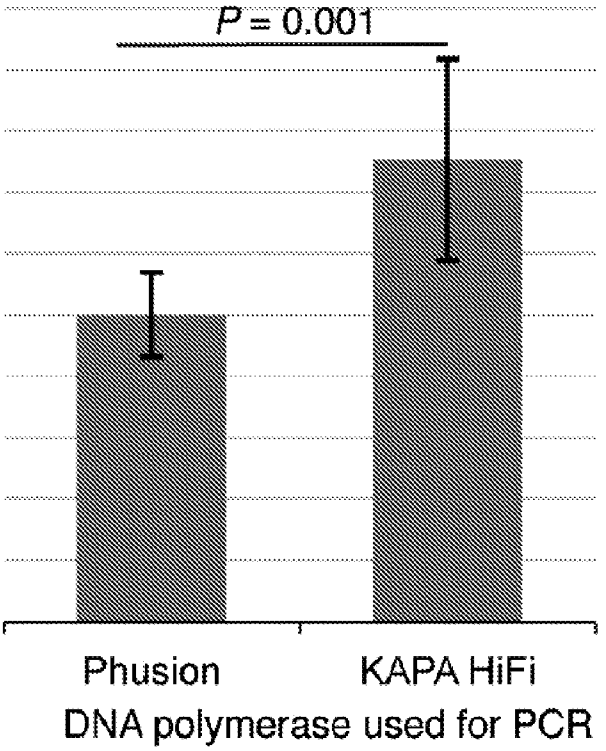


Figure 11F

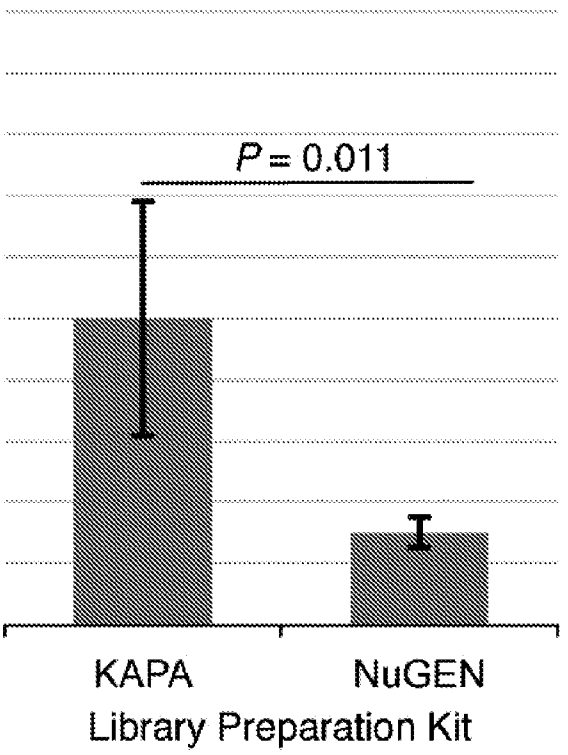
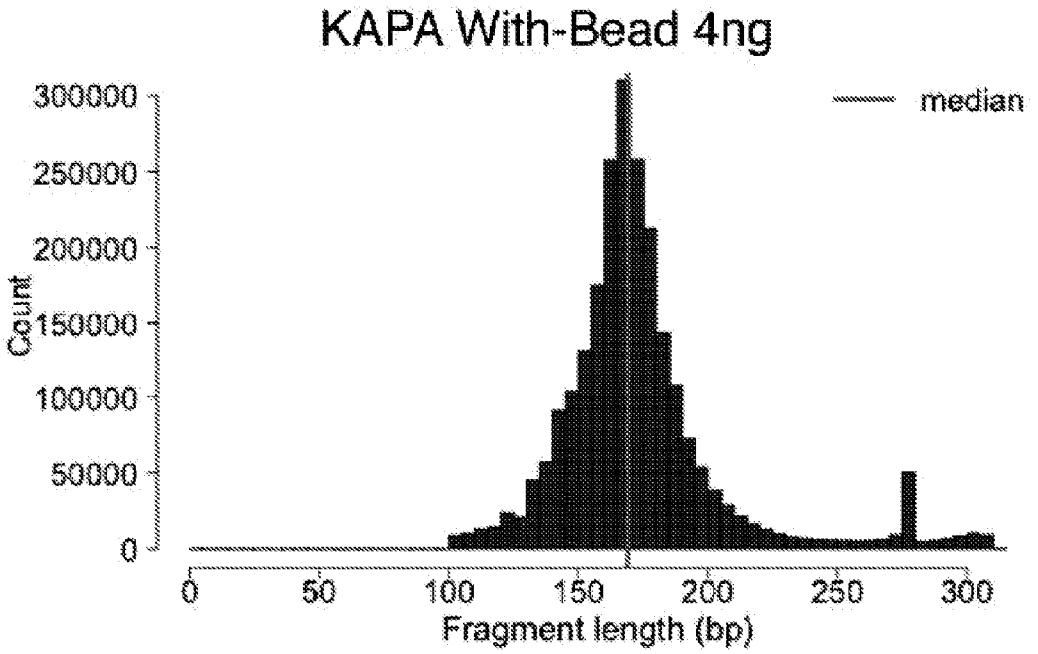
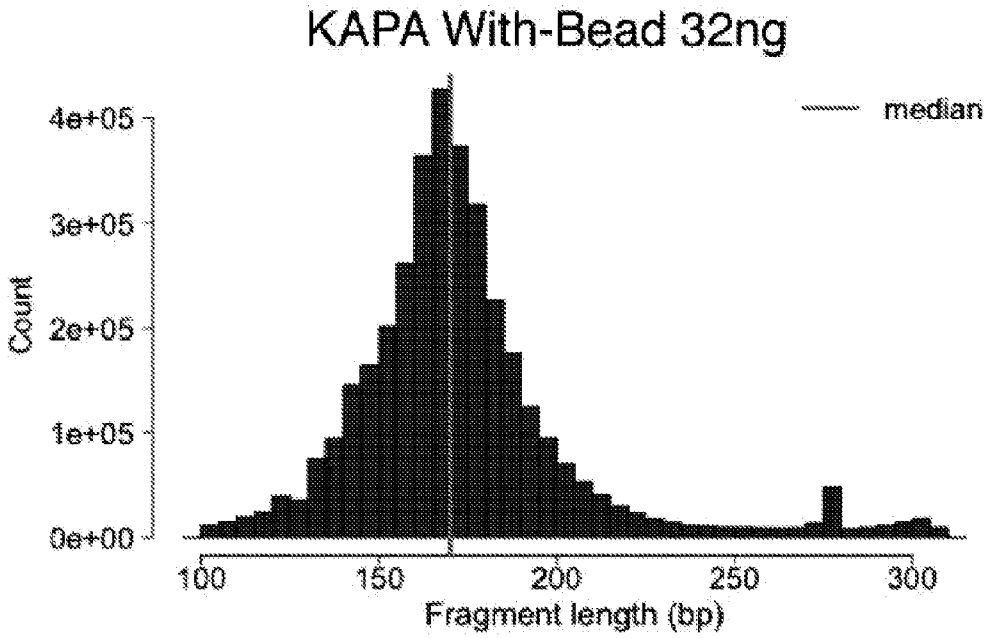


Figure 12A


KAPA With-Bead 4ng
median
Count
300000
250000
200000
150000
100000
50000
0
0
50
100
150
200
250
300
Fragment length (bp)


KAPA With-Bead 32ng
median
Count
4e+05
3e+05
2e+05
1e+05
0e+00
100
150
200
250
300
Fragment length (bp)

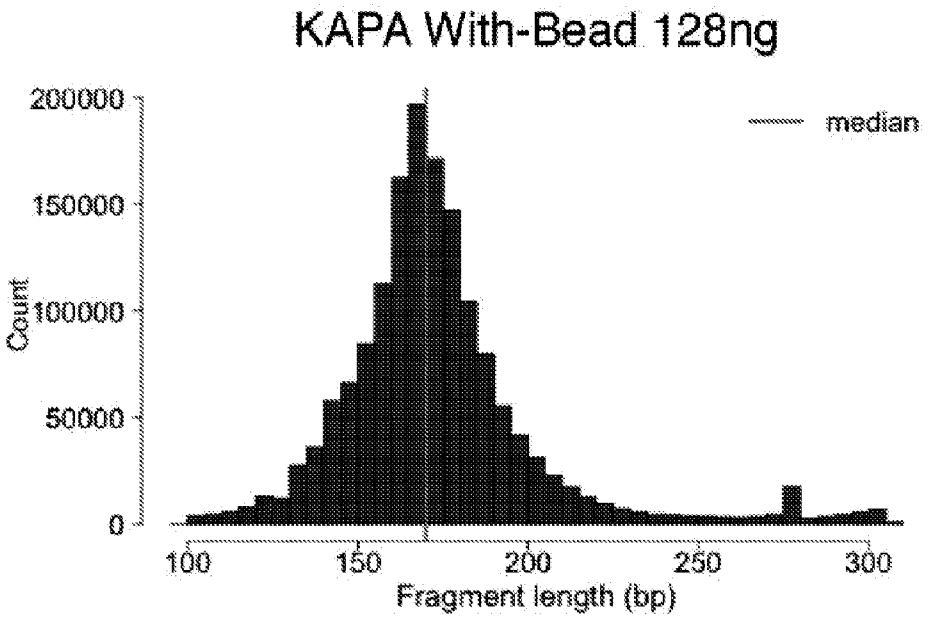
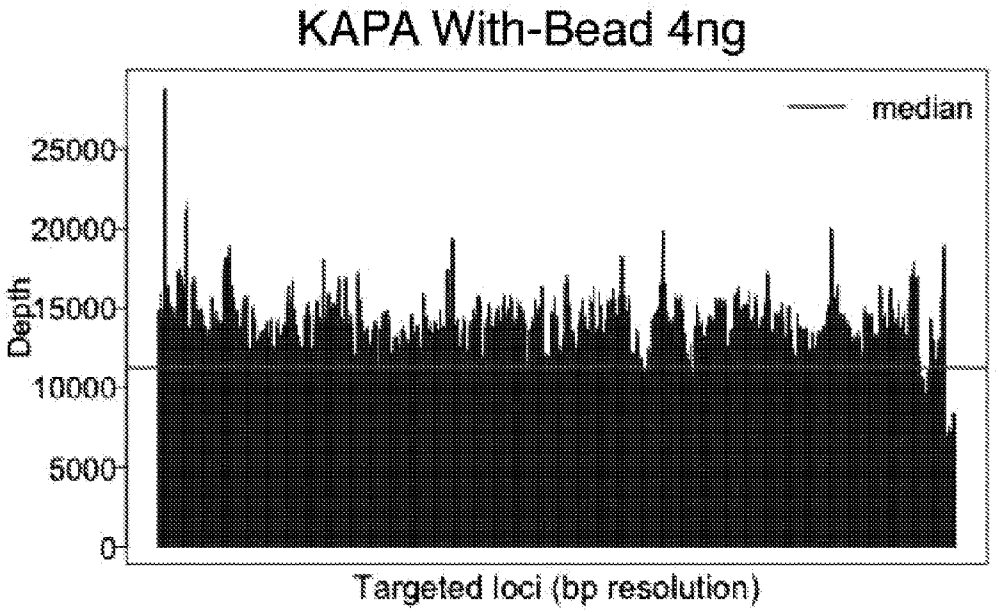
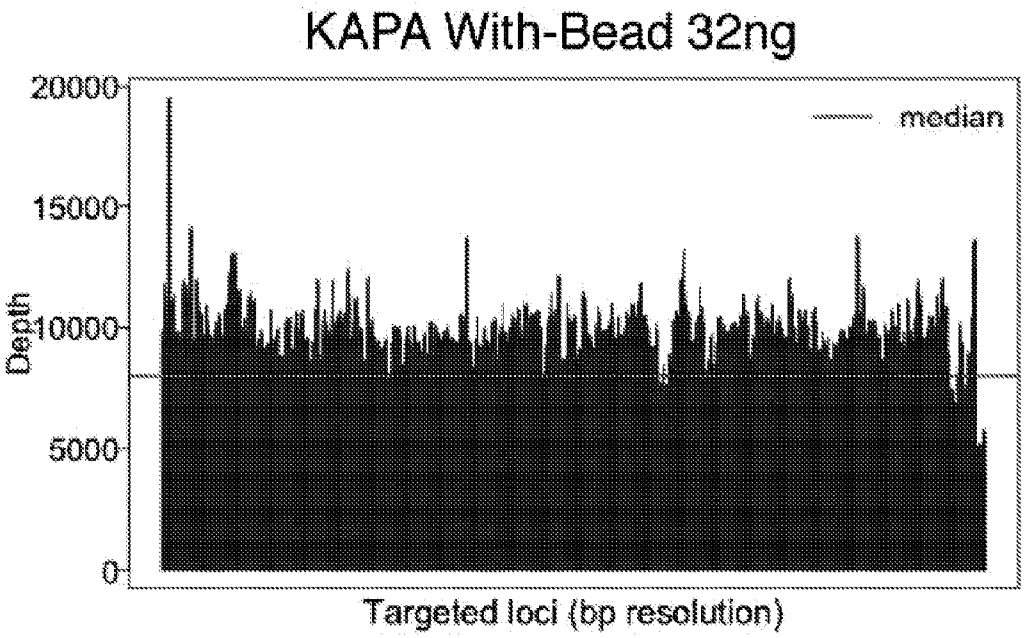


Figure 12A
(Continued)

Figure 12B





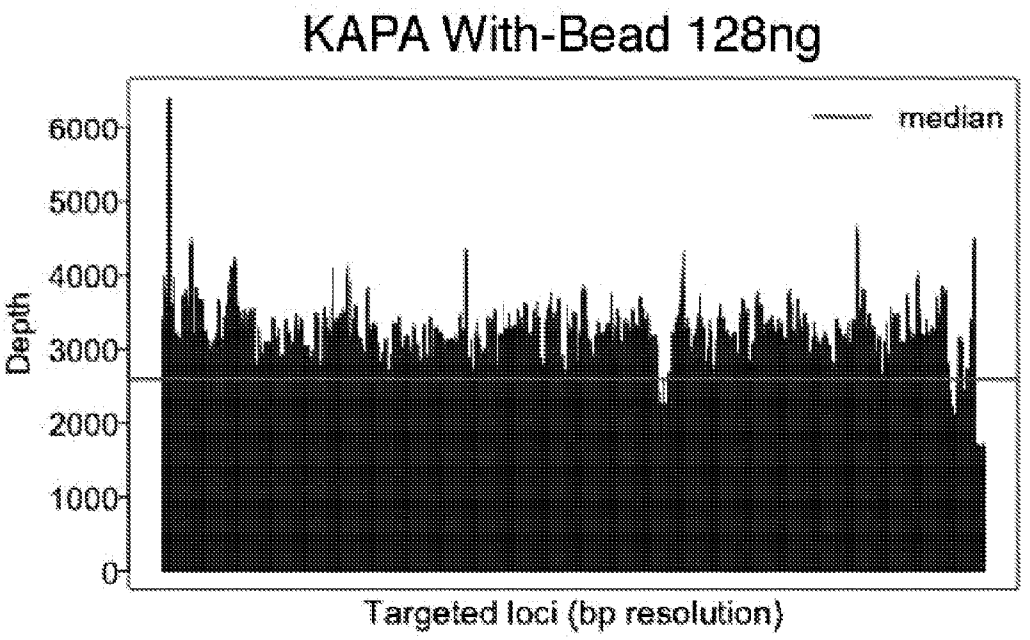


Figure 12B
(Continued)

Figure 12C

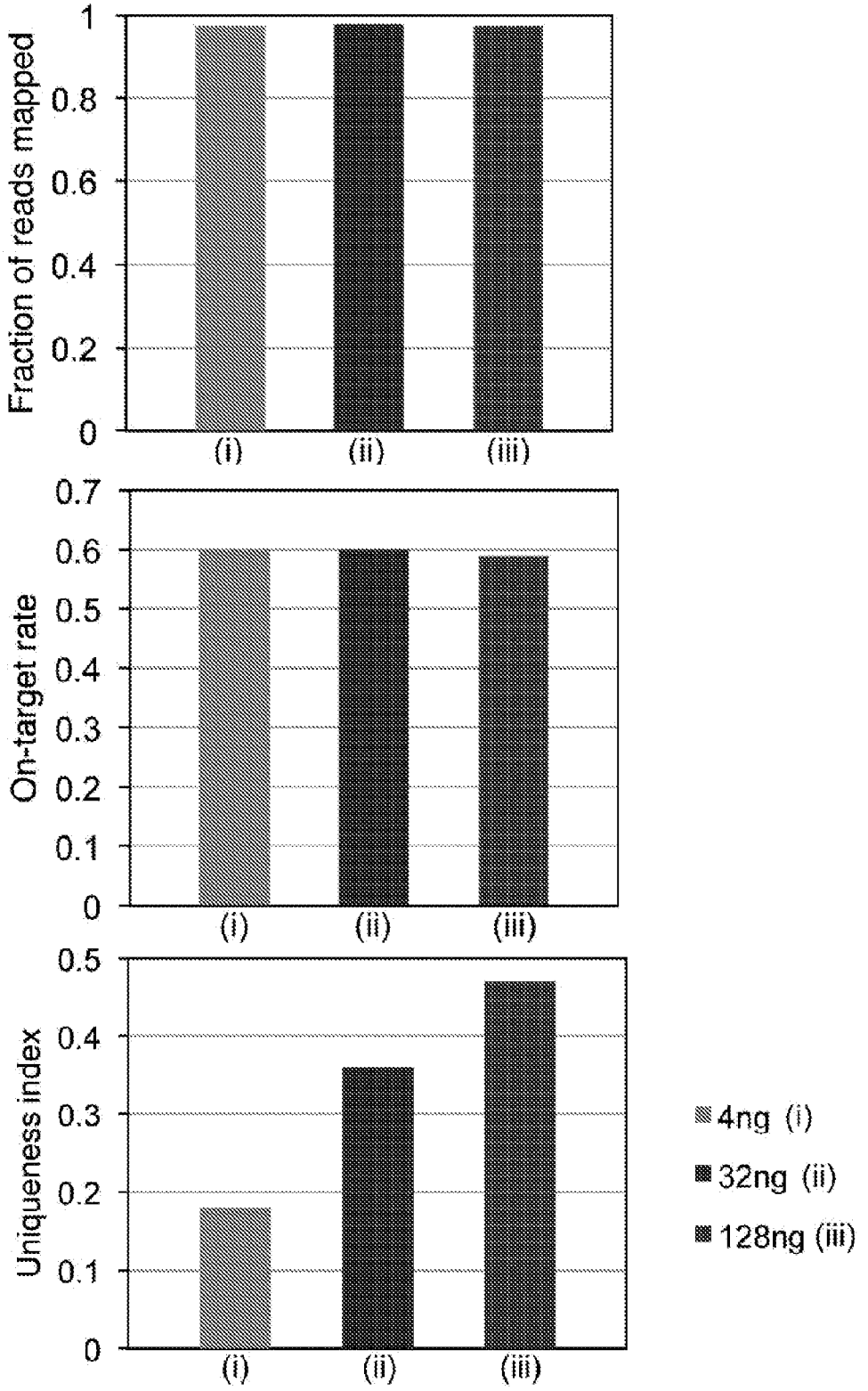

Fraction of reads mapped
1
0.8
0.6
0.4
0.2
0
(i)
(ii)
(iii)
On-target rate
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
(i)
(ii)
(iii)
Uniqueness index
0.5
0.4
0.3
0.2
0.1
0
(i)
(ii)
(iii)
4ng (i)
32ng (ii)
128ng (iii)

Figure 13A

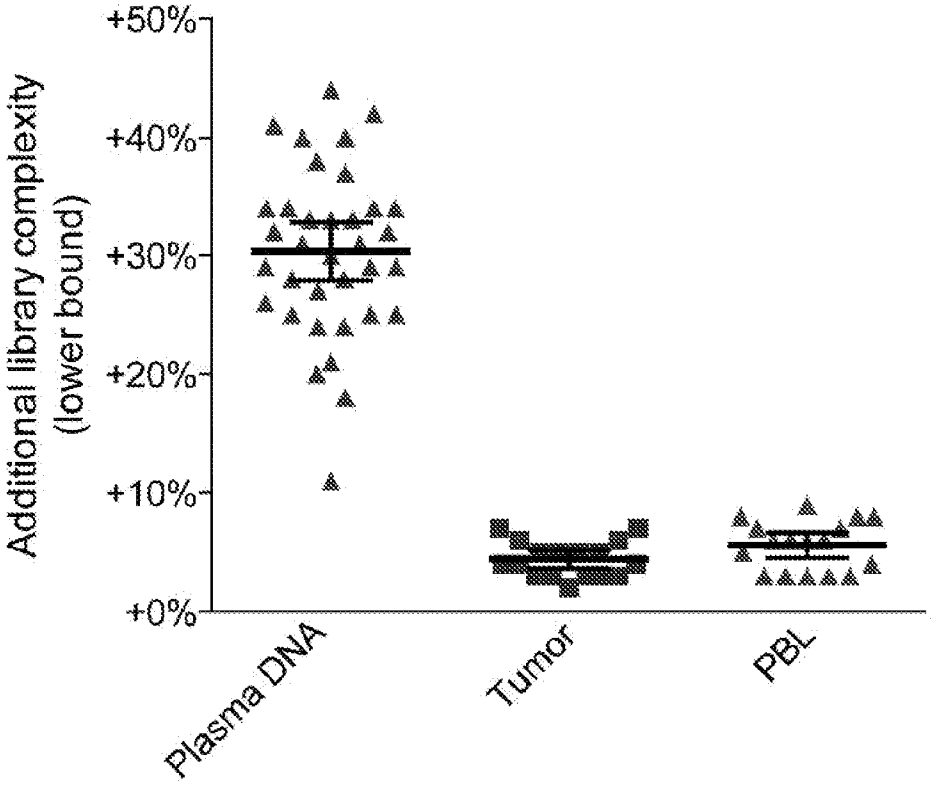
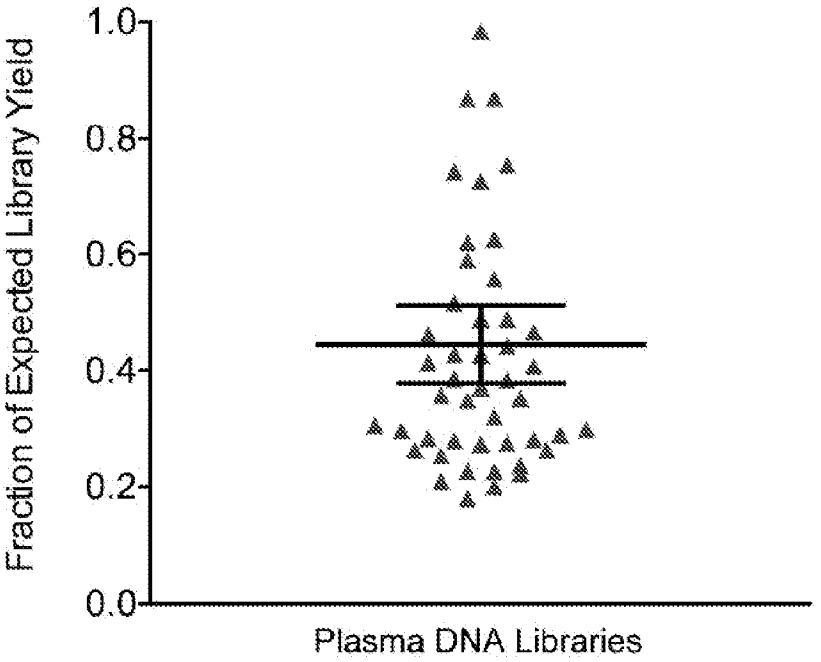


Figure 13B



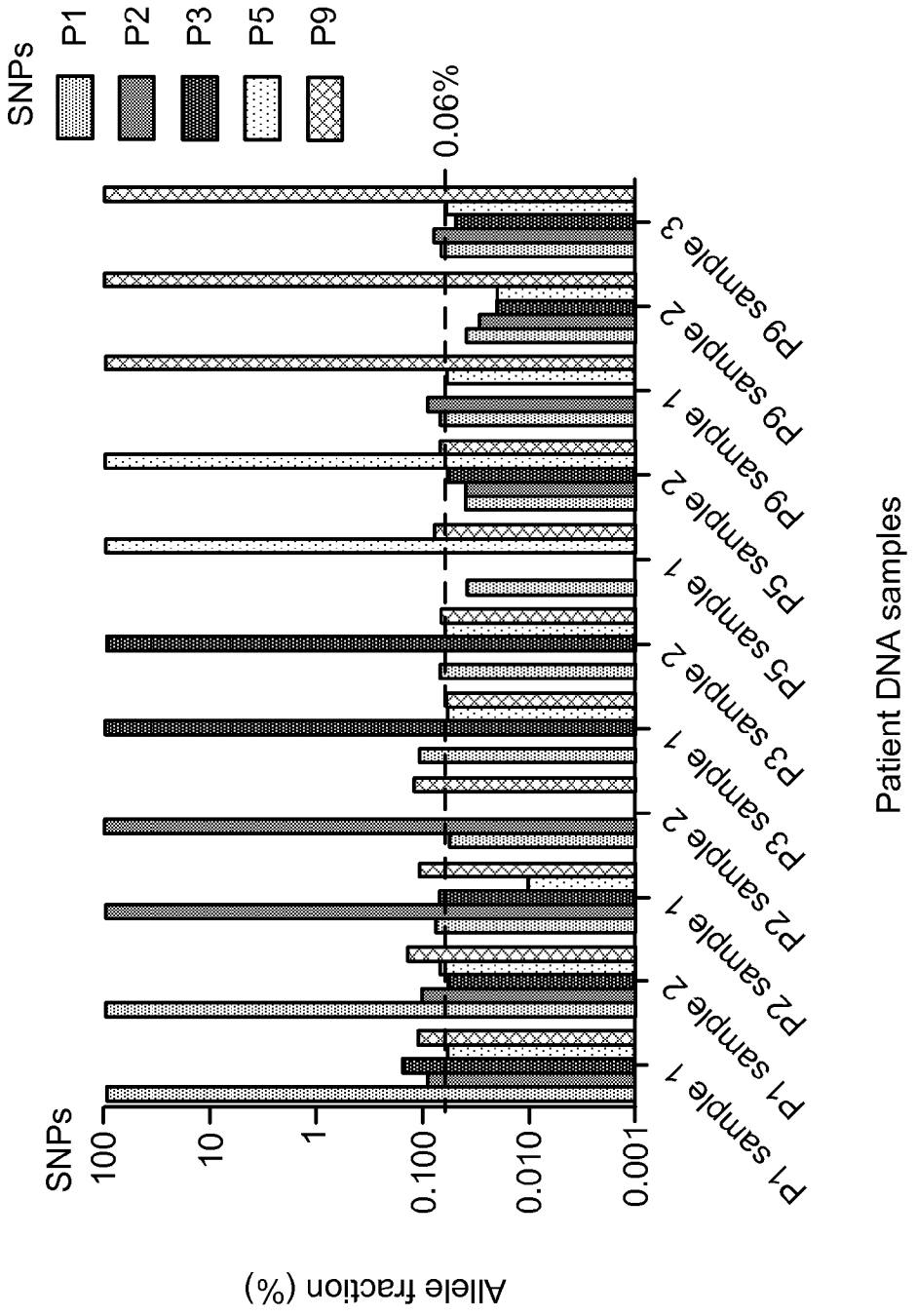

SNPs
P1
P2
P3
P5
P9
0.06%
SNPs
100
10
1
0.100
0.010
0.001
Allele fraction (%)
P1 sample 1
P1 sample 2
P2 sample 1
P2 sample 2
P3 sample 1
P3 sample 2
P5 sample 1
P5 sample 2
P9 sample 1
P9 sample 2
P9 sample 3
Patient DNA samples

Figure 14

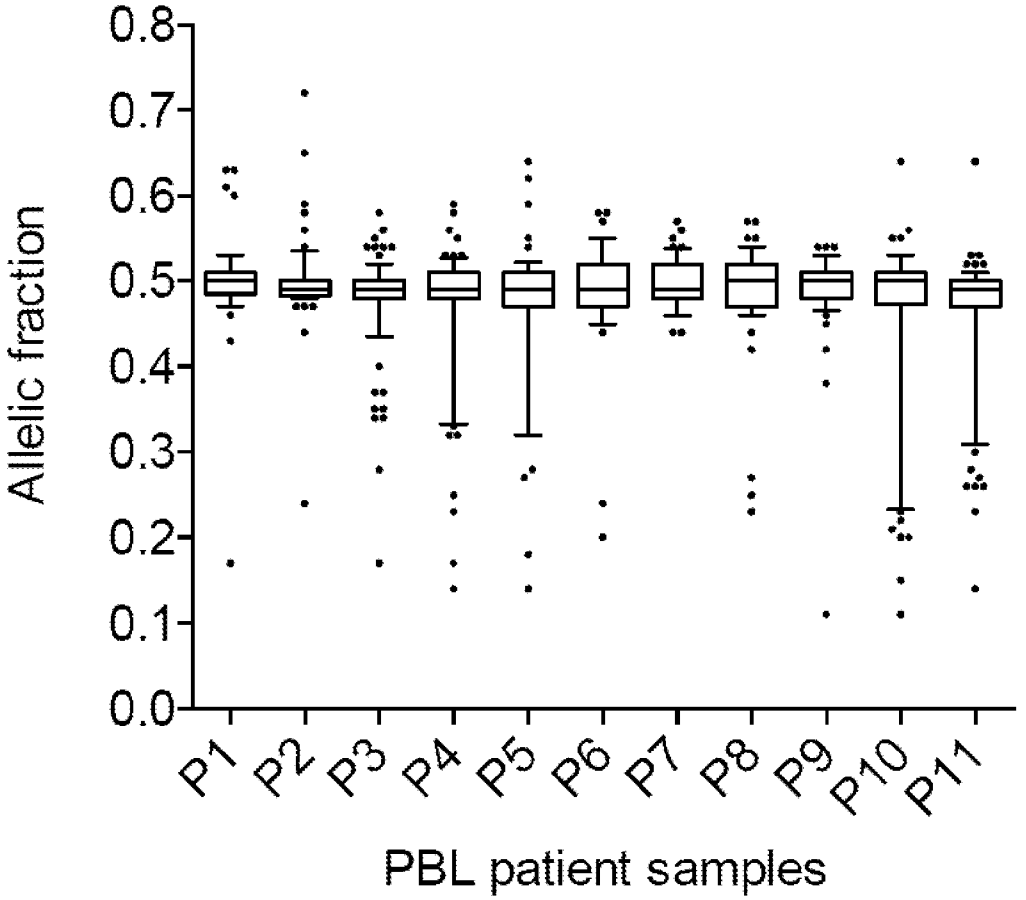

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Allelic fraction
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
PBL patient samples

Figure 15

Figure 16A

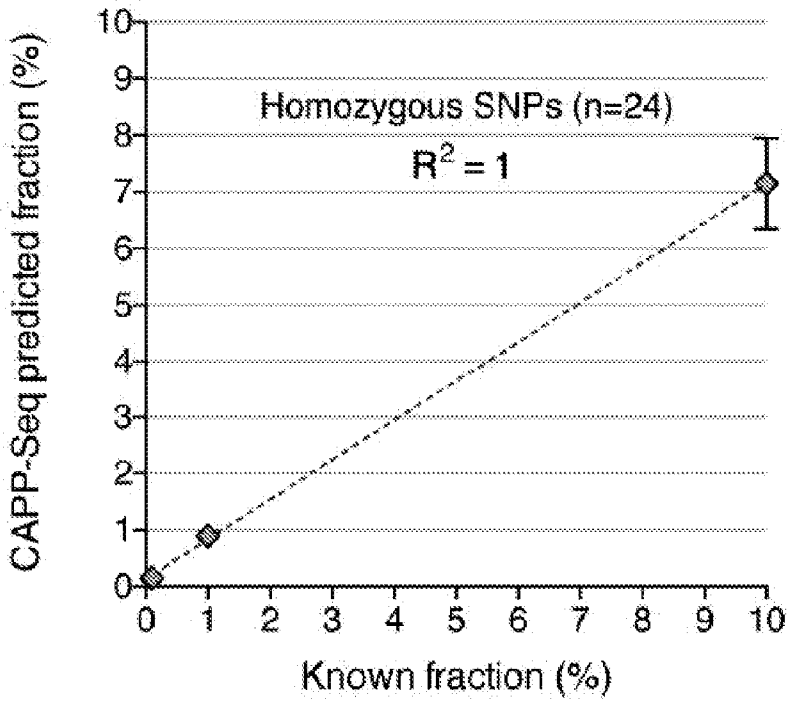


Figure 16B

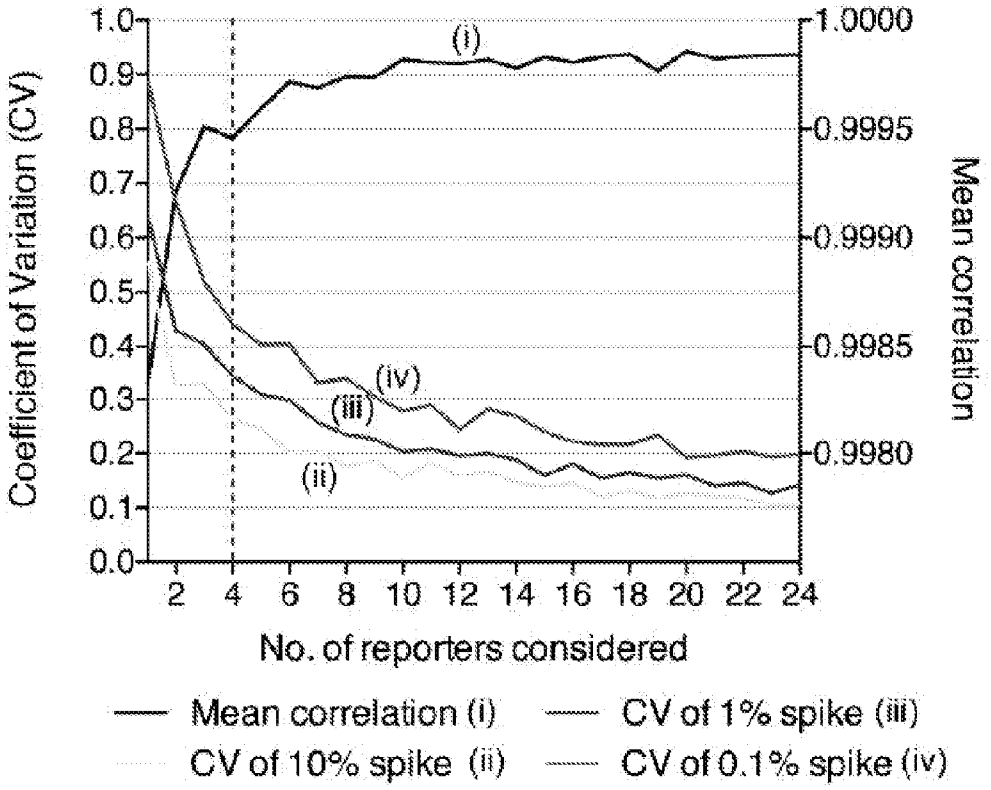


Figure 16C

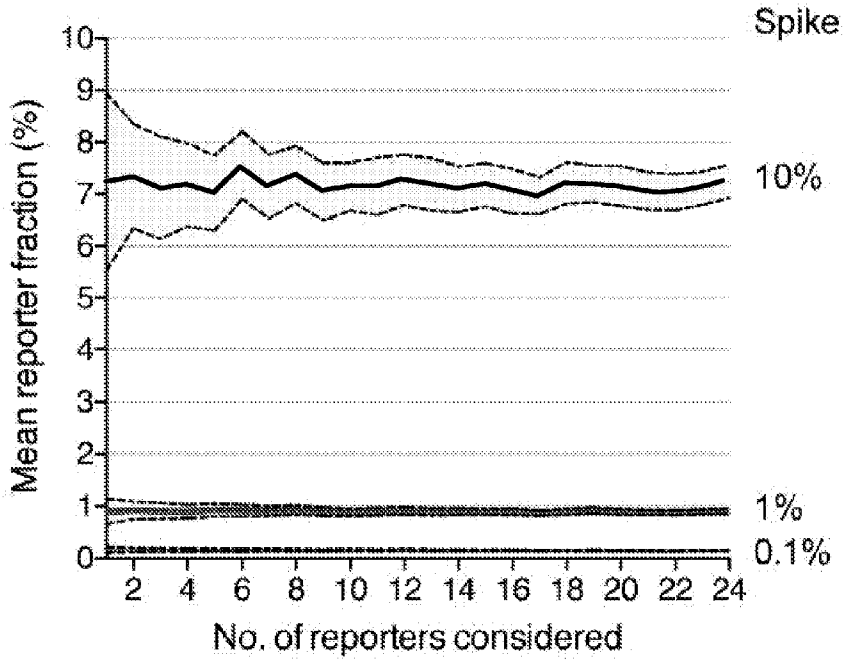


Figure 16D

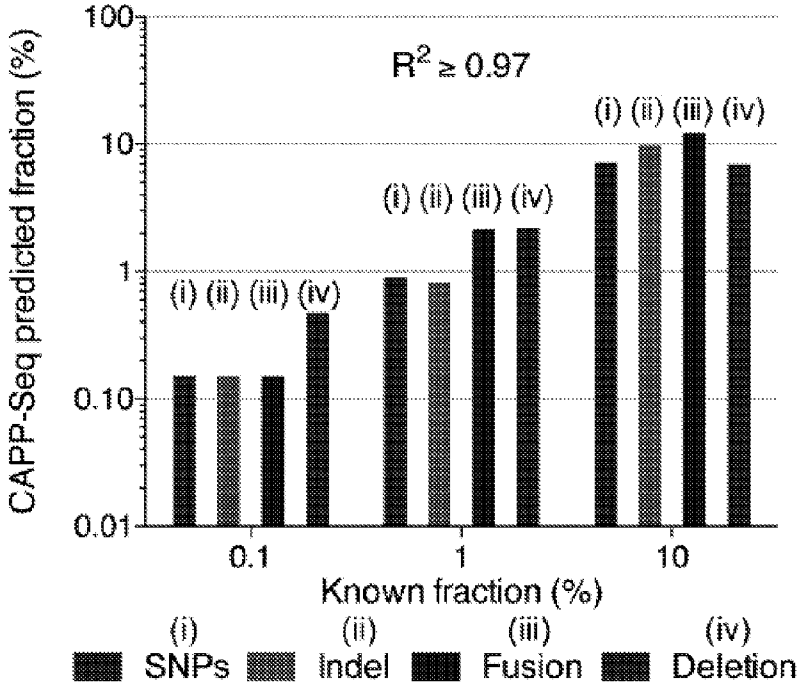
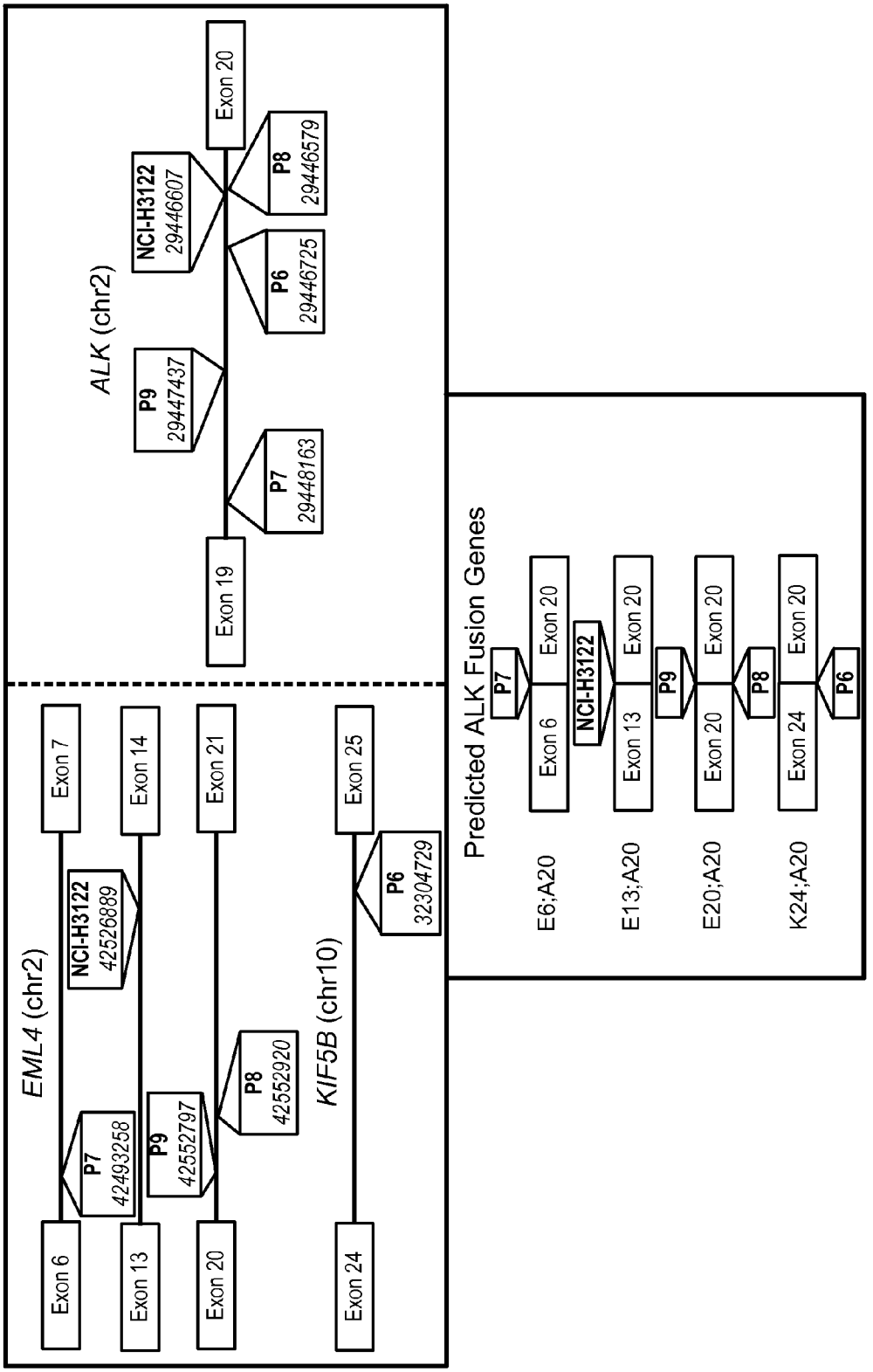


Figure 17A


ALK (chr2)
Exon 19
P7
29448163
P9
29447437
NCI-H3122
29446607
P6
29446725
P8
29446579
Exon 20
EML4 (chr2)
Exon 6
P7
42493258
Exon 7
NCI-H3122
42526889
Exon 13
P9
42552797
Exon 14
Exon 20
P8
42552920
Exon 21
KIF5B (chr10)
Exon 24
P6
32304729
Exon 25
Predicted ALK Fusion Genes
E6;A20
P7
Exon 6
Exon 20
E13;A20
NCI-H3122
Exon 13
Exon 20
E20;A20
P9
Exon 20
Exon 20
K24;A20
P8
Exon 24
Exon 20
P6

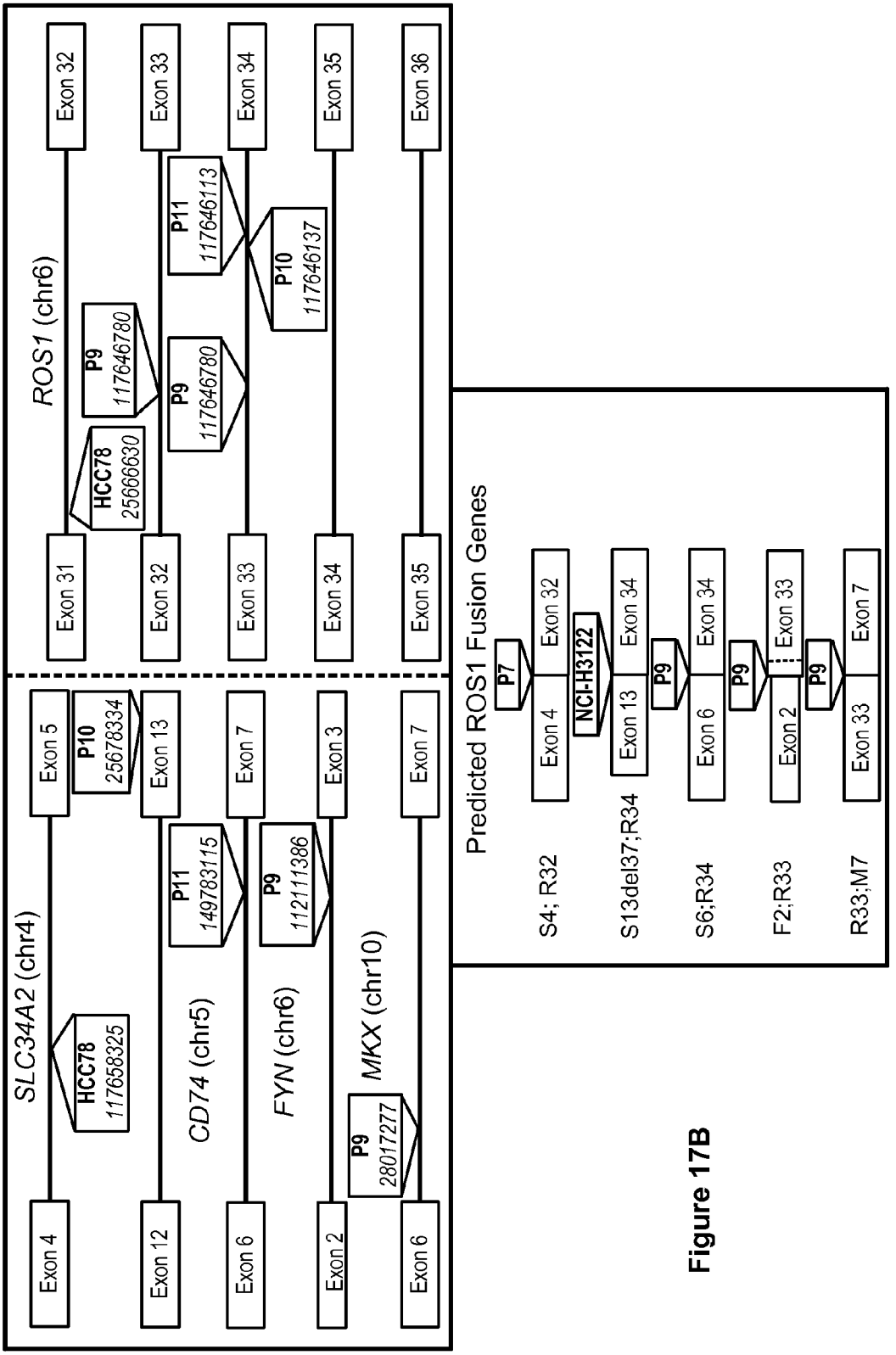

ROS1 (chr6)
Exon 31
Exon 32
Exon 33
Exon 34
Exon 35
Exon 36
HCC78
25666630
P9
117646780
P9
117646780
P11
117646113
P10
117646137
SLC34A2 (chr4)
Exon 4
Exon 5
HCC78
117658325
P10
25678334
CD74 (chr5)
Exon 12
Exon 13
P11
149783115
FYN (chr6)
Exon 6
Exon 7
P9
112111386
MKX (chr10)
Exon 2
Exon 3
P9
28017277
Exon 6
Exon 7
Predicted ROS1 Fusion Genes
S4; R32
Exon 4
P7
Exon 32
S13del37;R34
Exon 13
NCI-H3122
Exon 34
S6;R34
Exon 6
P9
Exon 34
F2;R33
Exon 2
P9
Exon 33
R33;M7
Exon 33
P9
Exon 7

Figure 17B

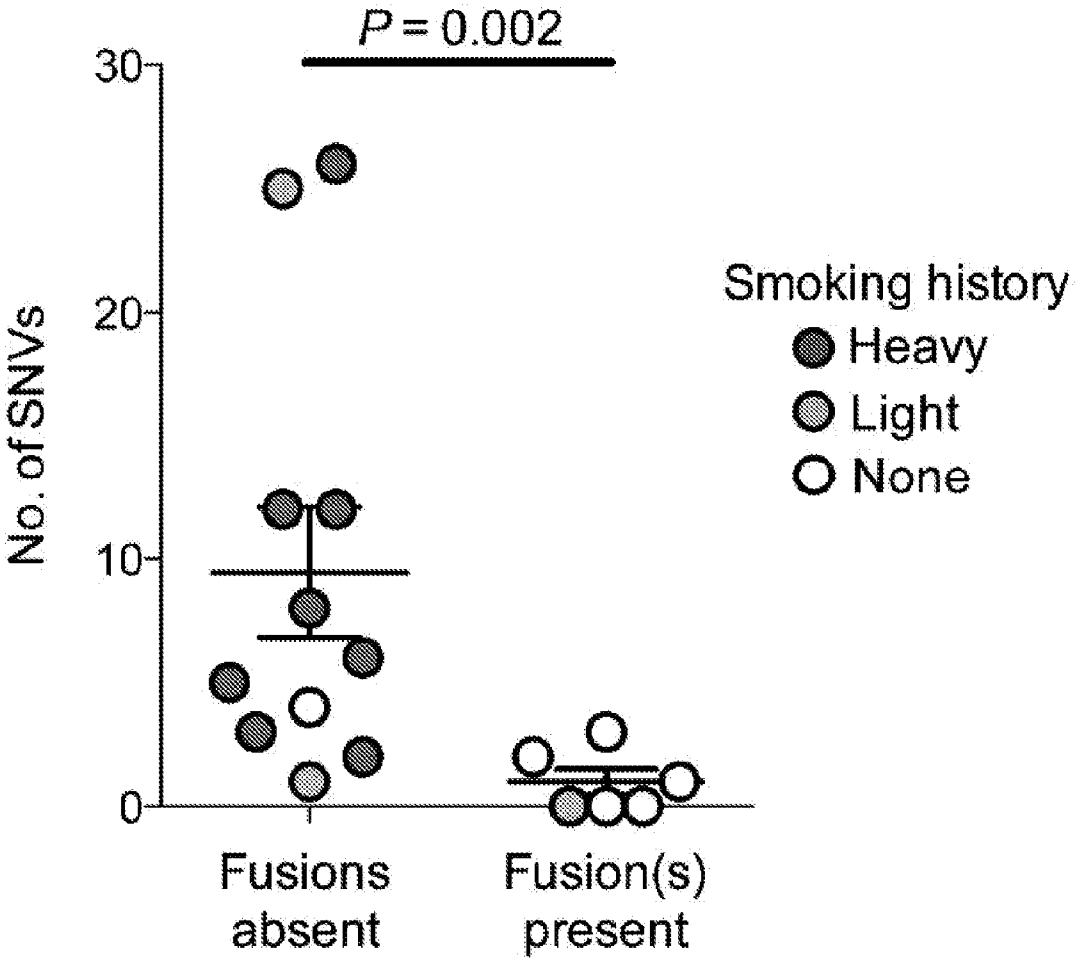

P = 0.002
30
20
10
0
No. of SNVs
Smoking history
Heavy
Light
None
Fusions absent
Fusion(s) present

Figure 18

Figure 19A

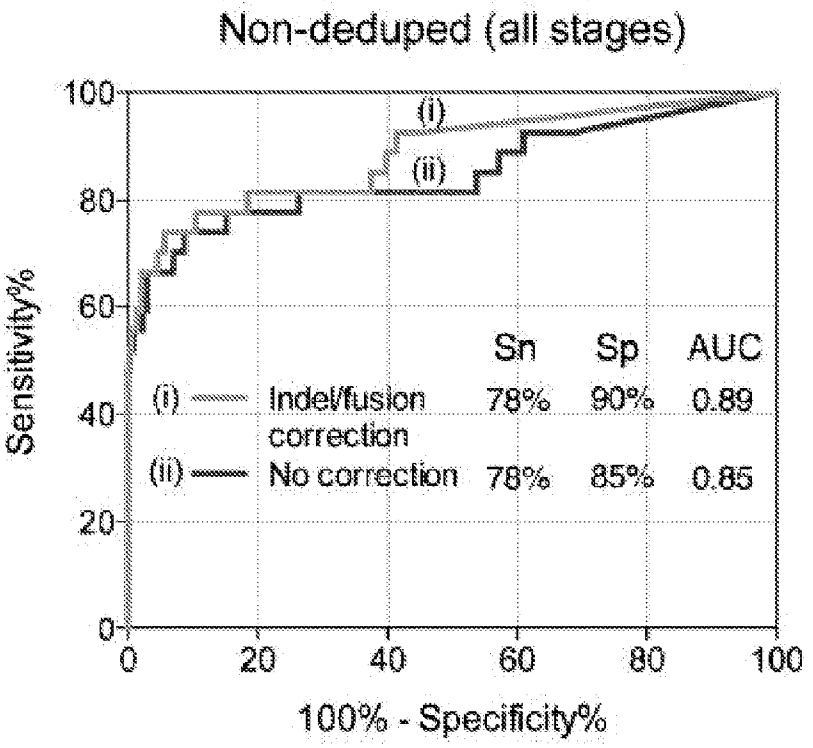


Figure 19B

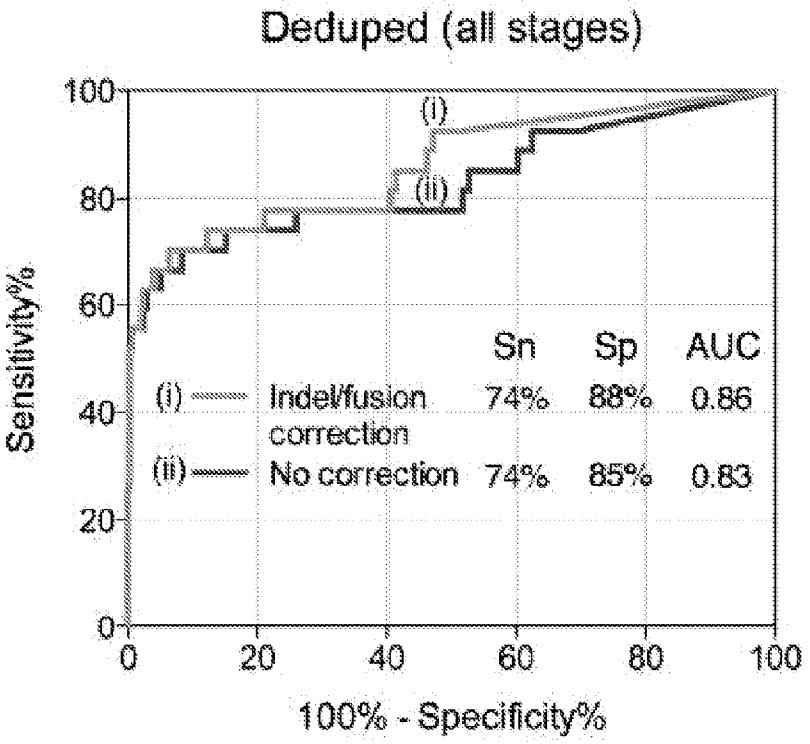


Figure 19C

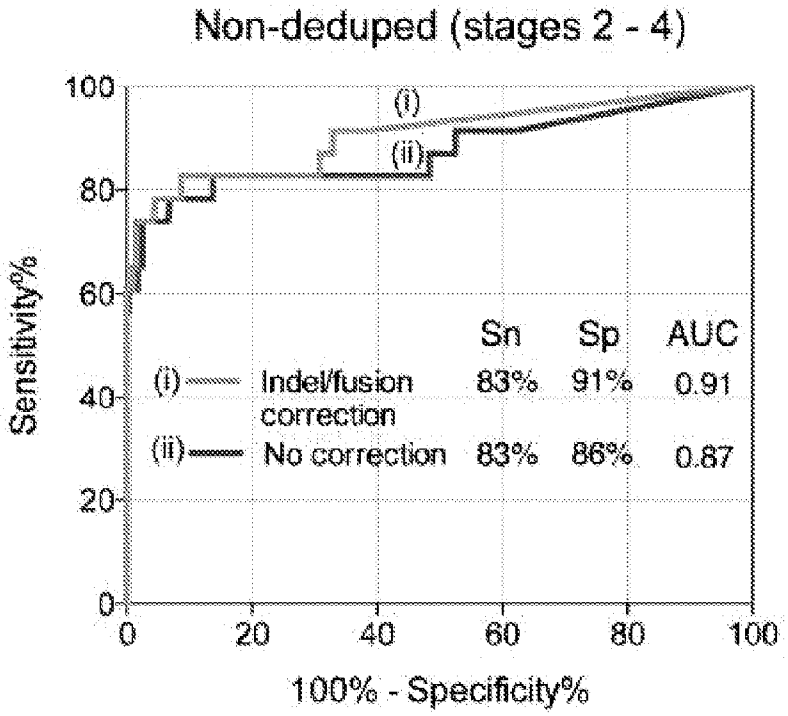
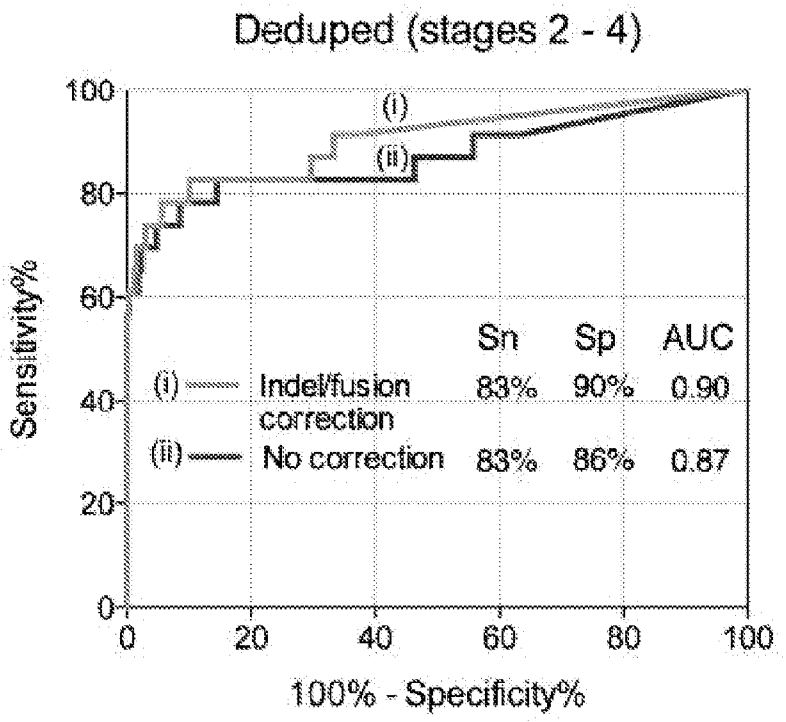
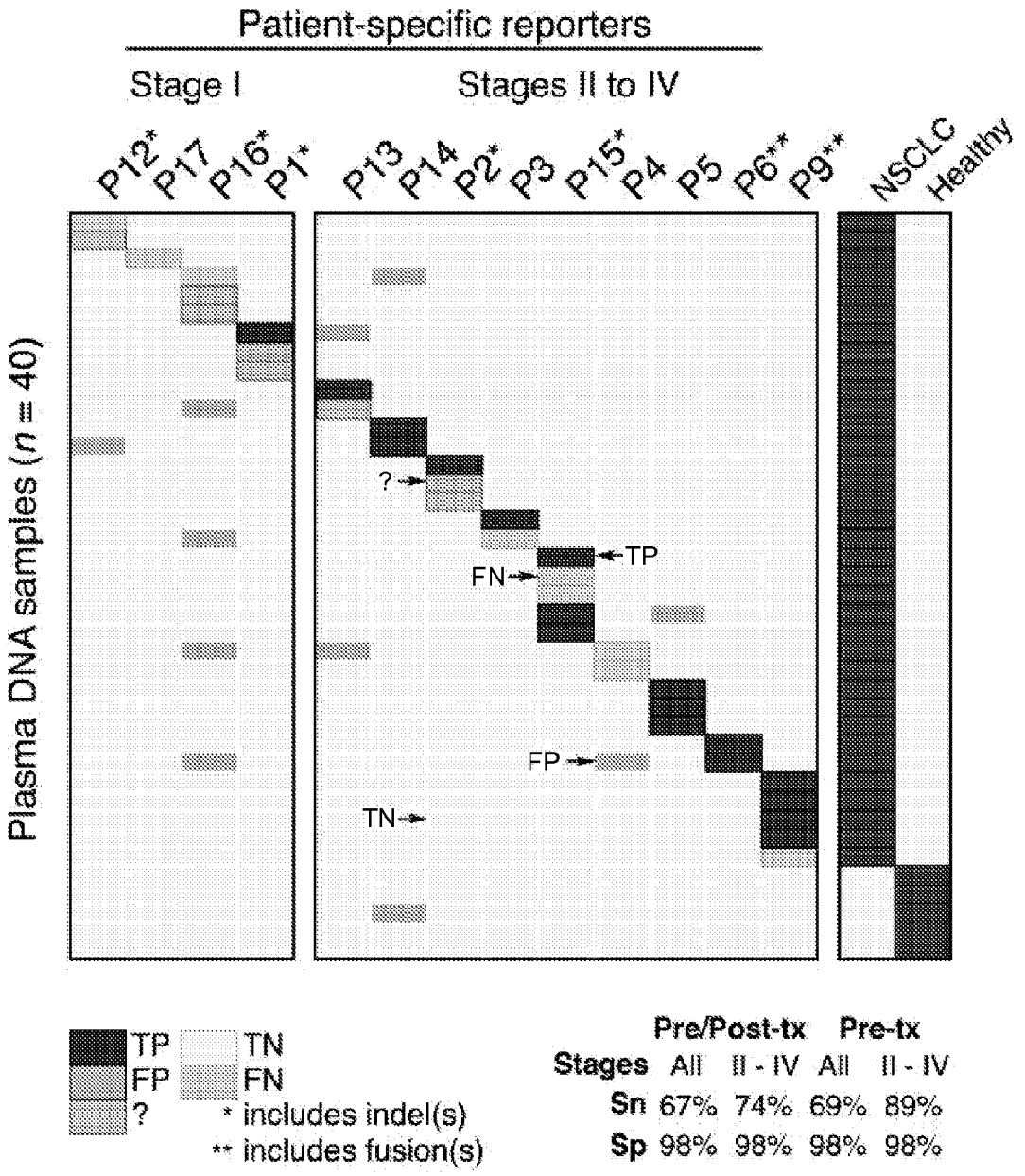


Figure 19D





	Pre/Post-tx		Pre-tx	
Stages	All	II - IV	All	II - IV
Sn	67%	74%	69%	89%
Sp	98%	98%	98%	98%

Figure 20

Figure 21A

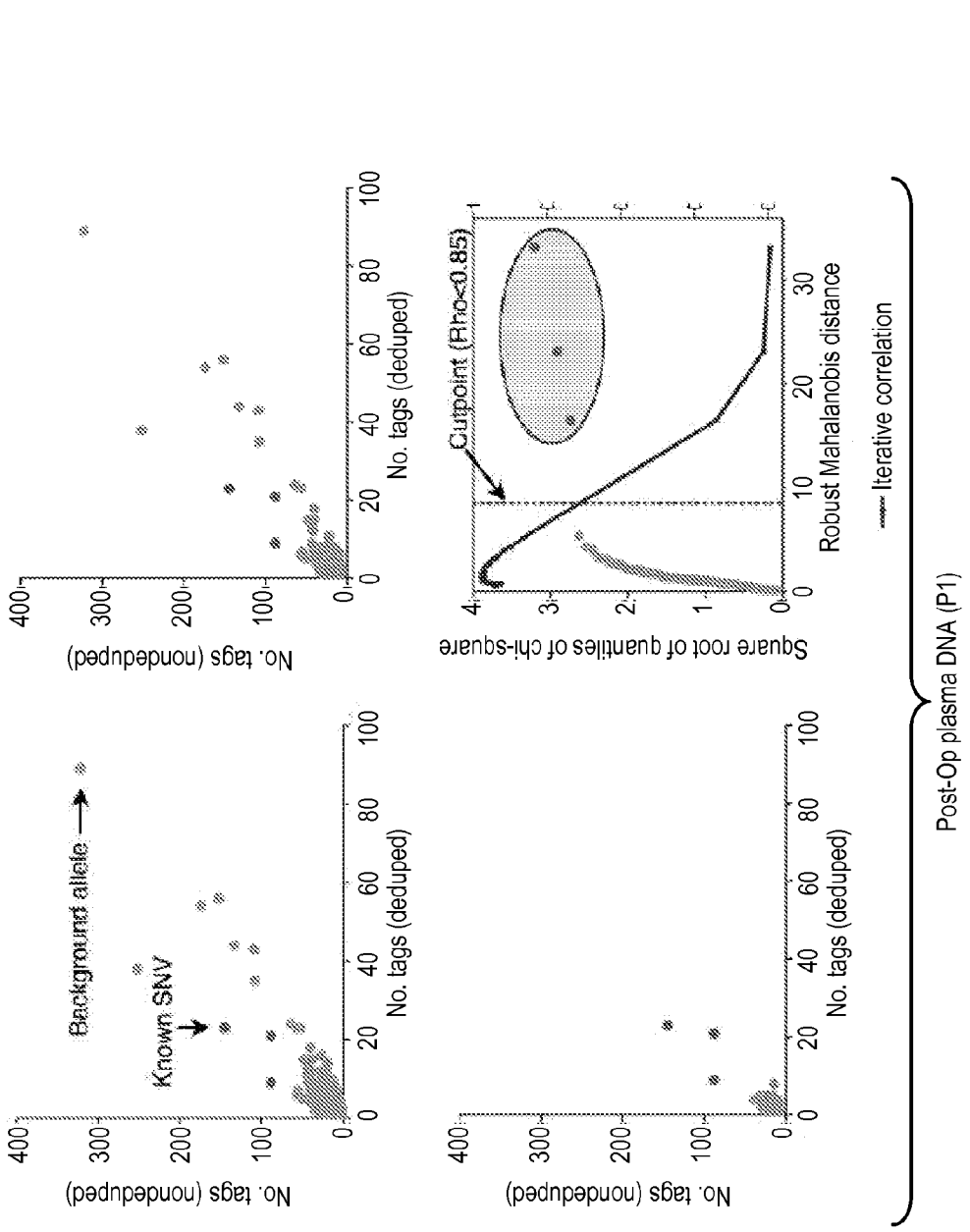


Figure 21B

Figure 21C

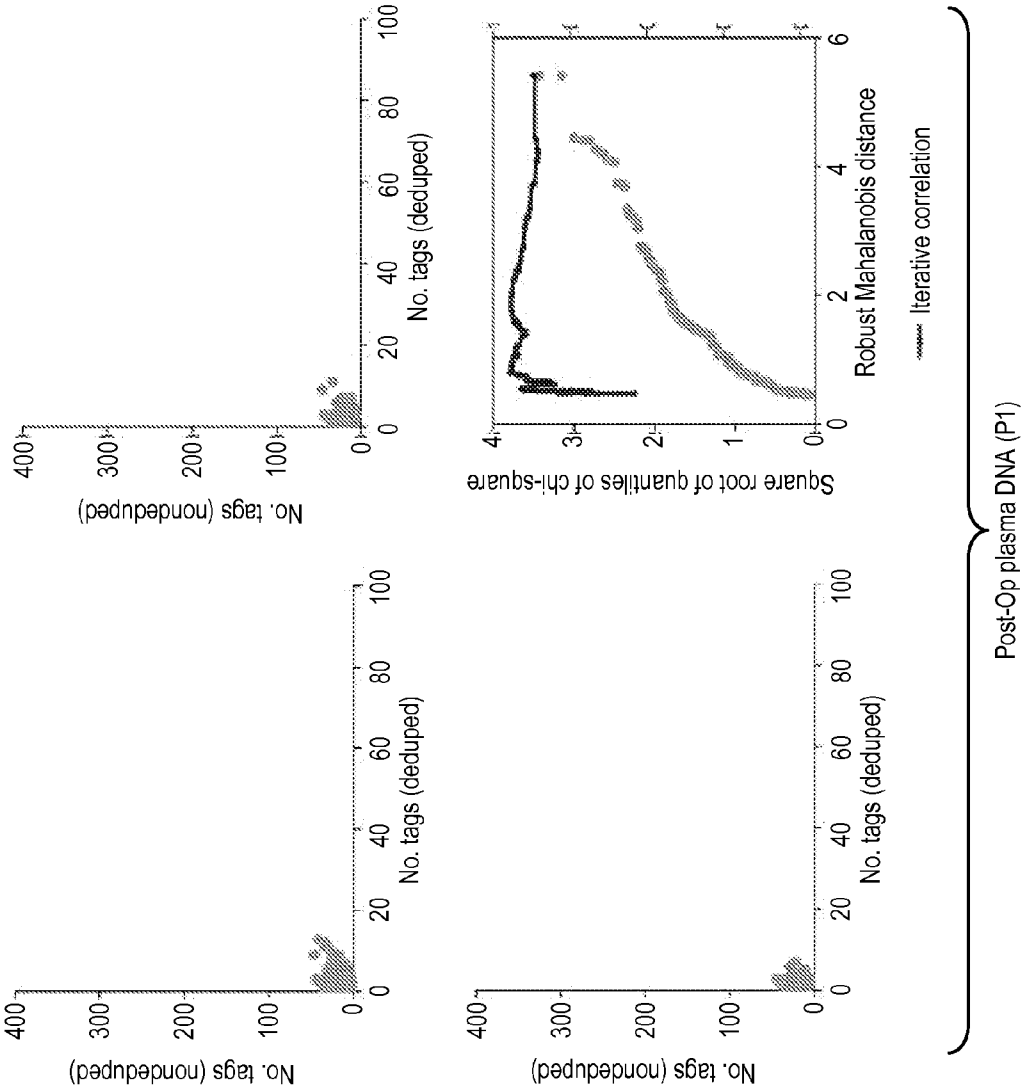
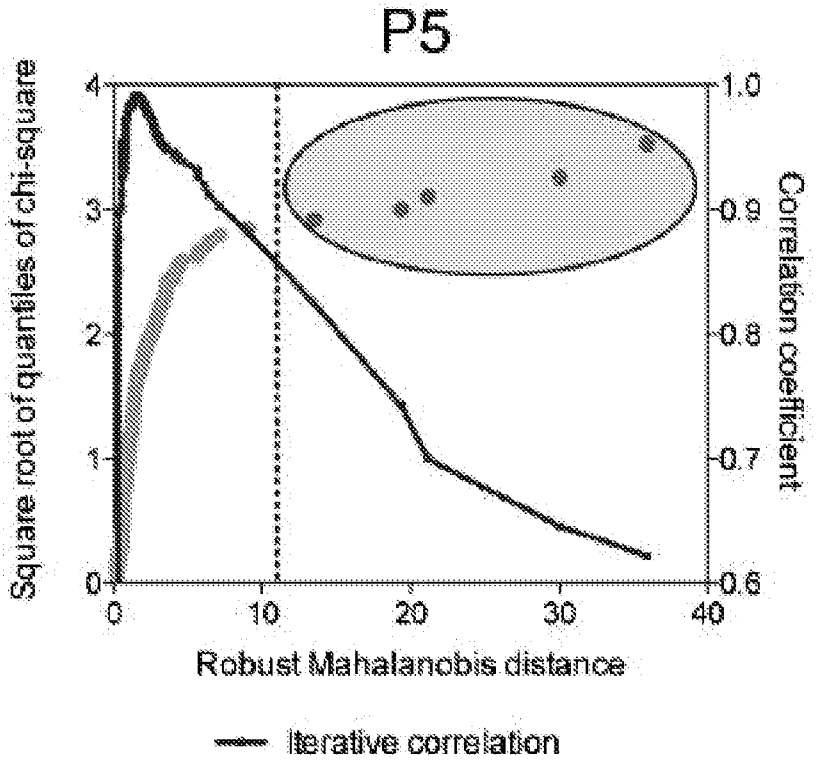
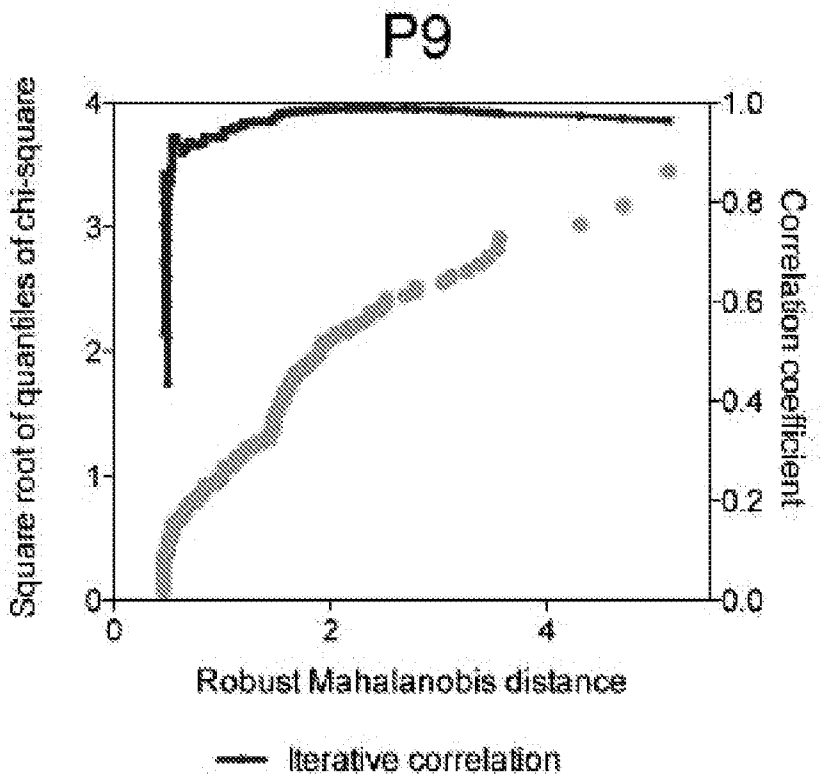

No. tags (nondeduped)
No. tags (deduped)
Square root of quantiles of chi-square
Robust Mahalanobis distance
Iterative correlation
Post-Op plasma DNA (P1)

Figure 21D


P5
Square root of quantiles of chi-square
Correlation coefficient
Robust Mahalanobis distance
Iterative correlation
0
1
2
3
4
0
10
20
30
40
0.6
0.7
0.8
0.9
1.0


P9
Square root of quantiles of chi-square
Correlation coefficient
Robust Mahalanobis distance
Iterative correlation
0
1
2
3
4
0
2
4
0.0
0.2
0.4
0.6
0.8
1.0

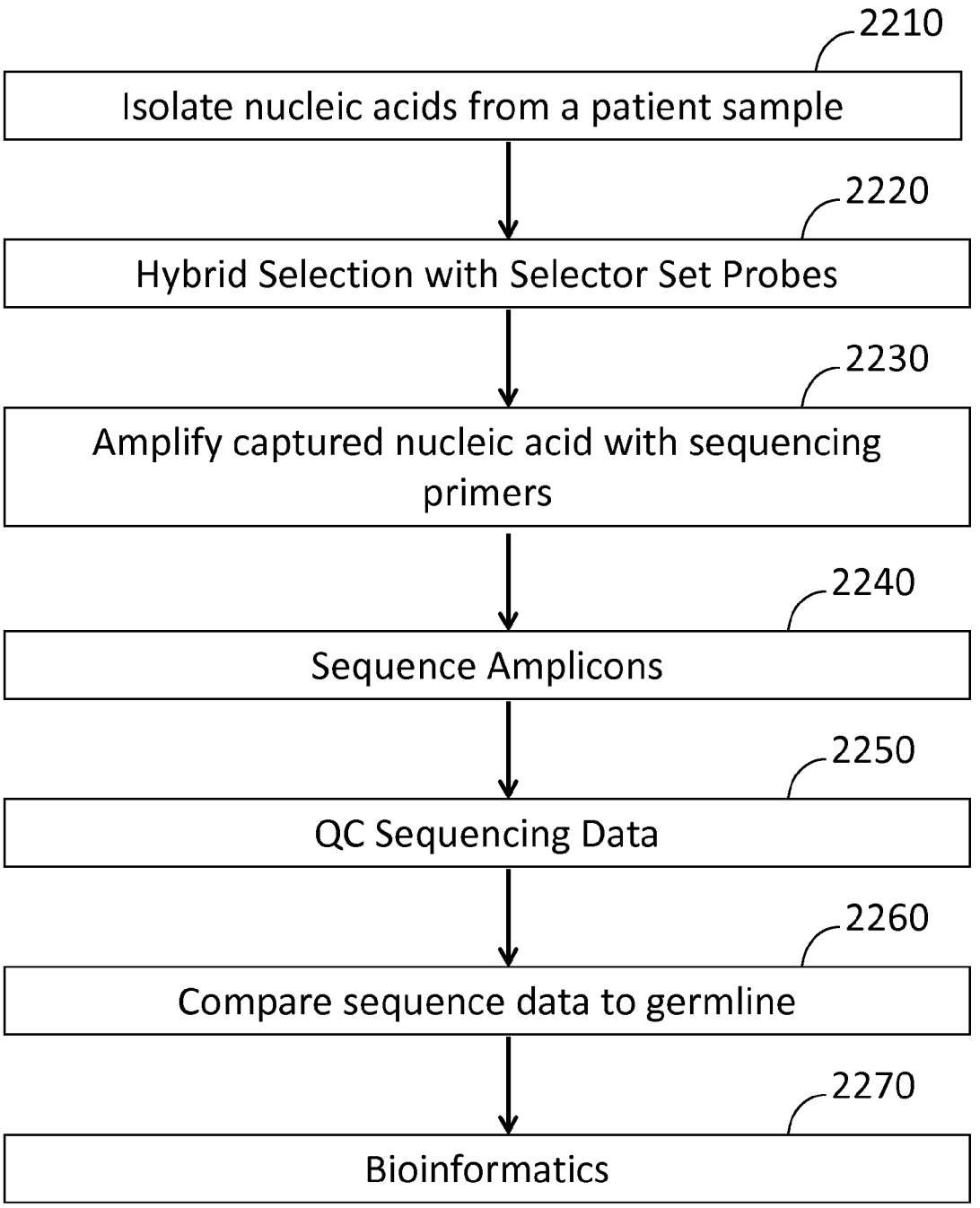

2210
Isolate nucleic acids from a patient sample
2220
Hybrid Selection with Selector Set Probes
2230
Amplify captured nucleic acid with sequencing primers
2240
Sequence Amplicons
2250
QC Sequencing Data
2260
Compare sequence data to germline
2270
Bioinformatics

Figure 22

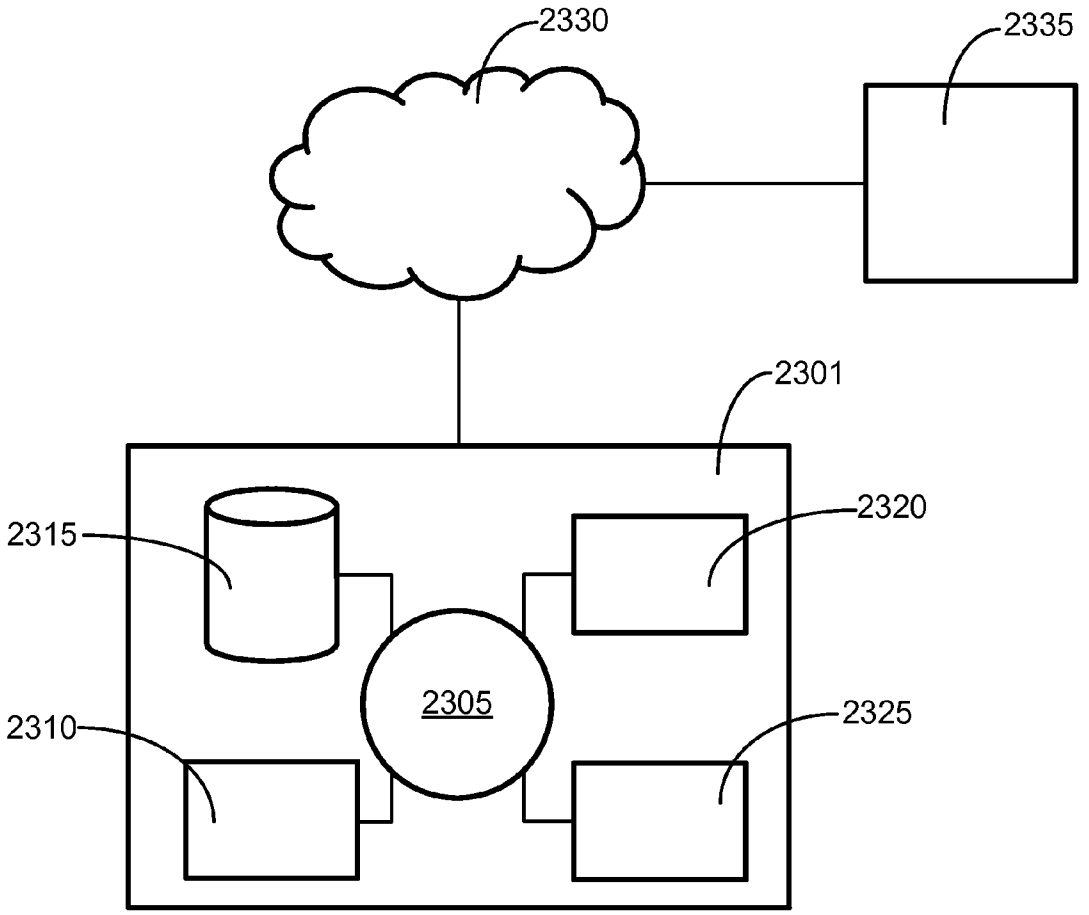


Figure 23

IDENTIFICATION AND USE OF CIRCULATING NUCLEIC ACID TUMOR MARKERS

STATEMENT OF GOVERNMENTAL SUPPORT

[0001] This invention was made with Government support under contract W81XWH-12-1-0285 awarded by the Department of Defense. The Government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0002] Tumors continually shed DNA into the circulation, where it is readily accessible (Stroun et al. (1987) *Eur J Cancer Clin Oncol* 23:707-712). Analysis of such cancer-derived cell-free DNA (cfDNA) has the potential to revolutionize detection and monitoring of cancer. Noninvasive access to malignant DNA is particularly attractive for solid tumors, which cannot be repeatedly sampled without invasive procedures. In non-small cell lung cancer (NSCLC), PCR-based assays have been used previously to detect recurrent point mutations in genes such as KRAS or EGFR in plasma DNA (Taniguchi et al. (2011) Clin. Cancer Res. 17:7808-7815; Gautschi et al. (2007) Cancer Lett. 254:265-273; Kuang et al. (2009) Clin. Cancer Res. 15:2630-2636; Rosell et al. (2009) N. Engl. J. Med. 361:958-967), but the majority of patients lack mutations in these genes.

[0003] Other studies have proposed identifying patient-specific chromosomal rearrangements in tumors via whole genome sequencing (WGS), followed by breakpoint qPCR from cfDNA (Leary et al. (2010) Sci. Transl. Med. 2:20ra14; McBride et al. (2010) Genes Chrom. Cancer 49:1062-1069). While sensitive, such methods require optimization of molecular assays for each patient, limiting their widespread clinical application. More recently, several groups have reported amplicon-based deep sequencing methods to detect cfDNA mutations in up to 6 recurrently mutated genes (Forshew et al. (2012) Sci. Transl. Med. 4:136ra168; Narayan et al. (2012) Cancer Res. 72:3492-3498; Kinde et al. (2011) Proc. Natl Acad. Sci. USA 108:9530-9535). While powerful, these approaches are limited by the number of mutations that can be interrogated (Rachlin et al. (2005) BMC Genomics 6:102) and the inability to detect genomic fusions.

[0004] PCT International Patent Publication No. 2011/103236 describes methods for identifying personalized tumor markers in a cancer patient using "mate-paired" libraries. The methods are limited to monitoring somatic chromosomal rearrangements, however, and must be personalized for each patient, thus limiting their applicability and increasing their cost.

[0005] U.S. Patent Application Publication No. 2010/0041048 A1 describes the quantitation of tumor-specific cell-free DNA in colorectal cancer patients using the "BEAMing" technique (Beads, Emulsion, Amplification, and Magnetics). While this technique provides high sensitivity and specificity, this method is for single mutations and thus any given assay can only be applied to a subset of patients and/or requires patient-specific optimization. U.S. Patent Application Publication No. 2012/0183967 A1 describes additional methods to identify and quantify genetic variations, including the analysis of minor variants in a DNA population, using the "BEAMing" technique.

[0006] U.S. Patent Application Publication No. 2012/0214678 A1 describes methods and compositions for detecting fetal nucleic acids and determining the fraction of cell-free fetal nucleic acid circulating in a maternal sample. While sensitive, these methods analyze polymorphisms occurring between maternal and fetal nucleic acids rather than polymorphisms that result from somatic mutations in tumor cells. In addition, methods that detect fetal nucleic acids in maternal circulation require much less sensitivity than methods that detect tumor nucleic acids in cancer patient circulation, because fetal nucleic acids are much more abundant than tumor nucleic acids.

[0007] U.S. Patent Application Publication Nos. 2012/0237928 A1 and 2013/0034546 describe methods for determining copy number variations of a sequence of interest in a test sample comprising a mixture of nucleic acids. While potentially applicable to the analysis of cancer, these methods are directed to measuring major structural changes in nucleic acids, such as translocations, deletions, and amplifications, rather than single nucleotide variations.

[0008] U.S. Patent Application Publication No. 2012/0264121 A1 describes methods for estimating a genomic fraction, for example, a fetal fraction, from polymorphisms such as small base variations or insertions-deletions. These methods do not, however, make use of optimized libraries of polymorphisms, such as, for example, libraries containing recurrently-mutated genomic regions.

[0009] U.S. Patent Application Publication No. 2013/0024127 A1 describes computer-implemented methods for calculating a percent contribution of cell-free nucleic acids from a major source and a minor source in a mixed sample. The methods do not, however, provide any advantages in identifying or making use of optimized libraries of polymorphisms in the analysis.

[0010] PCT International Publication No. WO 2010/141955 A2 describes methods of detecting cancer by analyzing panels of genes from a patient-obtained sample and determining the mutational status of the genes in the panel. The methods rely on a relatively small number of known cancer genes, however, and they do not provide any ranking of the genes according to effectiveness in detection of relevant mutations. In addition, the methods were unable to detect the presence of mutations in the majority of serum samples from actual cancer patients.

[0011] There is thus a need for new and improved methods to detect and monitor tumor-related nucleic acids in cancer patients.

SUMMARY OF THE INVENTION

[0012] Compositions and methods, including methods of bioinformatic analysis, are provided for the highly sensitive analysis of circulating tumor DNA (ctDNA), e.g. DNA sequences present in the blood of an individual that are derived from tumor cells. The methods of the invention may be referred to as CAncer Personalized Profiling by Deep Sequencing (CAPP-Seq). Tumors of particular interest are solid tumors, including without limitation carcinomas, sarcomas, gliomas, lymphomas, melanomas, etc., although hematologic cancers, such as leukemias, are not excluded.

[0013] The methods of the invention combine optimized library preparation methods with a multi-phase bioinformatics approach to design a "selector" population of DNA oligonucleotides, which correspond to recurrently mutated regions in the cancer of interest. The selector population of DNA oligonucleotides, which may be referred to as a selector set, comprises probes for a plurality of genomic regions, and is

designed such that at least one mutation within the plurality of genomic regions is present in a majority of all subjects with the specific cancer; and in preferred embodiments multiple mutations are present in a majority of all subjects with the specific cancer.

[0014] In some embodiments of the invention, methods are provided for the identification of a selector set appropriate for a specific tumor type. Also provided are oligonucleotide compositions of selector sets, which may be provided adhered to a solid substrate, tagged for affinity selection, etc.; and kits containing such selector sets. Included, without limitation, is a selector set suitable for analysis of non-small cell lung carcinoma (NSCLC). Such kits may include executable instructions for bioinformatics analysis of the CAPP-Seq data.

[0015] In other embodiments, methods are provided for the use of a selector set in the diagnosis and monitoring of cancer in an individual patient. In such embodiments the selector set is used to enrich, e.g. by hybrid selection, for ctDNA that corresponds to the regions of the genome that are most likely to contain tumor-specific somatic mutations. The "selected" ctDNA is then amplified and sequenced to determine which of the selected genomic regions are mutated in the individual tumor. An initial comparison is optionally made with the individual's germline DNA sequence and/or a tumor biopsy sample from the individual. These somatic mutations provide a means of distinguishing ctDNA from germline DNA, and thus provide useful information about the presence and quantity of tumor cells in the individual.

[0016] In some embodiments, the ctDNA content in an individual's blood, or blood derivative, sample is determined at one or more time points, optionally in conjunction with a therapeutic regimen. The presence of the ctDNA correlates with tumor burden, and is useful in monitoring response to therapy, monitoring residual disease, monitoring for the presence of metastases, monitoring total tumor burden, and the like. Although not required, for some methods CAPP-Seq may be performed in conjunction with tumor imaging methods, e.g. PET/CT scans and the like.

[0017] In other embodiments, CAPP-seq is used for cancer screening and biopsy-free tumor genotyping, where a patient ctDNA sample is analyzed without reference to a biopsy sample. In some such embodiments, where CAPP-Seq identifies a mutation in a clinically actionable target from a ctDNA sample, the methods include providing a therapy appropriate for the target. Such mutations include, without limitation, rearrangements and other mutations involving oncogenes, receptor tyrosine kinases, etc. Actionable targets may include, for example, ALK, ROS1, RET, EGFR, KRAS, and the like.

[0018] The CAPP-Seq methods may include steps of data analysis, which may be provided as a program of instructions executable by computer and performed by means of software components loaded into the computer. Such methods include the design for identification selector set for a cancer of interest. Other bioinformatics methods are provided for determining and quantitating when circulating tumor DNA is detectable above background, e.g. using an approach that integrates information content and classes of mutation into a detection index.

[0019] Disclosed herein is a method for determining the presence of tumor nucleic acids (tNA) in a cell-free nucleic acids (cfNA) sample from an individual by detection of somatic mutations. The method may comprise (a) obtaining a cfNA sample; (b) selecting the cfNA for sequences corresponding to a plurality of regions of mutations in a cancer of interest; (c) sequencing the selected cfNA; (d) determining the presence of somatic mutations, wherein the presence of the somatic mutations may be indicative of tumor cells present in the individual; and (e) providing the individual with an assessment of the presence of tumor cells.

[0020] The cell-free nucleic acid may be cell-free DNA (cfDNA). The cell-free nucleic acid may be cell-free RNA (cfRNA). The cell-free nucleic acids may be a mixture of cell-free DNA (cfDNA) and cell-free RNA (cfRNA). The tumor nucleic acid may be a nucleic acid originating from a tumor cell. The tumor nucleic acid may be tumor-derived DNA (tDNA). The tumor nucleic acid may be a circulating tumor DNA (ctDNA). The tumor nucleic acid may be tumor-derived RNA (tRNA). The tumor nucleic acid may be a circulating tumor RNA (ctRNA). The tumor nucleic acids may be a mixture of tumor-derived DNA and tumor-derived RNA. The tumor nucleic acids may be a mixture of ctDNA and ctRNA.

[0021] Selecting the cfNA may comprise (i) hybridizing the cell-free nucleic acid sample to a plurality of selector set probes comprising a specific binding member; (ii) binding hybridized nucleic acids to a complementary specific binding member; and (iii) washing away unbound DNA.

[0022] The cfNA sample may be compared to a known tumor DNA sequence from the individual.

[0023] The cfNA sample may be de novo analyzed for the presence of somatic mutations.

[0024] The somatic mutations may include single nucleotide variants, insertions, deletions, copy number variations, and rearrangements.

[0025] The plurality of regions of mutations may comprise at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175 or 200 different genomic regions. The plurality of regions of mutations may comprise at least 500 different genomic regions. The plurality of genomic regions of mutations may comprise a total of from 100 to 500 kb of sequence.

[0026] At least one somatic mutation may be present in at least 60%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% of individuals in a patient population for the cancer of interest.

[0027] The cancer of interest may be a leukemia. The cancer of interest may be a solid tumor. The cancer may be a carcinoma. The carcinoma may be an adenocarcinoma or a squamous cell carcinoma. The carcinoma may be non-small cell lung cancer.

[0028] The individual may be not previously diagnosed with cancer. The individual may be undergoing treatment for cancer.

[0029] Two or more samples may be obtained from the individual over a period of time and compared for residual disease or tumor burden.

[0030] The method may further comprise treating the individual in accordance with the analysis of the presence of tumor cells. The method may further comprise treating the individual based on the detection of the somatic mutations.

[0031] Determining the presence of somatic mutations may comprise: (i) integrating cfDNA fractions across all somatic SNVs; (ii) performing a position-specific background adjustment; and (iii) evaluating statistical significance by Monte Carlo sampling of background alleles across the selector, wherein steps (i)-(iii) are embodied as a program of instruc-

tions executable by computer and performed by means of software components loaded into the computer.

[0032] The method may further comprise analysis of insertions and/or deletions by comparing its fractional abundance in a given cfDNA sample against its fractional abundance in a cohort. The method may further comprise combining the fractional abundance into a single Z-score.

[0033] The method may further comprise integrating different mutation types to estimate the significance of tumor burden quantitation.

[0034] Determining the presence of somatic mutations may be identification of genomic fusion events and breakpoints by the method comprising: (i) identification of discordant reads; (ii) detection of breakpoints at base pair-resolution, and (iii) in silico validation of candidate fusions, wherein steps (i)-(iii) are embodied as a program of instructions executable by computer and performed by means of software components loaded into the computer.

[0035] Determining the presence of somatic mutation may comprise the steps of (i) taking allele frequencies from a single cfDNA sample and selecting high quality data; (ii) testing whether a given input cfDNA allele may be significantly different from the corresponding paired germline allele; (iii) assembling a database of cfDNA background allele frequencies by binomial distribution; (iv) testing whether a given input allele differs significantly from cfDNA background at the same position, and selecting those with an average background frequency of a predetermined threshold; and (v) distinguishing tumor-derived SNVs from remaining background noise by outlier analysis, wherein steps (i)-(v) may be embodied as a program of instructions executable by computer and performed by means of software components loaded into the computer.

[0036] The selector set probes may comprise sequences corresponding to a mutated genomic regions identified by the method comprising identifying a plurality of genomic regions from a group of genomic regions that may be mutated in a specific cancer.

[0037] Identifying the plurality of genomic regions may comprise for each genomic region in the plurality of genomic regions, ranking the genomic region to maximize the number of all subjects with the specific cancer having at least one mutation within the genomic region.

[0038] Identifying the plurality of genomic regions may comprise: (i) selecting genes known to be drivers in the cancer of interest to generate a pool of known drivers; (ii) selecting exons from known drivers with the highest recurrence index (RI) that identify at least one new patient compared to step (a); and repeating until no further exons meet these criteria; (iii) identifying remaining exons of known drivers with an RI≥30 and with SNVs covering ≥3 patients in the relevant database that result in the largest reduction in patients with only 1 SNV; and repeating until no further exons meet these criteria; (iv) repeating step (b) using RI≥20; (v) adding in all exons from additional genes previously predicted to harbor driver mutations; and (vi) adding for known recurrent rearrangement the introns most frequently implicated in the fusion event and the flanking exons, wherein steps (i)-(vi) are embodied as a program of instructions executable by computer and performed by means of software components loaded into the computer.

[0039] The plurality of regions of mutations in a cancer of interest may be selected from the regions set forth in Table 2.

[0040] The method of Claim **27**, wherein the plurality of regions of mutations may comprise at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 regions set forth in Table 2.

[0041] Further disclosed herein are compositions comprising selector set probes. The composition may comprise a set of selector set probes of at least about 25 nucleotides in length, comprising a specific binding member, and comprising sequences from at least 100 regions set forth in Table 2.

[0042] The set of selector probes may comprise oligonucleotides comprising sequences from at least 300 regions from Table 2. The set of selector probes may comprise oligonucleotides comprising sequences from at least 500 regions from Table 2.

[0043] Further disclosed herein are populations of cell-free DNA (cfDNA). The population of cfDNA may be an enriched population. The enriched population of cfDNA may be produced by hybrid selection. Hybrid selection may comprise of use of one or more selector set probes. The selector set probes may be attached to a solid or semi-solid support. The support may comprise an array. The support may comprise a bead. The bead may be a coated bead. The bead may be a streptavidin bead. The solid support may comprise a flat surface. The solid support may comprise a slide. The solid support may comprise a glass slide.

[0044] Further disclosed herein are methods for detecting, diagnosing, prognosing, or therapy selection for a subject suffering from a disease or condition. The method may comprise: (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; and (b) using sequence information derived from (a) to detect cell-free non-germline DNA (cfNG-DNA) in the sample, wherein the method may be capable of detecting a percentage of cfNG-DNA that may be less than 2% of total cfDNA.

[0045] The method may be capable of detecting a percentage of ctDNA that may be less than 1.5% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that may be less than 1% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that may be less than 0.5% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that may be less than 0.1% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that may be less than 0.01% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that may be less than 0.001% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that may be less than 0.0001% of the total cfDNA.

[0046] The sample may be a plasma or serum sample (sweat, breath, tears, saliva, urine, stool, amniotic fluid). The sample may be a cerebral spinal fluid sample. In some instances, the sample is not a pap smear fluid sample. In some instances, the sample is not a cyst fluid sample. In some instances, the sample is not a pancreatic fluid sample.

[0047] The sequence information may comprise information related to at least 10, 20, 30, 40, 100, 200, or 300 genomic regions. The genomic regions may comprise genes, exonic regions, intronic regions, untranslated regions, non-coding regions or a combination thereof. The genomic regions may comprise two or more of exonic regions, intronic regions, and untranslated regions. The genomic regions may comprise at least one exonic region and at least one intronic region. At

least 5% of the genomic regions may comprise intronic regions. At least about 20% of the genomic regions may comprise exonic regions.

[0048] The genomic regions may comprise less than 1.5 megabases (Mb) of the genome. The genomic regions may comprise less than 1 Mb of the genome. The genomic regions may comprise less than 500 kilobases (kb) of the genome. The genomic regions may comprise less than 50, 75, 100 or 350 kb of the genome. The genomic regions may comprise between 100 kb to 300 kb of the genome.

[0049] The sequence information may comprise information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to a plurality of genomic regions.

[0050] The plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects.

[0051] The total size of the genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The total size of the genomic regions of the selector set may be between 100 kb to 300 kb of the genome.

[0052] The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 2. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 6. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 7. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 8. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 9. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 10. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 11. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 12. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 13. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 14. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 15. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 16. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 17. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 18. In some instances, the subject is not suffering from a pancreatic cancer.

[0053] Obtaining sequence information of the cell-free DNA sample may comprise performing massively parallel sequencing. Massively parallel sequencing may be performed on a subset of a genome of cfDNA from the cfDNA sample. The subset of the genome may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The subset of the genome may comprise between 100 kb to 300 kb of the genome.

[0054] Obtaining sequence information of the cell-free DNA sample may comprise using single molecule barcoding. Using single molecule barcoding may comprise attaching barcodes comprising different sequences to nucleic acids from the cfDNA sample.

[0055] The sequence information may comprise sequence information pertaining to the adaptors. The sequence information may comprise sequence information pertaining to the molecular barcodes. The sequence information may comprise sequence information pertaining to the sample indexes.

[0056] The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from the subject. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more different subjects. The two or more samples may be the same type of sample. The two or more samples may be two different types of sample. The two or more samples may be obtained from the subject at the same time point. The two or more samples may be obtained from the subject at two or more time points. The samples from two or more different subjects may be indexed and pooled together prior to sequencing.

[0057] Using the sequence information may comprise detecting one or more mutations. The one or more mutations may comprise one or more SNVs, indels, fusions, breakpoints, structural variants, variable number of tandem repeats, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, copy number variants or a combination thereof in selected regions of the subject's genome. Using the sequence information may comprise detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting at least one SNV, indel, copy number variant, and rearrangement in selected regions of the subject's genome.

[0058] In some instances, detecting the one or more mutations does not involve performing digital PCR (dPCR).

[0059] Detecting the one or more mutations may comprise applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of

genomic regions comprising one or more mutations present in at least about 60% of cancer subjects from population of cancer subjects.

[0060] The cfNG-DNA may be derived from a tumor in the subject. The method may further comprise detecting a cancer in the subject based on the detection of the cfNG-DNA. The method may further comprise diagnosing a cancer in the subject based on the detection of the cfNG-DNA. Diagnosing the cancer may have a sensitivity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the cancer may have a specificity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise prognosing a cancer in the subject based on the detection of the cfNG-DNA. Prognosing the cancer may have a sensitivity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Prognosing the cancer may have a specificity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise determining a therapeutic regimen for the subject based on the detection of the cfNG-DNA. The method may further comprise administering an anti-cancer therapy to the subject based on the detection of the cfNG-DNA.

[0061] The cfNG-DNA may be derived from a fetus in the subject. The method may further comprise diagnosing a disease or condition in the fetus based on the detection of the cfNG-DNA. Diagnosing the disease or condition in the fetus may have a sensitivity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the disease or condition in the fetus may have a specificity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%.

[0062] The cfNG-DNA may be derived from a transplanted organ, cell or tissue in the subject. The method may further comprise diagnosing an organ transplant rejection in the subject based on the detection of the cfNG-DNA. Diagnosing the organ transplant rejection may have a sensitivity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the organ transplant rejection may have a specificity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise prognosing a risk of organ transplant rejection in the subject based on the detection of the cfNG-DNA. Prognosing the risk of organ transplant rejection may have a sensitivity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Prognosing the risk of organ transplant rejection may have a specificity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise determining an immunosuppresive therapy for the subject based on the detection of the cfNG-DNA. The method may further comprise administering an immunosuppresive therapy to the subject based on the detection of the cfNG-DNA.

[0063] Further disclosed herein are methods of diagnosing a cancer. The method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information may be derived from regions that are mutated in at least 80% of a population of subjects afflicted with a cancer; and (b) diagnosing a cancer selected from a group consisting of lung cancer, breast cancer, colorectal cancer and prostate cancer in the subject based on the sequence information, wherein the method has a sensitivity of at least 80%.

[0064] The regions that are mutated may comprise a total size of less than 1.5 Mb of the genome. The regions that are mutated may comprise a total size of less than 1 Mb of the genome. The regions that are mutated may comprise a total size of less than 500 kb of the genome. The regions that are mutated may comprise a total size of less than 350 kb of the genome. The regions that are mutated may comprise a total size of less than 300 kb of the genome. The regions that are mutated may comprise a total size of less than 250 kb of the genome. The regions that are mutated may comprise a total size of less than 200 kb of the genome. The regions that are mutated may comprise a total size of less than 150 kb of the genome. The regions that are mutated may comprise a total size of less than 100 kb of the genome. The regions that are mutated may comprise a total size of less than 50 kb of the genome. The regions that are mutated may comprise a total size of less than 40 kb of the genome. The regions that are mutated may comprise a total size of less than 30 kb of the genome. The regions that are mutated may comprise a total size of less than 20 kb of the genome. The regions that are mutated may comprise a total size of less than 10 kb of the genome.

[0065] The regions that are mutated may comprise a total size between 100 kb-300 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-200 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-150 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-100 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-75 kb of the genome. The regions that are mutated may comprise a total size between 1 kb-50 kb of the genome.

[0066] The sequence information may be derived from 2 or more regions. The sequence information may be derived from 3 or more regions. The sequence information may be derived from 4 or more regions. The sequence information may be derived from 5 or more regions. The sequence information may be derived from 6 or more regions. The sequence information may be derived from 7 or more regions. The sequence information may be derived from 8 or more regions. The sequence information may be derived from 9 or more regions. The sequence information may be derived from 10 or more regions. The sequence information may be derived from 20 or more regions. The sequence information may be derived from 30 or more regions. The sequence information may be derived from 40 or more regions. The sequence information may be derived from 50 or more regions. The sequence information may be derived from 60 or more regions. The sequence information may be derived from 70 or more regions. The sequence information may be derived from 80 or more

regions. The sequence information may be derived from 90 or more regions. The sequence information may be derived from 100 or more regions.

[0067] The population of subjects afflicted with the cancer may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA).

[0068] The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 60% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 70% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 80% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 90% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 95% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 99% of the population of subjects afflicted with the cancer.

[0069] The sequence information may be derived from regions that may be mutated in at least 65% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 70% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 75% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 80% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 85% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 90% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 95% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 99% of the population of subjects afflicted with the cancer.

[0070] Obtaining the sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more lncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof.

[0071] Alternatively, or additionally, obtaining the sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof.

[0072] In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1.

[0073] The method may further comprise detecting mutations in the regions based on the sequencing information. Diagnosing the cancer may be based on the detection of the mutations. The detection of at least 3 mutations may be indicative of the cancer. The detection of one or more mutations in three or more regions may be indicative of the cancer.

[0074] The breast cancer may be a BRCA1 cancer.

[0075] The method may have a sensitivity of at least 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

[0076] The method may have a specificity of at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

[0077] The method may further comprise providing a computer-generated report comprising the diagnosis of the cancer.

[0078] Further disclosed herein are methods of determining a prognosis of a condition or disease in a subject in need thereof. The method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information may be derived from regions that are mutated in at least 80% of a population of subjects afflicted with a condition; and (b) determining a prognosis of a condition or disease in the subject based on the sequence information.

[0079] The regions that are mutated may comprise a total size of less than 1.5 Mb of the genome. The regions that are mutated may comprise a total size of less than 1 Mb of the genome. The regions that are mutated may comprise a total size of less than 500 kb of the genome. The regions that are mutated may comprise a total size of less than 350 kb of the genome. The regions that are mutated may comprise a total size of less than 300 kb of the genome. The regions that are mutated may comprise a total size of less than 250 kb of the genome. The regions that are mutated may comprise a total size of less than 200 kb of the genome. The regions that are mutated may comprise a total size of less than 150 kb of the genome. The regions that are mutated may comprise a total size of less than 100 kb of the genome. The regions that are mutated may comprise a total size of less than 50 kb of the genome. The regions that are mutated may comprise a total size of less than 40 kb of the genome. The regions that are mutated may comprise a total size of less than 30 kb of the genome. The regions that are mutated may comprise a total size of less than 20 kb of the genome. The regions that are mutated may comprise a total size of less than 10 kb of the genome.

[0080] The regions that are mutated may comprise a total size between 100 kb-300 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-200 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-150 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-100 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-75 kb of the genome. The regions that are mutated may comprise a total size between 1 kb-50 kb of the genome.

[0081] The sequence information may be derived from 2 or more regions. The sequence information may be derived from 3 or more regions. The sequence information may be derived from 4 or more regions. The sequence information may be derived from 5 or more regions. The sequence information may be derived from 6 or more regions. The sequence information may be derived from 7 or more regions. The sequence information may be derived from 8 or more regions. The sequence information may be derived from 9 or more regions. The sequence information may be derived from 10 or more regions. The sequence information may be derived from 20 or more regions. The sequence information may be derived from 30 or more regions. The sequence information may be derived from 40 or more regions. The sequence information may be derived from 50 or more regions. The sequence information may be derived from 60 or more regions. The sequence information may be derived from 70 or more regions. The sequence information may be derived from 80 or more regions. The sequence information may be derived from 90 or more regions. The sequence information may be derived from 100 or more regions.

[0082] The population of subjects afflicted with the cancer may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA).

[0083] The sequence information may be derived from regions that may be mutated in at least 65% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 70% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 75% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 80% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 85% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 90% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 95% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 99% of the population of subjects afflicted with the cancer.

[0084] Obtaining the sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more lncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof.

[0085] Alternatively, or additionally, obtaining the sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof.

[0086] In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1.

[0087] The method may further comprise detecting mutations in the regions based on the sequencing information. Prognosing the condition or disease may be based on the detection of the mutations. The detection of at least 3 mutations may be indicative of an outcome of the condition or disease. The detection of one or more mutations in three or more regions may be indicative of an outcome of the condition or disease.

[0088] The condition may be a cancer. The cancer may be a solid tumor. The solid tumor may be non-small cell lung cancer (NSCLC). The cancer may be a breast cancer. The breast cancer may be a BRCA1 cancer. The cancer may be a lung cancer, colorectal cancer, prostate cancer, ovarian cancer, esophageal cancer, breast cancer, lymphoma, or leukemia.

[0089] The method may have a sensitivity of at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

[0090] The method may have a specificity of at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

[0091] The method may further comprise providing a computer-generated report comprising the prognosis of the condition.

[0092] Further disclosed herein are methods of diagnosing, prognosing, or determining a therapeutic regimen for a subject afflicted with or susceptible of having a cancer. The method may comprise (a) obtaining sequence information for selected regions of genomic DNA from a cell-free DNA sample from the subject; (b) using the sequence information to determine the presence or absence of one or more mutations in the selected regions, wherein at least 70% of a population of subjects afflicted with the cancer have mutation(s) in the regions; and (c) providing a report with a diagnosis, prognosis or treatment regimen to the subject, based on the presence or absence of the one or more mutations.

[0093] The selected regions may comprise a total size of less than 1.5 Mb of the genome. The selected regions may comprise a total size of less than 1 Mb of the genome. The selected regions may comprise a total size of less than 500 kb of the genome. The selected regions may comprise a total size of less than 350 kb of the genome. The selected regions may comprise a total size of less than 300 kb of the genome. The selected regions may comprise a total size of less than 250 kb of the genome. The selected regions may comprise a total size of less than 200 kb of the genome. The selected regions may comprise a total size of less than 150 kb of the genome. The selected regions may comprise a total size of less than 100 kb of the genome. The selected regions may comprise a total size of less than 50 kb of the genome. The selected regions may comprise a total size of less than 40 kb of the genome. The selected regions may comprise a total size of less than 30 kb of the genome. The selected regions may comprise a total size of less than 20 kb of the genome. The selected regions may comprise a total size of less than 10 kb of the genome.

[0094] The selected regions may comprise a total size between 100 kb-300 kb of the genome. The selected regions may comprise a total size between 5 kb-200 kb of the genome.

The selected regions may comprise a total size between 5 kb-150 kb of the genome. The selected regions may comprise a total size between 5 kb-100 kb of the genome. The selected regions may comprise a total size between 5 kb-75 kb of the genome. The selected regions may comprise a total size between 1 kb-50 kb of the genome.

[0095] The sequence information may be derived from 2 or more regions. The sequence information may be derived from 3 or more regions. The sequence information may be derived from 4 or more regions. The sequence information may be derived from 5 or more regions. The sequence information may be derived from 6 or more regions. The sequence information may be derived from 7 or more regions. The sequence information may be derived from 8 or more regions. The sequence information may be derived from 9 or more regions. The sequence information may be derived from 10 or more regions. The sequence information may be derived from 20 or more regions. The sequence information may be derived from 30 or more regions. The sequence information may be derived from 40 or more regions. The sequence information may be derived from 50 or more regions. The sequence information may be derived from 60 or more regions. The sequence information may be derived from 70 or more regions. The sequence information may be derived from 80 or more regions. The sequence information may be derived from 90 or more regions. The sequence information may be derived from 100 or more regions.

[0096] The population of subjects afflicted with the cancer may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA).

[0097] The sequence information may be derived from regions that may be mutated in at least 65% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 70% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 75% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 80% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 85% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 90% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 95% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 99% of the population of subjects afflicted with the cancer.

[0098] Obtaining the sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more lncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof.

[0099] Alternatively, or additionally, obtaining the sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof.

[0100] In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1.

[0101] Detection of at least 3 mutations may be indicative of an outcome of the cancer. Detection of at least 4 mutations may be indicative of an outcome of the cancer. Detection of at least 5 mutations may be indicative of an outcome of the cancer. Detection of at least 6 mutations may be indicative of an outcome of the cancer.

[0102] Detection of one or more mutations in three or more regions may be indicative of an outcome of the cancer. Detection of one or more mutations in four or more regions may be indicative of an outcome of the cancer. Detection of one or more mutations in five or more regions may be indicative of an outcome of the cancer. Detection of one or more mutations in six or more regions may be indicative of an outcome of the cancer.

[0103] The cancer may be non-small cell lung cancer (NSCLC). The cancer may be a breast cancer. The breast cancer may be a BRCA1 cancer. The cancer may be a lung cancer, colorectal cancer, prostate cancer, ovarian cancer, esophageal cancer, breast cancer, lymphoma, or leukemia.

[0104] The method of diagnosing or prognosing the cancer may have a sensitivity of at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method of diagnosing or prognosing the cancer may have a specificity of at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

[0105] The may further comprise administering a therapeutic drug to the subject. The may further comprise modifying a therapeutic regimen. Modifying the therapeutic regimen may comprise terminating the therapeutic regimen. Modifying the therapeutic regimen may comprise increasing a dosage or frequency of the therapeutic regimen. Modifying the therapeutic regimen may comprise decreasing a dosage or frequency of the therapeutic regimen. Modifying the therapeutic regimen may comprise starting the therapeutic regimen.

[0106] Further disclosed herein are methods of determining a therapeutic region for the treatment of a condition in a subject in need thereof. The method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information may be derived from regions that are mutated in at least 80% of a population of subjects afflicted with a condition; and (b) determining a therapeutic regimen for a condition in the subject based on the sequence information.

[0107] The regions that are mutated may comprise a total size of less than 1.5 Mb of the genome. The regions that are mutated may comprise a total size of less than 1 Mb of the genome. The regions that are mutated may comprise a total size of less than 500 kb of the genome. The regions that are mutated may comprise a total size of less than 350 kb of the genome. The regions that are mutated may comprise a total

size of less than 300 kb of the genome. The regions that are mutated may comprise a total size of less than 250 kb of the genome. The regions that are mutated may comprise a total size of less than 200 kb of the genome. The regions that are mutated may comprise a total size of less than 150 kb of the genome. The regions that are mutated may comprise a total size of less than 100 kb of the genome. The regions that are mutated may comprise a total size of less than 50 kb of the genome. The regions that are mutated may comprise a total size of less than 40 kb of the genome. The regions that are mutated may comprise a total size of less than 30 kb of the genome. The regions that are mutated may comprise a total size of less than 20 kb of the genome. The regions that are mutated may comprise a total size of less than 10 kb of the genome.

[0108] The regions that are mutated may comprise a total size between 100 kb-300 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-200 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-150 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-100 kb of the genome. The regions that are mutated may comprise a total size between 5 kb-75 kb of the genome. The regions that are mutated may comprise a total size between 1 kb-50 kb of the genome.

[0109] The sequence information may be derived from 2 or more regions. The sequence information may be derived from 3 or more regions. The sequence information may be derived from 4 or more regions. The sequence information may be derived from 5 or more regions. The sequence information may be derived from 6 or more regions. The sequence information may be derived from 7 or more regions. The sequence information may be derived from 8 or more regions. The sequence information may be derived from 9 or more regions. The sequence information may be derived from 10 or more regions. The sequence information may be derived from 20 or more regions. The sequence information may be derived from 30 or more regions. The sequence information may be derived from 40 or more regions. The sequence information may be derived from 50 or more regions. The sequence information may be derived from 60 or more regions. The sequence information may be derived from 70 or more regions. The sequence information may be derived from 80 or more regions. The sequence information may be derived from 90 or more regions. The sequence information may be derived from 100 or more regions.

[0110] The population of subjects afflicted with the cancer may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA).

[0111] The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 60% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 70% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 80% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 90% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 95% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 99% of the population of subjects afflicted with the cancer.

[0112] The sequence information may be derived from regions that may be mutated in at least 65% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 70% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 75% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 80% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 85% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 90% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 95% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that may be mutated in at least 99% of the population of subjects afflicted with the cancer.

[0113] Obtaining the sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more lncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof.

[0114] Alternatively, or additionally, obtaining the sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof.

[0115] In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1.

[0116] The method may further comprise detecting mutations in the regions based on the sequencing information. Determining the therapeutic regimen may be based on the detection of the mutations.

[0117] The condition may be a cancer. The cancer may be a solid tumor. The solid tumor may be non-small cell lung cancer (NSCLC). The cancer may be a breast cancer. The breast cancer may be a BRCA1 cancer. The cancer may be a lung cancer, colorectal cancer, prostate cancer, ovarian cancer, esophageal cancer, breast cancer, lymphoma, or leukemia.

[0118] Further disclosed herein are methods of assessing tumor burden in a subject in need thereof. The method may comprise (a) obtaining sequence information on cell-free nucleic acids derived from a sample from the subject; (b) using a computer readable medium to determine quantities of

circulating tumor DNA (ctDNA) in the sample; (c) assessing tumor burden based on the quantities of ctDNA; and (d) reporting the tumor burden to the subject or a representative of the subject.

[0119] Determining quantities of ctDNA may comprise determining absolute quantities of ctDNA. Determining quantities of ctDNA may comprise determining relative quantities of ctDNA. Determining quantities of ctDNA may be performed by counting sequence reads pertaining to the ctDNA. Determining quantities of ctDNA may be performed by quantitative PCR. Determining quantities of ctDNA may be performed by digital PCR. Determining quantities of ctDNA may comprise counting sequencing reads of the ctDNA.

[0120] Determining quantities of ctDNA may be performed by molecular barcoding of the ctDNA. Molecular barcoding of the ctDNA may comprise attaching adaptors to one or more ends of the ctDNA. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0121] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence. The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0122] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0123] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0124] Adaptors may be attached to one end of a nucleic acid from a sample. The nucleic acids may be DNA. The DNA may be cell-free DNA (cfDNA). The DNA may be circulating tumor DNA (ctDNA). The nucleic acids may be RNA. Adaptors may be attached to both ends of the nucleic acid. Adaptors may be attached to one or more ends of a single-stranded nucleic acid. Adaptors may be attached to one or more ends of a double-stranded nucleic acid.

[0125] Adaptors may be attached to the nucleic acid by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the nucleic acid by primer extension. Adaptors may be attached to the nucleic acid by reverse transcription. Adaptors may be attached to the nucleic acids by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the nucleic acid. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the nucleic acid.

[0126] The sequence information may comprise information related to one or more genomic regions. The sequence information may comprise information related to at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 100, 200, 300 genomic regions. The genomic regions may comprise genes, exonic regions, intronic regions, untranslated regions, non-coding regions or a combination thereof.

[0127] The genomic regions may comprise two or more of exonic regions, intronic regions, and untranslated regions. The genomic regions may comprise at least one exonic region and at least one intronic region. At least 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, or 25% of the genomic regions may comprise intronic regions. At least 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, or 25% of the genomic regions may comprise untranslated regions. At least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may comprise exonic regions. At least less

than about 97%, 95%, 93%, 90%, 87%, 85%, 83%, 80%, 75%, 70%, 65%, 60%, 55%, 50% of the genomic regions may comprise exonic regions.

[0128] The genomic regions may comprise less than 1.5 megabases (Mb) of the genome. The genomic regions may comprise less than 1 Mb of the genome.

[0129] The genomic regions may comprise less than 500 kilobases (kb) of the genome.

[0130] The genomic regions may comprise less than 350 kb of the genome. The genomic regions may comprise less than 300 kb of the genome. The genomic regions may comprise less than 250 kb of the genome. The genomic regions may comprise less than 200 kb of the genome. The genomic regions may comprise less than 150 kb of the genome. The genomic regions may comprise less than 100 kb of the genome. The genomic regions may comprise less than 50 kb of the genome. The genomic regions may comprise less than 40 kb, 30 kb, 20 kb, or 10 kb of the genome.

[0131] The genomic regions may comprise between 100 kb to 300 kb of the genome. The genomic regions may comprise between 100 kb to 200 kb of the genome. The genomic regions may comprise between 10 kb to 300 kb of the genome. The genomic regions may comprise between 10 kb to 300 kb of the genome. The genomic regions may comprise between 10 kb to 200 kb of the genome. The genomic regions may comprise between 10 kb to 150 kb of the genome. The genomic regions may comprise between 10 kb to 100 kb of the genome. The genomic regions may comprise between 10 kb to 75 kb of the genome. The genomic regions may comprise between 5 kb to 70 kb of the genome. The genomic regions may comprise between 1 kb to 50 kb of the genome.

[0132] The sequence information may comprise information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions.

[0133] The sequence information may comprise information pertaining to a plurality of genomic regions.

[0134] The plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects.

[0135] The total size of the genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The total size of the genomic regions of the selector set may be between 100 kb to 300 kb of the genome.

[0136] The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 2.

[0137] Obtaining sequence information may comprise performing massively parallel sequencing. Massively parallel sequencing may be performed on a subset of a genome of the cell-free nucleic acids from the sample.

[0138] The subset of the genome may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, 150 kb, 100 kb, 75 kb, 50 kb, 40 kb, 30 kb, 20 kb, 10 kb, or 5 kb of the genome. The subset of the genome may comprise between 100 kb to 300 kb of the genome. The subset of the genome may comprise between 100 kb to 200 kb of the genome. The subset of the genome may comprise between 10 kb to 300 kb of the genome. The subset of the genome may comprise between 10 kb to 200 kb of the genome. The subset of the genome may comprise between 10 kb to 100 kb of the genome. The subset of the genome may comprise between 5 kb to 100 kb of the genome. The subset of the genome may comprise between 5 kb to 70 kb of the genome. The subset of the genome may comprise between 1 kb to 50 kb of the genome.

[0139] The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from the subject. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from two or more subjects. The two or more samples may be the same type of sample. The two or more samples may be two different types of sample. The two or more samples may be obtained at the same time point. The two or more samples may be obtained at two or more time points.

[0140] Determining the quantities of ctDNA may comprise detecting one or more mutations. Determining the quantities of ctDNA may comprise detecting two or more different types of mutations. The types of mutations include, but are not limited to, SNVs, indels, fusions, breakpoints, structural variants, variable number of tandem repeats, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, or a combination thereof in selected regions of the subject's genome. Determining the quantities of ctDNA may comprise detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Determining the quantities of ctDNA may comprise detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Determining the quantities of ctDNA may comprise detecting at least one SNV, indel, copy number variant, and rearrangement in selected regions of the subject's genome.

[0141] In some instances, determining the quantities of ctDNA does comprise performing digital PCR (dPCR). Determining the quantities of ctDNA may comprise applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set.

[0142] The selector set may comprise a plurality of genomic regions comprising one or more mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of genomic regions comprise two or more different types of mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in at least about 60% of cancer subjects from population of cancer subjects.

[0143] The representative of the subject may be a healthcare provider. The healthcare provider may be a nurse, physician, medical technician, or hospital personnel. The representative of the subject may be a family member of the subject. The representative of the subject may be a legal guardian of the subject.

[0144] Further disclosed herein are methods of determining a disease state of a cancer in a subject. The method may comprise (a) obtaining a quantity of circulating tumor DNA

(ctDNA) in a sample from the subject; (b) obtaining a volume of a tumor in the subject; and (c) determining a disease state of a cancer in the subject based on a ratio of the quantity of ctDNA to the volume of the tumor. A high ctDNA to volume ratio may be indicative of radiographically occult disease. A low ctDNA to volume ratio may be indicative of non-malignant state.

[0145] The method may further comprise modifying a diagnosis or prognosis of the cancer based on the ratio of the quantity of the ctDNA to the volume of the tumor. The method may comprise diagnosing a stage of the cancer based on the ratio of the quantity of the ctDNA to the volume of the tumor. Modifying the diagnosis may comprise changing the stage of the cancer based on the ratio of the quantity of the ctDNA to the volume of the tumor. For example, a subject may be diagnosed with a stage III cancer. However, a low ratio of the quantity of the ctDNA to the volume of the tumor may result in adjusting the diagnosis of the cancer to a stage I or II cancer. Modifying a prognosis of the cancer may comprise changing the predicted outcome or status of the cancer. For example, a doctor may predict that a cancer in the subject is in remission based on the tumor volume. However, a high ratio of the quantity of the ctDNA to the volume of the tumor may result in a prediction that the cancer is recurrent.

[0146] Obtaining the volume of the tumor may comprise obtaining an image of the tumor. Obtaining the volume of the tumor may comprise obtaining a CT scan of the tumor.

[0147] Obtaining the quantity of ctDNA may comprise PCR. Obtaining the quantity of ctDNA may comprise digital PCR. Obtaining the quantity of ctDNA may comprise quantitative PCR.

[0148] Obtaining the quantity of ctDNA may comprise obtaining sequencing information on the ctDNA. The sequencing information may comprise information relating to one or more genomic regions based on a selector set.

[0149] Obtaining the quantity of ctDNA may comprise hybridization of the ctDNA to an array. The array may comprise a plurality of probes for selective hybridization of one or more genomic regions based on a selector set. The selector set may comprise one or more genomic regions from Table 2. The selector set may comprise one or more genomic regions comprising one or more mutations, wherein the one or more mutations may be present in a population of subjects suffering from a cancer. The selector set may comprise a plurality of genomic regions comprising a plurality of mutations, wherein the plurality of mutations may be present in at least 60% of a population of subjects suffering from a cancer.

[0150] Further disclosed herein are methods of detecting stage I cancer in a subject in need thereof. The method may comprise (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced may be based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA; and (c) detecting a stage I cancer in the sample based on the quantity of the cell-free DNA.

[0151] Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR.

[0152] Determining quantities of cell-free DNA (cfDNA) may be performed by molecular barcoding of the cfDNA. Molecular barcoding of the cfDNA may comprise attaching adaptors to one or more ends of the cfDNA. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0153] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence. The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0154] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The

sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0155] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0156] Adaptors may be attached to one end of the cfDNA. Adaptors may be attached to both ends of the cfDNA. Adaptors may be attached to one or more ends of a single-stranded cfDNA. Adaptors may be attached to one or more ends of a double-stranded cfDNA.

[0157] Adaptors may be attached to the cfDNA by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the cfDNA by primer extension. Adaptors may be attached to the cfDNA by reverse transcription. Adaptors may be attached to the cfDNA by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the cfDNA. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the cfDNA.

[0158] Sequencing may comprise massively parallel sequencing. Sequencing may comprise shotgun sequencing.

[0159] The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2.

[0160] At least 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2.

[0161] The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer.

[0162] The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 100 kb, 90 kb, 80 kb, 70 kb, 60 kb, 50 kb, 40 kb, 30 kb, 20 kb, 10 kb, 5 kb, or 1 kb of a genome.

[0163] The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 75 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 50 kb of a genome.

[0164] The method of detecting the stage I cancer may have a sensitivity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage I cancer may have a sensitivity of at least 60%. The method of detecting the stage I cancer may have a sensitivity of at least 70%. The method of detecting the stage I cancer may have a sensitivity of at least 80%. The method of detecting the stage I cancer may have a sensitivity of at least 90%. The method of detecting the stage I cancer may have a sensitivity of at least 95%.

[0165] The method of detecting the stage I cancer may have a specificity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage I cancer may have a specificity of at least 60%. The method of detecting the stage I cancer may have a specificity of at least 70%. The method of detecting the stage I cancer may have a specificity of at least 80%. The method of detecting the stage I cancer may have a specificity of at least 90%. The method of detecting the stage I cancer may have a specificity of at least 95%.

[0166] The method may detect at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage I cancer. The method may detect at least 50% or more of stage I cancer. The method may detect at least 60% or more of stage I cancer. The method may detect at least 70% or more of stage I cancer. The method may detect at least 75% or more of stage I cancer.

[0167] Further disclosed herein are methods of detecting stage II cancer. The method may comprise (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced may be based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA; and (c) detecting a stage II cancer in the sample based on the quantity of the cell-free DNA.

[0168] Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR.

[0169] Determining quantities of cell-free DNA (cfDNA) may be performed by molecular barcoding of the cfDNA. Molecular barcoding of the cfDNA may comprise attaching adaptors to one or more ends of the cfDNA. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0170] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence.

The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0171] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0172] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0173] Adaptors may be attached to one end of the cfDNA. Adaptors may be attached to both ends of the cfDNA. Adaptors may be attached to one or more ends of a single-stranded cfDNA. Adaptors may be attached to one or more ends of a double-stranded cfDNA.

[0174] Adaptors may be attached to the cfDNA by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the cfDNA by primer extension. Adaptors may be attached to the cfDNA by reverse transcription. Adaptors may be attached to the cfDNA by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the cfDNA. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the cfDNA.

[0175] Sequencing may comprise massively parallel sequencing. Sequencing may comprise shotgun sequencing.

[0176] The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2.

[0177] At least 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2.

[0178] The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer.

[0179] The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 100 kb, 90 kb, 80 kb, 70 kb, 60 kb, 50 kb, 40 kb, 30 kb, 20 kb, 10 kb, 5 kb, or 1 kb of a genome.

[0180] The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 75 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 50 kb of a genome.

[0181] The method of detecting the stage II cancer may have a sensitivity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage II cancer may have a sensitivity of at least 60%. The method of detecting the stage II cancer may have a sensitivity of at least 70%. The method of detecting the stage II cancer may have a sensitivity of at least 80%. The method of detecting the stage II cancer may have a sensitivity of at least 90%. The method of detecting the stage II cancer may have a sensitivity of at least 95%.

[0182] The method of detecting the stage II cancer may have a specificity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage II cancer may have a specificity of at least 60%. The method of detecting the stage II cancer may have a specificity of at least 70%. The method of detecting the stage II cancer may have a specificity of at least 80%. The method of detecting the stage II cancer may have a specificity of at least 90%. The method of detecting the stage II cancer may have a specificity of at least 95%.

[0183] The method may detect at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%,

85%, 87%, 90%, 92%, 95%, 97% or more of stage II cancer. The method may detect at least 50% or more of stage II cancer. The method may detect at least 60% or more of stage II cancer. The method may detect at least 70% or more of stage II cancer. The method may detect at least 75% or more of stage II cancer. The method may detect at least 80% or more of stage II cancer. The method may detect at least 85% or more of stage II cancer. The method may detect at least 90% or more stage II cancer.

[0184] Further disclosed herein are methods of detecting stage III cancer in a subject in need thereof. The method may comprise (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced may be based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA; and (c) detecting a stage III cancer in the sample based on the quantity of the cell-free DNA.

[0185] Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR.

[0186] Determining quantities of cell-free DNA (cfDNA) may be performed by molecular barcoding of the cfDNA. Molecular barcoding of the cfDNA may comprise attaching adaptors to one or more ends of the cfDNA. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0187] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence. The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0188] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0189] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0190] Adaptors may be attached to one end of the cfDNA. Adaptors may be attached to both ends of the cfDNA. Adaptors may be attached to one or more ends of a single-stranded cfDNA. Adaptors may be attached to one or more ends of a double-stranded cfDNA.

[0191] Adaptors may be attached to the cfDNA by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the cfDNA by primer extension. Adaptors may be attached to the cfDNA by reverse transcription. Adaptors may be attached to the cfDNA by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the cfDNA. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the cfDNA.

[0192] Sequencing may comprise massively parallel sequencing. Sequencing may comprise shotgun sequencing.

[0193] The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2.

[0194] At least 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2.

[0195] The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%,

70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer.

[0196] The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 100 kb, 90 kb, 80 kb, 70 kb, 60 kb, 50 kb, 40 kb, 30 kb, 20 kb, 10 kb, 5 kb, or 1 kb of a genome.

[0197] The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 75 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 50 kb of a genome.

[0198] The method of detecting the stage III cancer may have a sensitivity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage III cancer may have a sensitivity of at least 60%. The method of detecting the stage III cancer may have a sensitivity of at least 70%. The method of detecting the stage III cancer may have a sensitivity of at least 80%. The method of detecting the stage III cancer may have a sensitivity of at least 90%. The method of detecting the stage III cancer may have a sensitivity of at least 95%.

[0199] The method of detecting the stage III cancer may have a specificity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage III cancer may have a specificity of at least 60%. The method of detecting the stage III cancer may have a specificity of at least 70%. The method of detecting the stage III cancer may have a specificity of at least 80%. The method of detecting the stage III cancer may have a specificity of at least 90%. The method of detecting the stage III cancer may have a specificity of at least 95%.

[0200] The method may detect at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage III cancer. The method may detect at least 50% or more of stage III cancer. The method may detect at least 60% or more of stage III cancer. The method may detect at least 70% or more of stage III cancer. The method may detect at least 75% or more of stage III cancer. The method may detect at least 80% or more of stage III cancer. The method may detect at least 85% or more of stage III cancer. The method may detect at least 90% or more of stage III cancer.

[0201] Further disclosed herein is a method of detecting stage IV cancer in a subject in need thereof. The method may comprise (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced may be based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA; and (c) detecting a stage IV cancer in the sample based on the quantity of the cell-free DNA.

[0202] Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR.

[0203] Determining quantities of cell-free DNA (cfDNA) may be performed by molecular barcoding of the cfDNA. Molecular barcoding of the cfDNA may comprise attaching adaptors to one or more ends of the cfDNA. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0204] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence. The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0205] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be

optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0206] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0207] Adaptors may be attached to one end of the cfDNA. Adaptors may be attached to both ends of the cfDNA. Adaptors may be attached to one or more ends of a single-stranded cfDNA. Adaptors may be attached to one or more ends of a double-stranded cfDNA.

[0208] Adaptors may be attached to the cfDNA by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the cfDNA by primer extension. Adaptors may be attached to the cfDNA by reverse transcription. Adaptors may be attached to the cfDNA by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the cfDNA. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the cfDNA.

[0209] Sequencing may comprise massively parallel sequencing. Sequencing may comprise shotgun sequencing.

[0210] The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2.

[0211] At least 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2.

[0212] The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer.

[0213] The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 100 kb, 90 kb, 80 kb, 70 kb, 60 kb, 50 kb, 40 kb, 30 kb, 20 kb, 10 kb, 5 kb, or 1 kb of a genome.

[0214] The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 10 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 75 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 50 kb of a genome.

[0215] The method of detecting the stage IV cancer may have a sensitivity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage IV cancer may have a sensitivity of at least 60%. The method of detecting the stage IV cancer may have a sensitivity of at least 70%. The method of detecting the stage IV cancer may have a sensitivity of at least 80%. The method of detecting the stage IV cancer may have a sensitivity of at least 90%. The method of detecting the stage IV cancer may have a sensitivity of at least 95%.

[0216] The method of detecting the stage IV cancer may have a specificity of at least 60%, 65%, 70%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method of detecting the stage IV cancer may have a specificity of at least 60%. The method of detecting the stage IV cancer may have a specificity of at least 70%. The method of detecting the stage IV cancer may have a specificity of at least 80%. The method of detecting the stage IV cancer may have a specificity of at least 90%. The method of detecting the stage IV cancer may have a specificity of at least 95%.

[0217] The method may detect at least 50%, 52%, 55%, 57%, 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage IV cancer. The method may detect at least 50% or more of stage IV cancer. The method may detect at least 60% or more of stage IV cancer. The method may detect at least 70% or more of stage IV cancer. The method may detect at least 75% or more of stage IV cancer. The method may detect at least 80% or more of stage IV cancer. The method may detect at least 85% or more of stage IV cancer. The method may detect at least 90% or more of stage IV cancer.

[0218] Further disclosed herein are methods of producing a selector set. The method may comprise (a) identifying genomic regions comprising mutations in one or more subjects from a population of subjects suffering from the cancer; (b) ranking the genomic regions based on a Recurrence Index (RI), wherein the RI of the genomic region is determined by dividing a number of subjects or tumors with mutations in the genomic region by a size of the genomic region; and (c) producing a selector set comprising one or more genomic regions based on the RI.

[0219] At least a subset of the genomic regions that are ranked may be exon regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions that are ranked may comprise exon regions. At least 30% of the genomic regions that are ranked may comprise exon regions. At least 40% of the genomic regions that are ranked may comprise exon regions. At least 50% of the genomic regions that are ranked may comprise exon regions. At least 60% of the genomic regions that are ranked may comprise exon regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions that are ranked

may comprise exon regions. Less than 97% of the genomic regions that are ranked may comprise exon regions. Less than 92% of the genomic regions that are ranked may comprise exon regions. Less than 84% of the genomic regions that are ranked may comprise exon regions. Less than 75% of the genomic regions that are ranked may comprise exon regions. Less than 65% of the genomic regions that are ranked may comprise exon regions.

[0220] At least a subset of the genomic regions of the selector set may comprise exon regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions of the selector set may comprise exon regions. At least 30% of the genomic regions of the selector set may comprise exon regions. At least 40% of the genomic regions of the selector set may comprise exon regions. At least 50% of the genomic regions of the selector set may comprise exon regions. At least 60% of the genomic regions of the selector set may comprise exon regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions of the selector set may comprise exon regions. Less than 97% of the genomic regions of the selector set may comprise exon regions. Less than 92% of the genomic regions of the selector set may comprise exon regions. Less than 84% of the genomic regions of the selector set may comprise exon regions. Less than 75% of the genomic regions of the selector set may comprise exon regions. Less than 65% of the genomic regions of the selector set may comprise exon regions.

[0221] At least a subset of the genomic regions that are ranked may be intron regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions that are ranked may comprise intron regions. At least 30% of the genomic regions that are ranked may comprise intron regions. At least 40% of the genomic regions that are ranked may comprise intron regions. At least 50% of the genomic regions that are ranked may comprise intron regions. At least 60% of the genomic regions that are ranked may comprise intron regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions that are ranked may comprise intron regions. Less than 97% of the genomic regions that are ranked may comprise intron regions. Less than 92% of the genomic regions that are ranked may comprise intron regions. Less than 84% of the genomic regions that are ranked may comprise intron regions. Less than 75% of the genomic regions that are ranked may comprise intron regions. Less than 65% of the genomic regions that are ranked may comprise intron regions.

[0222] At least a subset of the genomic regions of the selector set may comprise intron regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions of the selector set may comprise intron regions. At least 30% of the genomic regions of the selector set may comprise intron regions. At least 40% of the genomic regions of the selector set may comprise intron regions. At least 50% of the genomic regions of the selector set may comprise intron regions. At least 60% of the genomic regions of the selector set may comprise intron regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions of the selector set may comprise intron regions. Less than 97% of the genomic regions of the selector set may comprise intron regions. Less than 92% of the genomic regions of the selector set may comprise intron regions. Less than 84% of the genomic regions of the selector set may comprise intron regions. Less than 75% of the genomic regions of the selector set may comprise intron regions. Less than 65% of the genomic regions of the selector set may comprise intron regions.

[0223] At least a subset of the genomic regions that are ranked may be untranslated regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions that are ranked may comprise untranslated regions. At least 30% of the genomic regions that are ranked may comprise untranslated regions. At least 40% of the genomic regions that are ranked may comprise untranslated regions. At least 50% of the genomic regions that are ranked may comprise untranslated regions. At least 60% of the genomic regions that are ranked may comprise untranslated regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions that are ranked may comprise untranslated regions. Less than 97% of the genomic regions that are ranked may comprise untranslated regions. Less than 92% of the genomic regions that are ranked may comprise untranslated regions. Less than 84% of the genomic regions that are ranked may comprise untranslated regions. Less than 75% of the genomic regions that are ranked may comprise untranslated regions. Less than 65% of the genomic regions that are ranked may comprise untranslated regions.

[0224] At least a subset of the genomic regions of the selector set may comprise untranslated regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions of the selector set may comprise untranslated regions. At least 30% of the genomic regions of the selector set may comprise untranslated regions. At least 40% of the genomic regions of the selector set may comprise untranslated regions. At least 50% of the genomic regions of the selector set may comprise untranslated regions. At least 60% of the genomic regions of the selector set may comprise untranslated regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions of the selector set may comprise untranslated regions. Less than 97% of the genomic regions of the selector set may comprise untranslated regions. Less than 92% of the genomic regions of the selector set may comprise untranslated regions. Less than 84% of the genomic regions of the selector set may comprise untranslated regions. Less than 75% of the genomic regions of the selector set may comprise untranslated regions. Less than 65% of the genomic regions of the selector set may comprise untranslated regions.

[0225] At least a subset of the genomic regions that are ranked may be non-coding regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions that are ranked may comprise non-coding regions. At least 30% of the genomic regions that are ranked may comprise non-coding regions. At least 40% of the genomic regions that are ranked may comprise non-coding regions. At least 50% of the genomic regions that are ranked may comprise non-coding regions. At least 60% of the genomic regions that are ranked may comprise non-coding regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%,

67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions that are ranked may comprise non-coding regions. Less than 97% of the genomic regions that are ranked may comprise non-coding regions. Less than 92% of the genomic regions that are ranked may comprise non-coding regions. Less than 84% of the genomic regions that are ranked may comprise non-coding regions. Less than 75% of the genomic regions that are ranked may comprise non-coding regions. Less than 65% of the genomic regions that are ranked may comprise non-coding regions.

[0226] At least a subset of the genomic regions of the selector set may comprise non-coding regions. At least 20%, 2%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97% of the genomic regions of the selector set may comprise non-coding regions. At least 30% of the genomic regions of the selector set may comprise non-coding regions. At least 40% of the genomic regions of the selector set may comprise non-coding regions. At least 50% of the genomic regions of the selector set may comprise non-coding regions. At least 60% of the genomic regions of the selector set may comprise non-coding regions. Less than 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50%, 45%, or 40% of the genomic regions of the selector set may comprise non-coding regions. Less than 97% of the genomic regions of the selector set may comprise non-coding regions. Less than 92% of the genomic regions of the selector set may comprise non-coding regions. Less than 84% of the genomic regions of the selector set may comprise non-coding regions. Less than 75% of the genomic regions of the selector set may comprise non-coding regions. Less than 65% of the genomic regions of the selector set may comprise non-coding regions.

[0227] Producing the selector set based on the RI may comprise selecting genomic regions that have a recurrence index in the top 60th, 65th, 70th, 72nd, 75th, 77th, 80th, 82nd, 85th, 87th, 90th, 92nd, 95th, or 97th or greater percentile. Producing the selector set based on the RI may comprise selecting genomic regions that have a recurrence index in the top 80th or greater percentile. Producing the selector set based on the RI may comprise selecting genomic regions that have a recurrence index in the top 70th or greater percentile. Producing the selector set based on the RI may comprise selecting genomic regions that have a recurrence index in the top 90th or greater percentile.

[0228] Producing the selector set further may comprise selecting genomic regions that result in the largest reduction in a number of subjects with one mutation in the genomic region.

[0229] Producing the selector set may comprise applying an algorithm to a subset of the ranked genomic regions. The algorithm may be applied 2, 3, 4, 5, 6, 7, 8, 9, 10 or more times. The algorithm may be applied two or more times. The algorithm may be applied three or more times.

[0230] Producing the selector set may comprise selecting genomic regions that maximize a median number of mutations per subject of the selector set. Producing the selector set may comprise selecting genomic regions that maximize the number of subjects in the selector set.

[0231] Producing the selector set may comprise selecting genomic regions that minimize the total size of the genomic regions.

[0232] The selector set may comprise information pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The selector set may comprise information pertaining to a plurality of genomic regions comprising 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more mutations present in at least one subject suffering from a cancer. The selector set may comprise information pertaining to a plurality of genomic regions comprising 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or more mutations present in at least one subject suffering from a cancer.

[0233] The selector set may comprise information pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more subjects suffering from a cancer. The one or more mutations within the genomic regions may be present in at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or more subjects suffering from a cancer.

[0234] The selector set may comprise information pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 1%, 2%, 3%, 4%, 5%, 6%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20% or more subjects from a population of subjects suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more subjects from a population of subjects suffering from a cancer.

[0235] The selector set may comprise sequence information pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The selector set may comprise sequence information pertaining to a plurality of genomic regions comprising 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more mutations present in at least one subject suffering from a cancer. The selector set may comprise sequence information pertaining to a plurality of genomic regions comprising 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or more mutations present in at least one subject suffering from a cancer.

[0236] The selector set may comprise sequence information pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more subjects suffering from a cancer. The one or more mutations within the genomic regions may be present in at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or more subjects suffering from a cancer.

[0237] The selector set may comprise sequence information pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 1%, 2%, 3%, 4%, 5%, 6%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20% or more subjects from a population of subjects suffering from a cancer. The one or more

mutations within the plurality of genomic regions may be present in at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more subjects from a population of subjects suffering from a cancer.

[0238] The selector set may comprise genomic coordinates pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The selector set may comprise genomic coordinates pertaining to a plurality of genomic regions comprising 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more mutations present in at least one subject suffering from a cancer. The selector set may comprise genomic coordinates pertaining to a plurality of genomic regions comprising 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or more mutations present in at least one subject suffering from a cancer.

[0239] The selector set may comprise genomic coordinates pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more subjects suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or more subjects suffering from a cancer.

[0240] The selector set may comprise genomic coordinates pertaining to a plurality of genomic regions comprising one or more mutations present in at least one subject suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 1%, 2%, 3%, 4%, 5%, 6%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20% or more subjects from a population of subjects suffering from a cancer. The one or more mutations within the plurality of genomic regions may be present in at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more subjects from a population of subjects suffering from a cancer.

[0241] The selector set may comprise genomic regions comprising one or more types of mutations. The selector set may comprise genomic regions comprising two or more types of mutations. The selector set may comprise genomic regions comprising three or more types of mutations. The selector set may comprise genomic regions comprising four or more types of mutations. The types of mutations may include, but are not limited to, single nucleotide variants (SNVs), insertions/deletions (indels), rearrangements, and copy number variants (CNVs).

[0242] The selector set may comprise genomic regions comprising two or more different types of mutations selected from a group consisting of single nucleotide variants (SNVs), insertions/deletions (indels), rearrangements, and copy number variants (CNVs). The selector set may comprise genomic regions comprising three or more different types of mutations selected from a group consisting of single nucleotide variants (SNVs), insertions/deletions (indels), rearrangements, and copy number variants (CNVs). The selector set may comprise genomic regions comprising four or more different types of mutations selected from a group consisting of single nucleotide variants (SNVs), insertions/deletions (indels), rearrangements, and copy number variants (CNVs).

[0243] The selector set may comprise a genomic region comprising at least one SNV and a genomic region comprising at least one other type of mutation. The selector set may comprise a genomic region comprising at least one SNV and a genomic region comprising at least one indel. The selector set may comprise a genomic region comprising at least one SNV and a genomic region comprising at least one rearrangement. The selector set may comprise a genomic region comprising at least one SNV and a genomic region comprising at least one CNV.

[0244] The selector set may comprise a genomic region comprising at least one indel and a genomic region comprising at least one other type of mutation. The selector set may comprise a genomic region comprising at least one indel and a genomic region comprising at least one SNV. The selector set may comprise a genomic region comprising at least one indel and a genomic region comprising at least one rearrangement. The selector set may comprise a genomic region comprising at least one indel and a genomic region comprising at least one CNV.

[0245] The selector set may comprise a genomic region comprising at least one rearrangement. The selector set may comprise a genomic region comprising at least one rearrangement and a genomic region comprising at least one other type of mutation. The selector set may comprise a genomic region comprising at least one rearrangement and a genomic region comprising at least one SNV. The selector set may comprise a genomic region comprising at least one rearrangement and a genomic region comprising at least one indel. The selector set may comprise a genomic region comprising at least one rearrangement and a genomic region comprising at least one CNV.

[0246] The selector set may comprise a genomic region comprising at least one CNV and a genomic region comprising at least one other type of mutation. The selector set may comprise a genomic region comprising at least one CNV and a genomic region comprising at least one SNV. The selector set may comprise a genomic region comprising at least one CNV and a genomic region comprising at least one indel. The selector set may comprise a genomic region comprising at least one CNV and a genomic region comprising at least one rearrangement.

[0247] At least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, or 20% of the genomic regions of the selector set may comprise a SNV. At least about 25%, 30%, 35%, 40%, 45%, 50%, 55%, or 60% of the genomic regions of the selector set may comprise a SNV. At least about 10% of the genomic regions of the selector set may comprise a SNV. At least about 15% of the genomic regions of the selector set may comprise a SNV. At least about 20% of the genomic regions of the selector set may comprise a SNV. At least about 30% of the genomic regions of the selector set may comprise a SNV. At least about 40% of the genomic regions of the selector set may comprise a SNV. At least about 50% of the genomic regions of the selector set may comprise a SNV. At least about 60% of the genomic regions of the selector set may comprise a SNV.

[0248] Less than 99%, 98%, 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50% of the genomic regions of the selector set may comprise a SNV. Less than 97% of the genomic regions of the selector set may comprise a SNV. Less than 95% of the genomic regions of the selector set may comprise a SNV. Less than 90% of the genomic regions of the selector set may comprise a SNV. Less than 85% of the

genomic regions of the selector set may comprise a SNV. Less than 77% of the genomic regions of the selector set may comprise a SNV.

[0249] The genomic regions of the selector set may comprise between about 10% to about 95% SNVs. The genomic regions of the selector set may comprise between about 10% to about 90% SNVs. The genomic regions of the selector set may comprise between about 15% to about 95% SNVs. The genomic regions of the selector set may comprise between about 20% to about 95% SNVs. The genomic regions of the selector set may comprise between about 30% to about 95% SNVs. The genomic regions of the selector set may comprise between about 30% to about 90% SNVs. The genomic regions of the selector set may comprise between about 30% to about 85% SNVs. The genomic regions of the selector set may comprise between about 30% to about 80% SNVs.

[0250] At least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, or 20% of the genomic regions of the selector set may comprise an indel. At least about 25%, 30%, 35%, 40%, 45%, 50%, 55%, or 60% of the genomic regions of the selector set may comprise an indel. At least about 1% of the genomic regions of the selector set may comprise an indel. At least about 3% of the genomic regions of the selector set may comprise an indel. At least about 5% of the genomic regions of the selector set may comprise an indel. At least about 8% of the genomic regions of the selector set may comprise an indel. At least about 10% of the genomic regions of the selector set may comprise an indel. At least about 15% of the genomic regions of the selector set may comprise an indel. At least about 30% of the genomic regions of the selector set may comprise an indel.

[0251] Less than 99%, 98%, 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50% of the genomic regions of the selector set may comprise an indel. Less than 97% of the genomic regions of the selector set may comprise an indel. Less than 95% of the genomic regions of the selector set may comprise an indel. Less than 90% of the genomic regions of the selector set may comprise an indel. Less than 85% of the genomic regions of the selector set may comprise an indel. Less than 77% of the genomic regions of the selector set may comprise an indel.

[0252] The genomic regions of the selector set may comprise between about 10% to about 95% indels. The genomic regions of the selector set may comprise between about 10% to about 90% indels. The genomic regions of the selector set may comprise between about 10% to about 85% indels. The genomic regions of the selector set may comprise between about 10% to about 80% indels. The genomic regions of the selector set may comprise between about 10% to about 75% indels. The genomic regions of the selector set may comprise between about 10% to about 70% indels. The genomic regions of the selector set may comprise between about 10% to about 60% indels. The genomic regions of the selector set may comprise between about 10% to about 50% indels.

[0253] At least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, or 20% of the genomic regions of the selector set may comprise a rearrangement. At least about 1% of the genomic regions of the selector set may comprise a rearrangement. At least about 2% of the genomic regions of the selector set may comprise a rearrangement. At least about 3% of the genomic regions of the selector set may comprise a rearrangement. At least about 4% of the genomic regions of the selector set may comprise a rearrangement. At least about 5% of the genomic regions of the selector set may comprise a rearrangement.

[0254] At least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, or 20% of the genomic regions of the selector set may comprise a CNV. At least about 25%, 30%, 35%, 40%, 45%, 50%, 55%, or 60% of the genomic regions of the selector set may comprise a CNV. At least about 1% of the genomic regions of the selector set may comprise a CNV. At least about 3% of the genomic regions of the selector set may comprise a CNV. At least about 5% of the genomic regions of the selector set may comprise a CNV. At least about 8% of the genomic regions of the selector set may comprise a CNV. At least about 10% of the genomic regions of the selector set may comprise a CNV. At least about 15% of the genomic regions of the selector set may comprise a CNV. At least about 30% of the genomic regions of the selector set may comprise a CNV.

[0255] Less than 99%, 98%, 97%, 95%, 92%, 90%, 87%, 85%, 82%, 80%, 77%, 75%, 72%, 70%, 67%, 65%, 62%, 60%, 57%, 55%, 52%, 50% of the genomic regions of the selector set may comprise a CNV. Less than 97% of the genomic regions of the selector set may comprise a CNV. Less than 95% of the genomic regions of the selector set may comprise a CNV. Less than 90% of the genomic regions of the selector set may comprise a CNV. Less than 85% of the genomic regions of the selector set may comprise a CNV. Less than 77% of the genomic regions of the selector set may comprise a CNV.

[0256] The genomic regions of the selector set may comprise between about 5% to about 80% CNVs. The genomic regions of the selector set may comprise between about 5% to about 70% CNVs. The genomic regions of the selector set may comprise between about 5% to about 60% CNVs. The genomic regions of the selector set may comprise between about 5% to about 50% CNVs. The genomic regions of the selector set may comprise between about 5% to about 40% CNVs. The genomic regions of the selector set may comprise between about 5% to about 35% CNVs. The genomic regions of the selector set may comprise between about 5% to about 30% CNVs. The genomic regions of the selector set may comprise between about 5% to about 25% CNVs.

[0257] The selector set may be used to classify a sample from a subject. The selector set may be used to classify 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 or more samples from a subject. The selector set may be used to classify two or more samples from a subject.

[0258] The selector set may be used to classify one or more samples from one or more subjects. The selector set may be used to classify two or more samples from two or more subjects. The selector set may be used to classify a plurality of samples from 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or more subjects.

[0259] The samples may be the same type of sample. The samples may be two or more different types of samples. The sample may be a plasma sample. The sample may be a tumor sample. The sample may be a germline sample. The sample may comprise tumor-derived molecules. The sample may comprise non-tumor-derived molecules.

[0260] The selector set may classify the sample as tumor-containing. The selector set may classify the sample as tumor-free.

[0261] The selector set may be a personalized selector set. The selector set may be used to diagnose a cancer in a subject

in need thereof. The selector set may be used to prognosticate a status or outcome of a cancer in a subject in need thereof. The selector set may be used to determine a therapeutic regimen for treating a cancer in a subject in need thereof.

[0262] Alternatively, the selector set may be a universal selector set. The selector set may be used to diagnose a cancer in a plurality of subjects in need thereof. The selector set may be used to prognosticate a status or outcome of a cancer in a plurality of subjects in need thereof. The selector set may be used to determine a therapeutic regimen for treating a cancer in a plurality of subjects in need thereof.

[0263] The plurality of subjects may comprise 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, or 100 or more subjects. The plurality of subjects may comprise 5 or more subjects. The plurality of subjects may comprise 10 or more subjects. The plurality of subjects may comprise 25 or more subjects. The plurality of subjects may comprise 50 or more subjects. The plurality of subjects may comprise 75 or more subjects. The plurality of subjects may comprise 100 or more subjects.

[0264] The selector set may be used to classify one or more subjects based on one or more samples from the one or more subjects. The selector set may be used to classify a subject as a responder to a therapy. The selector set may be used to classify a subject as a non-responder to a therapy.

[0265] The selector set may be used to design a plurality of oligonucleotides. The plurality of oligonucleotides may selectively hybridize to one or more genomic regions identified by the selector set. At least two oligonucleotides may selectively hybridize to one genomic region. At least three oligonucleotides may selectively hybridize to one genomic region. At least four oligonucleotides may selectively hybridize to one genomic region.

[0266] An oligonucleotide of the plurality of oligonucleotides may be at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 nucleotides in length. An oligonucleotide may be at least about 20 nucleotides in length. An oligonucleotide may be at least about 30 nucleotides in length. An oligonucleotide may be at least about 40 nucleotides in length. An oligonucleotide may be at least about 45 nucleotides in length. An oligonucleotide may be at least about 50 nucleotides in length.

[0267] An oligonucleotide of the plurality of oligonucleotides may be less than or equal to 300, 275, 250, 225, 200, 190, 180, 170, 160, 150, 140, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, or 70 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be less than or equal to 200 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be less than or equal to 150 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be less than or equal to 110 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be less than or equal to 100 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be less than or equal to 80 nucleotides in length.

[0268] An oligonucleotide of the plurality of oligonucleotides may be between about 20 to 200 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 20 to 170 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 20 to 150 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 20 to 130 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 20 to 120 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 30 to 150 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 30 to 120 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 40 to 150 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 40 to 120 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 50 to 150 nucleotides in length. An oligonucleotide of the plurality of oligonucleotides may be between about 50 to 120 nucleotides in length.

[0269] An oligonucleotide of the plurality of oligonucleotides may be attached to a solid support. The solid support may be a bead. The bead may be a coated bead. The bead may be a streptavidin coated bead. The solid support may be an array. The solid support may be a glass slide.

[0270] Further disclosed herein are methods of producing a personalized selector set. The method may comprise (a) obtaining a genotype of a tumor in a subject; (b) identifying genomic regions comprising one or more mutations based on the genotype of the tumor; and (c) producing a selector set comprising at least one genomic region.

[0271] Obtaining the genotype of the tumor in the subject may comprise conducting a sequencing reaction on a sample from the subject. Sequencing may comprise whole genome sequencing. Sequencing may comprise whole exome sequencing.

[0272] Sequencing may comprise use of one or more adaptors. The adaptors may be attached to one or more nucleic acids from the sample. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0273] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence. The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular

barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0274] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0275] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0276] Adaptors may be attached to one end of a nucleic acid from a sample. The nucleic acids may be DNA. The DNA may be cell-free DNA (cfDNA). The DNA may be circulating tumor DNA (ctDNA). The nucleic acids may be RNA. Adaptors may be attached to both ends of the nucleic acid. Adaptors may be attached to one or more ends of a single-stranded nucleic acid. Adaptors may be attached to one or more ends of a double-stranded nucleic acid.

[0277] Adaptors may be attached to the nucleic acid by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the nucleic acid by primer extension. Adaptors may be attached to the nucleic acid by reverse transcription. Adaptors may be attached to the nucleic acids by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the nucleic acid. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the nucleic acid.

[0278] Identifying genomic regions comprising one or more mutations based on the genotype of the tumor may comprise determining a consensus sequence for the genomic region comprising the one or more mutations. Determining the consensus sequence may be based on the adaptors. Determining the consensus sequence may be based on the molecular barcode portion of the adaptor. Determining the consensus sequence may comprise analyzing sequence reads pertaining to a molecular barcode. Determining the consensus sequence may comprise determining a percentage of sequence reads with identical sequences based on the molecular barcode. Identifying genomic regions comprising one or more mutations may comprise producing a list of genomic regions based on a percentage of the consensus sequence. Producing the list of genomic regions may comprise selecting genomic regions with at least 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% consensus based on the molecular barcode. For example, sequence information may be arranged into molecular barcode families (e.g., sequences with identical molecular barcodes are grouped together). Analysis of a molecular barcode family may reveal two different sequences. 1000 sequence reads may be associated with a first sequence and 10 sequence reads may be associated with a second sequence. The dominant sequence (e.g., the first sequence) may have a consensus of 99% (e.g., (1000 divided by 1010) times 100%). The list of genomic regions may comprise the dominant sequence of the genomic region. The list of genomic regions may comprise genomic regions with 90% consensus based on the molecular barcode. The list of genomic regions may comprise genomic regions with 95% consensus based on the molecular barcode. The list of genomic regions may comprise genomic regions with 98% consensus based on the molecular barcode. The list of genomic regions may comprise genomic regions with 100% sequence consensus based on the molecular barcode. Identifying genomic regions comprising one or more mutations based on the genotype of the tumor may comprise producing a list of genomic regions ranked by a percentage of their sequence consensus.

[0279] Identifying genomic regions comprising one or more mutations based on the genotype of the tumor may comprise calculating a fractional abundance of the genomic region. Identifying genomic regions comprising one or more mutations based on the genotype of the tumor may comprise calculating a fractional abundance of the genomic region from the list of genomic regions ranked by the percentage of their sequence consensus. The fractional abundance may be calculated by dividing a number of sequence reads that pertain to a genomic region with the one or more mutations by a total number of sequence reads for the genomic regions. For example, a genomic region may comprise exon 2 of gene X. A total number of sequence reads pertaining to the genomic region may be 1000, with 100 of the sequence reads containing an insertion in exon 2 of gene X. The fractional abundance of the genomic region containing the insertion in exon 2 of gene X would be 0.1 (e.g., 100 sequence reads divided by 1000). Identifying genomic regions comprising one or more mutations based on the genotype of the tumor may comprise producing a list of genomic regions ranked by their fractional abundance.

[0280] Producing the selector set may comprise selecting one or more genomic regions from the list of genomic regions ranked by their fractional abundance. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of less than 50%, 47%, 45%, 42%, 40%, 37%, 35%, 34%, 33%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of less than 37%. Producing the selector set

may comprise selecting one or more genomic regions with a fractional abundance of less than 33%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of less than 30%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of less than 27%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of less than 25%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of between about 0.00001% to about 35%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of between about 0.00001% to about 30%. Producing the selector set may comprise selecting one or more genomic regions with a fractional abundance of between about 0.00001% to about 27%.

[0281] The selector set may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more genomic regions. The selector set may comprise one genomic region. The selector set may comprise at least 2 genomic regions. The selector set may comprise at least 3 genomic regions.

[0282] The genomic regions of the selector set may comprise one or more previously unidentified mutations. The genomic regions of the selector set may comprise 2 or more previously unidentified mutations. The genomic regions of the selector set may comprise 3 or more previously unidentified mutations. The genomic regions of the selector set may comprise 4 or more previously unidentified mutations.

[0283] The genomic regions may comprise one or more mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs. The genomic regions may comprise two or more mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs. The genomic regions may comprise three or more mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs. The genomic regions may comprise four or more mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs.

[0284] The genomic regions may comprise one or more types of mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs. The genomic regions may comprise two or more types of mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs. The genomic regions may comprise three or more types of mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs. The genomic regions may comprise four or more types of mutations selected from a group consisting of SNVs, indels, rearrangements, and CNVs.

[0285] Further disclosed herein are computer readable media for use in the methods disclosed herein. The computer readable medium may comprise sequence information for two or more genomic regions wherein (a) the genomic regions may comprise one or more mutations in greater than 80% of tumors from a population of subjects afflicted with a cancer; (b) the genomic regions represent less than 1.5 Mb of the genome; and (c) one or more of the following (i) the condition may be not hairy cell leukemia, ovarian cancer, Waldenstrom's macroglobulinemia; (ii) a genomic region may comprise at least one mutation in at least one subject afflicted with the cancer; (iii) the cancer includes two or more different types of cancer; (iv) the two or more genomic regions may be derived from two or more different genes; (v) the genomic regions may comprise two or more mutations; or (vi) the two or more genomic regions may comprise at least 10 kb.

[0286] In some instances, the condition is not hairy cell leukemia.

[0287] The genomic regions may comprise one or more mutations in greater than 60% of tumors from an additional population of subjects afflicted with another type of cancer.

[0288] The genomic regions may be derived from 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more different genes. The genomic regions may be derived from 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more different genes.

[0289] The genomic regions may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, or 50 kb. The genomic regions may comprise at least 5 kb. The genomic regions may comprise at least 10 kb. The genomic regions may comprise at least 50 kb.

[0290] The sequence information may comprise genomic coordinates pertaining to the 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more genomic regions. The sequence information may comprise genomic coordinates pertaining to the 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more genomic regions. The sequence information may comprise genomic coordinates pertaining to the 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 or more genomic regions.

[0291] The sequence information may comprise a nucleic acid sequence pertaining to the 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more genomic regions. The sequence information may comprise a nucleic acid sequence pertaining to the 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more genomic regions. The sequence information may comprise a nucleic acid sequence pertaining to the 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 or more genomic regions.

[0292] The sequence information may comprise a length of the 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more genomic regions. The sequence information may comprise a length of the 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more genomic regions. The sequence information may comprise a length of the 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 or more genomic regions.

[0293] Further disclosed herein are compositions for use in the methods and systems disclosed herein. The composition may comprise a set of oligonucleotides that selectively hybridize to a plurality of genomic regions, wherein (a) greater than 80% of tumors from a population of cancer subjects include one or more mutations in the genomic regions; (b) the plurality of genomic regions represent less than 1.5 Mb of the genome; and (c) the set of oligonucleotides may comprise 5 or more different oligonucleotides that selectively hybridize to the plurality of genomic regions.

[0294] An oligonucleotide of the set of oligonucleotides may comprise a tag. The tag may be biotin. The tag may be a label. The label may be a fluorescent label or dye. The tag may be an adaptor.

[0295] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 2. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 525 regions from those identified in Table 2. The genomic

regions may comprise at least 2 regions from those identified in Table 2. The genomic regions may comprise at least 20 regions from those identified in Table 2. The genomic regions may comprise at least 60 regions from those identified in Table 2. The genomic regions may comprise at least 100 regions from those identified in Table 2. The genomic regions may comprise at least 300 regions from those identified in Table 2. The genomic regions may comprise at least 400 regions from those identified in Table 2. The genomic regions may comprise at least 500 regions from those identified in Table 2.

[0296] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 2. At least about 5% of the genomic regions may be regions identified in Table 2. At least about 10% of the genomic regions may be regions identified in Table 2. At least about 20% of the genomic regions may be regions identified in Table 2. At least about 30% of the genomic regions may be regions identified in Table 2. At least about 40% of the genomic regions may be regions identified in Table 2.

[0297] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 6. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, or 830 regions from those identified in Table 6. The genomic regions may comprise at least 2 regions from those identified in Table 6. The genomic regions may comprise at least 20 regions from those identified in Table 6. The genomic regions may comprise at least 60 regions from those identified in Table 6. The genomic regions may comprise at least 100 regions from those identified in Table 6. The genomic regions may comprise at least 300 regions from those identified in Table 6. The genomic regions may comprise at least 600 regions from those identified in Table 6. The genomic regions may comprise at least 800 regions from those identified in Table 6.

[0298] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 6. At least about 5% of the genomic regions may be regions identified in Table 6. At least about 10% of the genomic regions may be regions identified in Table 6. At least about 20% of the genomic regions may be regions identified in Table 6. At least about 30% of the genomic regions may be regions identified in Table 6. At least about 40% of the genomic regions may be regions identified in Table 6.

[0299] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 7. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, or 450 regions from those identified in Table 7. The genomic regions may comprise at least 2 regions from those identified in Table 7. The genomic regions may comprise at least 20 regions from those identified in Table 7. The genomic regions may comprise at least 60 regions from those identified in Table 7. The genomic regions may comprise at least 100 regions from those identified in Table 7. The genomic regions may comprise at least 200 regions from those identified in Table 7. The genomic regions may comprise at least 300 regions from those identified in Table 7. The genomic regions may comprise at least 400 regions from those identified in Table 7.

[0300] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 7. At least about 5% of the genomic regions may be regions identified in Table 7. At least about 10% of the genomic regions may be regions identified in Table 7. At least about 20% of the genomic regions may be regions identified in Table 7. At least about 30% of the genomic regions may be regions identified in Table 7. At least about 40% of the genomic regions may be regions identified in Table 7.

[0301] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 8. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 8. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or 1050 regions from those identified in Table 8. The genomic regions may comprise at least 2 regions from those identified in Table 8. The genomic regions may comprise at least 20 regions from those identified in Table 8. The genomic regions may comprise at least 60 regions from those identified in Table 8. The genomic regions may comprise at least 100 regions from those identified in Table 8. The genomic regions may comprise at least 300 regions from those identified in Table 8. The genomic regions may comprise at least 600 regions from those identified in Table 8. The genomic regions may comprise at least 800 regions from those identified in Table 8. The genomic regions may comprise at least 1000 regions from those identified in Table 8.

[0302] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 8. At least about 5% of the genomic regions may be regions identified in Table 8. At least about 10% of the genomic regions may be regions identified in Table 8. At least about 20% of the genomic regions may be regions identified in Table 8. At least about 30% of the genomic regions may be regions identified in Table 8. At least about 40% of the genomic regions may be regions identified in Table 8.

[0303] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 9. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 9. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, or 1500 regions from those identified in Table 9. The genomic regions may comprise at least 2 regions from those identified in Table 9. The genomic regions may comprise at least 20 regions from those identified in Table 9. The genomic regions may comprise at least 60 regions from those identified in Table 9. The genomic regions may comprise at least 100 regions from those identified in Table 9. The genomic regions may comprise at least 300 regions from those identified in Table 9. The genomic regions may comprise at least 500 regions from those identified in Table 9. The genomic regions may comprise at least 1000 regions from those identified in Table 9. The genomic regions may comprise at least 1300 regions from those identified in Table 9.

[0304] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 9. At least about 5% of the genomic regions may be regions identified in Table 9. At least about 10% of the genomic regions may be regions identified in Table 9. At least about 20% of the genomic regions may be regions identified in Table 9. At least about 30% of the genomic regions may be regions identified in Table 9. At least about 40% of the genomic regions may be regions identified in Table 9.

[0305] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 10. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 10. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, or 330 regions from those identified in Table 10. The genomic regions may comprise at least 2 regions from those identified in Table 10. The genomic regions may comprise at least 20 regions from those identified in Table 10. The genomic regions may comprise at least 60 regions from those identified in Table 10.

[0306] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 10. At least about 5% of the genomic regions may be regions identified in Table 10. At least about 10% of the genomic regions may be regions identified in Table 10. At least about 20% of the genomic regions may be regions identified in Table 10. At least about 30% of the genomic regions may be regions identified in Table 10. At least about 40% of the genomic regions may be regions identified in Table 10.

[0307] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 11. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 11. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, or 460 regions from those identified in Table 11. The genomic regions may comprise at least 2 regions from those identified in Table 11. The genomic regions may comprise at least 20 regions from those identified in Table 11. The genomic regions may comprise at least 60 regions from those identified in Table 11. The genomic regions may comprise at least 100 regions from those identified in Table 11. The genomic regions may comprise at least 200 regions from those identified in Table 11. The genomic regions may comprise at least 300 regions from those identified in Table 11. The genomic regions may comprise at least 400 regions from those identified in Table 11.

[0308] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 11. At least about 5% of the genomic regions may be regions identified in Table 11. At least about 10% of the genomic regions may be regions identified in Table 11. At least about 20% of the genomic regions may be regions identified in Table 11. At least about 30% of the genomic regions may be regions identified in Table 11. At least about 40% of the genomic regions may be regions identified in Table 11.

[0309] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 12. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 12. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480 or 500 regions from those identified in Table 12. The genomic regions may comprise at least 2 regions from those identified in Table 12. The genomic regions may comprise at least 20 regions from those identified in Table 12. The genomic regions may comprise at least 60 regions from those identified in Table 12. The genomic regions may comprise at least 100 regions from those identified in Table 12. The genomic regions may comprise at least 200 regions from those identified in Table 12. The genomic regions may comprise at least 300 regions from those identified in Table 12. The genomic regions may comprise at least 400 regions from those identified in Table 12. The genomic regions may comprise at least 500 regions from those identified in Table 12.

[0310] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 12. At least about 5% of the genomic regions may be regions identified in Table 12. At least about 10% of the genomic regions may be regions identified in Table 12. At least about 20% of the genomic regions may be regions identified in Table 12. At least about 30% of the genomic regions may be regions identified in Table 12. At least about 40% of the genomic regions may be regions identified in Table 12.

[0311] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 13. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 13. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, or 1450 regions from those identified in Table 13. The genomic regions may comprise at least 2 regions from those identified in Table 13. The genomic regions may comprise at least 20 regions from those identified in Table 13. The genomic regions may comprise at least 60 regions from those identified in Table 13. The genomic regions may comprise at least 100 regions from those identified in Table 13. The genomic regions may comprise at least 300 regions from those identified in Table 13. The genomic regions may comprise at least 500 regions from those identified in Table 13. The genomic regions may comprise at least 1000 regions from those identified in Table 13. The genomic regions may comprise at least 1300 regions from those identified in Table 13.

[0312] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 13. At least about 5% of the genomic regions may be regions identified in Table 13. At least about 10% of the genomic regions may be regions identified in

Table 13. At least about 20% of the genomic regions may be regions identified in Table 13. At least about 30% of the genomic regions may be regions identified in Table 13. At least about 40% of the genomic regions may be regions identified in Table 13.

[0313] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 14. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 14. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1210, 1220, 1230, or 1240 regions from those identified in Table 14. The genomic regions may comprise at least 2 regions from those identified in Table 14. The genomic regions may comprise at least 20 regions from those identified in Table 14. The genomic regions may comprise at least 60 regions from those identified in Table 14. The genomic regions may comprise at least 100 regions from those identified in Table 14. The genomic regions may comprise at least 300 regions from those identified in Table 14. The genomic regions may comprise at least 500 regions from those identified in Table 14. The genomic regions may comprise at least 1000 regions from those identified in Table 14. The genomic regions may comprise at least 1100 regions from those identified in Table 14. The genomic regions may comprise at least 1200 regions from those identified in Table 14.

[0314] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 14. At least about 5% of the genomic regions may be regions identified in Table 14. At least about 10% of the genomic regions may be regions identified in Table 14. At least about 20% of the genomic regions may be regions identified in Table 14. At least about 30% of the genomic regions may be regions identified in Table 14. At least about 40% of the genomic regions may be regions identified in Table 14.

[0315] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 15. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, or 170 regions from those identified in Table 15. The genomic regions may comprise at least 2 regions from those identified in Table 15. The genomic regions may comprise at least 20 regions from those identified in Table 15. The genomic regions may comprise at least 60 regions from those identified in Table 15. The genomic regions may comprise at least 100 regions from those identified in Table 15. The genomic regions may comprise at least 120 regions from those identified in Table 15. The genomic regions may comprise at least 150 regions from those identified in Table 15.

[0316] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 15. At least about 5% of the genomic regions may be regions identified in Table 15. At least about 10% of the genomic regions may be regions identified in Table 15. At least about 20% of the genomic regions may be regions identified in Table 15. At least about 30% of the genomic regions may be regions identified in Table 15. At least about 40% of the genomic regions may be regions identified in Table 15.

[0317] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 16. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 16. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, or 2050 regions from those identified in Table 16. The genomic regions may comprise at least 2 regions from those identified in Table 16. The genomic regions may comprise at least 20 regions from those identified in Table 16. The genomic regions may comprise at least 60 regions from those identified in Table 16. The genomic regions may comprise at least 100 regions from those identified in Table 16. The genomic regions may comprise at least 300 regions from those identified in Table 16. The genomic regions may comprise at least 500 regions from those identified in Table 16. The genomic regions may comprise at least 1000 regions from those identified in Table 16. The genomic regions may comprise at least 1200 regions from those identified in Table 16. The genomic regions may comprise at least 1500 regions from those identified in Table 16. The genomic regions may comprise at least 1700 regions from those identified in Table 16. The genomic regions may comprise at least 2000 regions from those identified in Table 16.

[0318] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 16. At least about 5% of the genomic regions may be regions identified in Table 16. At least about 10% of the genomic regions may be regions identified in Table 16. At least about 20% of the genomic regions may be regions identified in Table 16. At least about 30% of the genomic regions may be regions identified in Table 16. At least about 40% of the genomic regions may be regions identified in Table 16.

[0319] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 17. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 17. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, or 1080 regions from those identified in Table 17. The genomic regions may comprise at least 2 regions from those identified in Table 17. The genomic regions may comprise at least 20 regions from those identified in Table 17. The genomic regions may comprise at least 60 regions from those identified in Table 17. The genomic regions may comprise at least 100 regions from those identified in Table 17. The genomic regions may comprise at least 300 regions from those identified in Table 17. The genomic regions may comprise at least 500 regions from those identified in Table 17. The genomic regions may comprise at least 1000 regions from those identified in Table 17. The genomic regions may comprise at least 1050 regions from those identified in Table 17.

[0320] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions

identified in Table 17. At least about 5% of the genomic regions may be regions identified in Table 17. At least about 10% of the genomic regions may be regions identified in Table 17. At least about 20% of the genomic regions may be regions identified in Table 17. At least about 30% of the genomic regions may be regions identified in Table 17. At least about 40% of the genomic regions may be regions identified in Table 17.

[0321] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 18. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 18. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480, 500, 520, 540, or 555 regions from those identified in Table 18. The genomic regions may comprise at least 2 regions from those identified in Table 18. The genomic regions may comprise at least 20 regions from those identified in Table 18. The genomic regions may comprise at least 60 regions from those identified in Table 18. The genomic regions may comprise at least 100 regions from those identified in Table 18. The genomic regions may comprise at least 200 regions from those identified in Table 18. The genomic regions may comprise at least 300 regions from those identified in Table 18. The genomic regions may comprise at least 400 regions from those identified in Table 18. The genomic regions may comprise at least 500 regions from those identified in Table 18.

[0322] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 18. At least about 5% of the genomic regions may be regions identified in Table 18. At least about 10% of the genomic regions may be regions identified in Table 18. At least about 20% of the genomic regions may be regions identified in Table 18. At least about 30% of the genomic regions may be regions identified in Table 18. At least about 40% of the genomic regions may be regions identified in Table 18.

[0323] The set of oligonucleotides may hybridize to less than 1.5, 1.45, 1.4, 1.35, 1.3, 1.25, 1.2, 1.15, 1.1, 1.05, or 1.0 Megabases (Mb) of the genome. The set of oligonucleotides may hybridize to less than 1000, 900, 800, 700, 600, 550, 500, 450, 400, 350, 300, 250, 200, 150, or 100 kb of the genome. The set of oligonucleotides may hybridize to less than 1.5 Megabases (Mb) of the genome. The set of oligonucleotides may hybridize to less than 1.25 Megabases (Mb) of the genome. The set of oligonucleotides may hybridize to less than 1 Megabases (Mb) of the genome. The set of oligonucleotides may hybridize to less than 1000 kb of the genome. The set of oligonucleotides may hybridize to less than 500 kb of the genome. The set of oligonucleotides may hybridize to less than 300 kb of the genome. The set of oligonucleotides may hybridize to less than 100 kb of the genome. The set of oligonucleotides may be capable of hybridizing to greater than 50 kb of the genome.

[0324] The set of oligonucleotides may be capable of hybridizing to 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, or 500 or more different genomic regions. The set of oligonucleotides may be capable of hybridizing to 5 or more different genomic regions. The set of oligonucleotides may be capable of hybridizing to 20 or more different genomic regions. The set of oligonucleotides may be capable of hybridizing to 50 or more different genomic regions. The set of oligonucleotides may be capable of hybridizing to 100 or more different genomic regions.

[0325] The plurality of genomic regions may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more different protein-coding regions. The protein-coding regions may comprise an exon, intron, untranslated region, or a combination thereof.

[0326] The plurality of genomic regions may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more different non-coding regions. The non-coding regions may comprise a non-coding RNA, ribosomal RNA (rRNA), transfer RNA (tRNA), or a combination thereof.

[0327] The oligonucleotides may be attached to a solid support. The solid support may be a bead. The bead may be a coated bead. The bead may be a streptavidin bead. The solid support may be an array. The solid support may be a glass slide.

[0328] Disclosed herein are populations of circulating tumor DNA (ctDNA) for use in any of the methods or systems disclosed herein. A population of circulating tumor DNA (ctDNA) may comprise ctDNA enriched by hybrid selection using any of the compositions comprising the set of oligonucleotides disclosed herein. A population of ctDNA may comprise ctDNA enriched by selective hybridization of the ctDNA using the set of oligonucleotides based on the selector sets disclosed herein. A population of ctDNA may comprise ctDNA enriched by selective hybridization using a set of oligonucleotides based on any of Tables 2 and 6-18.

[0329] Further disclosed herein are arrays for use in any of the methods and systems disclosed herein. The array may comprise a plurality of oligonucleotides to selectively capture genomic regions, wherein the genomic regions may comprise a plurality of mutations present in greater 60% of a population of subjects suffering from a cancer.

[0330] The plurality of mutations may be present in greater 60% of an additional population of subjects suffering from an additional type of cancer. The plurality of mutations may be present in greater 60% of an additional population of subjects suffering from two or more additional types of cancer. The plurality of mutations may be present in greater 60% of an additional population of subjects suffering from three or more additional types of cancer. The plurality of mutations may be present in greater 60% of an additional population of subjects suffering from four or more additional types of cancer.

[0331] An oligonucleotide of the set of oligonucleotides may comprise a tag. The tag may be biotin. The tag may comprise a label. The label may be a fluorescent label or dye. The tag may be an adaptor. The adaptor may comprise a molecular barcode. The adaptor may comprise a sample index.

[0332] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 2. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 525 regions from those identified in Table 2. The genomic regions may comprise at least 2 regions from those identified in Table 2. The genomic regions may comprise at least 20 regions from those identified in Table 2. The genomic regions may comprise at least 60 regions from those identified in Table 2. The genomic regions may comprise at least 100

regions from those identified in Table 2. The genomic regions may comprise at least 300 regions from those identified in Table 2. The genomic regions may comprise at least 400 regions from those identified in Table 2. The genomic regions may comprise at least 500 regions from those identified in Table 2.

[0333] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 2. At least about 5% of the genomic regions may be regions identified in Table 2. At least about 10% of the genomic regions may be regions identified in Table 2. At least about 20% of the genomic regions may be regions identified in Table 2. At least about 30% of the genomic regions may be regions identified in Table 2. At least about 40% of the genomic regions may be regions identified in Table 2.

[0334] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 6. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, or 830 regions from those identified in Table 6. The genomic regions may comprise at least 2 regions from those identified in Table 6. The genomic regions may comprise at least 20 regions from those identified in Table 6. The genomic regions may comprise at least 60 regions from those identified in Table 6. The genomic regions may comprise at least 100 regions from those identified in Table 6. The genomic regions may comprise at least 300 regions from those identified in Table 6. The genomic regions may comprise at least 600 regions from those identified in Table 6. The genomic regions may comprise at least 800 regions from those identified in Table 6.

[0335] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 6. At least about 5% of the genomic regions may be regions identified in Table 6. At least about 10% of the genomic regions may be regions identified in Table 6. At least about 20% of the genomic regions may be regions identified in Table 6. At least about 30% of the genomic regions may be regions identified in Table 6. At least about 40% of the genomic regions may be regions identified in Table 6.

[0336] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 7. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, or 450 regions from those identified in Table 7. The genomic regions may comprise at least 2 regions from those identified in Table 7. The genomic regions may comprise at least 20 regions from those identified in Table 7. The genomic regions may comprise at least 60 regions from those identified in Table 7. The genomic regions may comprise at least 100 regions from those identified in Table 7. The genomic regions may comprise at least 200 regions from those identified in Table 7. The genomic regions may comprise at least 300 regions from those identified in Table 7. The genomic regions may comprise at least 400 regions from those identified in Table 7.

[0337] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 7. At least about 5% of the genomic regions may be regions identified in Table 7. At least about 10% of the genomic regions may be regions identified in Table 7. At least about 20% of the genomic regions may be regions identified in Table 7. At least about 30% of the genomic regions may be regions identified in Table 7. At least about 40% of the genomic regions may be regions identified in Table 7.

[0338] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 8. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 8. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or 1050 regions from those identified in Table 8. The genomic regions may comprise at least 2 regions from those identified in Table 8. The genomic regions may comprise at least 20 regions from those identified in Table 8. The genomic regions may comprise at least 60 regions from those identified in Table 8. The genomic regions may comprise at least 100 regions from those identified in Table 8. The genomic regions may comprise at least 300 regions from those identified in Table 8. The genomic regions may comprise at least 600 regions from those identified in Table 8. The genomic regions may comprise at least 800 regions from those identified in Table 8. The genomic regions may comprise at least 1000 regions from those identified in Table 8.

[0339] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 8. At least about 5% of the genomic regions may be regions identified in Table 8. At least about 10% of the genomic regions may be regions identified in Table 8. At least about 20% of the genomic regions may be regions identified in Table 8. At least about 30% of the genomic regions may be regions identified in Table 8. At least about 40% of the genomic regions may be regions identified in Table 8.

[0340] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 9. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 9. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, or 1500 regions from those identified in Table 9. The genomic regions may comprise at least 2 regions from those identified in Table 9. The genomic regions may comprise at least 20 regions from those identified in Table 9. The genomic regions may comprise at least 60 regions from those identified in Table 9. The genomic regions may comprise at least 100 regions from those identified in Table 9. The genomic regions may comprise at least 300 regions from those identified in Table 9. The genomic regions may comprise at least 500 regions from those identified in Table 9. The genomic regions may comprise at least 1000 regions from those identified in Table 9. The genomic regions may comprise at least 1300 regions from those identified in Table 9.

[0341] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 9. At least about 5% of the genomic regions may be regions identified in Table 9. At least about 10% of the genomic regions may be regions identified in Table 9. At least about 20% of the genomic regions may be regions identified

in Table 9. At least about 30% of the genomic regions may be regions identified in Table 9. At least about 40% of the genomic regions may be regions identified in Table 9.

[0342] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 10. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 10. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, or 330 regions from those identified in Table 10. The genomic regions may comprise at least 2 regions from those identified in Table 10. The genomic regions may comprise at least 20 regions from those identified in Table 10. The genomic regions may comprise at least 60 regions from those identified in Table 10.

[0343] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 10. At least about 5% of the genomic regions may be regions identified in Table 10. At least about 10% of the genomic regions may be regions identified in Table 10. At least about 20% of the genomic regions may be regions identified in Table 10. At least about 30% of the genomic regions may be regions identified in Table 10. At least about 40% of the genomic regions may be regions identified in Table 10.

[0344] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 11. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 11. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, or 460 regions from those identified in Table 11. The genomic regions may comprise at least 2 regions from those identified in Table 11. The genomic regions may comprise at least 20 regions from those identified in Table 11. The genomic regions may comprise at least 60 regions from those identified in Table 11. The genomic regions may comprise at least 100 regions from those identified in Table 11. The genomic regions may comprise at least 200 regions from those identified in Table 11. The genomic regions may comprise at least 300 regions from those identified in Table 11. The genomic regions may comprise at least 400 regions from those identified in Table 11.

[0345] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 11. At least about 5% of the genomic regions may be regions identified in Table 11. At least about 10% of the genomic regions may be regions identified in Table 11. At least about 20% of the genomic regions may be regions identified in Table 11. At least about 30% of the genomic regions may be regions identified in Table 11. At least about 40% of the genomic regions may be regions identified in Table 11.

[0346] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 12. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 12. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480 or 500 regions from those identified in Table 12. The genomic regions may comprise at least 2 regions from those identified in Table 12. The genomic regions may comprise at least 20 regions from those identified in Table 12. The genomic regions may comprise at least 60 regions from those identified in Table 12. The genomic regions may comprise at least 100 regions from those identified in Table 12. The genomic regions may comprise at least 200 regions from those identified in Table 12. The genomic regions may comprise at least 300 regions from those identified in Table 12. The genomic regions may comprise at least 400 regions from those identified in Table 12. The genomic regions may comprise at least 500 regions from those identified in Table 12.

[0347] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 12. At least about 5% of the genomic regions may be regions identified in Table 12. At least about 10% of the genomic regions may be regions identified in Table 12. At least about 20% of the genomic regions may be regions identified in Table 12. At least about 30% of the genomic regions may be regions identified in Table 12. At least about 40% of the genomic regions may be regions identified in Table 12.

[0348] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 13. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 13. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, or 1450 regions from those identified in Table 13. The genomic regions may comprise at least 2 regions from those identified in Table 13. The genomic regions may comprise at least 20 regions from those identified in Table 13. The genomic regions may comprise at least 60 regions from those identified in Table 13. The genomic regions may comprise at least 100 regions from those identified in Table 13. The genomic regions may comprise at least 300 regions from those identified in Table 13. The genomic regions may comprise at least 500 regions from those identified in Table 13. The genomic regions may comprise at least 1000 regions from those identified in Table 13. The genomic regions may comprise at least 1300 regions from those identified in Table 13.

[0349] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 13. At least about 5% of the genomic regions may be regions identified in Table 13. At least about 10% of the genomic regions may be regions identified in Table 13. At least about 20% of the genomic regions may be regions identified in Table 13. At least about 30% of the genomic regions may be regions identified in Table 13. At least about 40% of the genomic regions may be regions identified in Table 13.

[0350] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more

regions from those identified in Table 14. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 14. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1210, 1220, 1230, or 1240 regions from those identified in Table 14. The genomic regions may comprise at least 2 regions from those identified in Table 14. The genomic regions may comprise at least 20 regions from those identified in Table 14. The genomic regions may comprise at least 60 regions from those identified in Table 14. The genomic regions may comprise at least 100 regions from those identified in Table 14. The genomic regions may comprise at least 300 regions from those identified in Table 14. The genomic regions may comprise at least 500 regions from those identified in Table 14. The genomic regions may comprise at least 1000 regions from those identified in Table 14. The genomic regions may comprise at least 1100 regions from those identified in Table 14. The genomic regions may comprise at least 1200 regions from those identified in Table 14.

[0351] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 14. At least about 5% of the genomic regions may be regions identified in Table 14. At least about 10% of the genomic regions may be regions identified in Table 14. At least about 20% of the genomic regions may be regions identified in Table 14. At least about 30% of the genomic regions may be regions identified in Table 14. At least about 40% of the genomic regions may be regions identified in Table 14.

[0352] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 15. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, or 170 regions from those identified in Table 15. The genomic regions may comprise at least 2 regions from those identified in Table 15. The genomic regions may comprise at least 20 regions from those identified in Table 15. The genomic regions may comprise at least 60 regions from those identified in Table 15. The genomic regions may comprise at least 100 regions from those identified in Table 15. The genomic regions may comprise at least 120 regions from those identified in Table 15. The genomic regions may comprise at least 150 regions from those identified in Table 15.

[0353] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 15. At least about 5% of the genomic regions may be regions identified in Table 15. At least about 10% of the genomic regions may be regions identified in Table 15. At least about 20% of the genomic regions may be regions identified in Table 15. At least about 30% of the genomic regions may be regions identified in Table 15. At least about 40% of the genomic regions may be regions identified in Table 15.

[0354] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 16. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 16. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, or 2050 regions from those identified in Table 16. The genomic regions may comprise at least 2 regions from those identified in Table 16. The genomic regions may comprise at least 20 regions from those identified in Table 16. The genomic regions may comprise at least 60 regions from those identified in Table 16. The genomic regions may comprise at least 100 regions from those identified in Table 16. The genomic regions may comprise at least 300 regions from those identified in Table 16. The genomic regions may comprise at least 500 regions from those identified in Table 16. The genomic regions may comprise at least 1000 regions from those identified in Table 16. The genomic regions may comprise at least 1200 regions from those identified in Table 16. The genomic regions may comprise at least 1500 regions from those identified in Table 16. The genomic regions may comprise at least 1700 regions from those identified in Table 16. The genomic regions may comprise at least 2000 regions from those identified in Table 16.

[0355] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 16. At least about 5% of the genomic regions may be regions identified in Table 16. At least about 10% of the genomic regions may be regions identified in Table 16. At least about 20% of the genomic regions may be regions identified in Table 16. At least about 30% of the genomic regions may be regions identified in Table 16. At least about 40% of the genomic regions may be regions identified in Table 16.

[0356] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 17. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 17. The genomic regions may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, or 1080 regions from those identified in Table 17. The genomic regions may comprise at least 2 regions from those identified in Table 17. The genomic regions may comprise at least 20 regions from those identified in Table 17. The genomic regions may comprise at least 60 regions from those identified in Table 17. The genomic regions may comprise at least 100 regions from those identified in Table 17. The genomic regions may comprise at least 300 regions from those identified in Table 17. The genomic regions may comprise at least 500 regions from those identified in Table 17. The genomic regions may comprise at least 1000 regions from those identified in Table 17. The genomic regions may comprise at least 1050 regions from those identified in Table 17.

[0357] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 17. At least about 5% of the genomic regions may be regions identified in Table 17. At least about 10% of the genomic regions may be regions identified in Table 17. At least about 20% of the genomic regions may be regions identified in Table 17. At least about 30% of the genomic regions may be regions identified in Table 17. At least about 40% of the genomic regions may be regions identified in Table 17.

[0358] The genomic regions may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 18. The genomic regions may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 18. The genomic regions may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480, 500, 520, 540, or 555 regions from those identified in Table 18. The genomic regions may comprise at least 2 regions from those identified in Table 18. The genomic regions may comprise at least 20 regions from those identified in Table 18. The genomic regions may comprise at least 60 regions from those identified in Table 18. The genomic regions may comprise at least 100 regions from those identified in Table 18. The genomic regions may comprise at least 200 regions from those identified in Table 18. The genomic regions may comprise at least 300 regions from those identified in Table 18. The genomic regions may comprise at least 400 regions from those identified in Table 18. The genomic regions may comprise at least 500 regions from those identified in Table 18.

[0359] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions may be regions identified in Table 18. At least about 5% of the genomic regions may be regions identified in Table 18. At least about 10% of the genomic regions may be regions identified in Table 18. At least about 20% of the genomic regions may be regions identified in Table 18. At least about 30% of the genomic regions may be regions identified in Table 18. At least about 40% of the genomic regions may be regions identified in Table 18.

[0360] The oligonucleotides may selectively capture 5, 10, 15, 20, 25, or 30 or more different genomic regions.

[0361] The oligonucleotides may hybridize to less than 1.5, 1.47, 1.45, 1.42, 1.40, 1.37, 1.35, 1.32, 1.30, 1.27, 1.25, 1.22, 1.20, 1.17, 1.15, 1.12, 1.10, 1.07, 1.05, 1.02, or 1.0 Megabases (Mb) of the genome. The oligonucleotides may hybridize to less than 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20, or 10 kb of the genome.

[0362] The oligonucleotides may be capable of hybridizing to greater than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, or 50 kb of the genome. The oligonucleotides may be capable of hybridizing to greater than 5 kb of the genome. The oligonucleotides may be capable of hybridizing to greater than 10 kb of the genome. The oligonucleotides may be capable of hybridizing to greater than 30 kb of the genome. The oligonucleotides may be capable of hybridizing to greater than 50 kb of the genome.

[0363] The plurality of genomic regions may comprise 2 or more different protein-coding regions. The plurality of genomic regions may comprise at least 3 different protein-coding regions. The protein-coding regions may comprise an exon, intron, untranslated region, or a combination thereof.

[0364] The plurality of genomic regions may comprise at least one non-coding region. The non-coding region may comprise a non-coding RNA, ribosomal RNA (rRNA), transfer RNA (tRNA), or a combination thereof.

[0365] Further disclosed herein are methods of determining a quantity of circulating tumor DNA (ctDNA). The method may comprise (a) ligating one or more adaptors to cell-free DNA (cfDNA) derived from a sample from a subject to produce one or more adaptor-ligated cfDNA; (b) performing sequencing on the one or more adaptor-ligated cfDNA, wherein the adaptor-ligated cfDNA to be sequenced are based on a selector set comprising a plurality of genomic regions; and (c) using a computer readable medium to determine a quantity of cfDNA originating from a tumor based on the sequencing information obtained from the adaptor-ligated cfDNA.

[0366] In some instances, sequencing does not comprise whole genome sequencing. In some instances, sequencing does not comprise whole exome sequencing. Sequencing may comprise massively parallel sequencing.

[0367] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 525 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 2.

[0368] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 2. At least about 5% of the genomic regions of the selector set may be regions identified in Table 2. At least about 10% of the genomic regions of the selector set may be regions identified in Table 2. At least about 20% of the genomic regions of the selector set may be regions identified in Table 2. At least about 30% of the genomic regions of the selector set may be regions identified in Table 2. At least about 40% of the genomic regions of the selector set may be regions identified in Table 2.

[0369] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, or 830 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 600 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 800 regions from those identified in Table 6.

[0370] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%,

85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 6. At least about 5% of the genomic regions of the selector set may be regions identified in Table 6. At least about 10% of the genomic regions of the selector set may be regions identified in Table 6. At least about 20% of the genomic regions of the selector set may be regions identified in Table 6. At least about 30% of the genomic regions of the selector set may be regions identified in Table 6. At least about 40% of the genomic regions of the selector set may be regions identified in Table 6.

[0371] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, or 450 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 7.

[0372] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 7. At least about 5% of the genomic regions of the selector set may be regions identified in Table 7. At least about 10% of the genomic regions of the selector set may be regions identified in Table 7. At least about 20% of the genomic regions of the selector set may be regions identified in Table 7. At least about 30% of the genomic regions of the selector set may be regions identified in Table 7. At least about 40% of the genomic regions of the selector set may be regions identified in Table 7.

[0373] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or 1050 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 600 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 800 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 8.

[0374] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 8. At least about 5% of the genomic regions of the selector set may be regions identified in Table 8. At least about 10% of the genomic regions of the selector set may be regions identified in Table 8. At least about 20% of the genomic regions of the selector set may be regions identified in Table 8. At least about 30% of the genomic regions of the selector set may be regions identified in Table 8. At least about 40% of the genomic regions of the selector set may be regions identified in Table 8.

[0375] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, or 1500 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 1300 regions from those identified in Table 9.

[0376] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 9. At least about 5% of the genomic regions of the selector set may be regions identified in Table 9. At least about 10% of the genomic regions of the selector set may be regions identified in Table 9. At least about 20% of the genomic regions of the selector set may be regions identified in Table 9. At least about 30% of the genomic regions of the selector set may be regions identified in Table 9. At least about 40% of the genomic regions of the selector set may be regions identified in Table 9.

[0377] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, or 330 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 20 regions

from those identified in Table 10. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 10.

[0378] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 10. At least about 5% of the genomic regions of the selector set may be regions identified in Table 10. At least about 10% of the genomic regions of the selector set may be regions identified in Table 10. At least about 20% of the genomic regions of the selector set may be regions identified in Table 10. At least about 30% of the genomic regions of the selector set may be regions identified in Table 10. At least about 40% of the genomic regions of the selector set may be regions identified in Table 10.

[0379] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, or 460 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 11.

[0380] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 11. At least about 5% of the genomic regions of the selector set may be regions identified in Table 11. At least about 10% of the genomic regions of the selector set may be regions identified in Table 11. At least about 20% of the genomic regions of the selector set may be regions identified in Table 11. At least about 30% of the genomic regions of the selector set may be regions identified in Table 11. At least about 40% of the genomic regions of the selector set may be regions identified in Table 11.

[0381] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480 or 500 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 12.

[0382] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 12. At least about 5% of the genomic regions of the selector set may be regions identified in Table 12. At least about 10% of the genomic regions of the selector set may be regions identified in Table 12. At least about 20% of the genomic regions of the selector set may be regions identified in Table 12. At least about 30% of the genomic regions of the selector set may be regions identified in Table 12. At least about 40% of the genomic regions of the selector set may be regions identified in Table 12.

[0383] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, or 1450 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 1300 regions from those identified in Table 13.

[0384] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 13. At least about 5% of the genomic regions of the selector set may be regions identified in Table 13. At least about 10% of the genomic regions of the selector set may be regions identified in Table 13. At least about 20% of the genomic regions of the selector set may be regions identified in Table 13. At least about 30% of the genomic regions of the selector set may be regions identified in Table 13. At least about 40% of the genomic regions of the selector set may be regions identified in Table 13.

[0385] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 14.

The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1210, 1220, 1230, or 1240 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 1100 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 1200 regions from those identified in Table 14.

[0386] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 14. At least about 5% of the genomic regions of the selector set may be regions identified in Table 14. At least about 10% of the genomic regions of the selector set may be regions identified in Table 14. At least about 20% of the genomic regions of the selector set may be regions identified in Table 14. At least about 30% of the genomic regions of the selector set may be regions identified in Table 14. At least about 40% of the genomic regions of the selector set may be regions identified in Table 14.

[0387] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, or 170 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 120 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 150 regions from those identified in Table 15.

[0388] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 15. At least about 5% of the genomic regions of the selector set may be regions identified in Table 15. At least about 10% of the genomic regions of the selector set may be regions identified in Table 15. At least about 20% of the genomic regions of the selector set may be regions identified in Table 15. At least about 30% of the genomic regions of the selector set may be regions identified in Table 15. At least about 40% of the genomic regions of the selector set may be regions identified in Table 15.

[0389] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, or 2050 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1200 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1500 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1700 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 2000 regions from those identified in Table 16.

[0390] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 16. At least about 5% of the genomic regions of the selector set may be regions identified in Table 16. At least about 10% of the genomic regions of the selector set may be regions identified in Table 16. At least about 20% of the genomic regions of the selector set may be regions identified in Table 16. At least about 30% of the genomic regions of the selector set may be regions identified in Table 16. At least about 40% of the genomic regions of the selector set may be regions identified in Table 16.

[0391] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, or 1080 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 500 regions

from those identified in Table 17. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 1050 regions from those identified in Table 17.

[0392] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 17. At least about 5% of the genomic regions of the selector set may be regions identified in Table 17. At least about 10% of the genomic regions of the selector set may be regions identified in Table 17. At least about 20% of the genomic regions of the selector set may be regions identified in Table 17. At least about 30% of the genomic regions of the selector set may be regions identified in Table 17. At least about 40% of the genomic regions of the selector set may be regions identified in Table 17.

[0393] The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480, 500, 520, 540, or 555 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 18.

[0394] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 18. At least about 5% of the genomic regions of the selector set may be regions identified in Table 18. At least about 10% of the genomic regions of the selector set may be regions identified in Table 18. At least about 20% of the genomic regions of the selector set may be regions identified in Table 18. At least about 30% of the genomic regions of the selector set may be regions identified in Table 18. At least about 40% of the genomic regions of the selector set may be regions identified in Table 18.

[0395] The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer. The plurality of genomic regions may comprise one or more mutations present in at least 60% or more of a population of subjects suffering from the cancer. The plurality of genomic regions may comprise one or more mutations present in at least 72% or more of a population of subjects suffering from the cancer. The plurality of genomic regions may comprise one or more mutations present in at least 80% or more of a population of subjects suffering from the cancer.

[0396] The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 Mb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 1 Mb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 500 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 100, 90, 80, 70, 60, 50, 40, 30, 20, 10 or 5 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 75 kb of a genome. The total size of the plurality of genomic regions of the selector set may comprise less than 50 kb of a genome.

[0397] The total size of the plurality of genomic regions of the selector set may be between 100 kb to 1000 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 500 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 500 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 300 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 5 kb to 200 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 1 kb to 100 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 1 kb to 50 kb of a genome.

[0398] Further disclosed herein are methods of preparing a library for sequencing. The method may comprise (a) conducting an amplification reaction on cell-free DNA (cfDNA) derived from a sample to produce a plurality of amplicons, wherein the amplification reaction may comprise 20 or fewer amplification cycles; and (b) producing a library for sequencing, the library comprising the plurality of amplicons.

[0399] The amplification reaction may comprise 19, 18, 17, 16, 15, 14, 13, 12, 11, or 10 or fewer amplification cycles. The amplification reaction may comprise 15 or fewer amplification cycles.

[0400] The method may further comprise attaching adaptors to one or more ends of the cfDNA. The adaptor may comprise a plurality of oligonucleotides. The adaptor may comprise one or more deoxyribonucleotides. The adaptor may comprise ribonucleotides. The adaptor may be single-stranded. The adaptor may be double-stranded. The adaptor may comprise double-stranded and single-stranded portions. For example, the adaptor may be a Y-shaped adaptor. The adaptor may be a linear adaptor. The adaptor may be a circular adaptor. The adaptor may comprise a molecular barcode, sample index, primer sequence, linker sequence or a combination thereof. The molecular barcode may be adjacent to the sample index. The molecular barcode may be adjacent to the primer sequence. The sample index may be adjacent to the

primer sequence. A linker sequence may connect the molecular barcode to the sample index. A linker sequence may connect the molecular barcode to the primer sequence. A linker sequence may connect the sample index to the primer sequence.

[0401] The adaptor may comprise a molecular barcode. The molecular barcode may comprise a random sequence. The molecular barcode may comprise a predetermined sequence. Two or more adaptors may comprise two or more different molecular barcodes. The molecular barcodes may be optimized to minimize dimerization. The molecular barcodes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first molecular barcode may introduce a single base error. The first molecular barcode may comprise greater than a single base difference from the other molecular barcodes. Thus, the first molecular barcode with the single base error may still be identified as the first molecular barcode. The molecular barcode may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The molecular barcode may comprise at least 3 nucleotides. The molecular barcode may comprise at least 4 nucleotides. The molecular barcode may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The molecular barcode may comprise less than 10 nucleotides. The molecular barcode may comprise less than 8 nucleotides. The molecular barcode may comprise less than 6 nucleotides. The molecular barcode may comprise 2 to 15 nucleotides. The molecular barcode may comprise 2 to 12 nucleotides. The molecular barcode may comprise 3 to 10 nucleotides. The molecular barcode may comprise 3 to 8 nucleotides. The molecular barcode may comprise 4 to 8 nucleotides. The molecular barcode may comprise 4 to 6 nucleotides.

[0402] The adaptor may comprise a sample index. The sample index may comprise a random sequence. The sample index may comprise a predetermined sequence. Two or more sets of adaptors may comprise two or more different sample indexes. Adaptors within a set of adaptors may comprise identical sample indexes. The sample indexes may be optimized to minimize dimerization. The sample indexes may be optimized to enable identification even with amplification or sequencing errors. For examples, amplification of a first sample index may introduce a single base error. The first sample index may comprise greater than a single base difference from the other sample indexes. Thus, the first sample index with the single base error may still be identified as the first molecular barcode. The sample index may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides. The sample index may comprise at least 3 nucleotides. The sample index may comprise at least 4 nucleotides. The sample index may comprise less than 20, 19, 18, 17, 16, or 15 nucleotides. The sample index may comprise less than 10 nucleotides. The sample index may comprise less than 8 nucleotides. The sample index may comprise less than 6 nucleotides. The sample index may comprise 2 to 15 nucleotides. The sample index may comprise 2 to 12 nucleotides. The sample index may comprise 3 to 10 nucleotides. The sample index may comprise 3 to 8 nucleotides. The sample index may comprise 4 to 8 nucleotides. The sample index may comprise 4 to 6 nucleotides.

[0403] The adaptor may comprise a primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer.

[0404] Adaptors may be attached to one end of a nucleic acid from a sample. The nucleic acids may be DNA. The DNA may be cell-free DNA (cfDNA). The DNA may be circulating tumor DNA (ctDNA). The nucleic acids may be RNA. Adaptors may be attached to both ends of the nucleic acid. Adaptors may be attached to one or more ends of a single-stranded nucleic acid. Adaptors may be attached to one or more ends of a double-stranded nucleic acid.

[0405] Adaptors may be attached to the nucleic acid by ligation. Ligation may be blunt end ligation. Ligation may be sticky end ligation. Adaptors may be attached to the nucleic acid by primer extension. Adaptors may be attached to the nucleic acid by reverse transcription. Adaptors may be attached to the nucleic acids by hybridization. Adaptors may comprise a sequence that is at least partially complementary to the nucleic acid. Alternatively, in some instances, adaptors do not comprise a sequence that is complementary to the nucleic acid.

[0406] The method may further comprise fragmenting the cfDNA. The method may further comprise end-repairing the cfDNA. The method may further comprise A-tailing the cfDNA.

[0407] Further disclosed herein are methods of determining a statistical significance of a selector set. The method may comprise (a) detecting a presence of one or more mutations in one or more samples from a subject, wherein the one or more mutations may be based on a selector set comprising genomic regions comprising the one or more mutations; (b) determining a mutation type of the one or more mutations present in the sample; and (c) determining a statistical significance of the selector set by calculating a ctDNA detection index based on a p-value of the mutation type of mutations present in the one or more samples.

[0408] In some instances, if a rearrangement is observed in two or more samples from the subject, then the ctDNA detection index is 0. At least one of the two or more samples may be a plasma sample. At least one of the two or more samples may be a tumor sample. The rearrangement may be a fusion or a breakpoint.

[0409] In some instances, if one type of mutation is present, then the ctDNA detection index is the p-value of the one type of mutation.

[0410] In some instances, if (i) two or more types of mutations are present in the sample; (ii) the p-values of the two or more types mutations are less than 0.1; and (iii) a rearrangement is not one of the types of mutations, then the ctDNA detection is calculated based on the combined p-values of the two or more mutations. The p-values of the two or more mutations may be combined according to Fisher's method. One of the two or more types of mutations may be a SNV. The p-value of the SNV may be determined by Monte Carlo sampling. One of the two or more types of mutations may be an indel.

[0411] In some instances, if (i) two or more types of mutations are present in the sample; (ii) a p-value of at least one of the two or more types of mutations are greater than 0.1; and (iii) a rearrangement is not one of the types of mutations, then the ctDNA detection is calculated based on the p-value of one of the two or more types mutations. One of the two or more types of mutations may be a SNV. The ctDNA detection index may be calculated based on the p-value of the SNV. One of the two or more types of mutations may be an indel.

[0412] Further disclosed herein are methods of identifying rearrangements in one or more nucleic acids. The method may comprise (a) obtaining sequencing information pertaining to a plurality of genomic regions; (b) producing a list of

genomic regions, wherein the genomic regions may be adjacent to one or more candidate rearrangement sites or the genomic regions may comprise one or more candidate rearrangement sites; and (c) applying an algorithm to the list of genomic regions to validate candidate rearrangement sites, thereby identifying rearrangements.

[0413] The sequencing information may comprise an alignment file. The alignment file may comprise an alignment file of pair-end reads, exon coordinates, and a reference genome.

[0414] The sequencing information may be obtained from a database. The database may comprise sequencing information pertaining to a population of subjects suffering from a disease or condition. The disease or condition may be a cancer.

[0415] The sequencing information may be obtained from one or more samples from one or more subjects.

[0416] Producing the list of genomic regions may comprise identifying discordant read pairs based on the sequencing information. The discordant read-pair may refer to a read and its mate, where: (i) the insert size may be not equal to the expected distribution of the dataset; or (ii) the mapping orientation of the reads may be unexpected.

[0417] Producing the list of genomic regions may comprise classifying the discordant read pairs based on the sequencing information. Producing the list of genomic regions further may comprise ranking the genomic regions. The genomic regions may be ranked in decreasing order of discordant read depth.

[0418] Producing the list of genomic regions may comprise selecting genomic regions with a minimum user-defined read depth.

[0419] The minimum user-defined read depth may be at least 2×, 3×, 4×, 5×, 6×, 7×, 8×, 9×, 10× or more.

[0420] The method may further comprise eliminating duplicate fragments.

[0421] Producing the list of genomic regions may comprise use of one or more algorithms. The algorithm may analyze properly paired reads in which one of the paired reads may be truncated to produce a soft-clipped read. The algorithm may analyze the soft-clipped reads based on a pattern. The pattern may be based on x number of skipped bases (Sx) and on y number of contiguous mapped bases (My). The pattern may be MySx or SxMy.

[0422] Applying the algorithm to validate the candidate rearrangement sites may comprise deleting candidate rearrangements with a read frequency of less than 2. Applying the algorithm to validate the candidate rearrangement sites may comprise ranking the candidate rearrangements based on their read frequency.

[0423] Applying the algorithm to validate the candidate rearrangement sites may comprise comparing two or more reads of the candidate rearrangement. Applying the algorithm to validate the candidate rearrangement sites may comprise identifying the candidate rearrangement as a rearrangement if the two or more reads have a sequence alignment.

[0424] Applying the algorithm to validate the candidate rearrangement sites may comprise evaluating inter-read concordance. Evaluating inter-read concordance may comprise dividing a first sequencing read of the candidate rearrangement site into a plurality of subsequences of length l. Evaluating inter-read concordance may comprise dividing a second sequencing read of the candidate rearrangement site into a plurality of subsequences of length l. Evaluating inter-read concordance may comprise comparing the subsequences of the first sequencing read to the subsequences of the second sequencing read. The first and second sequencing reads may be considered concordant if a minimum matching threshold may be achieved.

[0425] Applying the algorithm to validate the candidate rearrangement sites may comprise in silico validation of the candidate rearrangement sites. In silico validation may comprise aligning sequencing reads of the candidate rearrangement site to a reference rearrangement sequence. The reference rearrangement sequence may be obtained from a reference genome. The candidate rearrangement site may be identified as a rearrangement if the reads map to the reference rearrangement sequence with an identity of at least 70%, 75%, 80%, 85%, 90%, 95%, 97% or more.

[0426] The candidate rearrangement site may be identified as a rearrangement if the length of the aligned sequences may be at least 70%, 75%, 80%, 85%, 90%, or 95% or more of the read length of the candidate rearrangement site.

[0427] Further disclosed herein are methods of identifying tumor-derived single nucleotide variations (SNVs). The method may comprise (a) obtaining a sample from a subject suffering from a cancer or suspected of suffering from a cancer; (b) conducting a sequencing reaction on the sample to produce sequencing information; (c) applying an algorithm to the sequencing information to produce a list of candidate tumor alleles based on the sequencing information from step (b), wherein a candidate tumor allele may comprise a non-dominant base that may be not a germline SNP; and (d) identifying tumor-derived SNVs based on the list of candidate tumor alleles.

[0428] Producing the list of candidate tumor alleles may comprise ranking the tumor alleles by their fractional abundance. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a fractional abundance in the top 70th 75th, 80th, 85th, 87th, 90th, $\mathbf{92}^{nd}$, 95th, or 97th percentile. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a fractional abundance of less than 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 0.1% of the total alleles in the sample from the subject.

[0429] Producing the list of candidate tumor alleles may comprise ranking the tumor alleles based on their sequencing depth. Producing the list of candidate tumor alleles may comprise selecting tumor alleles that meet a minimum sequencing depth. The minimum sequencing depth may be at least 100×, 200×, 300×, 400×, 500×, 600×, 700×, 800×, 900×, 1000× or more.

[0430] Producing the list of candidate tumor alleles may comprise calculating a strand bias percentage of a tumor allele. Producing the list of candidate tumor alleles may comprise ranking the tumor alleles based on their strand bias percentage. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a user-defined strand bias percentage. The user-defined strand bias percentage may be less than or equal to 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97%.

[0431] Producing the list of candidate tumor alleles may comprise comparing the sequence of the tumor allele to a reference tumor allele. Producing the list of candidate tumor alleles further may comprise identifying tumor alleles that are different from the reference tumor allele.

[0432] Identifying the tumor alleles that are different from the reference tumor allele may comprise use of one or more statistical analyses. The one or more statistical analyses may

comprise using Bonferroni correction to calculate a Bonferroni-adjusted binomial probability for the tumor allele.

[0433] Producing the list of candidate tumor alleles may comprise selecting tumor alleles based on the Bonferroni-adjusted binomial probability. The Bonferroni-adjusted binomial probability of a candidate tumor allele may be less than or equal to 3×10^{-8}, 2.9×10^{-8}, 2.8×10^{-8}, 2.7×10^{-8}, 2.6×10^{-8}, 2.5×10^{-8}, 2.3×10^{-8}, 2.2×10^{-8}, 2.1×10^{-8}, 2.09×10^{-8}, 2.08×10^{-8}, 2.07×10^{-8}, 2.06×10^{-8}, 2.05×10^{-8}, 2.04×10^{-8}, 2.03×10^{-8}, 2.02×10^{-8}, 2.01×10^{-8} or 2×10^{-8}. The Bonferroni-adjusted binomial probability of a candidate tumor allele may be less than or equal to 2.08×10^{-8}.

[0434] Identifying the tumor alleles that are different from the reference tumor allele further may comprise applying a Z-test to the Bonferroni-adjusted binomial probability to produce a Bonferroni-adjusted single-tailed Z-score for the tumor allele. A tumor allele with a Bonferroni-adjusted single-tailed Z-score of greater than or equal to 6, 5.9, 5.8, 5.7, 5.6, 5.5, 5.4, 5.3, 5.2, 5.1, or 5.0 may be considered to be different from the reference tumor allele.

[0435] The sample may be a blood sample. The sample may be a paired sample.

[0436] Further disclosed herein are methods of producing a selector set. The method may comprise (a) obtaining sequencing information of a tumor sample from a subject suffering from a cancer; (b) comparing the sequencing information of the tumor sample to sequencing information from a non-tumor sample from the subject to identify one or more mutations specific to the sequencing information of the tumor sample; and (c) producing a selector set comprising one or more genomic regions comprising the one or more mutations specific to the sequencing information of the tumor sample.

[0437] The selector set may comprise sequencing information pertaining to the one or more genomic regions. The selector set may comprise genomic coordinates pertaining to the one or more genomic regions.

[0438] The selector set may be used to produce a plurality of oligonucleotides that selectively hybridize the one or more genomic regions. The plurality of oligonucleotides may be biotinylated.

[0439] The one or more mutations may comprise SNVs. The one or more mutations may comprise indels. The one or more mutations may comprise rearrangements.

[0440] Producing the selector set may comprise identifying tumor-derived SNVs using the methods disclosed herein.

[0441] Producing the selector set may comprise identifying tumor-derived rearrangements using the method disclosed herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0442] FIG. **1**A-**1**D: Development of CAncer Personalized Profiling by Deep Sequencing (CAPP-Seq). (FIG. **1**A) Schematic depicting design of CAPP-Seq selectors and their application for assessing circulating tumor DNA. (FIG. **1**B) Multi-phase design of the NSCLC selector. Phase 1: Genomic regions harboring known/suspected driver mutations in NSCLC are captured. Phases 2-4: Addition of exons containing recurrent SNVs using WES data from lung adenocarcinomas and squamous cell carcinomas from TCGA (n=407). Regions were selected iteratively to maximize the number of mutations per tumor while minimizing selector size. Recurrence index=total unique patients with mutations covered per kb of exon. Phases 5-6: Exons of predicted NSCLC drivers and introns/exons harboring breakpoints in rearrangements involving ALK, ROS1, and RET were added. Bottom: increase of selector length during each design phase. (FIG. **1**C) Analysis of the number of SNVs per lung adenocarcinoma covered by the NSCLC selector in the TCGA WES cohort (Training; n=229) and an independent lung adenocarcinoma WES data set (Validation; n=183). Results are compared to selectors randomly sampled from the exome ($P<1.0\times10^{-6}$ for the difference between random selectors and the NSCLC selector). (FIG. **1**D) Number of SNVs per patient identified by the NSCLC selector in WES data from three adenocarcinomas from TCGA, colon (COAD), rectal (READ), and endometrioid (UCEC) cancers.

[0443] FIG. **2**A-**2**I: Analytical performance. (FIG. **2**A-**2**C) Quality parameters from a representative CAPP-Seq analysis of plasma cfDNA, including length distribution of sequenced cfDNA fragments (FIG. **2**A), and depth of sequencing coverage across all genomic regions in the selector]FIG. **2**B). (FIG. **2**C) Variation in sequencing depth across cfDNA samples from 4 patients. Orange envelope represents s.e.m. (FIG. **2**D) Analysis of background rate for 40 plasma cfDNA samples collected from 13 NSCLC patients and 5 healthy individuals. (FIG. **2**E) Analysis of biological background in d focusing on 107 recurrent somatic mutations from a previously reported SNaPshot panel. Mutations found in a given patient's tumor were excluded. The mean frequency over all subjects was ~0.01%. A single outlier mutation (TP53 R175H) is indicated by an orange diamond. (FIG. **2**F) Individual mutations from e ranked by most to least recurrent, according to mean frequency across the 40 cfDNA samples. The p-value threshold of 0.01 (horizontal line) corresponds to the 99th percentile of global selector background in d. (FIG. **2**G) Dilution series analysis of expected versus observed frequencies of mutant alleles using CAPP-Seq. Dilution series were generated by spiking fragmented HCC78 DNA into control cfDNA. (FIG. **2**H) Analysis of the effect of the number of SNVs considered on the estimates of fractional abundance (95% confidence intervals shown in gray). (FIG. **2**I) Analysis of the effect of the number of SNVs considered on the mean correlation coefficient between expected and observed cancer fractions (blue dashed line) using data from panel h. 95% confidence intervals are shown for e-f. Statistical variation for g is shown as s.e.m.

[0444] FIG. **3**A-**3**C: Sensitivity and specificity analysis. (FIG. **3**A) Receiver Operating

[0445] Characteristic (ROC) analysis of cfDNA samples from pre-treatment samples and healthy controls, divided into all stages (n=13 patients) and stages II-IV (n=9 patients). Area Under the Curve (AUC) values are significant at $P<0.0001$. Sn, sensitivity; Sp, specificity. (FIG. **3**B) Raw data related to a. TP, true positive; FP, false positive; TN, true negative; FN, false negative. (FIG. **3**C) Concordance between tumor volume, measured by CT or PET/CT, and pg per mL of ctDNA from pretreatment samples (n=9), measured by CAPP-Seq. Patients P6 and P9 were excluded due to inability to accurately assess tumor volume and differences related to the capture of fusions, respectively. Of note, linear regression was performed in non-log space; the log-log axes and dashed diagonal line are for display purposes only.

[0446] FIG. **4**A-**4**I: Noninvasive detection and monitoring of circulating tumor DNA. (FIG. **4**A-**4**H) Disease monitoring using CAPP-Seq. (FIG. **4**A-**4**B) Disease burden changes in response to treatment in a stage III NSCLC patient using SNVs and an indel (FIG. **4**A), and a stage IV NSCLC patient using three rearrangement breakpoints (FIG. **4**B). (FIG. **4**C)

Concordance between different reporters (SNVs and a fusion) in a stage IV NSCLC patient. (FIG. **4**D) Detection of a subclonal EGFR T790M resistance mutation in a patient with stage IV NSCLC. The fractional abundance of the dominant clone and T790M-containing clone are shown in the primary tumor (left) and plasma samples (right). (FIG. **4**E-**4**F) CAPP-Seq results from post-treatment cfDNA samples are predictive of clinical outcomes in a stage IIB NSCLC patient FIG. **4**E and Stage IIIB NSCLC patient (FIG. **4**F). (FIG. **4**G-**4**H) Monitoring of tumor burden following complete tumor resection (FIG. **4**G) and Stereotactic Ablative Radiotherapy (SABR) (FIG. **4**H) for two stage IB NSCLC patients. (FIG. **4**I) Exploratory analysis of the potential application of CAPP-Seq for biopsy-free tumor genotyping or cancer screening. All plasma cfDNA samples from patients in Table 1 were examined for the presence of mutant allele outliers without knowledge of the primary tumor mutations; samples with detectable mutations are shown, along with two samples determined to be cancer-negative (P1-2 and P16-3) and a sample without tumor-derived SNVs (P9-5; see Table 1). The lowest mutant allele fraction detected was ~0.5% (dashed horizontal line). Error bars in d represent s.e.m. Tu, tumor; Ef, pleural effusion; SD, stable disease; PD, progressive disease; PR, partial response; CR, complete response; DOD, dead of disease.

[0447] FIG. **5**A-**5**B: Comparison to other methods for detection of ctDNA in plasma. (FIG. **5**A) Analytical modeling of CAPP-Seq, WES, and WGS for different detection limits of tumor cfDNA in plasma. Calculations are based on the median number of mutations detected per NSCLC for CAPP-Seq (e.g., 4) and the reported number of mutations in NSCLC exomes and genomes. The vertical dotted line represents the median fraction of tumor-derived cfDNA in plasma from NSCLC patients in this study (see below). (FIG. **5**B) Costs for WES and WGS to achieve the same theoretical detection limit as CAPP-Seq (shown as a dark solid line in FIG. **5**A).

[0448] FIG. **6**: CAPP-Seq computational pipeline. Major steps of the bioinformatics pipeline for mutation discovery and quantitation in plasma are schematically illustrated.

[0449] FIG. **7**A-**7**B: Statistical enrichment of recurrently mutated NSCLC exons captures known drivers. We employed two metrics to prioritize exons with recurrent mutations for inclusion in the CAPP-Seq NSCLC selector. The first, termed Recurrence Index (RI), is defined as the number of unique patients (e.g. tumors) with somatic mutations per kilobase of a given exon and the second metric is based on the minimum number of unique patients (e.g. tumors) with mutations in a given kb of exon. We analyzed exons containing at least one non-silent SNV genotyped by TCGA (n=47,769) in a combined cohort of 407 lung adenocarcinoma (LUAD) and squamous cell carcinoma (SCC) patients. (FIG. **7**A Known/suspected NSCLC drivers are highly enriched at RI≥30 (inset), comprising 1.8% (n=861) of analyzed exons. (FIG. **7**B) Known/suspected NSCLC drivers are highly enriched at ≥3 patients with mutations per exon (inset), encompassing 16% of analyzed exons.

[0450] FIG. **8**A-**8**E: FACTERA analytical pipeline for breakpoint mapping. Major steps used by FACTERA to precisely identify genomic breakpoints from aligned paired-end sequencing data are anecdotally illustrated using two hypothetical genes, w and v. (FIG. **8**A) Improperly paired, or "discordant," reads (indicated in yellow) are used to locate genes involved in a potential fusion (in this case, w and v). (FIG. **8**B) Because truncated (e.g., soft-clipped) reads may indicate a fusion breakpoint, any such reads within genomic regions delineated by w and v are also further analyzed. (FIG. **8**C) Consider soft-clipped reads, R1 and R2, whose non-clipped segments map to w and v, respectively. If R1 and R2 derive from a fragment encompassing a true fusion between w and v, then the mapped portion of R1 should match the soft-clipped portion of R2, and vice versa. This is assessed by FACTERA using fast k-mer indexing and comparison. (FIG. **8**D) Four possible orientations of R1 and R2 are depicted. However, only Cases 1a and 2a can generate valid fusions. Thus, prior to k-mer comparison (FIG. **8**C), the reverse complement of R1 is taken for Cases 1b and 2b, respectively, converting them into Cases 1a and 2a. (FIG. **8**E) In some cases, short sequences immediately flanking the breakpoint are identical, preventing unambiguous determination of the breakpoint. Let iterators i and j denote the first matching sequence positions between R1 and R2. To reconcile sequence overlap, FACTERA arbitrarily adjusts the breakpoint in R2 (e.g., bp2) to match R1 (e.g., bp1) using the sequence offset determined by differences in distance between bp2 and i, and bp1 and j. Two cases are illustrated, corresponding to sequence orientations described in FIG. **8**D.

[0451] FIG. **9**A-**9**B: Application of FACTERA to NSCLC cell lines NCI-H3122 and HCC78, and Sanger-validation of breakpoints. (FIG. **9**A) Pile-up of a subset of soft-clipped reads mapping to the EML4-ALK fusion identified in NCI-H3122 along with the corresponding Sanger chromatogram (from top to bottom SEQ ID NOs:1-11). (FIG. **9**B) Same as a, but for the SLC34A2-ROS1 translocation identified in HCC78 (from top to bottom SEQ ID NOs:12-22).

[0452] FIG. **10**A-**10**C: Improvements in CAPP-Seq performance with optimized library preparation procedures. Using 32 ng of input cfDNA from plasma, we compared standard versus 'with bead'[5] library preparation methods, as well as two commercially available DNA polymerases (Phusion and KAPA HiFi). We also compared template pre-amplification by Whole Genome Amplification (WGA) using Degenerate Oligonucleotide PCR (DOP). Indices considered for these comparisons included (FIG. **10**A) length of the captured cfDNA fragments sequenced, (FIG. **10**B) depth and uniformity of sequencing coverage across all genomic regions in the selector, and (FIG. **10**C) sequence mapping and capture statistics, including uniqueness. Collectively, these comparisons identified KAPA HiFi polymerase and a "with bead" protocol as having most robust and uniform performance.

[0453] FIG. **11**A-**11**F: Optimizing allele recovery from low input cfDNA during Illumina library preparation. Bars reflect the relative yield of CAPP-Seq libraries constructed from 4 ng cfDNA, calculated by averaging quantitative PCR measurements of n=4 pre-selected reporters within CAPP-Seq with pre-defined amplification efficiencies. (FIG. **11**A) Sixteen hour ligation at 16° C. increases ligation efficiency and reporter recovery. (FIG. **11**B) Adapter ligation volume did not have a significant effect on ligation efficiency and reporter recovery. (FIG. **11**C) Performing enzymatic reactions "with-bead" to minimize tube transfer steps increases reporter recovery. (FIG. **11**D) Increasing adapter concentration during ligation increases ligation efficiency and reporter recovery. Reporter recovery is also higher when using KAPA HiFi DNA polymerase compared to Phusion DNA polymerase (FIG. **11**E) and when using the KAPA Library Preparation Kit with the modifications in a-d compared to the NuGEN SP Ovation Ultralow Library System with automation on a Mon-

drian SP Workstation (FIG. **11**F). Relative reporter abundance was determined by qPCR using the $2^{-\Delta Ct}$ method. A two-sided t test with equal variance was used to test the statistical significance between groups. All values are presented as means±s.d. N.S., not significant. Based on these results, we estimate that combining the methodological modifications in FIG. **11**A and FIG. **11**C-**11**E improves yield in NGS libraries by 3.3-fold.

[0454] FIG. **12**A-**12**C: CAPP-Seq performance with various amounts of input cfDNA. (FIG. **12**A) Length of the captured cfDNA fragments sequenced. (FIG. **12**B) Depth of sequencing coverage across all genomic regions in the selector (pre-duplicate removal). (FIG. **12**C) Sequence mapping and capture statistics. As expected, more input cfDNA mass correlates with more unique fragments sequenced.

[0455] FIG. **13**A-**13**B. Analysis of library complexity and molecule recovery. (FIG. **13**A) The expected proportion of additional library complexity present in post-duplicate reads is plotted for all patient and control samples, including plasma cfDNA (n=40) and paired tumor/PBL specimens (n=17 each). Because of the highly stereotyped size of cfDNA fragments occurring naturally in blood plasma, when compared with genomic DNA shorn by sonication, any two fragments of DNA circulating in plasma are inherently more likely by chance to have arisen from different original molecules, whether considering tumor or non-tumor cells as the source of this cfDNA. To estimate this "missing" complexity, we reasoned that two DNA fragments (e.g., paired end reads) with identical start/end coordinates that differ by a single a priori defined germline variant (e.g. one maternal and one paternal allele) represent two unique and independent starting molecules rather than technical artifacts (e.g. PCR duplicates). Therefore, the number of fragments sharing identical start/end coordinates with both maternal and paternal germline alleles of heterozygous SNPs were used to estimate additional library complexity. Library complexity estimates updated to factor in these data are also provided in Tables 3, 20 and 21 and determined as described herein. (FIG. **13**B) Empirical assessment of molecule recovery in cfDNA (n=40) by determination of the mass of DNA produced compared to the expected library yield based on mass input, number of PCR cycles, and efficiency (mean=46%). (FIG. **13**A-**13**B) Values are presented as means±95% confidence intervals.

[0456] FIG. **14**. Analysis of library cross-contamination. Allelic fractions of patient-specific homozygous germline SNPs were assessed in cfDNA samples multiplexed on the same lane. SNPs were called as described in the Methods. The mean "cross-contamination" rate in cfDNA samples was 0.06%, shown by the horizontal dotted line. This level of contamination is too low to affect our estimates of tumor burden given the low fraction of tumor-derived cfDNA in plasma of NSCLC patients (median of ~0.1%; FIG. **5***a*) (e.g., 0.06×0.1=0.006% of a given sample would on average represent contamination from ctDNA of another sample). Of note, to minimize the risk of inter-sample contamination, we use aerosol barrier tips, work in hoods, and do not multiplex tumor and plasma libraries in the same lane.

[0457] FIG. **15**. Analysis of selector-wide bias in captured sequence. Because the NSCLC selector was designed to target the hg19 reference genome, we reasoned that selector bias for SNVs, if any, should be discernable as a systematically lower ratio of non-reference to reference alleles in heterozygous germline SNPs. Therefore, we analyzed high confidence SNPs detected by VarScan in patient PBL samples, where high confidence was defined as variants with a non-reference fraction >10% present in the common SNPs subset of dbSNP (version 137.0). As shown, we detected a very small skew toward reference (8 of 11 samples have a median non-reference allelic frequency of 49%; the remaining 3 samples are unbiased). Importantly, such bias appears too small to significantly affect our results. Boxes represent the interquartile range, and whiskers encapsulate the 10th to 90th percentiles. Germline SNPs were identified using VarScan 2.

[0458] FIG. **16**A-**16**D: Empirical spiking analysis of CAPP-Seq using two NSCLC cell lines. (FIG. **16**A) Expected and observed (by CAPP-Seq) fractions of NCI-H3122 DNA spiked into control HCC78 DNA are linear for all fractions tested (0.1%, 1%, and 10%; R^2=1). Using data from FIG. **16**B, analysis of the effect of the number of SNVs considered on the estimates of fractional abundance (95% confidence intervals shown in gray). (FIG. **16**C) Analysis of the effect of the number of SNVs considered on the mean correlation coefficient and coefficient of variation between expected and observed cancer fractions (blue dashed line) using data from panel a. (FIG. **16**D) Expected and observed fractions of the EML4-ALK fusion present in HCC78 are linear (R^2=0.995) over all spiking concentrations tested (see FIG. **9**B for breakpoint verification). The observed EML4-ALK fractions were normalized based on the relative abundance of the fusion in 100% H3122 DNA. Moreover, both a single heterozygous insertion ('Indel'; chr7: 107416855, +T) and a 4.9 kb homozygous deletion ('Deletion', chr17: 29422259-29592392) in NCI-H3122 were concordant with defined concentrations. Values in a are presented as means±s.e.m.

[0459] FIG. **17**A-**17**B: Base-pair resolution breakpoint mapping for all patients and cell lines enumerated by FACTERA. Gene fusions involving ALK (FIG. **17**A) and ROS1 (FIG. **17**B) are graphically depicted. Schematics in the top panels indicate the exact genomic positions (HG19 NCBI Build 37.1/GRCh37) of the breakpoints in ALK, ROS1, EML4, KIF5B, SLC34A2, CD74, MKX, and FYN. Bottom panels depict exons flanking the predicted gene fusions with notation indicating the 5' fusion partner gene and last fused exon followed by the 3' fusion partner gene and first fused exon. For example, in S13del37;R34 exons 1-13 of SLC34A2 (excluding the 3' 37 nucleotides of exon 13) are fused to exons 34-43 of ROS1. Exons in FYN are from its 5'UTR and precede the first coding exon. The green dotted line in the predicted FYN-ROS1 fusion indicates the first in-frame methionine in ROS1 exon 33, which preserves an open reading frame encoding the ROS1 kinase domain. All rearrangements were each independently confirmed by PCR and/or FISH.

[0460] FIG. **18**: Presence of fusions is inversely related to the number of SNVs detected by CAPP-Seq. For each patient listed in Table 1 the number of identified SNVs versus the presence (n=11) or absence (n=6) of detected genomic fusions is plotted. Statistical significance was determined using a two-sided Wilcoxon rank sum test, and summarized values are presented as means±s.e.m.

[0461] FIG. **19**A-**19**D. Receiver Operating Curve (ROC) analysis of CAPP-Seq performance including both pre- and post-treatment samples. Comparison of sensitivity and specificity achieved for non-deduped (FIGS. **19**A and **19**C) and deduped (post PCR duplicate removal) data (FIGS. **19**B and **19**D). In addition, all stages (FIG. **19**A-**19**B) are compared with intermediate to advanced stages (stages II-IV, FIGS. **19**C and **19**D). Finally, for all ROC analyses, the effect of the

indel/fusion filter on sensitivity/specificity is shown. Reporter fractions for both non-deduped and deduped cfDNA samples are provided in Table 4.

[0462] FIG. **20**. CAPP-Seq sensitivity and specificity over all patient reporters and sequenced plasma cfDNA samples. All values shown reflect a ctDNA detection index of 0.03. See Methods for details on detection metrics, and determination of cancer-positive, cancer-negative, and unknown categories.

[0463] FIG. **21**A-**21**D. Non-invasive cancer screening with CAPP-Seq, related to FIG. **4**I. (FIG. **21**A) Steps to identify candidate SNVs in plasma cfDNA demonstrated using a patient sample with NSCLC (P6, see Table 4). Following stepwise filtration, outlier detection is applied. (FIG. **21**B) Same as a, but using a plasma cfDNA sample from a patient who had their tumor surgically removed. No SNVs are identified, as expected. (FIG. **21**C, **21**D) Three additional representative samples applying retrospective screening to patients analyzed in this study. P2 and P5 samples have confirmed tumor-derived SNVs, while P9 is cancer positive but lacks tumor-derived SNVs. Red points, confirmed tumor-derived SNVs; Green points, background noise.

[0464] FIG. **22**. depicts a flow chart of patient analysis.

[0465] FIG. **23**. shows a system for implementing the methods of the disclosure.

DETAILED DESCRIPTION OF THE INVENTION

[0466] It is characteristic of cancer cells that due to somatic mutation the genome sequence of the cancer cell is changed from the genome sequence of the individual from which it is derived. Most human cancers are relatively heterogeneous for somatic mutations in individual genes. Specifically, in most human tumors, recurrent somatic alterations of single genes account for a minority of patients, and only a minority of tumor types can be defined using a small number of recurrent mutations at predefined positions. The present invention solves this problem by use of enrichment of tumor-derived nucleic acid molecules from total genomic nucleic acids with a selector set. The design of the selector is vital because (1) it dictates which mutations can be detected in with high probability for a patient with a given cancer, and (2) the selector size (in kb) directly impacts the cost and depth of sequence coverage.

[0467] While the specific genetic changes differ from individual to individual and between types of cancer, there are regions of the genome that show recurrent changes. In those regions there is an increased probability that any given individual cancer will show genetic variation. The genetic changes in cancer cells provide a means by which cancer cells can be distinguished from normal (e.g., non-cancer) cells. Cell-free DNA, for example the DNA fragments found in blood samples, can be analyzed for the presence of genetic variation distinctive of tumor cells. However, the absolute levels of tumor DNA in such samples is often small, and the genetic variation may represent only a very small portion of the entire genome. The present invention addresses this issue by providing methods for selective detection of mutated regions associated with cancer, thereby allowing accurate detection of cancer cell DNA or RNA from the background of normal cell DNA or RNA. Although the methods disclosed herein may specifically refer to DNA (e.g., cell-free DNA, circulating tumor DNA), it should be understood that the methods, compositions, and systems disclosed herein are applicable to all types of nucleic acids (e.g., RNA, DNA, RNA/DNA hybrids).

[0468] Provided herein are methods for the ultrasensitive detection of a minority nucleic acid in a heterogeneous sample. The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from a subject; and (b) using sequence information derived from (a) to detect cell-free minority nucleic acids in the sample, wherein the method is capable of detecting a percentage of the cell-free minority nucleic acids that is less than 2% of total cfDNA. The minority nucleic acid may refer to a nucleic acid that originated from a cell or tissue that is different from a normal cell or tissue from the subject. For example, the subject may be infected with a pathogen such as a bacteria and the minority nucleic acid may be a nucleic acid from the pathogen. In another example, the subject is a recipient of a cell, tissue or organ from a donor and the minority nucleic acid may be a nucleic acid originating from the cell, tissue or organ from the donor. In another example, the subject is a pregnant subject and the minority nucleic acid may be a nucleic acid originating from a fetus. The method may comprise using the sequence information to detect one or more somatic mutations in the fetus. The method may comprise using the sequence information to detect one or more post-zygotic mutations in the fetus. Alternatively, the subject may be suffering from a cancer and the minority nucleic acid may be a nucleic acid originating from a cancer cell.

[0469] Provided herein are methods for the ultrasensitive detection of circulating tumor DNA in a sample. The method may be called CAncer Personalized Profiling by Deep Sequencing (CAPP-Seq). The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from a subject; and (b) using sequence information derived from (a) to detect cell-free tumor DNA (ctDNA) in the sample, wherein the method is capable of detecting a percentage of ctDNA that is less than 2% of total cfDNA. CAPP-Seq may accurately quantify cell-free tumor DNA from early and advanced stage tumors. CAPP-Seq may identify mutant alleles down to 0.025% with a detection limit of <0.01%. Tumor-derived DNA levels often paralleled clinical responses to diverse therapies and CAPP-Seq may identify actionable mutations. CAPP-Seq may be routinely applied to noninvasively detect and monitor tumors, thus facilitating personalized cancer therapy.

[0470] Disclosed herein are methods for determining a quantity of circulating tumor DNA (ctDNA) in a sample. The method may comprise (a) ligating one or more adaptors to cell-free DNA (cfDNA) derived from a sample from a subject to produce one or more adaptor-ligated cfDNA; (b) performing sequencing on the one or more adaptor-ligated cfDNA, wherein the adaptor-ligated cfDNA to be sequenced is based on a selector set comprising a plurality of genomic regions; and (c) using a computer readable medium to determine a quantity of cfDNA originating from a tumor based on the sequencing information obtained from the adaptor-ligated cfDNA.

[0471] Further disclosed herein are methods of detecting, diagnosing, or prognosing a status or outcome of a cancer in a subject. The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; (b) using sequence information derived from (a) to detect cell-free tumor DNA (ctDNA) in the sample wherein the method is capable of detecting a percentage of ctDNA that is less than 2% of total cfDNA.

[0472] Further disclosed herein are methods of diagnosing a status or outcome of a cancer in a subject. The method may

comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information is derived from genomic regions that are mutated in at least 80% of a population of subjects afflicted with a cancer; and (b) diagnosing a cancer selected from a group consisting of lung cancer, breast cancer, colorectal cancer and prostate cancer in the subject based on the sequence information, wherein the method has a sensitivity of 80%.

[0473] Further disclosed herein are methods of prognosing a status or outcome of a cancer in a subject. The method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information is derived from regions that are mutated in at least 80% of a population of subjects afflicted with a condition; and (b) determining a prognosis of a condition in the subject based on the sequence information.

[0474] Further disclosed herein are methods of selecting a therapy for a subject suffering from a cancer. The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; (b) using sequence information derived from (a) to detect cell-free tumor DNA (ctDNA) in the sample wherein the method is capable of detecting a percentage of ctDNA that is less than 2% of total cfDNA.

[0475] Alternatively, the method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information is derived from regions that are mutated in at least 80% of a population of subjects afflicted with a condition; and (b) determining a therapeutic regimen of a condition in the subject based on the sequence information.

[0476] Further disclosed herein are methods for diagnosing, prognosing, or determining a therapeutic regimen for a subject afflicted with or suspected of having a cancer. The method may comprise (a) obtaining sequence information for selected regions of genomic DNA from a cell-free DNA sample from the subject; (b) using the sequence information to determine the presence or absence of one or more mutations in the selected regions, wherein at least 70% of a population of subjects afflicted with the cancer have mutation(s) in the regions; and (c) providing a report with a diagnosis, prognosis or treatment regimen to the subject, based on the presence or absence of the one or more mutations.

[0477] Further disclosed herein are methods for assessing tumor burden in a subject. The method may comprise (a) obtaining sequence information on cell-free nucleic acids derived from a sample from the subject; (b) using a computer readable medium to determine quantities of circulating tumor DNA (ctDNA) in the sample; (c) assessing tumor burden based on the quantities of ctDNA; and (d) reporting the tumor burden to the subject or a representative of the subject.

[0478] Further disclosed herein are methods for determining a disease state of a cancer in a subject. The method may comprise (a) obtaining a quantity of circulating tumor DNA (ctDNA) in a sample from the subject; (b) obtaining a volume of a tumor in the subject; and (c) determining a disease state of a cancer in the subject based on a ratio of the quantity of ctDNA to the volume of the tumor.

[0479] Disclosed herein are methods for detecting at least 50% of stage I cancer with a specificity of greater than 90%. The method may comprise (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage I cancer in the sample based on the quantity of the cell-free DNA.

[0480] Disclosed herein are methods for detecting at least 60% of stage II cancer with a specificity of greater than 90% comprising (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage II cancer in the sample based on the quantity of the cell-free DNA.

[0481] Disclosed herein are methods for detecting at least 60% of stage III cancer with a specificity of greater than 90% comprising (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage III cancer in the sample based on the quantity of the cell-free DNA.

[0482] Disclosed herein are methods for detecting at least 60% of stage IV cancer with a specificity of greater than 90% comprising (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage IV cancer in the sample based on the quantity of the cell-free DNA.

[0483] Also provided are selector sets for use in the methods disclosed herein. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in a population of subjects suffering from a cancer. The selector set may be a library of recurrently mutated genomic regions used in the CAPP-Seq methods. The targeting of recurrently mutated genomic regions may allow a distinction between tumor cell DNA and normal DNA. In addition, the targeting of recurrently mutated genomic region may provide for simultaneous detection of point mutations, copy number variation, insertions/deletions, and rearrangements.

[0484] The selector set may be a computer readable medium. The computer readable medium may comprise nucleic acid sequence information for two or more genomic DNA regions wherein (a) the genomic regions comprise one or more mutations in >80% of tumors from a population of subjects afflicted with a cancer; (b) the genomic DNA regions represent less than 1.5 Mb of the genome; and (c) one or more of the following: (i) the condition is not hairy cell leukemia, ovarian cancer, Waldenstrom's macroglobulinemia; (ii) each of the genomic DNA regions comprises at least one mutation in at least one subject afflicted with the cancer; (iii) the cancer includes two or more different types of cancer; (iv) the two or more genomic regions are derived from two or more different genes; (v) the genomic regions comprise two or more mutations; or (vi) the two or more genomic regions comprise at least 10 kb.

[0485] The selector set may provide, for example, oligonucleotides useful in selective amplification of tumor-derived nucleic acids. The selector set may provide, for example, oligonucleotides useful in selective capture or enrichment of tumor-derived nucleic acids. Disclosed herein are compositions comprising a set of oligonucleotides based on the selector set. The composition may comprise a set of oligonucleotides that selectively hybridize to a plurality of genomic DNA regions, wherein (a) >80% of tumors from a population of cancer subjects include one or more mutations in the genomic DNA regions; (b) the plurality of genomic DNA regions represent less than 1.5 Mb of the genome; and (c) the set of oligonucleotides comprise 5 or more different oligonucleotides that selectively hybridize to the plurality of genomic DNA regions.

[0486] The composition may comprise oligonucleotides that selectively hybridize to a plurality of genomic regions, wherein the genomic regions comprise a plurality of mutations present in >60% of a population of subjects suffering from a cancer.

[0487] Further disclosed herein is an array comprising a plurality of oligonucleotides to selectively capture genomic regions, wherein the genomic regions comprise a plurality of mutations present in >60% of a population of subjects suffering from a cancer.

[0488] Further disclosed herein are methods of producing a selector set for a cancer. The method of producing a selector set for a cancer may comprise (a) identifying recurrently mutated genomic DNA regions of the selected cancer; and (b) prioritizing regions using one or more of the following criteria (i) a Recurrence Index (RI) for the genomic region(s), wherein the RI is the number of unique patients or tumors with somatic mutations per length of a genomic region; and (ii) a minimum number of unique patients or tumors with mutations in a length of genomic region.

[0489] Disclosed herein are methods of enriching for circulating tumor DNA from a sample.

[0490] The method may comprise contacting cell-free nucleic acids from a sample with a plurality of oligonucleotides, wherein the plurality of oligonucleotides selectively hybridize to a plurality of genomic regions comprising a plurality of mutations present in >60% of a population of subjects suffering from a cancer.

[0491] Alternatively, the method may comprise contacting cell-free nucleic acids from a sample with a set of oligonucleotides, wherein the set of oligonucleotides selectively hybridize to a plurality of genomic regions, wherein (a) >80% of tumors from a population of cancer subjects include one or more mutations in the genomic regions; (b) the plurality of genomic regions represent less than 1.5 Mb of the genome; and (c) the set of oligonucleotides comprise 5 or more different oligonucleotides that selectively hybridize to the plurality of genomic regions.

[0492] Further disclosed herein are methods of preparing a nucleic acid sample for sequencing.

[0493] The method may comprise (a) conducting an amplification reaction on cell-free DNA (cfDNA) derived from a sample to produce a plurality of amplicons, wherein the amplification reaction comprises 20 or fewer amplification cycles; and (b) producing a library for sequencing, the library comprising the plurality of amplicons.

[0494] Further disclosed herein are systems for implementing one or more of the methods or steps of the methods disclosed herein. FIG. **23** shows a computer system (also "system" herein) **2301** programmed or otherwise configured for implementing the methods of the disclosure, such as producing a selector set and/or data analysis. The system **2301** includes a central processing unit (CPU, also "processor" and "computer processor" herein) **2305**, which can be a single core or multi core processor, or a plurality of processors for parallel processing. The system **2301** also includes memory **2310** (e.g., random-access memory, read-only memory, flash memory), electronic storage unit **2315** (e.g., hard disk), communications interface **2320** (e.g., network adapter) for communicating with one or more other systems, and peripheral devices **2325**, such as cache, other memory, data storage and/or electronic display adapters. The memory **2310**, storage unit **2315**, interface **2320** and peripheral devices **2325** are in communication with the CPU **2305** through a communications bus (solid lines), such as a motherboard. The storage unit **2315** can be a data storage unit (or data repository) for storing data. The system **2301** is operatively coupled to a computer network ("network") **2330** with the aid of the communications interface **2320**. The network **2330** can be the Internet, an internet and/or extranet, or an intranet and/or extranet that is in communication with the Internet. The network **2330** in some cases is a telecommunication and/or data network. The network **2330** can include one or more computer servers, which can enable distributed computing, such as cloud computing. The network **2330** in some cases, with the aid of the system **2301**, can implement a peer-to-peer network, which may enable devices coupled to the system **2301** to behave as a client or a server.

[0495] The system **2301** is in communication with a processing system **2335**. The processing system **2335** can be configured to implement the methods disclosed herein. In some examples, the processing system **2335** is a nucleic acid sequencing system, such as, for example, a next generation sequencing system (e.g., Illumina sequencer, Ion Torrent sequencer, Pacific Biosciences sequencer). The processing system **2335** can be in communication with the system **2301** through the network **2330**, or by direct (e.g., wired, wireless) connection. The processing system **2335** can be configured for analysis, such as nucleic acid sequence analysis.

[0496] Methods as described herein can be implemented by way of machine (or computer processor) executable code (or software) stored on an electronic storage location of the system **2301**, such as, for example, on the memory **2310** or electronic storage unit **2315**. During use, the code can be executed by the processor **2305**. In some examples, the code can be retrieved from the storage unit **2315** and stored on the memory **2310** for ready access by the processor **2305**. In some situations, the electronic storage unit **2315** can be precluded, and machine-executable instructions are stored on memory **2310**.

[0497] Disclosed herein is a computer-implemented system for calculating a recurrence index for one or more genomic regions. The computer-implemented system may comprise (a) a digital processing device comprising an operating system configured to perform executable instructions and a memory device; and (b) a computer program including instructions executable by the digital processing device to create a recurrence index, the computer program comprising (i) a first software module configured to receive data pertaining to a plurality of mutations; (ii) a second software module configured to relate the plurality of mutations to one or more genomic regions and/or one or more subjects; and (iii) a third software module configured to calculate a recurrence index of

one or more genomic regions, wherein the recurrence index is based on a number of mutations per subject per kilobase of nucleotide sequence.

[0498] Selector Set

[0499] The methods, kits, and systems disclosed herein may comprise one or more selector sets or uses thereof. A selector set may be a bioinformatics construct comprising the sequence information for regions of the genome (e.g., genomic regions) associated with one or more cancers of interest. A selector set may be a bioinformatics construct comprising genomic coordinates for one or more genomic regions. The genomic regions may comprise one or more recurrently mutated regions. The genomic regions may comprise one or more mutations associated with one or more cancers of interest.

[0500] The number of genomic regions in a selector set may vary depending on the nature of the cancer. The inclusion of larger numbers of genomic regions may generally increase the likelihood that a unique somatic mutation will be identified. Including too many genomic regions in the library is not without a cost, however, since the number of genomic regions is directly related to the length of nucleic acids that must be sequenced in the analysis. At the extreme, the entire genome of a tumor sample and a genomic sample could be sequenced, and the resulting sequences could be compared to note any differences.

[0501] The selector sets of the invention may address this problem by identifying genomic regions that are recurrently mutated in a particular cancer, and then ranking those regions to maximize the likelihood that the region will include a distinguishing somatic mutation in a particular tumor. The library of recurrently mutated genomic regions, or "selector set", can be used across an entire population for a given cancer or class of cancers, and does not need to be optimized for each subject.

[0502] The selector set may comprise at least about 2, 3, 4, 5, 6, 7, 8, or 9 different genomic regions. The selector set may comprise at least about 10 different genomic regions; at least about 25, at least about 50, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 500, at least about 600, at least about 700, at least about 800, at least about 900, at least about 1000 or more different genomic regions.

[0503] The selector set may comprise between about 10 to about 1000 different genomic regions. The selector set may comprise between about 10 to about 900 different genomic regions. The selector set may comprise between about 10 to about 800 different genomic regions. The selector set may comprise between about 10 to about 700 different genomic regions. The selector set may comprise between about 20 to about 600 different genomic regions. The selector set may comprise between about 20 to about 500 different genomic regions. The selector set may comprise between about 20 to about 400 different genomic regions. The selector set may comprise between about 50 to about 500 different genomic regions. The selector set may comprise between about 50 to about 400 different genomic regions. The selector set may comprise between about 50 to about 300 different genomic regions.

[0504] The selector set may comprise a plurality of genomic regions. The plurality of genomic regions may comprise at most 5000 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 2000 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 1000 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 500 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 400 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 300 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 200 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 150 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 100 different genomic regions. In some embodiments, the plurality of genomic regions comprises at most 50 different genomic regions or even fewer.

[0505] A genomic region may comprise a protein-coding region, or portion thereof. A protein-coding region may refer to a region of the genome that encodes for a protein. A protein-coding region may comprise an intron, exon, and/or untranslated region (UTR). A genomic region may comprise two or more protein-coding regions, or portions thereof. For example, a genomic region may comprise a portion of an exon and a portion of an intron. A genomic region may comprise three or more protein-coding regions, or portions thereof. For example, a genomic region may comprise a portion of a first exon, a portion of an intron, and a portion of a second exon. Alternatively, or additionally, a genomic region may comprise a portion of an exon, a portion of an intron, and a portion of an untranslated region.

[0506] A genomic region may comprise a gene. A genomic region may comprise only a portion of a gene. A genomic region may comprise an exon of a gene. A genomic region may comprise an intron of a gene. A genomic region may comprise an untranslated region (UTR) of a gene. In some instances, a genomic region does not comprise an entire gene. A genomic region may comprise less than 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5% of a gene. A genomic region may comprise less than 60% of a gene.

[0507] A genomic region may comprise a nonprotein-coding region. A nonprotein-coding region may also be referred to as a noncoding region. A nonprotein-coding region may refer to a region of the genome that does not encode for a protein. A nonprotein-coding region may be transcribed into a noncoding RNA (ncRNA). The noncoding RNA may have a known function. For example, the noncoding RNA may be a transfer RNA (tRNA), ribosomal RNA (rRNA), and/or regulatory RNA. The noncoding RNA may have an unknown function. Examples of ncRNA include, but are not limited to, tRNA, rRNA, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), microRNA, small interfering RNA (siRNAs), Piwi-interacting RNA (piRNA), and long ncRNA (e.g., Xist, HOTAIR). A genomic region may comprise a pseudogene, transposon and/or retrotransposon.

[0508] A genomic region may comprise a recurrently mutated region. A recurrently mutated region may refer to a region of the genome, usually the human genome, in which there is an increased probability of genetic mutation in a cancer of interest, relative to the genome as a whole. A recurrently mutation region may refer to a region of the genome that contains one or more mutations that is recurrent in the population. For example, a recurrently mutation region may refer to a region of the genome that contains a mutation that is

present in two or more subjects in a population. A recurrently mutated region may be characterized by a "Recurrence Index" (RI). The RI generally refers to the number of individual subjects (e.g., cancer patients) with a mutation that occurs within a given kilobase of genomic sequence (e.g., number of patients with mutations/genomic region length in kb). A genomic region may also be characterized by the number of patients with a mutation per exon. Thresholds for each metric (e.g. RI and patients per exon or genomic region) may be selected to statistically enrich for known/suspected drivers of the cancer of interest. A known/suspected driver of the cancer of interest may be a gene. In non-small cell lung carcinoma (NSCLC), these metrics may enrich for known/suspected drivers (see genes listed in Table 2). Thresholds can also be selected by arbitrarily choosing the top percentile for each metric.

[0509] A selector set may comprise a genomic region comprising a mutation that is not recurrent in the population. For example, a genomic region may comprise one or more mutations that are present in a given subject. In some instances, a genomic region that comprises one or more mutations in a subject may be used to produce a personalized selector set for the subject.

[0510] The term "mutation" may refer to a genetic alteration in the genome of an organism. For the purposes of the invention, mutations of interest are typically changes relative to the germline sequence, e.g. cancer cell specific changes. Mutations may include single nucleotide variants (SNV), copy number variants (CNV), insertions, deletions and rearrangements (e.g., fusions). The selector set may comprise one or more genomic regions comprising one or more mutations selected from a group consisting of SNV, CNV, insertions, deletions, and rearrangements. The selector set may comprise a plurality of genomic regions comprising two or more mutations selected from a group consisting of SNV, CNV, insertions, deletions, and rearrangements. The selector set may comprise a plurality of genomic regions comprising three or more mutations selected from a group consisting of SNV, CNV, insertions, deletions, and rearrangements. The selector set may comprise a plurality of genomic regions comprising four or more mutations selected from a group consisting of SNV, CNV, insertions, deletions, and rearrangements. The selector set may comprise a plurality of genomic regions comprising five or more mutations selected from a group consisting of SNV, CNV, insertions, deletions, and rearrangements. The selector set may comprise a plurality of genomic regions comprising at least one SNV, insertion, and deletion. The selector set may comprise a plurality of genomic regions comprising at least one SNV and rearrangement. The selector set may comprise a plurality of genomic regions comprising at least one insertion, deletion, and rearrangement. The selector set may comprise a plurality of genomic regions comprising at least one deletion and rearrangement. The selector set may comprise a plurality of genomic regions comprising at least one insertion and rearrangement. The selector set may comprise a plurality of genomic regions comprising at least one SNV, insertion, deletion, and rearrangement. The selector set may comprise a plurality of genomic regions comprising at least one rearrangement and at least one mutation selected from a group consisting of SNV, insertion, and deletion. The selector set may comprise a plurality of genomic regions comprising at least one rearrangement and at least one mutation selected from a group consisting of SNV, CNV, insertion, and deletion.

[0511] A selector set may comprise a mutation in a genomic region known to be associated with a cancer. The mutation in a genomic region known to be associated with a cancer may be referred to as a "known somatic mutation." A known somatic mutation may be a mutation located in one or more genes known to be associated with a cancer. A known somatic mutation may be a mutation located in one or more oncogenes. For example, known somatic mutations may include one or more mutations located in p53, EGFR, KRAS and/or BRCA1.

[0512] A selector set may comprise a mutation in a genomic region predicted to be associated with a cancer. A selector set may comprise a mutation in a genomic region that has not been reported to be associated with a cancer.

[0513] A genomic region may comprise a sequence of the human genome of sufficient size to capture one or more recurrent mutations. The methods of the invention may be directed at cfDNA, which is generally less than about 200 bp in length, and thus a genomic region may be generally less than about 10 kb. The length of genomics region in a selector set may be on average around about 100 bp, about 125 bp, about 150 bp, 175 bp, about 200 bp, about 225 bp, about 250 bp, about 275 bp, or around about 300 bp. Generally the genomic region for a SNV can be quite short, from about 45 to about 500 bp in length, while the genomic region for a fusion or other genomic rearrangement may be longer, from around about 1 Kbp to about 10 Kbp in length. A genomic region in a selector set may be less than about 10 Kbp, 9 Kbp, 8 Kbp, 7 Kbp, 6 Kbp, 5 Kbp, 4 Kbp, 3 Kbp, 2 Kbp, or 1 Kbp in length. A genomic region in a selector set may be less than about 1000 bp, 900 bp, 800 bp, 700 bp, 600 bp, 500 bp, 400 bp, 300 bp, 200 bp, or 100 bp. A genomic region may be said to "identify" a mutation when the mutation is within the sequence of that genomic region.

[0514] In some embodiments, the total sequence covered by the selector set is less than about 1.5 megabase pairs (Mbp), 1.4 Mbp, 1.3 Mbp, 1.2 Mbp, 1.1 Mbp, 1 Mbp. The total sequence covered by the selector set may be less than about 1000 kb, less than about 900 kb, less than about 800 kb, less than about 700 kb, less than about 600 kb, less than about 500 kb, less than about 400 kb, less than about 350 kb, less than about 300 kb, less than about 250 kb, less than about 200 kb, or less than about 150 kb. The total sequence covered by the selector set may be between about 100 kb to 500 kb. The total sequence covered by the selector set may be between about 100 kb to 350 kb. The total sequence covered by the selector set may be between about 100 kb to 150 kb.

[0515] The selector set may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more mutations in a plurality of genomic regions. The selector set may comprise 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more mutations in a plurality of genomic regions. The selector set may comprise 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 or more mutations in a plurality of genomic regions.

[0516] At least a portion of the mutations may be within the same genomic region. At least about 2, 3, 4, 5, 6, 7, 8, 9, 10 or more mutations may be within the same genomic region. At least about 2 mutations may be within the same genomic region. At least about 3 mutations may be within the same genomic region.

[0517] At least a portion of the mutations may be within different genomic regions. At least about 2, 3, 4, 5, 6, 7, 8, 9, 10 or more mutations may be within two or more different

genomic regions. At least about 2 mutations may be within two or more different genomic regions. At least about 3 mutations may be within two or more different genomic regions.

[0518] Two or more mutations may be in two or more different genomic regions of the same noncoding region. Two or more mutations may be in two or more different genomic regions of the same protein-coding region. Two or more mutations may be in two or more different genomic regions of the same gene. For example, a first mutation may be located in a first genomic region comprising a first exon of a first gene and a second mutation may be located in a second genomic region comprising a second exon of the first gene. In another example, a first mutation may be located in a first genomic region comprising a first portion of a first long noncoding RNA and a second mutation may be located in a second genomic region comprising a second portion of the first long noncoding RNA.

[0519] Alternatively, or additionally, two or more mutations may be in two or more different genomic regions of two or more different noncoding regions, protein-coding regions, and/or genes. For example, a first mutation may be located in a first genomic region comprising a first exon of a first gene and a second mutation may be located in a second genomic region comprising a second exon of a second gene. In another example, a first mutation may be located in a first genomic region comprising a first exon of a first gene and a second mutation may be located in a second genomic region comprising a portion of a microRNA.

[0520] The selector set may identify a median of at least 2, usually at least 3, and preferably at least 4 different mutations per individual subject. The selector set may identify a median of at least 5, 6, 7, 8, 9, 10, 11, 12, 13 or more different mutations per individual subject. The different mutations may be in one or more genomic regions. The different mutations may be in 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more genomic regions. The different mutations may be in 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more recurrently mutated regions.

[0521] The median number of mutations identified by the selector set may be determined in a population of up to 10, up to 25, up to 25, up to 50, up to 87, up to 100 or more subjects. The median number of mutations identified by the selector set may be determined in a population of up to 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400 or more subjects. In such a population, a selector set of interest may identify one or more mutations in at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 85%, at least 87%, at least 90%, at least 92%, at least 95% or more of the subjects.

[0522] The total mutations identified by the selector set may be present in at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 85%, at least 87%, at least 90%, at least 92%, at least 95% or more of subjects in a population. For example, the selector set may identify a first mutation present in 20% of the subjects and second mutation in 80% of the subjects, thus the total mutations identified by the selector set may be present in 80% to 100% of the subjects in the population.

[0523] In addition to a bioinformatics construct, a selector set can be used to generate an oligonucleotide or set of oligonucleotides for specific capture, sequencing and/or amplification of cfDNA corresponding to a genomic region. The set of oligonucleotides may include at least one oligonucleotide for each genomic region that is to be targeted. Oligonucleotides may have the general characteristic of sufficient length to uniquely identify the genomic region, e.g. usually at least about 15 nucleotides, at least about 16, 17, 18, 19, 20 nucleotides in length. An oligonucleotide may further comprise an adapter for the sequencing system; a tag for sorting; a specific binding tag, e.g. biotin, FITC, etc. Oligonucleotides for amplification may comprise a pair of sequences flanking the region of interest, and of opposite orientation. The oligonucleotide may comprise a primer sequence. The oligonucleotide may comprise a sequence that is complementary to at least a portion of the genomic region.

[0524] The methods set forth herein may generate a bioinformatics construct comprising the selector set sequence information. In order to use the selector set for patient diagnostic and prognostic methods, a set of selector probes may be generated from the selector set library. The set of selector probes may comprise a sequence from at least about 20 genomic regions, at least about 30 genomic regions, at least about 40 genomic regions, at least about 50 genomic regions, at least about 60 genomic regions, at least about 70 genomic regions, at least about 80 genomic regions, at least about 90 genomic regions, at least about 100 genomic regions, at least about 200 genomic regions, at least about 300 genomic regions, at least about 400 genomic regions, or at least about 500 genomic regions. The genomic regions may be selected from the genomic regions set forth in any one of Tables 2 and 6-18. The selection may be based on bioinformatics criteria, including the additional value provided by the region, the RI, etc. In some embodiments a pre-set coverage of patients is used as a cut-off, for example where at least 90% have one or more of the SNV, where at least 95% have one or more of the SNV, where at least 98% have one or more of the SNV.

[0525] The selector set may comprise one or more genomic regions identified by Table 2. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 525 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 2. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 2.

[0526] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 2. At least about 5% of the genomic regions of the selector set may be regions identified in Table 2. At least about 10% of the genomic regions of the selector set may be regions identified in Table 2. At least about 20% of the genomic regions of the selector set may be regions identified in Table 2. At least about 30% of the genomic regions of the selector set may be regions identified in Table

2. At least about 40% of the genomic regions of the selector set may be regions identified in Table 2.

[0527] The selector set may comprise one or more genomic regions identified by Table 6. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, or 830 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 600 regions from those identified in Table 6. The genomic regions of the selector set may comprise at least 800 regions from those identified in Table 6.

[0528] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 6. At least about 5% of the genomic regions of the selector set may be regions identified in Table 6. At least about 10% of the genomic regions of the selector set may be regions identified in Table 6. At least about 20% of the genomic regions of the selector set may be regions identified in Table 6. At least about 30% of the genomic regions of the selector set may be regions identified in Table 6. At least about 40% of the genomic regions of the selector set may be regions identified in Table 6.

[0529] The selector set may comprise one or more genomic regions identified by Table 7. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, or 450 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 7. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 7.

[0530] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 7. At least about 5% of the genomic regions of the selector set may be regions identified in Table 7. At least about 10% of the genomic regions of the selector set may be regions identified in Table 7. At least about 20% of the genomic regions of the selector set may be regions identified in Table 7. At least about 30% of the genomic regions of the selector set may be regions identified in Table 7. At least about 40% of the genomic regions of the selector set may be regions identified in Table 7.

[0531] The selector set may comprise one or more genomic regions identified by Table 8. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or 1050 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 600 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 800 regions from those identified in Table 8. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 8.

[0532] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 8. At least about 5% of the genomic regions of the selector set may be regions identified in Table 8. At least about 10% of the genomic regions of the selector set may be regions identified in Table 8. At least about 20% of the genomic regions of the selector set may be regions identified in Table 8. At least about 30% of the genomic regions of the selector set may be regions identified in Table 8. At least about 40% of the genomic regions of the selector set may be regions identified in Table 8.

[0533] The selector set may comprise one or more genomic regions identified by Table 9. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, or 1500 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 9. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 9.

The genomic regions of the selector set may comprise at least 1300 regions from those identified in Table 9.

[0534] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 9. At least about 5% of the genomic regions of the selector set may be regions identified in Table 9. At least about 10% of the genomic regions of the selector set may be regions identified in Table 9. At least about 20% of the genomic regions of the selector set may be regions identified in Table 9. At least about 30% of the genomic regions of the selector set may be regions identified in Table 9. At least about 40% of the genomic regions of the selector set may be regions identified in Table 9.

[0535] The selector set may comprise one or more genomic regions identified by Table 10. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, or 330 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 10. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 10.

[0536] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 10. At least about 5% of the genomic regions of the selector set may be regions identified in Table 10. At least about 10% of the genomic regions of the selector set may be regions identified in Table 10. At least about 20% of the genomic regions of the selector set may be regions identified in Table 10. At least about 30% of the genomic regions of the selector set may be regions identified in Table 10. At least about 40% of the genomic regions of the selector set may be regions identified in Table 10.

[0537] The selector set may comprise one or more genomic regions identified by Table 11. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, or 460 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 11. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 11.

[0538] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 11. At least about 5% of the genomic regions of the selector set may be regions identified in Table 11. At least about 10% of the genomic regions of the selector set may be regions identified in Table 11. At least about 20% of the genomic regions of the selector set may be regions identified in Table 11. At least about 30% of the genomic regions of the selector set may be regions identified in Table 11. At least about 40% of the genomic regions of the selector set may be regions identified in Table 11.

[0539] The selector set may comprise one or more genomic regions identified by Table 12. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480 or 500 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 12. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 12.

[0540] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 12. At least about 5% of the genomic regions of the selector set may be regions identified in Table 12. At least about 10% of the genomic regions of the selector set may be regions identified in Table 12. At least about 20% of the genomic regions of the selector set may be regions identified in Table 12. At least about 30% of the genomic regions of the selector set may be regions identified in Table 12. At least about 40% of the genomic regions of the selector set may be regions identified in Table 12.

[0541] The selector set may comprise one or more genomic regions identified by Table 13. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800,

850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, or 1450 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 13. The genomic regions of the selector set may comprise at least 1300 regions from those identified in Table 13.

[0542] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 13. At least about 5% of the genomic regions of the selector set may be regions identified in Table 13. At least about 10% of the genomic regions of the selector set may be regions identified in Table 13. At least about 20% of the genomic regions of the selector set may be regions identified in Table 13. At least about 30% of the genomic regions of the selector set may be regions identified in Table 13. At least about 40% of the genomic regions of the selector set may be regions identified in Table 13.

[0543] The selector set may comprise one or more genomic regions identified by Table 14. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1210, 1220, 1230, or 1240 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 1100 regions from those identified in Table 14. The genomic regions of the selector set may comprise at least 1200 regions from those identified in Table 14.

[0544] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 14. At least about 5% of the genomic regions of the selector set may be regions identified in Table 14. At least about 10% of the genomic regions of the selector set may be regions identified in Table 14. At least about 20% of the genomic regions of the selector set may be regions identified in Table 14. At least about 30% of the genomic regions of the selector set may be regions identified in Table 14. At least about 40% of the genomic regions of the selector set may be regions identified in Table 14.

[0545] The selector set may comprise one or more genomic regions identified by Table 15. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, or 170 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 120 regions from those identified in Table 15. The genomic regions of the selector set may comprise at least 150 regions from those identified in Table 15.

[0546] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 15. At least about 5% of the genomic regions of the selector set may be regions identified in Table 15. At least about 10% of the genomic regions of the selector set may be regions identified in Table 15. At least about 20% of the genomic regions of the selector set may be regions identified in Table 15. At least about 30% of the genomic regions of the selector set may be regions identified in Table 15. At least about 40% of the genomic regions of the selector set may be regions identified in Table 15.

[0547] The selector set may comprise one or more genomic regions identified by Table 16. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, or 2050 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1200 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 1500 regions from those identified in Table 16. The genomic regions of the selector set

may comprise at least 1700 regions from those identified in Table 16. The genomic regions of the selector set may comprise at least 2000 regions from those identified in Table 16.

[0548] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 16. At least about 5% of the genomic regions of the selector set may be regions identified in Table 16. At least about 10% of the genomic regions of the selector set may be regions identified in Table 16. At least about 20% of the genomic regions of the selector set may be regions identified in Table 16. At least about 30% of the genomic regions of the selector set may be regions identified in Table 16. At least about 40% of the genomic regions of the selector set may be regions identified in Table 16.

[0549] The selector set may comprise one or more genomic regions identified by Table 17. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, or 1080 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 1000 regions from those identified in Table 17. The genomic regions of the selector set may comprise at least 1050 regions from those identified in Table 17.

[0550] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 17. At least about 5% of the genomic regions of the selector set may be regions identified in Table 17. At least about 10% of the genomic regions of the selector set may be regions identified in Table 17. At least about 20% of the genomic regions of the selector set may be regions identified in Table 17. At least about 30% of the genomic regions of the selector set may be regions identified in Table 17. At least about 40% of the genomic regions of the selector set may be regions identified in Table 17.

[0551] The selector set may comprise one or more genomic regions identified by Table 18. The genomic regions of the selector set may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 375, 400, 420, 440, 460, 480, 500, 520, 540, or 555 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 2 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 20 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 60 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 100 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 200 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 300 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 400 regions from those identified in Table 18. The genomic regions of the selector set may comprise at least 500 regions from those identified in Table 18.

[0552] At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions of the selector set may be regions identified in Table 18. At least about 5% of the genomic regions of the selector set may be regions identified in Table 18. At least about 10% of the genomic regions of the selector set may be regions identified in Table 18. At least about 20% of the genomic regions of the selector set may be regions identified in Table 18. At least about 30% of the genomic regions of the selector set may be regions identified in Table 18. At least about 40% of the genomic regions of the selector set may be regions identified in Table 18.

[0553] Selector set probes may be at least about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 nucleotides in length. Selector set probes may be at least about 20 nucleotides in length. Selector set probes may be at least about 30 nucleotides in length. Selector set probes may be at least about 40 nucleotides in length. Selector set probes may be at least about 50 nucleotides in length.

[0554] Selector probes may be of about 15 to about 250 nucleotides in length. Selector set probes may be about 15 to about 200 nucleotides in length. Selector set probes may be about 15 to about 170 nucleotides in length. Selector set probes may be about 15 to about 150 nucleotides in length. Selector set probes may be about 25 to about 200 nucleotides in length. Selector set probes may be about 25 to about 150 nucleotides in length. Selector set probes may be about 50 to about 150 nucleotides in length. Selector set probes may be about 50 to about 125 nucleotides in length.

[0555] 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more selector set probes may correspond to one genomic region. Two or more selector set probes may correspond to one genomic region. Three or more selector set probes may correspond to one genomic region. A set of selector set probes therefore may have the complexity of the selector set from which it is obtained. Selector probes may be synthesized using conventional methods, or generated by any other suitable molecular biology approach. Selector probes may be hybridized to cfDNA for hybrid capture, as described herein. Selector probes may comprise a binding moiety that allows capture of the hybrid. Various binding moieties (e.g., tags) useful for this purpose are known in the art, including without limitation biotin, HIS tags, MYC tags, FITC, and the like.

[0556] Exemplary selector sets are provided in Tables 2, and 6-18. The selector set comprising one or more genomic regions identified in Table 2 may be useful for non-small cell lung carcinoma (NSCLC). The selector set comprising one or more genomic regions identified in Table 6 may be useful for

breast cancer. The selector set comprising one or more genomic regions identified in Table 7 may be useful for colorectal cancer. The selector set comprising one or more genomic regions identified in Table 8 may be useful for diffuse large B-cell lymphoma (DLBCL). The selector set comprising one or more genomic regions identified in Table 9 may be useful for Ehrlich ascites carcinoma (EAC). The selector set comprising one or more genomic regions identified in Table 10 may be useful for follicular lymphoma (FL). The selector set comprising one or more genomic regions identified in Table 11 may be useful for head and Neck squamous cell carcinoma (HNSC). The selector set comprising one or more genomic regions identified in Table 12 may be useful for NSCLC. The selector set comprising one or more genomic regions identified in Table 13 may be useful for NSCLC. The selector set comprising one or more genomic regions identified in Table 14 may be useful for ovarian cancer. The selector set comprising one or more genomic regions identified in Table 15 may be useful for ovarian cancer. The selector set comprising one or more genomic regions identified in Table 16 may be useful for pancreatic cancer. The selector set comprising one or more genomic regions identified in Table 17 may be useful for prostate adenocarcinoma. The selector set comprising one or more genomic regions identified in Table 18 may be useful for skin cutaneous melanoma. The selector set of any one of Tables 2 and 6-18 may be useful for carcinomas and sub-generically for adenocarcinomas or squamous cell carcinomas.

[0557] Methods for Producing a Selector Set

[0558] Disclosed herein are methods of producing a selector set. One objective in designing a selector set may comprise maximizing the fraction of patients covered and the number of mutations per patient covered while minimizing selector size. Evaluating all possible combinations of genomic regions to build such a selector set may be an exponentially large problem (e.g., 2^n possible exon combinations given n exons), rendering the use of an approximation algorithm critical. Thus, a heuristic strategy may be used to produce a selector set.

[0559] The selector sets disclosed herein may be rationally designed for a given ctDNA detection limit, sequencing cost, and/or DNA input mass. Such a selector set may be designed using a selector design calculator. A selector design calculator may be based on the following analytical model: the probability P of recovering at least 1 read of a single mutant allele in plasma for a given sequencing read depth and detection limit of ctDNA in plasma may be modeled by a binomial distribution. Given P, the probability of detecting all identified tumor mutations in plasma may be modeled by a geometric distribution. With this design calculator, one can first estimate how many tumor reporters will be needed to achieve a desired sensitivity, and can then target a selector size that balances this number with considerations of cost and DNA mass input. FIG. **5***a* shows a graphical representation of the probability P of detecting ctDNA in plasma for different detection limits of ctDNA in plasma for CAPP-Seq (dark, thick line), whole exome sequence (i and ii), and whole genome sequence (iii).

[0560] The method of producing a selector set may comprise (a) calculating a recurrence index of a genomic region of a plurality of genomic regions by dividing a number of subjects that have one or more mutations in the genomic region by a length of the genomic region; and (b) producing a selector set comprising one or more genomic regions of the plurality of genomic regions by selecting genomic regions based on the recurrence index. For example, 10 subjects may contain one or more mutations in a genomic region comprising 100 bases. The recurrence index could be calculated by dividing the number of subjects containing mutations in the one or more genomic regions by the length of the genomic region. In this example, the recurrence index for this genomic region would be 10 subjects divided by 100 bases, which equals 0.1 subjects per base.

[0561] The method may further comprise ranking genomic regions of the plurality of genomic regions by the recurrence index. Producing the selector set based on the recurrence index may comprise selecting genomic regions that have a recurrence index in the top 70th, 75th, 80th, 85th, $\mathbf{90}^{th}$, or 95th or greater percentile. Producing the selector set based on the recurrence index may comprise selecting genomic regions that has a recurrence index in the top 90th percentile. For example, a first genomic region may have a recurrence index in the top 80th percentile and a second genomic region may have a recurrence index in the bottom 20th percentile. The selector set based on genomic regions with a recurrence index in the top 75th percentile may comprise the first genomic region, but not the second genomic region.

[0562] The method may further comprise ranking the genomic regions by the number of subjects having one or more mutations in the genomic region. Producing the selector set may further comprise selecting genomic regions in the top 70th, 75th, 80th, 85th, 90th, or 95th or greater percentile of number of subjects having one or more mutations in the genomic region. Producing the selector set may further comprise selecting genomic regions in the top 90th or greater percentile of number of subjects having one or more mutations in the genomic region.

[0563] The length of the genomic region may be in kilobases. The length of the genomic region may be in bases. For genomic regions containing known somatic mutations associated with a cancer, the length of the genomic region may consist essentially on the subsequence of the known mutation. For genomic regions containing known somatic mutations associated with a cancer, the length of the genomic region may consist essentially on the subsequence of the known mutation and one or more bases flanking the subsequence of the known mutation. For genomic regions containing known somatic mutations associated with a cancer, the length of the genomic region may consist essentially on the subsequence of the known mutation and 1 to 5 bases flanking the subsequence of the known mutation. For genomic regions containing known somatic mutations associated with a cancer, the length of the genomic region may consist essentially on the subsequence of the known mutation and 5 or fewer bases flanking the subsequence of the known mutation. The recurrence index for a genomic region comprising a known somatic mutation may be recalculated based on the length of the subsequence of the known mutation or the length of the subsequence of the known mutation with additional bases flanking the subsequence of the known mutation. For example, a genomic region may comprise 200 bases and the known somatic mutation within the genomic region may comprise 100 bases. The recurrence index may be calculated by dividing the number of subjects containing one or more mutations in the genomic region divided by the length of the somatic mutation with the genomic region (e.g., 100 bases).

[0564] Further disclosed herein is a method of producing a selector set comprising (a) identifying, with the aid of a

computer processor, a plurality of genomic regions comprising one or more mutations by analyzing data pertaining to the plurality of genomic regions from a population of subjects suffering from a cancer; and (b) applying an algorithm to the data to produce a selector set comprising two or more genomic regions of the plurality of genomic regions, wherein the algorithm is used to maximize a median number of mutations in the genomic regions of the selector set in the population of subjects.

[0565] Identifying the plurality of genomic regions may comprise calculating a recurrence index of one or more genomic regions of the plurality of genomic regions. The algorithm may be applied to the data pertaining to genomic regions with a recurrence index in the top 40th, 45th, 50th, 55th, 57th, 60th, 63rd, or 65th or higher percentile. The algorithm may be applied to data pertaining to genomic regions having a recurrence index of at least about 15, 20, 25, 30, 35, 40, 45, or 50 or more.

[0566] Identifying the plurality of genomic regions may comprise determining a number of subjects having one or more mutations in a genomic region. The algorithm may be applied to the data pertaining to genomic regions in the top 40th, 45th, 50th, 55th, 57th, 60th, 63rd, or 65th or greater percentile of number of subjects having one or more mutations in the genomic region

[0567] The algorithm may maximize the median number of mutations by identifying genomic regions that result in the largest reduction in subjects with one mutation in the genomic region. Producing the selector set may comprise selecting genomic regions that result in the largest reduction in subjects with one mutation in the genomic region.

[0568] The algorithm may be applied to the data pertaining to genomic regions meeting a minimum threshold. The minimum threshold may pertain to the recurrence index. For example, the algorithm may be applied to genomic regions having a recurrence index in the top 60th percentile. In another example, the algorithm may be applied to genomic regions that have a recurrence index of greater than or equal to 30. Alternatively, or additionally, the minimum threshold may pertain to genomic regions in the top 60th percentile of the number of subjects having one or more mutations in the genomic region.

[0569] The algorithm may be applied 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more times. The algorithm may be applied one or more times. The algorithm may be applied two or more times. The algorithm may be applied to a first set of genomic regions meeting a first minimum threshold. For example, the algorithm may be applied to a first set of genomic regions in the top 60th percentile of the recurrence index and the top 60th percentile of the number of subjects having one or more mutations in the genomic region. The algorithm may be applied a second set of genomic regions meeting a second minimum threshold. For example, the algorithm may be applied to a second set of genomic regions having a recurrence index of greater than or equal to 20.

[0570] The median number of mutations in the genomic regions in the population of subjects may be at least about 2, 3, 4, 5, 6, 7, 8, 9, 10 or more mutations. The median number of mutations in the genomic regions in the population of subjects may be at least about 2, 3, or 4 or more mutations.

[0571] The algorithm may further be used to maximize a number of subjects containing one or more mutations within the genomic regions in the selector set. The algorithm may further be used to maximize a percentage of subjects from the population containing the one or more mutations within the genomic regions in the selector set. The percentage of subjects from the population containing the one or more mutations within the genomic regions may be at least about 60%, 65%, 70%, 75%, 80%, 85%, 87%, 90%, 92%, 95%, or 97% or more.

[0572] Alternatively, the method of producing a selector set may comprise (a) obtaining data pertaining to a plurality of genomic regions from a population of subjects suffering from a cancer; and (b) applying an algorithm to the data to produce a selector set comprising two or more genomic regions of the plurality of genomic regions, wherein the algorithm is used to maximize a number of subjects containing one or more mutations within the genomic regions in the selector.

[0573] The algorithm may maximize the number of subjects containing the one or more mutations by calculating a recurrence index of the genomic regions. Producing the selector set may comprise selecting one or more genomic regions based on the recurrence index.

[0574] The algorithm may maximize the number of subjects containing the one or more mutations by identifying genomic regions comprising one or more mutations found in 2, 3, 4, 5, 6, 7, 8, 9, 10 or more subjects. The algorithm may maximize the number of subjects containing the one or more mutations by identifying genomic regions comprising one or more mutations found in 5 or more subjects. Producing the selector set may comprise selecting one or more genomic regions based on a frequency of the mutation within the genomic region in the population of subjects.

[0575] Producing the selector set may comprise iterative addition of the genomic regions to the selector set. Producing the selector set may comprise selecting one or more genomic regions that identify mutations in at least one new subject from the population of subjects. For example, a selector set may comprise genomic regions A, B, and C, which contain mutations observed in subjects 1, 2, 3, 4, 5, 6, 7 and 8. Genomic region D may contain a mutation observed in subjects 1-4 and 10. Genomic region E may contain a mutation observed in subjects 1-5. Genomic region D identified at least one additional subject (e.g., subject 10) and may be added to the selector set, whereas genomic region E did not identify an additional subject and is not added to the selector set.

[0576] Producing the selector set may comprise selecting one or more genomic regions based on minimizing overlap of subjects already identified by the selector. For example, a selector set may comprise genomic regions A, B, C, and D, which contain mutations observed in subjects 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. Genomic region E may contain a mutation observed in subjects 1-5, 11, and 13. Genomic region F may contain a mutation observed in subjects 12 and 15. Genomic region E had 5 subjects in common with the selector set, whereas genomic region F had no subjects in common with the selector set. Thus, genomic region F may be added to the selector set.

[0577] The algorithm may be used to maximize a percentage of subjects from the population containing the one or more mutations within the genomic regions in the selector. The percentage of subjects from the population containing the one or more mutations within the genomic regions may be at least about 60%, 65%, 70%, 75%, 80%, 85%, 87%, 90%, 92%, 95%, or 97% or more.

[0578] The algorithm may further be used to maximize a median number of mutations in the genomic regions in a subject of the population of subjects. The median number of

mutations in the genomic regions in the subject may be at least about 2, 3, 4, 5, 6, 7, 8, 9, 10 or more mutations. The median number of mutations in the genomic regions in the subject may be at least about 2, 3, or 4 or more mutations.

[0579] Producing the selector set may further comprise adding genomic regions comprising one or more mutations known to be associated with a cancer. Producing the selector set may further comprise adding genomic regions comprising one or more mutations predicted to be associated with a cancer. Producing the selector set may further comprise adding genomic regions comprising one or more rearrangements. Producing the selector set may further comprise adding genomic regions comprising one or more fusions.

[0580] The method may further comprise identifying one or more genomic regions that contain one or more recurrent mutations in a cancer. The identification of these recurrent mutations may benefit greatly from the availability of databases such as, for example, The Cancer Genome Atlas (TCGA) and its subsets. Such databases may serve as the starting point for identifying the recurrently mutated genomic regions of the selector sets. The databases may also provide a sample of mutations occurring within a given percentage of subjects with a specific cancer.

[0581] The method of producing a selector set may comprise (a) identifying a plurality of genomic regions; (b) prioritizing the plurality of genomic regions; and (c) selecting one or more genomic regions for inclusion in a selector set. The following design strategy can be used to identify and prioritize genomic regions for inclusion in a selector set. Three phases may incorporate known and suspected driver genes, as well as genomic regions known to participate in clinically actionable fusions, while another three phases may employ an algorithmic approach to maximize both the number of patients covered and SNVs per patient, utilizing the "Recurrence Index" (RI) as described herein. The strategy may utilize an initial patient database to evaluate the utility of including genomic regions in the selector set. A typical database for this purpose may include sequence information from at least 25, at least 50, at least 100, at least 200, at least 300 or more individual tumors. The method for producing a selector set may comprise one or more of the following phases:

[0582] Phase 1 (Known drivers). Genes known to be drivers in the cancer of interest are selected based on the pattern of SNVs previously identified in tumors.

[0583] Phase 2 (Maximize coverage). To maximize coverage, for each exon with SNVs covering ≥5 cancer patients in the starting database, select the exon with highest RI that identified at least 1 new patient when compared to the prior phase. Among exons with equally high RI, add the exon with minimum overlap among patients already captured by the selector. Repeat until no further exons met these criteria.

[0584] Phase 3 (RI≥30). For each remaining exon with an RI≥30 and with SNVs covering ≥3 patients in the relevant database, identify the exon that results in the largest reduction in patients with only 1 SNV. To break ties among equally best exons, the exon with highest RI was chosen. This was repeated until no additional exons satisfied these criteria.

[0585] Phase 4 (RI≥20). Repeat the procedure in Phase 3, but using RI≥20.

[0586] Phase 5 (Predicted drivers). Add in all exons from additional genes previously predicted to harbor driver mutations in the cancer of interest.

[0587] Phase 6 (Add fusions). Add in for known recurrent rearrangements the introns most frequently implicated in the fusion event and the flanking exons.

[0588] It should be understood, however, that the addition of known drivers, predicted drivers and fusions can be performed independently and in any order.

[0589] A method of producing a selector set may comprise (a) calculating a recurrence index for a plurality of genomic regions from a population of subjects suffering from a cancer by dividing a number of subjects containing one or more mutations in a genomic region of the plurality of genomic regions by a size of the genomic region; and (b) ranking the plurality of genomic regions based on their recurrence index.

[0590] A method of producing a selector set may comprise (a) calculating a recurrence index for a plurality of genomic regions from a population of subjects suffering from a cancer by dividing a number of subjects containing one or more mutations in a genomic region of the plurality of genomic regions by a size of the genomic region; and (b) producing a selector set comprising two or more genomic regions of the plurality of genomic regions by (i) using the recurrence index to maximize coverage of the selector set for the population of subjects; and/or (ii) using the recurrence index to maximize a median number of mutations per subject in the population of subjects.

[0591] Maximizing subject coverage may comprise use of a metric termed "Recurrence Index" (RI). The RI may refer to the number of subjects that harbor mutations (e.g., SNVs/indels) in a given kilobase of genomic sequence. This metric can be further normalized by the number of subjects per study to allow comparison of different studies and distinct cancers. A similar approach was used to produce a selector set for non-small cell lung cancer (NSCLC) (see FIG. **1***b*). For one exemplary NSCLC selector set, exons were the primary genomic unit and indels were not considered. A portion of an exon may contain known somatic mutations. In this case, the algorithm only includes the subsequence of the portion of the exon containing known lesions flanked by a user-defined buffer (by default, =1 base). RI may be recalculated for each exon following this adjustment. The algorithm may rank genomic regions by decreasing RI. The algorithm may consider a subset of the genomic regions. For example, the algorithm may only consider genomic regions in the top P percentile of both RI and/or the number of subjects per exon ($P=90^{th}$ percentile by default, but is user modifiable). Selector design may proceed by iteratively traversing the list of ranked genomic regions, selecting each genomic region that adds additional subject coverage with minimal additional space. This may continue until all genomic regions satisfying percentile filters have been evaluated and/or a user-defined maximum selector size has been reached.

[0592] Producing the selector set may comprise maximizing the median number of mutations per subject. Maximizing the median number of mutations per subject may comprise use of one or more algorithms. Maximizing the median number of mutations per subject may comprise use of one or more thresholds or filters to evaluate the genomic regions for inclusion in the selector set. The thresholds or filters may be based on the recurrence index. For example, the filter may be a percentile filter of the recurrence index. The percentile filters may be relaxed to permit the assessment of additional genomic regions for inclusion in the selector set. The percentile filter may be set at (⅔)×P, where P is a top percentile of RI. The threshold may be user-defined. The threshold may be

greater than or equal to ⅔. Alternatively, the threshold is less than or equal to ⅔. P may also be user-defined. The algorithm may proceed through the list of genomic regions ranked by decreasing RI, iteratively adding regions that maximally increase the median number of mutations per subject. The process may terminate after assessing all genomic regions that pass percentile filters, and/or if the desired selector size endpoint is reached. This process may be repeated for a third round or more by continuing to relax the percentile threshold. Maximizing the median number of mutations per subject may comprise (i) ranking two or more genomic regions based on their recurrence index; (ii) producing a list of genomic regions comprising a subset of the genomic regions, wherein the genomic regions in the list have a recurrence index in the top 60th percentile; and (iii) producing a preliminary selector set by adding genomic regions to the preliminary selector set and calculating a median number of mutations per subject in the preliminary selector set.

[0593] Further disclosed herein is a method of producing a selector set comprising (a) obtaining data pertaining to one or more genomic regions; (b) applying an algorithm to the data to determine for a genomic region: (i) a presence of one or more mutations in the genomic region; (ii) a number of subjects with mutations in that genomic region; and (iii) a recurrence index (RI), wherein the RI is determined by dividing the number of subjects with mutations in the genomic region by the size of genomic region; and (c) producing a selector set comprising one or more genomic regions based on the recurrence index of the one or more genomic regions.

[0594] The method may further comprise recalculating the recurrence index for one or more genomic regions comprising known mutations. The size of the known mutation may be less than the size of the genomic region. Recalculating the recurrence index may comprise dividing the number of subjects with known mutations in the genomic region by the size of the known mutation. For example, the size of a genomic region may be 200 basepairs and the size of the known mutation within the genomic region may be 100 basepairs. The recurrence index for the genomic region may be determined by dividing the number of subjects with the known mutation in the genomic region by the size of the known mutation (e.g., 100 base pairs) rather than dividing by the size of the entire genomic region (e.g., 200 base pairs).

[0595] The method may further comprise ranking the two or more genomic regions based on the recurrence index. The list of ranked genomic regions may comprise a subset of the genomic regions ranked by the recurrence index. The list of ranked genomic regions may comprise a subset of the genomic regions that satisfy one or more criteria. The one or more criteria may be based on the recurrence index. For example, the list of ranked genomic regions may comprise a subset of genomic regions that have a recurrence index in the top 90th percentile. Producing the selector set may comprise selecting the one or more genomic regions based on the recurrence index. Producing the selector set may comprise selecting the one or more genomic regions based on the rank of the two or more genomic regions. The two or more genomic regions may be ranked with the aid of an algorithm. The algorithm used to rank the two or more genomic regions based on the recurrence may be the same algorithm used to determine the recurrence index of the one or more genomic regions. The algorithm may be a different from the algorithm used to determine the recurrence index.

[0596] The method may further comprise iteratively traversing a list of ranked genomic regions and selecting genomic regions that provide additional subject coverage with minimal addition to the total size of the genomic regions of a proposed selector set. For example, a first genomic region may add two new subjects to the proposed selector set and the size of the proposed selector set may increase by 10 base pairs, whereas a second genomic region may add two new subjects to the proposed selector set and the size of the proposed selector set may increase by 100 base pairs. The first genomic region may be selected over the second genomic region for inclusion in the proposed selector set. The entire list of ranked genomic regions may be traversed. Alternatively, a portion of the list of ranked genomic regions may be traversed. For example, the traversal and selection of genomic regions may be based on a user-defined maximum selector size. Once the maximum selector size has been reached, the step of traversing the list of ranked genomic regions and selecting genomic regions may be terminated. An algorithm may be used to traverse the list of ranked genomic regions and to select genomic regions for inclusion in the selector set. The algorithm may be the same algorithm used to determine the recurrence index. The algorithm may be a different from the algorithm used to determine the recurrence index.

[0597] The method may further comprise iteratively traversing a list of ranked genomic regions and selecting genomic regions that maximize the median number of mutations per subject in the population of subjects of the selector set. The median number of mutations per subject for a proposed selector set may be determined by (a) counting a number of mutations N in each subject across all genomic regions for the proposed selector set; and (b) applying an algorithm to identify the median number of mutations by sorting the subjects by the number of mutations. For example, a proposed selector set may comprise 10 genomic regions comprising 20 mutations in a population of 9 subjects. A first subject may have 4 mutations, a second subject may have 2 mutations, a third subject may have 3 mutations, a fourth subject may have 6 mutations, a fifth subject have may 8 mutations, a sixth subject may have 6 mutations, a seventh subject may have eight mutations, an eighth subject may have 4 mutations, and a ninth subject may have two mutations. The median of {2, 2, 3, 4, 4, 6, 8, 8} is 4. A genomic region may be selected for inclusion in the selector set if the inclusion of the genomic region increases the median number of mutations per subject in the population of subjects in the selector set. For example, a first genomic region may contain one mutation present in two of the ten subjects and second genomic region may contain one mutation present in three of the ten subjects. The second genomic region may be selected for inclusion into the selector set over the first genomic region because addition of the second genomic region to the selector set would result in a greater increase the median number of mutations per subject than addition of the first genomic region. The entire list of ranked genomic regions may be traversed. Alternatively, a portion of the list of ranked genomic regions may be traversed. For example, the traversal and selection of genomic regions may be based on a user-defined maximum selector size. Once the maximum selector size has been reached, the step of traversing the list of ranked genomic regions and selecting genomic regions may be terminated.

[0598] Methods of producing a selector set may comprise: (a) obtaining sequencing information of a tumor sample from a subject suffering from a cancer; (b) comparing the sequenc-

ing information of the tumor sample to sequencing information from a non-tumor sample from the subject to identify one or more mutations specific to the sequencing information of the tumor sample; and (c) producing a selector set comprising one or more genomic regions comprising the one or more mutations specific to the sequencing information of the tumor sample. The selector set may comprise sequencing information pertaining to the one or more genomic regions. The selector set may comprise genomic coordinates pertaining to the one or more genomic regions. The selector set may comprise a plurality of oligonucleotides that selectively hybridize the one or more genomic regions. The plurality of oligonucleotides may be biotinylated. The one or more mutations comprise SNVs. The one or more mutations comprise indels. The one or more mutations comprise rearrangements. Producing the selector set may comprise identifying tumor-derived SNVs based on the methods disclosed herein. Producing the selector set may comprise identifying tumor-derived rearrangements based on the methods disclosed herein.

[0599] Application of the approaches described herein for mutated genomic regions in non-small cell lung cancer may result in the selector set shown in Table 2. The selector set created according to the methods of the invention may identify genomic regions that are highly likely to include identifiable mutations in tumor sequences. This selector set may include a relatively small total number of genomic regions and thus a relatively short cumulative length of genomic regions and yet may provide a high overall coverage of likely mutations in a population. The selector set does not, therefore, need to be optimized on a patient-by-patient basis. The relatively short cumulative length of genomic regions also means that the analysis of cancer-derived cell-free DNA using these libraries may be highly sensitive. The relatively short cumulative length of genomic regions may allow the sequencing of cell-free DNA to a great depth.

[0600] The selector sets comprising recurrently mutated genomic regions created according to the instant methods may enable the identification of patient-specific mutations and/or tumor-specific mutations within the genomic regions in a high percentage of subjects. Specifically, in these selector sets, at least one mutation within the plurality of genomic regions may be present in at least 60% of a population of subjects with the specific cancer. In some embodiments, at least two mutations within the plurality of genomic regions are present in at least 60% of a population of subjects with the specific cancer. In specific embodiments, at least three mutations, or even more, within the plurality of genomic regions are present in at least 60% of a population of subjects with the specific cancer.

[0601] The methods for creating a selector set, as disclosed herein, may be implemented by a programmed computer system. Therefore, according to another aspect, the instant disclosure provides computer systems for creating a selector set (e.g., library of recurrently mutated genomic regions). Such systems may comprise at least one processor and a non-transitory computer-readable medium storing computer-executable instructions that, when executed by the at least one processor, cause the computer system to carry out the methods described herein for creating a selector set (e.g., library).

[0602] ctDNA Detection Index

[0603] The methods, kits and systems disclosed herein may comprise a ctDNA detection index or use thereof. Generally, the ctDNA detection index is based on a p-value of one or more types of mutations present in a sample from a subject. The ctDNA detection index may comprise an integration of information content across a plurality of mutations and classes of somatic mutations. The ctDNA detection index may be analogous to a false positive rate. The ctDNA detection index may be based on a decision tree in which fusion breakpoints take precedence due to their nonexistent background and/or in which p-values from multiple classes of mutations may be integrated. The classes of mutations may include, but are not limited to, SNVs, indels, copy number variants, and rearrangements.

[0604] The ctDNA detection index may be used to assess the statistical significance of a selector set comprising genomic regions comprising multiple classes of mutations. For example, the ctDNA detection index may be used to assess the statistical significance of a selector set comprising genomic regions comprising SNVs and indels. In another example, the ctDNA detection index may be used to assess the statistical significance of a selector set comprising genomic regions comprising SNVs and rearrangements. In another example, the ctDNA detection index may be used to assess the statistical significance of a selector set comprising genomic regions comprising rearrangements and indels. In another example, the ctDNA detection index may be used to assess the statistical significance of a selector set comprising genomic regions comprising SNVs, indels, copy number variants, and rearrangements. The calculation of the ctDNA detection index may be based on the types (e.g., classes) of mutations within the genomic region of a selector set that are detected in a subject. For example, a selector set may comprise genomic regions comprising SNVs, indels, copy number variants, and rearrangements, however, the types of mutations for the selector that are detected in a subject may be SNVs and indels. The ctDNA detection index may be determined by combining a p-value of the SNVs and a p-value of the indels. Any method that is suitable for combining independent, partial tests may be used to combine the p-value of the SNVs and indels. Combining the p-values of the SNVs and indels may be based on Fisher's method.

[0605] A method of determining a ctDNA detection index may comprise (a) detecting a presence of one or more mutations in one or more samples from a subject, wherein the one or more mutations are based on a selector set comprising genomic regions comprising the one or more mutations; (b) determining a mutation type of the one or more mutations present in the sample; and (c) calculating a ctDNA detection index based on a p-value of the mutation type of mutations present in the one or more samples.

[0606] For instances in which a single type of mutation is present in the sample from the subject, the ctDNA detection index is based on the p-value of the single type of mutation. The p-value of the single type of mutation may be estimated by Monte Carlo sampling. Monte Carlo sampling may use a broad class of computational algorithms that rely on repeated random sampling to obtain a p-value. The ctDNA detection index may be equivalent to the p-value of the single type of mutation.

[0607] For instances in which a rearrangement (e.g., fusion) is detected in a tumor sample and a plasma sample from the subject, the ctDNA detection index is based on the p-value of the rearrangement. The p-value of the rearrangement may be 0. Thus, the ctDNA detection index is the p-value of the rearrangement, which is 0.

[0608] For instances in which a rearrangement (e.g., fusion) is detected in only a tumor sample from the subject

and not in a plasma sample from the subject, the ctDNA detection index is based on the p-value of the other types of mutations.

[0609] For instances in which (a) a SNV and indel are detected in a sample from the subject; (b) a p-value of the SNV is less than 0.1 and a p-value of the indel is less than 0.1; and (c) a rearrangement is not detected in a plasma sample from the subject, the ctDNA detection index is calculated based on the combined p-values of the SNV and indel. Any method that is suitable for combining independent, partial tests may be used to combine the p-value of the SNVs and indels. The p-values of the SNV and indel may be combined according to Fisher's method. Thus, the ctDNA detection index is the combined p-value of the SNV and indel.

[0610] For instances in which (a) a SNV and indel are detected in a sample from the subject; (b) a p-value of the SNV is not less than 0.1 or a p-value of the indel is not less than 0.1; and (c) a rearrangement is not detected in a plasma sample from the subject, the ctDNA detection index is based on the p-value of the SNV. Thus, the ctDNA detection index is the p-value of the SNV.

[0611] A ctDNA detection index may be significant if the ctDNA detection index is less than or equal to 0.10, 0.09, 0.08, 0.07, 0.06, 0.05, 0.04, 0.03, 0.02, or 0.01. A ctDNA detection index may be significant if the ctDNA detection index is less than or equal to 0.05. A ctDNA detection index may be significant if the ctDNA detection index is less than or equal to a false positive rate (FPR).

[0612] A ctDNA detection index may be calculated for a subject based on his or her array of reporters (e.g., mutations) using the following rules, executed in any order:

[0613] (i) For cases where only a single reporter type is present in a patient's tumor, the corresponding p-value is used (estimated by Monte Carlo sampling).

[0614] (ii) If SNV and indel reporters are detected, and if each independently has a p-value <0.1, their respective p-values are combined using Fisher's method. Otherwise, given the prioritization of SNVs in the selector design, the SNV p-value is used.

[0615] (iii) If a fusion breakpoint identified in a tumor sample (e.g., involving ROS1, ALK, or RET) is recovered in plasma DNA from the same patient, it trumps all other mutation types, and its p-value (~0) is used.

[0616] (iv) If a fusion detected in the tumor is not found in corresponding plasma (potentially due to hybridization inefficiency), the p-value for any remaining mutation type(s) is used.

[0617] The ctDNA detection index may be considered significant if the ctDNA detection index is ≤0.05 (≈false positive rate (FPR)≤5%), which is the threshold that maximized CAPP-Seq sensitivity and specificity in ROC analyses (determined by Euclidean distance to a perfect classifier; e.g., true positive report (TPR)=1 and FPR=0).

[0618] Calculating a ctDNA detection index may comprise determining a significance of SNVs. In some embodiments, to evaluate the significance SNVs, the strategy integrates cfDNA fractions across all somatic SNVs, performs a position-specific background adjustment, and evaluates statistical significance by Monte Carlo sampling of background alleles across the selector. This allows the quantitation of low levels of ctDNA with potentially high rates of allelic drop out. The method for evaluating the significance of SNVs may utilize the following steps:

[0619] adjusting the allelic fraction f for each of n SNVs from patient P for a given cfDNA sample θ by the operation $f^*=\max\{0, f-(e-\mu)\}$, where f is the raw allelic fraction in cfDNA, e is the position-specific error rate for the given allele across all cfDNA samples, and μ denotes the mean selector-wide background rate;

[0620] comparing with Monte Carlo simulation the adjusted mean SNV fraction $F^*(=(\Sigma f^*)/n)$ against the null distribution of background alleles across the selector;

[0621] determining a SNV p-value for patient P as the percentile of F* with respect to the null distribution of background alleles in θ.

[0622] Calculating a ctDNA detection index may comprise determining a significance of rearrangements. The recovery of a tumor-derived genomic fusion (rearrangement) can be assigned a p-value of ~0, due to the very low error rate.

[0623] Calculating a ctDNA detection index may comprise determining a significance of indels. The analysis of insertions and deletions (indels) may be separately evaluated utilizing the following steps:

[0624] For each indel in patient P compare its fraction in a given cfDNA sample θ against its fraction in every cfDNA sample in a cohort (excluding cfDNA samples from the same patient P) with a Z-test; where each read strand is optionally assessed separately and combined into a single Z-score;

[0625] if patient P has more than 1 indel, all indel-specific Z-scores are combined into a final Z statistic.

[0626] The p-values of the different mutation types may be integrated to estimate the statistical significance (e.g., p-value) of tumor burden quantitation. Thus, the ctDNA detection index, which integrates the p-values of different mutation types, may be used to estimate the statistical significance of tumor burden quantitation. For each sample, a ctDNA detection index may be calculated based on p-value integration from the plurality of somatic mutations that are detected. The ctDNA detection index may be determined based on the methods disclosed herein. For cases where only a single somatic mutation is present in a sample, the corresponding p-value may be used. If a fusion breakpoint identified in a tumor sample is recovered in cfDNA from the same patient, the p-value of the fusion breakpoint may be used. If SNV and indel somatic mutations are detected, and if each independently has a p-value <0.1, their respective p-values may be combined and the resulting p-value is used. If the ctDNA detection index is determined to be 0.05, then the p-value of the tumor burden quantitation is 0.05. A ctDNA detection index of ≤0.05 may suggest that a subject's mutations are significantly detectable in a sample from the subject. A ctDNA detection index that is less than the false positive rate (FPR) may suggest that a subject's mutations are significantly detectable in a sample from the subject.

[0627] Selector Set Sensitivity and Specificity

[0628] The selector set may be chosen to provide a desired sensitivity and/or specificity. As is known in the art, the relative sensitivity and/or specificity of a predictive model can be "tuned" to favor either the selectivity metric or the sensitivity metric, where the two metrics have an inverse relationship. One or both of sensitivity and specificity can be at least about at least about 0.6, at least about 0.65, at least about 0.7, at least about 0.75, at least about 0.8, at least about 0.85, at least about 0.9, or higher.

[0629] The sensitivity and specificity may be statistical measures of the performance of selector set to perform a function. For example, the sensitivity of the selector set may be used to assess the use of the selector set to correctly diagnose or prognosticate a status or outcome of a cancer in a subject. The sensitivity of the selector set may measure the proportion of subjects which are correctly identified as suffering from a cancer. The sensitivity of the selector set may also measure the use of the selector set to correctly screen for a cancer in a subject. The sensitivity of the selector set may also measure the use of the selector set to correctly diagnose a cancer in a subject. The sensitivity of the selector set may also measure the use of the selector set to correctly prognosticate a cancer in a subject. The sensitivity of the selector set may also measure the use of the selector set to correctly identify a subject as a responder to a therapeutic regimen. The sensitivity may be at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70% or greater. The sensitivity may be at least about 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or greater.

[0630] Sensitivity may vary according to the tumor stage. The sensitivity may be at least about 50%, at least about 52%, at least about 55%, at least about 57%, at least about 60%, at least about 62%, at least about 65%, at least about 67%, at least about 70%, at least about 72%, at least about 75%, at least about 77%, at least about 80%, at least about 85%, at least about 87%, at least about 90%, at least about 92%, at least about 95%, at least about 98%, at least about 99% or more for tumors at stage I. The sensitivity may be at least about 50% for tumors at stage I. The sensitivity may be at least about 65% for tumors at stage I. The sensitivity may be at least about 72% for tumors at stage I. The sensitivity may be at least about 75% for tumors at stage I The sensitivity may be at least about 85% for tumors at stage I The sensitivity may be at least about 92% for tumors at stage I.

[0631] The sensitivity may be at least about 50%, at least about 52%, at least about 55%, at least about 57%, at least about 60%, at least about 62%, at least about 65%, at least about 67%, at least about 70%, at least about 72%, at least about 75%, at least about 77%, at least about 80%, at least about 85%, at least about 87%, at least about 90%, at least about 92%, at least about 95%, at least about 98%, at least about 99% or more for tumors at stage II. The sensitivity may be at least about 60% for tumors at stage II. The sensitivity may be at least about 75% for tumors at stage II. The sensitivity may be at least about 85% for tumors at stage II. The sensitivity may be at least about 92% for tumors at stage II.

[0632] The sensitivity may be at least about 50%, at least about 52%, at least about 55%, at least about 57%, at least about 60%, at least about 62%, at least about 65%, at least about 67%, at least about 70%, at least about 72%, at least about 75%, at least about 77%, at least about 80%, at least about 85%, at least about 87%, at least about 90%, at least about 92%, at least about 95%, at least about 98%, at least about 99% or more for tumors at stage III. The sensitivity may be at least about 60% for tumors at stage III. The sensitivity may be at least about 75% for tumors at stage III. The sensitivity may be at least about 85% for tumors at stage III. The sensitivity may be at least about 92% for tumors at stage III.

[0633] The sensitivity may be at least about 50%, at least about 52%, at least about 55%, at least about 57%, at least about 60%, at least about 62%, at least about 65%, at least about 67%, at least about 70%, at least about 72%, at least about 75%, at least about 77%, at least about 80%, at least about 85%, at least about 87%, at least about 90%, at least about 92%, at least about 95%, at least about 98%, at least about 99% or more for tumors at stage IV. The sensitivity may be at least about 60% for tumors at stage IV. The sensitivity may be at least about 75% for tumors at stage IV. The sensitivity may be at least about 85% for tumors at stage IV. The sensitivity may be at least about 92% for tumors at stage IV.

[0634] The sensitivity may be at least about and may be at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 87%, at least about 90%, at least about 92%, at least about 95%, at least about 98%, at least about 99% or more with healthy controls.

[0635] The AUC value may also vary according to tumor stage. The AUC value may be at least about 0.50, at least about 0.52, at least about 0.55, at least about 0.57, at least about 0.60, at least about 0.62, at least about 0.65, at least about 0.67, at least about 0.70, at least about 0.72, at least about 0.75, at least about 0.77, at least about 0.80, at least about 0.82, at least about 0.85, at least about 0.87, at least about 0.90, at least about 0.92, at least about 0.95, at least about 0.97 or more for stage I cancer. The AUC value may be at least about 0.50 for stage I cancer. The AUC value may be at least about 0.55 for stage I cancer. The AUC value may be at least about 0.60 for stage I cancer. The AUC value may be at least about 0.70 for stage I cancer. The AUC value may be at least about 0.75 for stage I cancer. The AUC value may be at least about 0.80 for stage I cancer.

[0636] The AUC value may be at least about 0.50, at least about 0.52, at least about 0.55, at least about 0.57, at least about 0.60, at least about 0.62, at least about 0.65, at least about 0.67, at least about 0.70, at least about 0.72, at least about 0.75, at least about 0.77, at least about 0.80, at least about 0.82, at least about 0.85, at least about 0.87, at least about 0.90, at least about 0.92, at least about 0.95, at least about 0.97 or more for stage II cancer. The AUC value may be at least about 0.50 for stage II cancer. The AUC value may be at least about 0.55 for stage II cancer. The AUC value may be at least about 0.60 for stage II cancer. The AUC value may be at least about 0.70 for stage II cancer. The AUC value may be at least about 0.75 for stage II cancer. The AUC value may be at least about 0.80 for stage II cancer. The AUC value may be at least about 0.90 for stage II cancer. The AUC value may be at least about 0.95 for stage II cancer.

[0637] The AUC value may be at least about 0.50, at least about 0.52, at least about 0.55, at least about 0.57, at least about 0.60, at least about 0.62, at least about 0.65, at least about 0.67, at least about 0.70, at least about 0.72, at least about 0.75, at least about 0.77, at least about 0.80, at least about 0.82, at least about 0.85, at least about 0.87, at least about 0.90, at least about 0.92, at least about 0.95, at least about 0.97 or more for stage III cancer. The AUC value may be at least about 0.50 for stage III cancer. The AUC value may be at least about 0.55 for stage III cancer. The AUC value may be at least about 0.60 for stage III cancer. The AUC value may be at least about 0.70 for stage III cancer. The AUC value may be at least about 0.75 for stage III cancer. The AUC value may be at least about 0.80 for stage III cancer. The AUC value may be at least about 0.90 for stage III cancer. The AUC value may be at least about 0.95 for stage III cancer.

[0638] The AUC value may be at least about 0.50, at least about 0.52, at least about 0.55, at least about 0.57, at least about 0.60, at least about 0.62, at least about 0.65, at least about 0.67, at least about 0.70, at least about 0.72, at least

about 0.75, at least about 0.77, at least about 0.80, at least about 0.82, at least about 0.85, at least about 0.87, at least about 0.90, at least about 0.92, at least about 0.95, at least about 0.97 or more for stage IV cancer. The AUC value may be at least about 0.50 for stage IV cancer. The AUC value may be at least about 0.55 for stage IV cancer. The AUC value may be at least about 0.60 for stage IV cancer. The AUC value may be at least about 0.70 for stage IV cancer. The AUC value may be at least about 0.75 for stage IV cancer. The AUC value may be at least about 0.80 for stage IV cancer. The AUC value may be at least about 0.90 for stage IV cancer. The AUC value may be at least about 0.95 for stage IV cancer.

[0639] The AUC values may be at least about 0.70, at least about 0.75, at least about 0.80, at least about 0.85, at least about 0.90, at least about 0.95 for healthy controls.

[0640] The specificity of the selector may measure the proportion of subjects which are correctly identified as not suffering from a cancer. The specificity of the selector set may also measure the use of the selector set to correctly make a diagnosis of no cancer in a subject. The specificity of the selector set may also measure the use of the selector set to correctly identify a subject as a non-responder to a therapeutic regimen. The specificity may be at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70% or greater. The specificity may be at least about 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or greater.

[0641] The selector set may be used to detect, diagnose, and/or prognosticate a status or outcome of a cancer in a subject based on the detection of one or more mutations within one or more genomic regions in the selector set in a sample from the subject. The sensitivity and/or specificity of the selector set to detect, diagnose, and/or prognosticate the status or outcome of the cancer in the subject may be tuned (e.g., adjusted/modified) by the ctDNA detection index. The ctDNA detection index may be used to assess the significance of classes of mutations detected in the sample from the subject by the selector set. The ctDNA detection index may be used to determine whether the detection of one or more classes of mutations by the selector set is significant. For example, the ctDNA detection index may determine that the classes of mutations detected by the selector set in a first subject is statistically significant, which may result in a diagnosis of cancer in the first subject. The ctDNA detection index may determine that the classes of mutations detected by the selector set in a second subject is not statistically significant, which may result in a diagnosis of no cancer in the second subject. As such, the ctDNA detection index may affect the analysis of the specificity and/or sensitivity of the selector set to detect, diagnose, and/or prognosticate the status or outcome of the cancer in the subject.

[0642] Identification of Rearrangements

[0643] Further disclosed herein are methods of identifying rearrangements. The rearrangement may be a genomic fusion event and/or breakpoint. The method may be used for de novo analysis of cfDNA samples. Alternatively, the method may be used for analysis of known tumor/germline DNA samples. The method may comprise a heuristic approach. Generally, the method may comprise (a) obtaining an alignment file of pair-end reads, exon coordinates, a reference genome, or a combination thereof; and (b) applying an algorithm to information from the alignment file to identify one or more rearrangements. The algorithm may be applied to information pertaining to one or more genomic regions. The algorithm may be applied to information that overlaps with one or more genomic regions.

[0644] The method may be termed FACTERA (FACile Translocation Enumeration and Recovery Algorithm). As input, FACTERA may use an alignment file of paired-end reads, exon coordinates, and a reference genome. In addition, the analysis can be optionally restricted to reads that overlap particular genomic regions. FACTERA may process the input in three sequential phases: identification of discordant reads, detection of breakpoints at base pair-resolution, and in silico validation of candidate fusions.

[0645] Further disclosed herein is a method of identifying rearrangements comprising (a) obtaining sequencing information pertaining to a plurality of genomic regions; (b) producing a list of genomic regions adjacent to one or more candidate rearrangement sites; (c) applying an algorithm to validate candidate rearrangement sites, thereby identifying rearrangements.

[0646] The sequencing information may comprise an alignment file. The alignment file may comprise an alignment file of pair-end reads, exon coordinates, and a reference genome. The sequencing information may be obtained from a database. The database may comprise sequencing information pertaining to a population of subjects suffering from a disease or condition. The database may be a pharmacogenomics database. The sequencing information may be obtained from one or more samples from one or more subjects.

[0647] Producing the list of genomic regions adjacent to the one or more candidate rearrangement sites may comprise identifying discordant read pairs based on the sequencing information. A discordant read-pair may refer to a read and its mate, where the insert size is not equal to (e.g., greater or less than) the expected distribution of the dataset, or where the mapping orientation of the reads is unexpected (e.g. both on the same strand). Producing the list of genomic regions adjacent to the one or more candidate rearrangement sites may comprise classifying the discordant read pairs based on the sequencing information.

[0648] Discordant read pairs may be introduced by NGS library preparation and/or sequencing artifacts (e.g., jumping PCR). However, they are also likely to flank the breakpoints of bona fide fusion events. Producing a list of genomic regions adjacent to the one or more candidate rearrangement sites may further comprise ranking the genomic regions. The genomic regions may be ranked in decreasing order of discordant read depth. The method may further comprise eliminating duplicate fragments. Producing a list of genomic regions adjacent to the one or more candidate rearrangement sites may comprise selecting genomic regions with a minimum user-defined read depth. The read depth may be at least 2×, 3×, 4×, 5×, 6×, 7×, 8×, 9×, 10× or more. The read depth may be at least about 2×.

[0649] Producing the list of genomic regions adjacent to the one or more candidate fusion sites may comprise use of one or more algorithms. The algorithm may analyze properly paired reads in which one of the two reads is "soft-clipped," or truncated. Soft-clipping may refer to truncating one or more ends of the paired reads. Soft-clipping may truncate the one or more ends by removing less than or equal to 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 base or base pair from the paired reads. Soft-clipping may comprise removing at least one base or base pair from the paired reads. Soft-clipping may comprise removing at least

one base or base pair from one end of the paired reads. Soft-clipping may comprise removing at least one base or base pair from both ends of the paired reads. Soft-clipped reads may allow for precise breakpoint determination. The precise breakpoint may be identified by parsing the CIGAR string associated with each mapped read, which compactly specifies the alignment operation used on each base (e.g. My=y contiguous bases were mapped, Sx=x bases were skipped). The algorithm may analyze soft-clipped reads with a specific pattern. For example, the algorithm may analyze soft-clipped reads with the following patterns, SxMy or MySx. The number of skipped bases x may have a minimum requirement. By setting a minimum requirement for the number of skipped bases x, the impact of non-specific sequence alignments may be reduced. The number of skipped bases may be at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more. The number of skipped bases may be at least 16. The number of skipped bases may be user-defined. The number of contiguous bases y may also be used-defined.

[0650] An algorithm may be used to validate candidate rearrangement sites. The algorithm may determine the read frequency for the candidate rearrangement sites. The algorithm may eliminate candidate rearrangement sites that do not meet a minimum read frequency. The minimum read frequency may be user-defined. The minimum read frequency may be at least about 2, 3, 4, 5, 6, 7, 8, 9, 10 or more reads. The minimum read frequency may be at least about 2 reads. The algorithm may rank the candidate rearrangement sites based on the read frequency. A candidate rearrangement site may contain multiple soft-clipped reads. The algorithm may select a representative soft-clipped read for a candidate rearrangement site. Selection of the representative soft-clipped read may be based on selecting a soft-clipped read that has a length that is closest to half the read length. If the mapped region of the representative soft-clipped read matches the mapped region of another soft-clipped read of the candidate rearrangement site, the algorithm may annotate the candidate rearrangement site as a rearrangement event. If the mapped region of the representative soft-clipped read matches the mapped region of another soft-clipped read of the candidate rearrangement site, the algorithm may identify the candidate rearrangement site as a rearrangement. If the mapped region of the representative soft-clipped read matches the mapped region of another soft-clipped read of the candidate rearrangement site, the algorithm may annotate the candidate rearrangement site as a fusion event. Applying the algorithm to validate the candidate rearrangements may comprise identifying the candidate rearrangement as a rearrangement if the two or more reads have a sequence alignment.

[0651] Validating the candidate rearrangement sites may further comprise using an algorithm to assess inter-read concordance. The algorithm may assess inter-read concordance by dividing a first sequence read of a soft-clipped sequence of a candidate rearrangement site into multiple possible subsequences of a user-defined length k. A second sequence read of the soft-clipped sequence may be divided into subsequences of length k. Subsequences of size k of the second sequence read may be compared to the first sequencing read, and the concordance of the two reads may be determined. For example, the soft-clipped sequence of a candidate fusion may be 100 bases and the soft-clipped sequence may be subdivided into a user-defined length of 10 bases. The subsequences with a length of 10 may be extracted from the first read and stored. A second read may be compared to the first read by selecting subsequences of 10 bases in the second read. The user-defined lengths may allow parts of the second read to be merged with the soft-clipped (e.g., non-mapping) parts of the first read into a composite sequence which is then assessed for improved mapping properties. Validating the candidate rearrangement may comprise dividing a first read into subsequences of k-mers. A second read may be divided into k-mers in order to rapidly compare it to the first read. If any k-mers overlap the first read, they are counted and used to assess sequence similarity. The two reads may be considered concordant if a minimum matching threshold is achieved. The minimum matching threshold may be a user-defined value. The minimum matching threshold may be 50% of the shortest length of the two sequences being compared. For example, the first sequence read may be 100 bases and the second sequence read may be 130 bases. The minimum matching threshold may be 50 bases (e.g., 100 bases times 0.50). The minimum matching threshold may be at least 10%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, or 80% of the shortest length of the two sequences being compared. The algorithm may process 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000 or more putative breakpoint pairs for each discordant gene (or genomic region) pair. The number of putative breakpoint pairs that the algorithm processes may be user-defined. Moreover, for a gene pair, the algorithm may compare reads whose orientations are compatible with valid fusions. Such reads may have soft-clipped sequences facing opposite directions. When this condition is not satisfied, the algorithm may use the reverse complement of read 1 for k-mer analysis.

[0652] In some instances, genomic subsequences flanking the true breakpoint may be nearly or completely identical, causing the aligned portions of soft-clipped reads to overlap. This may prevent an unambiguous determination of the breakpoint. As such, an algorithm may be used to adjust the breakpoint in one read (e.g., read 2) to match the other (e.g., read 1). For a read, the algorithm may calculate the distance between the breakpoint and the read coordinate corresponding to the first k-mer match between reads. For example, let x be defined as the distance between the breakpoint coordinate of read 1 and the index of the first matching k-mer, j, and y be defined as the corresponding distance for read 2. Then, the offset is estimated as the difference in distances (x, y) between the two reads. Thus, for instances in which a fusion event cannot be unambiguously determined based on the sequence reads, an algorithm is used to determine a fusion site.

[0653] The method may further comprise in silico validation of candidate rearrangement sites. An algorithm may perform a local realignment of reads of the candidate rearrangement sites against a reference rearrangement sequence. The reference rearrangement sequence may be obtained from a reference genome. The local alignment may be of sequences flanking the candidate rearrangement site. The local alignment may be of sequences within 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 or more base pairs of the candidate rearrangement site. The local alignment may be of sequences within 500 base pairs of the candidate rearrangement site. BLAST may be used align the sequences. A BLAST database may be constructed by collecting reads that map to a candidate fusion sequence, including discordant reads and soft-clipped reads, as well as unmapped reads in the original input file. Reads that map to the reference rearrangement sequence with a user-defined identity (e.g., at least 95%)

and/or a length of the aligned sequences is a user-defined percentage (e.g., 90%) of the input read length. The reads that span or flank the breakpoint may be counted. The user-defined identity may be at least about 70%, 75%, 80%, 85%, 90%, 95%, 97% or more. The length of the aligned sequences may be at least about 70%, 75%, 80%, 85%, 90%, or 95% or more of the input read length (e.g., read length of the candidate rearrangement sequence). The output redundancies may be minimized by removing fusion sequences within an interval of at least 20 base pairs or more of a fusion sequence with greater read support and with the same sequence orientation (to avoid removing reciprocal fusions).

[0654] The method may further comprise producing an output pertaining to the rearrangement. The output may comprise one or more of the following gene pair, genomic coordinates of the rearrangement, the orientation of the rearrangement (e.g., forward-forward or forward-reverse), genomic sequences within 50 bp of the rearrangement, and depth statistics for reads spanning and flanking the rearrangement.

[0655] The method may further comprise enumerating a fusion allele frequency. For example, fusion allele frequency in sequenced cfDNA may be enumerated as described herein and in Example 1. The fusion allele frequency may be calculated as α/β, where α is the number of breakpoint-spanning reads, and β is the mean overall depth within a genomic region at a predefined distance around the breakpoint. Thus, the fusion allele frequency may be calculated by dividing the number of rearrangement-spanning reads by the mean overall depth within a genomic region at a predefined distance around the breakpoint.

[0656] The method of identifying rearrangements may be applied to whole genome sequencing data or other suitable next-generation sequencing datasets. The genomic regions comprising the rearrangements identified from this data may be used to design a selector set.

[0657] The method of identifying rearrangements may be applied to sequencing data from a subject. The method may identify subject-specific breakpoints in tumor genomic DNA captured by a selector set. The method may be used to determine whether the subject-specific breakpoints are present in corresponding plasma DNA sample from the subject.

[0658] Identification of Tumor-Derived SNVs

[0659] Further disclosed herein are non-invasive methods of identifying tumor-derived SNVs.

[0660] The tumor-derived SNVs may be identified without prior knowledge of somatic variants identified in a corresponding tumor biopsy sample. In some embodiments of the invention, cfDNA is analyzed without comparison to a known tumor DNA sample from the patient. In such embodiments, the presence of ctDNA utilizes iterative models for (i) background noise in paired germline DNA, (ii) base-pair resolution background frequencies in cfDNA across the selector set, and (iii) sequencing error in cfDNA. These methods may utilize the following steps, which can be iterated through data point to automatically call tumor-derived SNVs:

[0661] taking allele frequencies from a single cfDNA sample and selecting high quality data;

[0662] testing whether a given input cfDNA allele is significantly different from the corresponding paired germline allele;

[0663] assembling a database of cfDNA background allele frequencies;

[0664] testing whether a given input allele differs significantly from cfDNA background at the same position, and selecting those with an average background frequency of a predetermined threshold, e.g. 5% or greater; 2.5% or greater, etc.

[0665] distinguishing tumor-derived SNVs from remaining background noise by outlier analysis.

[0666] The non-invasive method of identifying tumor-derived SNVs may comprise (a) obtaining a sample from a subject suffering from a cancer or suspected of suffering from a cancer; (b) conducting a sequencing reaction on the sample to produce sequencing information; (c) applying an algorithm to the sequencing information to produce a list of candidate tumor alleles based on the sequencing information from step (b), wherein a candidate tumor allele comprises a non-dominant base that is not a germline SNP; and (d) identifying tumor-derived SNVs based on the list of candidate tumor alleles. The candidate tumor allele may refer to a genomic region comprising a candidate SNV.

[0667] The candidate tumor allele may be a high quality candidate tumor allele. A high quality background allele may refer to the non-dominant base with the highest fractional abundance, excluding germline SNPs. The fractional abundance of a candidate tumor allele may be calculated by dividing a number of supporting reads by a total sequencing depth at that genomic position. For example, for a candidate mutation in a first genomic region, twenty sequence reads may contain a first sequence with the candidate mutation and 100 sequence reads may contain a second sequence without the candidate mutation. The candidate tumor allele may be the first sequence containing the candidate mutation. Based on this example, the fractional abundance of the candidate tumor allele would be 20 divided by 120, which is ~17%. Producing the list of candidate tumor alleles may comprise ranking the tumor alleles based on their fractional abundance. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with the highest fractional abundance. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a fractional abundance in the top 70th, 75th, 80th, 85th, 87th, 90th, 92nd, 95th, or 97th percentile. A candidate tumor allele may have a fractional abundance of less than 35%, 30%, 27%, 25%, 20%, 18%, 15%, 13%, 10%, 9%, 8%, 7%, 6.5%, 6%, 5.5%, 5%, 4.5%, 4%, 3.5%, 3%, 2.5%, 2%, 1.75%, 1.50%, 1.25%, or 1% of the total alleles pertaining to the candidate tumor allele in the sample from the subject. A candidate tumor allele may have a fractional abundance of less than 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of the total alleles pertaining to the candidate tumor allele in the sample from the subject. The candidate tumor allele may have a fractional abundance of less than 0.5% of the total alleles in the sample from the subject. The sample may comprise paired samples from the subject. Thus, the fractional abundance may be based on paired samples from the subject. The paired samples may comprise a sample containing suspected tumor-derived nucleic acids and a sample containing non-tumor-derived nucleic acids. For example, the paired samples may comprise a plasma sample and a sample containing peripheral blood lymphocytes (PBLs) or peripheral blood mononuclear cells (PBMCs).

[0668] The candidate tumor allele may have a minimum sequencing depth. Producing the list of candidate tumor alleles may comprise ranking the tumor alleles based on their sequencing depth. Producing the list of candidate tumor alleles may comprise selecting tumor alleles that meet a minimum sequencing depth. The minimum sequencing depth may be at least 100×, 200×, 300×, 400×, 500×, 600×, 700×, 800×,

900×, 1000× or more. The minimum sequencing depth may be at least about 500×. The minimum sequencing depth may be user-defined.

[0669] The candidate tumor allele may have a strand bias percentage. Producing the list of candidate tumor alleles may comprise calculating the strand bias percentage of a tumor allele. Producing the list of candidate tumor alleles may comprise ranking the tumor alleles based on their strand bias percentage. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a strand bias percentage of less than or equal to 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97%. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a strand bias percentage of less than or equal to 90%. The strand bias percentage may be user-defined.

[0670] Producing the list of candidate tumor alleles may comprise comparing the sequence of the tumor allele to a reference tumor allele. The reference tumor allele may be a germline allele. Producing the list of candidate tumor alleles may comprise determining whether the candidate tumor allele is different from a reference tumor allele. Producing the list of candidate tumor alleles may comprise selecting tumor alleles that are different from the reference tumor allele.

[0671] Determining whether the tumor allele is different from the reference tumor allele may comprise use of one or more statistical analyses. The statistical analysis may comprise using Bonferroni correction to calculate a Bonferroni-adjusted binomial probability for a tumor allele. The Bonferroni-adjusted binomial probability may be calculated by dividing a desired p-value cutoff (alpha) by the number of hypotheses tested. The number of hypotheses tested may be calculated by multiplying the number of bases in a selector by the number of possible base changes. The Bonferroni-adjusted binomial probability may be calculated by dividing the desired p-value cutoff (alpha) by the number of bases in a selector multiplied by the number of possible base changes. The Bonferroni-adjusted binomial probability may be used to determine whether the tumor allele occurred by chance. Producing the list of candidate tumor alleles may comprise selecting tumor alleles based on the Bonferroni-adjusted binomial probability. A candidate tumor allele may have a Bonferroni-adjusted binomial probability of less than or equal to 3×10^{-8}, 2.9×10^{-8}, 2.8×10^{-8}, 2.7×10^{-8}, 2.6×10^{-8}, 2.5×10^{-8}, 2.3×10^{-8}, 2.2×10^{-8}, 2.1×10^{-8}, 2.09×10^{-8}, 2.08×10^{-8}, 2.07×10^{-8}, 2.06×10^{-8}, 2.05×10^{-8}, 2.04×10^{-8}, 2.03×10^{-8}, 2.02×10^{-8}, 2.01×10^{-8} or 2×10^{-8}. A candidate tumor allele may have a Bonferroni-adjusted binomial probability of less than or equal to 2.08×10^{-8}.

[0672] Determining whether the tumor allele is different from the reference tumor allele may comprise use of a binomial distribution. The binomial distribution may be used to assemble a database of candidate tumor allele frequencies. An algorithm, such as a Z-test, may be used to determine whether a candidate tumor allele differs significantly from a typical circulating allele at the same position. A significant difference may refer to a difference that is unlikely to have occurred by chance. The Z-test may be applied to the Bonferroni-adjusted bionomial probability of the tumor alleles to produce a Bonferroni-adjusted single-tailed Z-score. The Bonferroni-adjusted single-tailed Z-score may be determined by using a normal distribution. A tumor allele with a Bonferroni-adjusted single-tailed Z-score of greater than or equal to 6, 5.9, 5.8, 5.7, 5.6, 5.5, 5.4, 5.3, 5.2, 5.1, or 5.0 is considered to be different from the reference tumor allele. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a Bonferroni-adjusted single-tailed Z-score of greater than or equal to 6, 5.9, 5.8, 5.7, 5.6, 5.5, 5.4, 5.3, 5.2, 5.1, or 5.0. Producing the list of candidate tumor alleles may comprise selecting tumor alleles with a Bonferroni-adjusted single-tailed Z-score of greater than 5.6.

[0673] Candidate tumor alleles may be based on genomic regions from a selector set. The list of candidate tumor alleles may comprise candidate tumor alleles with a frequency of less than or equal to 10%, 9%, 8%, 7%, 6.5%, 6%, 5.5%, 5%, 4.5%, 4%, 3.5%, or 3%. The list of candidate tumor alleles may comprise candidate tumor alleles with a frequency of less than 5%.

[0674] Identifying tumor-derived SNVs based on the list of candidate tumor alleles may comprise testing the candidate tumor alleles from the list of candidate tumor alleles for sequencing errors. Testing the candidate tumor alleles for sequencing errors may be based on the duplication rate of the candidate tumor allele. The duplication rate may be determined by comparing the number of supporting reads for a candidate tumor allele for nondeduped data (e.g., all fragments meeting quality control criteria) and deduped data (e.g., unique fragments meeting quality control criteria). The candidate tumor alleles may be ranked based on their duplication rate. A tumor-derived SNV may be in a candidate tumor allele with a low duplication rate.

[0675] Identifying tumor-derived SNVs may further comprise use of an outlier analysis. The outlier analysis may be used to distinguish candidate tumor-derived SNVs from the remaining background noise. The outlier analysis may comprise comparing the square root of the robust distance Rd (Mahalanobis distance) to the square root of the quantiles of a chi-squared distribution Cs. Tumor-derived SNVs may be identified from the outliers in the outlier analysis.

[0676] The sequencing information may pertain to regions flanking one or more genomic regions from a selector set. The sequencing information may pertain to regions flanking genomic coordinates from a selector set. The sequencing information may pertain to regions within 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more base pairs of a genomic region from a selector set. The sequencing information may pertain to regions within 500 base pairs of a genomic region from a selector set. The sequencing information may pertain to regions within 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more base pairs of a genomic coordinate from a selector set. The sequencing information may pertain to regions within 500 base pairs of a genomic coordinate from a selector set.

[0677] Computer Program

[0678] The methods described herein may be performed by a computer program product that comprises a computer executable logic that is recorded on a computer readable medium. For example, the computer program can execute some or all of the following functions: (i) controlling isolation of nucleic acids from a sample, (ii) pre-amplifying nucleic acids from the sample or (iii) selecting, amplifying, sequencing or arraying specific regions in the sample, (iv) identifying and quantifying somatic mutations in a sample, (v) comparing data on somatic mutations detected from the sample with a predetermined threshold, (vi) determining the tumor load based on the presence of somatic mutations in the cfDNA, and (vii) declaring an assessment of tumor load, residual disease, response to therapy, or initial diagnosis. The computer program may calculate a recurrence index. The computer pro-

gram may rank genomic regions by the recurrence index. The computer program may select one or more genomic regions based on the recurrence index. The computer program may produce a selector set. The computer program may add genomic regions to the selector set. The computer program may maximize subject coverage of the selector set. The computer program may maximize a median number of mutations per subject in a population. The computer program may calculate a ctDNA detection index. The computer program may calculate a p-value of one or more types of mutations. The computer program may identify genomic regions comprising one or more mutations present in one or more subjects suffering from a cancer. The computer program may identify novel mutations present in one or more subjects suffering from a cancer. The computer program may identify novel fusions present in one or more subjects suffering from a cancer.

[0679] The computer executable logic can work in any computer that may be any of a variety of types of general-purpose computers such as a personal computer, network server, workstation, or other computer platform now or later developed. In some embodiments, a computer program product is described comprising a computer usable medium having the computer executable logic (computer software program, including program code) stored therein. The computer executable logic can be executed by a processor, causing the processor to perform functions described herein. In other embodiments, some functions are implemented primarily in hardware using, for example, a hardware state machine. Implementation of the hardware state machine so as to perform the functions described herein will be apparent to those skilled in the relevant arts.

[0680] The program can provide a method of evaluating the presence of tumor cells in an individual by accessing data that reflects the sequence of the selected cfDNA from the individual, and/or the quantitation of one or more nucleic acids from the cfDNA in the circulation of the individual. The one or more nucleic acids from the cfDNA in the circulation to be quantified may be based on genomic regions or genomic coordinates provided by a selector set.

[0681] In one embodiment, the computer executing the computer logic of the invention may also include a digital input device such as a scanner. The digital input device can provide information on a nucleic acid, e.g., polymorphism levels/quantity.

[0682] In some embodiments, the invention provides a computer readable medium comprising a set of instructions recorded thereon to cause a computer to perform the steps of (i) receiving data from one or more nucleic acids detected in a sample; and (ii) diagnosing or predicting tumor load, residual disease, response to therapy, or initial diagnosis based on the quantitation.

[0683] Sequencing

[0684] Genotyping ctDNA and/or detection, identification and/or quantitation of the ctDNA can utilize sequencing. Sequencing can be accomplished using high-throughput systems. In some cases, high throughput sequencing generates at least 1,000, at least 5,000, at least 10,000, at least 20,000, at least 30,000, at least 40,000, at least 50,000, at least 100,000 or at least 500,000 sequence reads per hour; with each read being at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120 or at least 150 bases per read. Sequencing can be performed using nucleic acids described herein such as genomic DNA, cDNA derived from RNA transcripts or RNA as a template. Sequencing may comprise massively parallel sequencing.

[0685] In some embodiments, high-throughput sequencing involves the use of technology available by Helicos BioSciences Corporation (Cambridge, Mass.) such as the Single Molecule Sequencing by Synthesis (SMSS) method. In some embodiments, high-throughput sequencing involves the use of technology available by 454 Lifesciences, Inc. (Branford, Conn.) such as the Pico Titer Plate device which includes a fiber optic plate that transmits chemiluminescent signal generated by the sequencing reaction to be recorded by a CCD camera in the instrument. This use of fiber optics allows for the detection of a minimum of 20 million base pairs in 4.5 hours.

[0686] In some embodiments, high-throughput sequencing is performed using Clonal Single Molecule Array (Solexa, Inc.) or sequencing-by-synthesis (SBS) utilizing reversible terminator chemistry. These technologies are described in part in U.S. Pat. Nos. 6,969,488; 6,897,023; 6,833,246; 6,787,308; and US Publication Application Nos. 20040106130; 20030064398; 20030022207; and Constans, A, The Scientist 2003, 17(13):36.

[0687] In some embodiments, high-throughput sequencing of RNA or DNA can take place using AnyDot.chips (Genovoxx, Germany), which allows for the monitoring of biological processes (e.g., miRNA expression or allele variability (SNP detection). In particular, the AnyDot-chips allow for 10×-50× enhancement of nucleotide fluorescence signal detection. Other high-throughput sequencing systems include those disclosed in Venter, J., et al. Science 16 Feb. 2001; Adams, M. et al, Science 24 Mar. 2000; and M. J, Levene, et al. Science 299:682-686, January 2003; as well as US Publication Application No. 20030044781 and 2006/0078937. The growing of the nucleic acid strand and identifying the added nucleotide analog may be repeated so that the nucleic acid strand is further extended and the sequence of the target nucleic acid is determined.

[0688] The methods disclosed herein may comprise conducting a sequencing reaction based on one or more genomic regions from a selector set. The selector set may comprise one or more genomic regions from Table 2. A sequencing reaction may be performed on 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set based on Table 2. A sequencing reaction may be performed on 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions from a selector set based on Table 2.

[0689] A sequencing reaction may be performed on a subset of genomic regions from a selector set. A sequencing reaction may be performed on 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300 or more genomic regions from a selector set. A sequencing reaction may be performed on 325, 350, 375, 400, 425, 450, 475, 500 or more genomic regions from a selector set.

[0690] A sequencing reaction may be performed on all of the genomic regions from a selector set. Alternatively, a sequencing reaction may be performed on 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more of the genomic regions from a selector set. A sequencing reaction may be performed on at least 10% of the genomic regions from a selector set. A sequencing reaction may be performed on at least 30% of the genomic regions from a selector set. A

sequencing reaction may be performed on at least 50% of the genomic regions from a selector set.

[0691] A sequencing reaction may be performed on less than 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% of the genomic regions from a selector set. A sequencing reaction may be performed on less than 10% of the genomic regions from a selector set. A sequencing reaction may be performed on less than 30% of the genomic regions from a selector set. A sequencing reaction may be performed on less than 50% of the genomic regions from a selector set.

[0692] The methods disclosed herein may comprise obtaining sequencing information for one or more genomic regions from a selector set. Sequencing information may be obtained for 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set based on Table 2. Sequencing information may be obtained for 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions from a selector set based on Table 2.

[0693] Sequencing information may be obtained for a subset of genomic regions from a selector set. Sequencing information may be obtained for 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300 or more genomic regions from a selector set. Sequencing information may be obtained for 325, 350, 375, 400, 425, 450, 475, 500 or more genomic regions from a selector set.

[0694] Sequencing information may be obtained for all of the genomic regions from a selector set. Alternatively, sequencing information may be obtained for 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions from a selector set. Sequencing information may be obtained for at least 10% of the genomic regions from a selector set. Sequencing information may be obtained for at least 30% of the genomic regions from a selector set.

[0695] Sequencing information may be obtained for less than 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the genomic regions from a selector set. Sequencing information may be obtained for less than 10% of the genomic regions from a selector set. Sequencing information may be obtained for less than 30% of the genomic regions from a selector set. Sequencing information may be obtained for less than 50% of the genomic regions from a selector set. Sequencing information may be obtained for less than 70% of the genomic regions from a selector set.

[0696] Amplification

[0697] The methods disclosed herein may comprise amplification of cell-free DNA (cfDNA) and/or of circulating tumor DNA (ctDNA). Amplification may comprise PCR-based amplification. Alternatively, amplification may comprise nonPCR-based amplification.

[0698] Amplification of cfDNA and/or ctDNA may comprise using bead amplification followed by fiber optics detection as described in Marguiles et al. "Genome sequencing in microfabricated high-density pricolitre reactors", Nature, doi: 10.1038/nature03959; and well as in US Publication Application Nos. 200200 12930; 20030058629; 20030 1001 02; 20030 148344; 20040248 161; 200500795 10,20050 124022; and 20060078909.

[0699] Amplification of the nucleic acid may comprise use of one or more polymerases. The polymerase may be a DNA polymerase. The polymerase may be a RNA polymerase. The polymerase may be a high fidelity polymerase. The polymerase may be KAPA HiFi DNA polymerase. The polymerase may be Phusion DNA polymerase.

[0700] Amplification may comprise 20 or fewer amplification cycles. Amplification may comprise 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, or 9 or fewer amplification cycles. Amplification may comprise 18 or fewer amplification cycles. Amplification may comprise 16 or fewer amplification cycles. Amplification may comprise 15 or fewer amplification cycles.

[0701] Sample

[0702] The methods, kits, and systems disclosed herein may comprise one or more samples or uses thereof. A "sample" may refer to any biological sample that is isolated from a subject. A sample can include, without limitation, an aliquot of body fluid, whole blood, platelets, serum, plasma, red blood cells, white blood cells or leucocytes, endothelial cells, tissue biopsies, synovial fluid, lymphatic fluid, ascites fluid, and interstitial or extracellular fluid. The term "sample" may also encompass the fluid in spaces between cells, including gingival crevicular fluid, bone marrow, cerebrospinal fluid (CSF), saliva, mucous, sputum, semen, sweat, urine, or any other bodily fluids. "Blood sample" can refer to whole blood or any fraction thereof, including blood cells, red blood cells, white blood cells or leucocytes, platelets, serum and plasma. The sample may be from a bodily fluid. The sample may be a plasma sample. The sample may be a serum sample. The sample may be a tumor sample. Samples can be obtained from a subject by means including but not limited to venipuncture, excretion, ejaculation, massage, biopsy, needle aspirate, lavage, scraping, surgical incision, or intervention or other means known in the art.

[0703] Samples useful for the methods of the invention may comprise cell-free DNA (cfDNA), e.g., DNA in a sample that is not contained within a cell. Typically such DNA may be fragmented, and may be on average about 170 nucleotides in length, which may coincide with the length of DNA around a single nucleosome. cfDNA may generally be a heterogeneous mixture of DNA from normal and tumor cells, and an initial sample of cfDNA may generally not be enriched for recurrently mutated regions of a cancer cell genome. The terms ctDNA, cell-free tumor DNA or "circulating tumor" DNA may be used to refer to the fraction of cfDNA in a sample that is derived from a tumor. One of skill in the art will understand that germline sequences may not be distinguished between a tumor source and a normal cell source, but sequences containing somatic mutations have a high probability of being derived from tumor DNA. A sample may be a control germline DNA sample. A sample may be a known tumor DNA sample. A sample may be cfDNA obtained from an individual suspected of having ctDNA in the sample.

[0704] The methods disclosed herein may comprise obtaining one or more samples from a subject. The one or more samples may be a tumor nucleic acid sample. Alternatively, or additionally, the one or more samples may be a genomic nucleic acid sample. It should be understood that the step of obtaining a tumor nucleic acid sample and a genomic nucleic acid sample from a subject with a specific cancer may occur in a single step. Alternatively, the step of obtaining a tumor nucleic acid sample and a genomic nucleic acid sample from a subject with a specific cancer may occur in separate steps. For example, it may be possible to obtain a single tissue sample from a patient, for example from a biopsy sample,

which includes both tumor nucleic acids and genomic nucleic acids. It is also within the scope of this step to obtain the tumor nucleic acid sample and the genomic nucleic acid sample from the subject in separate samples, in separate tissues, or even at separate times.

[0705] The sample may comprise nucleic acids. The nucleic acids may be cell-free nucleic acids. The nucleic acids may be circulating nucleic acids. The nucleic acids may be from a tumor. The nucleic acids may be circulating tumor DNA (ctDNA). The nucleic acids may be cell-free DNA (cfDNA). The nucleic acids may be genomic nucleic acids. The nucleic acids may be tumor nucleic acids.

[0706] The step of obtaining a tumor nucleic acid sample and a genomic nucleic acid sample from a subject with a specific cancer may also include the process of extracting a biological fluid or tissue sample from the subject with the specific cancer. These particular steps are well understood by those of ordinary skill in the medical arts, particularly by those working in the medical laboratory arts.

[0707] The step of obtaining a tumor nucleic acid sample and a genomic nucleic acid sample from a subject with a specific cancer may additionally include procedures to improve the yield or recovery of the nucleic acids in the sample. For example, the step may include laboratory procedures to separate the nucleic acids from other cellular components and contaminants that may be present in the biological fluid or tissue sample. As noted, such steps may improve the yield and/or may facilitate the sequencing reactions.

[0708] It should also be understood that the step of obtaining a tumor nucleic acid sample and a genomic nucleic acid sample from a subject with a specific cancer may be performed by a commercial laboratory that does not even have direct contact with the subject. For example, the commercial laboratory may obtain the nucleic acid samples from a hospital or other clinical facility where, for example, a biopsy or other procedure is performed to obtain tissue from a subject. The commercial laboratory may thus carry out all the steps of the instantly-disclosed methods at the request of, or under the instructions of, the facility where the subject is being treated or diagnosed.

[0709] A sample may be selected for DNA corresponding to regions of recurrent mutations, utilizing a selector set as described herein. In some embodiments, the selection process comprises the following method. DNA obtained from cellular sources may be fragmented to approximate the size of cfDNA, e.g. of from about 50 to about 1 KB in length. The DNA may then be denatured, and hybridized to a population of selector set probes comprising a specific binding member, e.g. biotin, etc. The composition of hybridized DNA may then be applied to a complementary binding member, e.g. avidin, streptavidin, an antibody specific for a tag, etc., and the unbound DNA washed free. The selected DNA population may then be washed free of the unbound DNA.

[0710] The captured DNA may then be sequenced by any suitable protocol. In some embodiments, the captured DNA is amplified prior to sequencing, where the amplification primers may utilize primers or oligonucleotides suitable for high throughput sequencing. The resulting product may be a set of DNA sequences enriched for sequences corresponding to regions of the genome that have recurrent mutations in the cancer of interest. The remaining analysis may utilize bioinformatics methods, which can vary with the type of somatic mutation, e.g. SNV, SNV, fusion, etc.

[0711] Further disclosed herein are methods of preparing a next-generation sequencing (NGS) library. The method may comprise (a) attaching adaptors to a plurality of nucleic acids to produce a plurality of adaptor-modified nucleic acids; and (b) amplifying the plurality of adaptor-modified nucleic acids, thereby producing a NGS library, wherein amplifying comprises 1 to 20 amplification cycles.

[0712] The methods disclosed herein may comprise attaching adaptors to nucleic acids. Attaching adaptors to nucleic acids may comprise ligating adaptors to nucleic acids. Attaching adaptors to nucleic acids may comprise hybridizing adaptors to nucleic acids. Attaching adaptors to nucleic acids may comprise primer extension.

[0713] The plurality of nucleic acids may be from a sample. Attaching the adaptors to the plurality of nucleic acids may comprise contacting the sample with the adaptors.

[0714] Attaching the adaptors to the nucleic acids may comprise incubating the adaptors and nucleic acids at a specific temperature or temperature range. Attaching the adaptors to the nucleic acids may comprise incubating the adaptors and nucleic acids at 20° C. Attaching the adaptors to the nucleic acids may comprise incubating the adaptors and nucleic acids at less 20° C. Attaching the adaptors to the nucleic acids may comprise incubating the adaptors and nucleic acids at 19° C., 18° C., 17° C., 16° C. or less. Alternatively, attaching the adaptors to the nucleic acids may comprise incubating the adaptors and nucleic acids at varying temperatures. For example, attaching the adaptors to the nucleic acids may comprise temperature cycling. Attaching the adaptors to the nucleic acids may comprise may comprise incubating the nucleic acids and adaptors at a first temperature for a first period of time, followed by incubation at one or more additional temperatures for one or more additional periods of time. The one or more additional temperatures may be greater than the first temperature or preceding temperature. Alternatively, or additionally, the one or more additional temperatures may be less than the first temperature or preceding temperature. For example, the nucleic acids and adaptors may be incubated at 10° C. for 30 second, followed by incubation at 30° C. for 30 seconds. The temperature cycling of 10° C. for 30 seconds and 30° C. for 30 second may be repeated multiple times. For example, attaching the adaptors to the nucleic acids by temperature cycling may comprise alternating the temperature from 10° C. to 30° C. in 30 second increments for a total time period of 12 to 16 hours.

[0715] The adaptors and nucleic acids may be incubated at a specified temperature or temperature range for a period of time. The adaptors and nucleic acid may be incubated at a specific temperature or temperature range for at least about 15 minutes. The adaptors and nucleic acid may be incubated at a specific temperature or temperature range for at least about 30 minutes, 60 minutes, 90 minutes, 120 minutes or more. The adaptors and nucleic acid may be incubated at a specific temperature or temperature range for at least about 1 hour, 2 hours, 3 hours, 4 hours, 5 hours, 6 hours, 7 hours, 8 hours, 9 hours, 10 hours, 12 hours, 14 hours, 16 hours, or more. The adaptors and nucleic acid may be incubated at a specific temperature or temperature range for at least about 16 hours.

[0716] The adaptors may be attached to the nucleic acid by incubating the nucleic acids and the adaptors at a temperature less than or equal to 20° C. for at least about 20, 30, 40, 50, 60, 70, 80, 90, 100 or more minutes. The adaptors may be attached to the nucleic acid by incubating the nucleic acids and the adaptors at a temperature less than or equal to 20, 19,

18, 17, 16° C. for at least about 1 hour. The adaptors may be attached to the nucleic acid by incubating the nucleic acids and the adaptors at a temperature less than or equal to 18° C. for at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more hours. The adaptors may be attached to the nucleic acid by incubating the nucleic acids and the adaptors at a temperature less than or equal to 20, 19, 18, 17, 16° C. for at least about 5 hours. The adaptors may be attached to the nucleic acid by incubating the nucleic acids and the adaptors at a temperature less than or equal to 16° C. for at least about 5 hours.

[0717] Attaching the adaptors to the nucleic acids may comprise use of one or more enzymes. The enzyme may be a ligase. The ligase may be a DNA ligase. The DNA ligase may be a T4 DNA ligase, *E. coli* DNA ligase, mammalian ligase, or a combination thereof. The mammalian ligase may be DNA ligase I, DNA ligase III, or DNA ligase IV. The ligase may be a thermostable ligase.

[0718] The adaptor may comprise a universal primer binding sequence. The adaptor may comprise a primer sequence. The primer sequence may enable sequencing of the adaptor-modified nucleic acids. The primer sequence may enable amplification of the adaptor-modified nucleic acids. The adaptor may comprise a barcode. The barcode may enable differentiation of two or more molecules of the same molecular species. The barcode may enable quantification of one or more molecules.

[0719] The method may further comprise contacting the plurality of nucleic acids with a plurality of beads to produce a plurality of bead-conjugated nucleic acids. The plurality of nucleic acids may be contacted with the plurality of beads after attaching the adaptors to the nucleic acids. Alternatively, or additionally, the plurality of nucleic acids may be contacted with the plurality of beads before amplification of the adaptor-modified nucleic acids. Alternatively, or additionally, the plurality of nucleic acids may be contacted with the plurality of beads after amplification of the adaptor-modified nucleic acids.

[0720] The beads may be magnetic beads. The beads may be coated beads. The beads may be antibody-coated beads. The beads may be protein-coated beads. The beads may be coated with one or more functional groups. The beads may be coated with one or more oligonucleotides.

[0721] Amplifying the plurality of adaptor-modified nucleic acids may comprise any method known in the art. For example, amplifying may comprise PCR-based amplification. Alternatively, amplifying may comprise nonPCR-based amplification. Amplifying may comprise any of the amplification methods disclosed herein.

[0722] Amplifying the plurality of adaptor-modified nucleic acids may comprise amplifying a product or derivative of the adaptor-modified nucleic acids. A product or derivative of the adaptor-ligated nucleic acids may comprise bead-conjugated nucleic acids, enriched-nucleic acids, fragmented nucleic acids, end-repaired nucleic acids, A-tailed nucleic acids, barcoded nucleic acids, or a combination thereof

[0723] Amplifying the adaptor-modified nucleic acids may comprise 1 to 20 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 1 to 18 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 1 to 17 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 1 to 16 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 2 to 20 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 2 to 18 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 2 to 16 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 3 to 20 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 3 to 19 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 3 to 17 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 4 to 20 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 4 to 18 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 4 to 16 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 5 to 20 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 5 to 19 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 5 to 18 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 5 to 17 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 5 to 16 amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 5 to 15 amplification cycles.

[0724] Amplifying the adaptor-modified nucleic acids may comprise 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, or 2 or fewer amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 20 or fewer amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 18 or fewer amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 16 or fewer amplification cycles. Amplifying the adaptor-modified nucleic acids may comprise 15 or fewer amplification cycles.

[0725] The method may further comprise fragmenting the plurality of nucleic acids to produce a plurality of fragmented nucleic acids. The plurality of nucleic acids may be fragmented prior to attaching the adaptors to the plurality of nucleic acids. The plurality of nucleic acids may be fragmented after attachment of the adaptors to the plurality of nucleic acids. The plurality of nucleic acids may be fragmented prior to amplification of the adaptor-modified nucleic acids. The plurality of nucleic acids may be fragmented after amplification of the adaptor-modified nucleic acids. Fragmenting the plurality of nucleic acids may comprise use of one or more restriction enzymes. Fragmenting the plurality of nucleic acids may comprise use of a sonicator. Fragmenting the plurality of nucleic acids may comprise shearing the nucleic acids.

[0726] The method may further comprise conducting an end repair reaction on the plurality of nucleic acids to produce a plurality of end repaired nucleic acids. The end repair reaction may be conducted prior to attaching the adaptors to the plurality of nucleic acids. The end repair reaction may be conducted after attaching the adaptors to the plurality of nucleic acids. The end repair reaction may be conducted prior to amplification of the adaptor-modified nucleic acids. The end repair reaction may be conducted after amplification of the adaptor-modified nucleic acids. The end repair reaction may be conducted prior to fragmenting the plurality of nucleic acids. The end repair reaction may be conducted after fragmenting the plurality of nucleic acids. Conducting the end repair reaction may comprise use of one or more end repair enzymes.

[0727] The method may further comprise conducting an A-tailing reaction on the plurality of nucleic acids to produce

a plurality of A-tailed nucleic acids. The A-tailing reaction may be conducted prior to attaching the adaptors to the plurality of nucleic acids. The A-tailing reaction may be conducted after attaching the adaptors to the plurality of nucleic acids. The A-tailing reaction may be conducted prior to amplification of the adaptor-modified nucleic acids. The A-tailing reaction may be conducted after amplification of the adaptor-modified nucleic acids. The A-tailing reaction may be conducted prior to fragmenting the plurality of nucleic acids. The A-tailing reaction may be conducted after fragmenting the plurality of nucleic acids. The A-tailing reaction may be conducted prior to end repair of the plurality of nucleic acids. The A-tailing reaction may be conducted after end repair of the plurality of nucleic acids. Conducting the A-tailing reaction may comprise use of one or more A-tailing enzymes.

[0728] The method may further comprise contacting the plurality of nucleic acids with a plurality of molecular barcodes to produce a plurality of barcoded nucleic acids. Producing the plurality of barcoded nucleic acids may occur prior to attaching the adaptors to the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur after attaching the adaptors to the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur prior to amplification of the adaptor-modified nucleic acids. Producing the plurality of barcoded nucleic acids may occur after amplification of the adaptor-modified nucleic acids. Producing the plurality of barcoded nucleic acids may occur prior to fragmenting the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur after fragmenting the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur prior to end repair of the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur after end repair the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur prior to A-tailing of the plurality of nucleic acids. Producing the plurality of barcoded nucleic acids may occur after A-tailing of the plurality of nucleic acids. The barcode may enable differentiation of two or more molecules of the same molecular species. The barcode may enable quantification of one or more molecules. The barcode may be a molecular barcode. The molecular barcode may be used to differentiate two or more molecules of the same molecular species. The molecular barcode may be used to differentiate two or more molecules of the same genomic region. The barcode may be a sample index. The sample index may be used to identify a sample from which the molecule (e.g., nucleic acid) originated from. For example, molecules from a first sample may be associated with a first sample index, whereas molecules from a second sample may be associated with a second sample index. The sample index from two or more samples may be different. The two or more samples may be from the same subject. The two or more samples may be from two or more subjects. The two or more samples may be obtained at the same time. Alternatively, or additionally, the two or more samples may be obtained at two or more time points.

[0729] The method may further comprise contacting the plurality of nucleic acids with a plurality of sequencing adaptors to produce a plurality of sequencer-adapted nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur prior to attaching the adaptors to the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur after attaching the adaptors to the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur prior to amplification of the adaptor-modified nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur after amplification of the adaptor-modified nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur prior to fragmenting the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur after fragmenting the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur prior to end repair of the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur after end repair the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur prior to A-tailing of the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur after A-tailing of the plurality of nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur prior to producing the barcoded nucleic acids. Producing the plurality of sequencer-adapted nucleic acids may occur after producing the barcoded nucleic acids. The sequencing adaptor may enable sequencing of the nucleic acids.

[0730] The method may further comprise contacting the plurality of nucleic acids with a plurality of primer adaptors to produce a plurality of primer-adapted nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to attaching the adaptors to the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after attaching the adaptors to the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to amplification of the adaptor-modified nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after amplification of the adaptor-modified nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to fragmenting the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after fragmenting the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to end repair of the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after end repair the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to A-tailing of the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after A-tailing of the plurality of nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to producing the barcoded nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after producing the barcoded nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur prior to producing the sequencer-adapted nucleic acids. Producing the plurality of primer-adapted nucleic acids may occur after producing the sequencer-adapted nucleic acids. Producing the plurality of primer-adapted nucleic acids may comprise ligating the primer adaptors to the nucleic acids. The primer adaptor may enable sequencing of the nucleic acids. The primer adaptor may enable amplification of the nucleic acids.

[0731] The method may further comprise conducting a hybridization reaction. The hybridization reaction may comprise use of a solid support. The hybridization reaction may comprise hybridizing the plurality of nucleic acids to the solid support. The hybridization reaction may comprise use of a

plurality of beads. The hybridization reaction may comprise hybridizing the plurality of nucleic acids to the plurality of beads. The method may further comprise conducting a hybridization reaction after an enzymatic reaction. The enzymatic reaction may comprise a ligation reaction. The enzymatic reaction may comprise a fragmentation reaction. The enzymatic reaction may comprise an end repair reaction. The enzymatic reaction may comprise an A-tailing reaction. The enzymatic reaction may comprise an amplification reaction. The method may further comprise conducting a hybridization reaction after one or more reactions selected from a group consisting of a ligation reaction, fragmentation reaction, end repair reaction, A-tailing reaction, and amplification reaction. The method may further comprise conducting a hybridization reaction after two or more reactions selected from a group consisting of a ligation reaction, fragmentation reaction, end repair reaction, A-tailing reaction, and amplification reaction. The method may further comprise conducting a hybridization reaction after three or more reactions selected from a group consisting of a ligation reaction, fragmentation reaction, end repair reaction, A-tailing reaction, and amplification reaction. The method may further comprise conducting a hybridization reaction after four or more reactions selected from a group consisting of a ligation reaction, fragmentation reaction, end repair reaction, A-tailing reaction, and amplification reaction. The hybridization reaction may be conducted after each reaction selected from a group consisting of ligation reaction, fragmentation reaction, end repair reaction, A-tailing reaction, and amplification reaction.

[0732] Nucleic Acid Detection Methods

[0733] Provided herein are methods for the ultrasensitive detection of a minority nucleic acid in a heterogeneous sample. The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from a subject; and (b) using sequence information derived from (a) to detect cell-free minority nucleic acids in the sample, wherein the method is capable of detecting a percentage of the cell-free minority nucleic acids that is less than 2% of total cfDNA. The minority nucleic acid may refer to a nucleic acid that originated from a cell or tissue that is different from a normal cell or tissue from the subject. For example, the subject may be infected with a pathogen such as a bacteria and the minority nucleic acid may be a nucleic acid from the pathogen. In another example, the subject is a recipient of a cell, tissue or organ from a donor and the minority nucleic acid may be a nucleic acid originating from the cell, tissue or organ from the donor. In another example, the subject is a pregnant subject and the minority nucleic acid may be a nucleic acid originating from a fetus. The method may comprise using the sequence information to detect one or more somatic mutations in the fetus. The method may comprise using the sequence information to detect one or more post-zygotic mutations in the fetus. Alternatively, the subject may be suffering from a cancer and the minority nucleic acid may be a nucleic acid originating from a cancer cell.

[0734] Provided herein are methods for the ultrasensitive detection of circulating tumor DNA in a sample. The method may be called CAncer Personalized Profiling by Deep Sequencing (CAPP-Seq). The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from a subject; and (b) using sequence information derived from (a) to detect cell-free tumor DNA (ctDNA) in the sample, wherein the method is capable of detecting a percentage of ctDNA that is less than 2% of total cfDNA. CAPP-Seq may accurately quantify cell-free tumor DNA from early and advanced stage tumors. CAPP-Seq may identify mutant alleles down to 0.025% with a detection limit of <0.01%. Tumor-derived DNA levels often paralleled clinical responses to diverse therapies and CAPP-Seq may identify actionable mutations. CAPP-Seq may be routinely applied to noninvasively detect and monitor tumors, thus facilitating personalized cancer therapy.

[0735] Disclosed herein are methods for determining a quantity of circulating tumor DNA (ctDNA) in a sample. The method may comprise (a) ligating one or more adaptors to cell-free DNA (cfDNA) derived from a sample from a subject to produce one or more adaptor-ligated cfDNA; (b) performing sequencing on the one or more adaptor-ligated cfDNA, wherein the adaptor-ligated cfDNA to be sequenced is based on a selector set comprising a plurality of genomic regions; and (c) using a computer readable medium to determine a quantity of cfDNA originating from a tumor based on the sequencing information obtained from the adaptor-ligated cfDNA. cfDNA originating from the tumor may be referred to as cell-free tumor DNA or circulating tumor DNA (ctDNA). The quantity of ctDNA may be a percentage. Determining the quantity of the ctDNA may comprise determining the sequence of one or more genomic regions from the selector set. Determining the quantity of the ctDNA may comprise determining a number of sequence reads that contain a sequence a mutation corresponding to one or more mutations in the one or more genomic regions based on the selector set. Determining the quantity of ctDNA may comprise determining a number of sequence reads that contain a sequence that does not contain a mutation corresponding to one or more mutations in the one or more genomic regions based on the selector set. Determining the quantity of ctDNA may comprise calculating a percentage of sequence reads that contain sequences with one or more mutations corresponding to one or more mutations in the one or more genomic regions based on the selector set. For example, a selector set may be used to obtain sequencing information for a first genomic region. The sequence information may comprise twenty sequencing reads pertaining to the first genomic region. Analysis of the sequencing information may determine that two of the sequencing reads contain a mutation corresponding to a first mutation in the first genomic region based on the selector set and eighteen of the sequencing reads do not contain a mutation corresponding to a mutation in the first genomic region based on the selector set. Thus, the quantity of the ctDNA may be equal to the percentage of sequencing reads with the mutation corresponding to a mutation in the first genomic region, which would be 10% (e.g., 2 reads divided by 20 reads times 100%). For sequence information pertaining to two or more genomic regions based on the selector set, determining the quantity of ctDNA may comprise calculating an average of the percentages the two or more genomic regions. For example, the percentage of sequencing reads containing a mutation corresponding to a first mutation in a first genomic region is 20% and the percentage of sequencing reads containing a mutation corresponding to a second mutation in a second genomic region is 40%; the quantity of ctDNA is the average of the percentages of the two genomic regions, which is 30% (e.g., (20%+40%) divided by 2). The quantity of ctDNA may be converted into a mass per unit volume value by multiplying the percentage of the ctDNA by the absolute concentration of the total cell-free DNA per unit volume. For example, the percentage of ctDNA may be 30% and the

concentration of the cell free DNA may be 10 nanograms per milliliter (ng/mL); the quantity of ctDNA may be 3 ng/mL (e.g., 0.30 times 10 ng/mL).

[0736] Alternatively, or additionally, determining the quantity of ctDNA may comprise use of adaptors comprising a barcode sequence. Two or more adaptors may contain two or more different barcode sequences. The barcode sequence may be a random sequence. A genomic region may be attached to an adaptor containing a barcode sequence. Identical genomic regions may be attached to adaptors containing different barcode sequences. Non-identical genomic regions may be attached to adaptors containing different barcode sequences. The barcode sequences may be used to count a number of occurrences of a genomic region. The quantity of the ctDNA may be based on counting a number of occurrences of genomic regions based on the selector set. Rather than basing the quantity of the ctDNA on the number of sequencing reads, the quantity of the ctDNA may be based on the number of different barcodes associated with one or more genomic regions. For example, ten different barcodes may be associated with sequences containing a mutation corresponding to a mutation in a first genomic region based on the selector set, resulting in a quantity of ctDNA of ten. For two or more genomic regions, the quantity of the ctDNA may be a sum of the quantity of the two or more genomic regions. For example, ten different barcodes may be associated with sequences containing a mutation corresponding to a mutation in a first genomic region and twenty different barcodes may be associated with sequences containing a mutation correspond to a mutation in a second genomic region, resulting in a quantity of ctDNA of 30. The quantity of the ctDNA may be a percentage of the total cell-free DNA. For example, ten different barcodes may be associated with sequences containing a mutation corresponding to a mutation in a first genomic region and forty different barcodes may be associated with sequences that do not contain a mutation corresponding to a mutation in the first genomic region, resulting in a quantity of ctDNA of 20% (e.g., (10 divided by 50) times 100%).

[0737] Disclosed herein are methods of enriching for circulating tumor DNA from a sample. The method may comprise contacting cell-free nucleic acids from a sample with a plurality of oligonucleotides, wherein the plurality of oligonucleotides selectively hybridize to a plurality of genomic regions comprising a plurality of mutations present in >60% of a population of subjects suffering from a cancer.

[0738] Alternatively, the method may comprise contacting cell-free nucleic acids from a sample with a set of oligonucleotides, wherein the set of oligonucleotides selectively hybridize to a plurality of genomic regions, wherein (a) >80% of tumors from a population of cancer subjects include one or more mutations in the genomic regions; (b) the plurality of genomic regions represent less than 1.5 Mb of the genome; and (c) the set of oligonucleotides comprise 5 or more different oligonucleotides that selectively hybridize to the plurality of genomic regions. The cell-free nucleic acids may be DNA. The cell-free nucleic acids may be RNA.

[0739] Applications

[0740] The selector sets created according to the methods described herein may be useful in the analysis of genetic alterations, particularly in comparing tumor and genomic sequences in a patient with cancer. As shown in FIG. **2**, a tissue biopsy sample from the patient may be used to discover mutations in the tumor by sequencing the genomic regions of the selector library in tumor and genomic nucleic acid samples and comparing the results. The selector sets may be designed to identify mutations in tumors from a large percentage of all patients, thus, it may not be necessary to optimize the library for each patient.

[0741] In some methods of the invention, the analysis of cfDNA for somatic mutations is compared to personalized tumor markers in an initial dataset developed from somatic mutations in a known tumor sample from an individual. To develop such a dataset, a sample of tumor cells or known tumor DNA may be obtained, which is compared to a germline sample. Preferably although not necessarily, a germline sample may be from the individual.

[0742] To "analyze" may include determining a set of values associated with a sample by determining a DNA sequence, and comparing the sequence against the sequence of a sample or set of samples from the same subject, from a control, from reference values, etc. as known in the art. To "analyze" can include performing a statistical analysis.

[0743] CAPP-seq may utilize hybrid selection of cfDNA corresponding to regions of recurrent mutation for diagnosis and monitoring of cancer in an individual patient. In such embodiments the selector set probes are used to enrich, e.g. by hybrid selection, for ctDNA that corresponds to the regions of the genome that are most likely to contain tumor-specific somatic mutations. The "selected" ctDNA is then amplified and sequenced to determine which of the selected genomic regions are mutated in the individual tumor. An initial comparison is optionally made with the individual's germline DNA sequence and/or a tumor biopsy sample from the individual. These somatic mutations provide a means of distinguishing ctDNA from germline DNA, and thus provide useful information about the presence and quantity of tumor cells in the individual. A flow chart for this process is provided in FIG. **22**.

[0744] In other embodiments, CAPP-seq is used for cancer screening and biopsy-free tumor genotyping, where a patient ctDNA sample is analyzed without reference to a biopsy sample. In some such embodiments, where CAPP-Seq identifies a mutation in a clinically actionable target from a ctDNA sample, the methods include providing a therapy appropriate for the target. Such mutations include, without limitation, rearrangements and other mutations involving oncogenes, receptor tyrosine kinases, etc.

[0745] Further disclosed herein is a method of detecting, diagnosing, prognosing, or therapy selection for a cancer subject comprising: (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; and (b) using sequence information derived from (a) to detect cell-free non-germline DNA (cfNG-DNA) in the sample, wherein the method is capable of detecting a percentage of cfNG-DNA that is less than 2% of total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 1.5% of the total cfDNA. The method may be capable of detecting a percentage of cfNG-DNA that is less than 1% of the total cfDNA. The method may be capable of detecting a percentage of cfNG-DNA that is less than 0.5% of the total cfDNA. The method may be capable of detecting a percentage of cfNG-DNA that is less than 0.1% of the total cfDNA. The method may be capable of detecting a percentage of cfNG-DNA that is less than 0.01% of the total cfDNA. The method may be capable of detecting a percentage of cfNG-DNA that is less than 0.001% of the total cfDNA. The method may be capable of detecting a percentage of cfNG-DNA that is less than 0.0001% of the total cfDNA. The sample may be

a plasma or serum sample. The sample may be a cerebral spinal fluid sample. In some instances, the sample is not a pap smear fluid sample. In some instances, the sample is a cyst fluid sample. In some instances, the sample is a pancreatic fluid sample. The sequence information may comprise information related to at least 10, 20, 30, 40, 100, 200, 300 genomic regions. The genomic regions may comprise genes, exonic regions, intronic regions, untranslated regions, non-coding regions or a combination thereof. The genomic regions may comprise two or more of exonic regions, intronic regions, and untranslated regions. The genomic regions may comprise at least one exonic region and at least one intronic region. At least 5% of the genomic regions may comprise intronic regions. At least about 20% of the genomic regions may comprise exonic regions. The genomic regions may comprise less than 1.5 megabases (Mb) of the genome. The genomic regions may comprise less than 1 Mb of the genome. The genomic regions may comprise less than 500 kilobases (kb) of the genome. The genomic regions may comprise less than 350 kb of the genome. The genomic regions may comprise between 100 kb to 300 kb of the genome. The sequence information may comprise information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to a plurality of genomic regions. The plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. The total size of the genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The total size of the genomic regions of the selector set may be between 100 kb to 300 kb of the genome. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 2. In some instances, the subject is not suffering from a pancreatic cancer. Obtaining sequence information may comprise performing massively parallel sequencing. Massively parallel sequencing may be performed on a subset of a genome of cfDNA from the cfDNA sample. The subset of the genome may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The subset of the genome may comprise between 100 kb to 300 kb of the genome. Obtaining sequence information may comprise using single molecule barcoding. Using single molecule barcoding may comprise attaching barcodes comprising different sequences to nucleic acids from the cfDNA sample. The sequence information may comprise sequence information pertaining to the barcodes. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from the subject. The two or more samples may be the same type of sample. The two or more samples may be two different types of sample. The two or more samples may be obtained from the subject at the same time point. The two or more samples may be obtained from the subject at two or more time points. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more different subjects. The samples from two or more different subjects may be indexed and pooled together prior to obtaining the sequencing information. Using the sequence information may comprise detecting one or more SNVs, indels, fusions, breakpoints, structural variants, variable number of tandem repeats, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, or a combination thereof in selected regions of the subject's genome. Using the sequence information may comprise detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting at least one SNV, indel, copy number variant, and rearrangement in selected regions of the subject's genome. In some instances, detecting does not involve performing digital PCR (dPCR). Detecting cell-free non-germline DNA may comprise applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in at least about 60% of cancer subjects from population of cancer subjects. The cfNG-DNA may be derived from a tumor in the subject. The method may further comprise detecting a cancer in the subject based on the detection of the cfNG-DNA. The method may further comprise diagnosing a cancer in the subject based on the detection of the cfNG-DNA. Diagnosing the cancer may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the cancer may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise prognosing a cancer in the subject based on the detection of the cfNG-DNA. Prognosing the cancer may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Prognosing the cancer may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise determining a therapeutic regimen for the subject based on the detection of the cfNG-DNA. The method may further comprise administering an anti-cancer therapy to the subject based on the detection of the cfNG-DNA. The cfNG-DNA may be derived from a fetus in the subject. The method may further comprise diagnosing a disease or condition in the fetus based on the detection of the cfNG-DNA. Diagnosing the disease or condition in the fetus may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the disease or condition in the fetus may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The cfNG-DNA may be derived from a transplanted organ, cell or tissue in the subject. The method may further comprise diagnosing

an organ transplant rejection in the subject based on the detection of the cfNG-DNA. Diagnosing the organ transplant rejection may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the organ transplant rejection may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise prognosing a risk of organ transplant rejection in the subject based on the detection of the cfNG-DNA. Prognosing the risk of organ transplant rejection may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Prognosing the risk of organ transplant rejection may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise determining an immunosuppresive therapy for the subject based on the detection of the cfNG-DNA. The method may further comprise administering an immunosuppresive therapy to the subject based on the detection of the cfNG-DNA.

[0746] Further disclosed herein are methods of detecting, diagnosing, or prognosing a status or outcome of a cancer in a subject. The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; (b) using sequence information derived from (a) to detect cell-free tumor DNA (ctDNA) in the sample wherein the method is capable of detecting a percentage of ctDNA that is less than 2% of total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 1.5% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 1% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.5% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.1% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.01% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.001% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.0001% of the total cfDNA. The sample may be a plasma or serum sample. The sample may be a cerebral spinal fluid sample. In some instances, the sample is not a pap smear fluid sample. In some instances, the sample is a cyst fluid sample. In some instances, the sample is a pancreatic fluid sample. The sequence information may comprise information related to at least 10, 20, 30, 40, 100, 200, 300 genomic regions. The genomic regions may comprise genes, exonic regions, intronic regions, untranslated regions, non-coding regions or a combination thereof. The genomic regions may comprise two or more of exonic regions, intronic regions, and untranslated regions. The genomic regions may comprise at least one exonic region and at least one intronic region. At least 5% of the genomic regions may comprise intronic regions. At least about 20% of the genomic regions may comprise exonic regions. The genomic regions may comprise less than 1.5 megabases (Mb) of the genome. The genomic regions may comprise less than 1 Mb of the genome. The genomic regions may comprise less than 500 kilobases (kb) of the genome. The genomic regions may comprise less than 350 kb of the genome. The genomic regions may comprise between 100 kb to 300 kb of the genome. The sequence information may comprise information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to a plurality of genomic regions. The plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. The total size of the genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The total size of the genomic regions of the selector set may be between 100 kb to 300 kb of the genome. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 2. In some instances, the subject is not suffering from a pancreatic cancer. Obtaining sequence information may comprise performing massively parallel sequencing. Massively parallel sequencing may be performed on a subset of a genome of cfDNA from the cfDNA sample. The subset of the genome may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The subset of the genome may comprise between 100 kb to 300 kb of the genome. Obtaining sequence information may comprise using single molecule barcoding. Using single molecule barcoding may comprise attaching barcodes comprising different sequences to nucleic acids from the cfDNA sample. The sequence information may comprise sequence information pertaining to the barcodes. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from the subject. The two or more samples may be the same type of sample. The two or more samples may be two different types of sample. The two or more samples may be obtained from the subject at the same time point. The two or more samples may be obtained from the subject at two or more time points. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more different subjects. The samples from two or more different subjects may be indexed and pooled together prior to obtaining the sequencing information. Using the sequence information may comprise detecting one or more SNVs, indels, fusions, breakpoints, structural variants, variable number of tandem repeats, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, or a combination thereof in selected regions of the subject's genome. Using the sequence information may comprise detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting at least one SNV, indel, copy number variant, and rearrangement in selected regions of the subject's genome. In some instances, detecting does not involve performing digital PCR (dPCR). Detecting ctDNA may com-

prise applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in at least about 60% of cancer subjects from population of cancer subjects. The ctDNA may be derived from a tumor in the subject. The method may further comprise detecting a cancer in the subject based on the detection of the ctDNA. The method may further comprise diagnosing a cancer in the subject based on the detection of the ctDNA. Diagnosing the cancer may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Diagnosing the cancer may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise prognosing a cancer in the subject based on the detection of the ctDNA. Prognosing the cancer may have a sensitivity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. Prognosing the cancer may have a specificity of at least about 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%. The method may further comprise determining a therapeutic regimen for the subject based on the detection of the ctDNA. The method may further comprise administering an anti-cancer therapy to the subject based on the detection of the ctDNA.

[0747] Further disclosed herein are methods of diagnosing a status or outcome of a cancer in a subject. The method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information is derived from genomic regions that are mutated in at least 80% of a population of subjects afflicted with a cancer; and (b) diagnosing a cancer selected from a group consisting of lung cancer, breast cancer, colorectal cancer and prostate cancer in the subject based on the sequence information, wherein the method has a sensitivity of 80%. The regions that are mutated may comprise a total size of less than 1.5 Mb of the genome. The regions that are mutated may comprise a total size of less than 1 Mb of the genome. The regions that are mutated may comprise a total size of less than 500 kb of the genome. The regions that are mutated may comprise a total size of less than 350 kb of the genome. The regions that are mutated may comprise a total size between 100 kb-300 kb of the genome. The sequence information may be derived from 2 or more regions. The sequence may be derived from 10 or more regions. The sequence may be derived from 50 or more regions. The population of subjects afflicted with the cancer may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA). The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 60% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 70% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 80% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 90% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 95% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 99% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 85% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 90% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 95% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 99% of the population of subjects afflicted with the cancer. The obtaining sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more 1ncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof. Obtaining sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof. In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. The method may further comprise detecting mutations in the regions based on the sequencing information. Diagnosing the cancer may be based on the detection of the mutations. The detection of at least 3 mutations may be indicative of the cancer. The detection of one or more mutations in three or more regions may be indicative of the cancer. The breast cancer may be a BRCA1 cancer. The method may have a sensitivity of at least 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method may have a specificity of at least 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method may further comprise providing a computer-generated report comprising the diagnosis of the cancer.

[0748] Further disclosed herein are methods of prognosing a status or outcome of a cancer in a subject. The method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information is derived from regions that are mutated in at least 80% of a population of subjects afflicted with a condition; and (b) determining a prognosis of a condition in the subject based on the sequence information. The regions that are mutated may comprise a total size of less than 1.5 Mb of the genome. The regions that are mutated may comprise a total size of less than 1 Mb of the genome. The regions that are mutated may comprise a total size of less than

500 kb of the genome. The regions that are mutated may comprise a total size of less than 350 kb of the genome. The regions that are mutated may comprise a total size between 100 kb-300 kb of the genome. The sequence information may be derived from 2 or more regions. The sequence may be derived from 10 or more regions. The sequence may be derived from 50 or more regions. The population of subjects afflicted with the condition may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA). The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 60% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 70% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 80% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 90% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 95% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 99% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 85% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 90% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 95% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 99% of the population of subjects afflicted with the condition. Obtaining sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more 1ncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof. Obtaining sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof. In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. The method may further comprise detecting mutations in the regions based on the sequencing information. Prognosing the condition may be based on the detection of the mutations. The detection of at least 3 mutations may be indicative of an outcome of the condition. The detection of one or more mutations in three or more regions may be indicative of an outcome of the condition. The condition may be a cancer. The cancer may be a solid tumor. The solid tumor may be non-small cell lung cancer (NSCLC). The cancer may be a breast cancer. The breast cancer may be a BRCA1 cancer. The cancer may be a lung cancer, colorectal cancer, prostate cancer, ovarian cancer, esophageal cancer, breast cancer, lymphoma, or leukemia. The method may have a sensitivity of at least 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method may have a specificity of at least 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method may further comprise providing a computer-generated report comprising the prognosis of the condition.

[0749] Disclosed herein are methods for detecting at least 50% of stage I cancer with a specificity of greater than 90%. The method may comprise (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage I cancer in the sample based on the quantity of the cell-free DNA. Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR. The quantity of the cell-free DNA may be determined by molecular barcoding of the cell-free DNA (cfDNA). Molecular barcoding of the cfDNA may comprise attaching barcodes to one or more ends of the cfDNA. The barcode may comprise a random sequence. Two or more barcodes may comprise two or more different random sequences. The barcode may comprise an adaptor sequence. Two or more barcodes may comprise the same adaptor sequence. The barcode may comprise a primer sequence. Two or more barcodes may comprise the same primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer. Attaching the barcodes to one or more ends of the ctDNA may comprise ligating the barcodes to the one or more ends of the ctDNA. Sequencing may comprise massively parallel sequencing. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2. At least 20%, 30%, 35%, 40%, 455, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set are based on genomic regions from Table 2. The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer. The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The method may have a sensitivity of at least 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method may detect at least 52%,

55%, 57%, 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage I cancer.

[0750] Disclosed herein are methods for detecting at least 60% of stage II cancer with a specificity of greater than 90% comprising (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage II cancer in the sample based on the quantity of the cell-free DNA. Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR. The quantity of the cell-free DNA may be determined by molecular barcoding of the cell-free DNA (cfDNA). Molecular barcoding of the cfDNA may comprise attaching barcodes to one or more ends of the cfDNA. The barcode may comprise a random sequence. Two or more barcodes may comprise two or more different random sequences. The barcode may comprise an adaptor sequence. Two or more barcodes may comprise the same adaptor sequence. The barcode may comprise a primer sequence. Two or more barcodes may comprise the same primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer. Attaching the barcodes to one or more ends of the ctDNA may comprise ligating the barcodes to the one or more ends of the ctDNA. Sequencing may comprise massively parallel sequencing. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2. At least 20%, 30%, 35%, 40%, 455, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2. The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer. The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The method may have a sensitivity of at least 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method may detect at least 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage II cancer.

[0751] Disclosed herein are methods for detecting at least 60% of stage III cancer with a specificity of greater than 90% comprising (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage III cancer in the sample based on the quantity of the cell-free DNA. Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR. The quantity of the cell-free DNA may be determined by molecular barcoding of the cell-free DNA (cfDNA). Molecular barcoding of the cfDNA may comprise attaching barcodes to one or more ends of the cfDNA. The barcode may comprise a random sequence. Two or more barcodes may comprise two or more different random sequences. The barcode may comprise an adaptor sequence. Two or more barcodes may comprise the same adaptor sequence. The barcode may comprise a primer sequence. Two or more barcodes may comprise the same primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer. Attaching the barcodes to one or more ends of the ctDNA may comprise ligating the barcodes to the one or more ends of the ctDNA. Sequencing may comprise massively parallel sequencing. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2. At least 20%, 30%, 35%, 40%, 455, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2. The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer. The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The method may have a sensitivity of at least 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method may detect at least 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage III cancer.

[0752] Disclosed herein are methods for detecting at least 60% of stage IV cancer with a specificity of greater than 90% comprising (a) performing sequencing on cell-free DNA derived from a sample, wherein the cell-free DNA to be sequenced is based on a selector set comprising a plurality of genomic regions; (b) using a computer readable medium to determine a quantity of the cell-free DNA based on the sequencing information of the cell-free DNA; and (c) detecting a stage IV cancer in the sample based on the quantity of the cell-free DNA. Determining the quantity of the cell-free DNA may comprise determining absolute quantities of the cell-free DNA. The quantity of the cell-free DNA may be determined by counting sequencing reads pertaining to the cell-free DNA. The quantity of the cell-free DNA may be determined by quantitative PCR. The quantity of the cell-free DNA may be determined by molecular barcoding of the cell-free DNA (cfDNA). Molecular barcoding of the cfDNA may comprise attaching barcodes to one or more ends of the cfDNA. The barcode may comprise a random sequence. Two or more barcodes may comprise two or more different random sequences. The barcode may comprise an adaptor sequence. Two or more barcodes may comprise the same adaptor sequence. The barcode may comprise a primer sequence. Two or more barcodes may comprise the same primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer. Attaching the barcodes to one or more ends of the ctDNA may

comprise ligating the barcodes to the one or more ends of the ctDNA. Sequencing may comprise massively parallel sequencing. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 or more genomic regions from Table 2. At least 20%, 30%, 35%, 40%, 455, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% or more of the genomic regions in the selector set may be based on genomic regions from Table 2. The plurality of genomic regions may comprise one or more mutations present in at least 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or 99% or more of a population of subjects suffering from the cancer. The total size of the plurality of genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of a genome. The total size of the plurality of genomic regions of the selector set may be between 100 kb to 300 kb of a genome. The method may have a sensitivity of at least 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97%, or 99% or more. The method may detect at least 60%, 62%, 65%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 92%, 95%, 97% or more of stage IV cancer.

[0753] Further disclosed herein are methods of selecting a therapy for a subject suffering from a cancer. The method may comprise (a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; (b) using sequence information derived from (a) to detect cell-free tumor DNA (ctDNA) in the sample; and (c) determining a therapy for the subject based on the detection of the ctDNA, wherein the method is capable of detecting a percentage of ctDNA that is less than 2% of total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 1.5% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 1% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.5% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.1% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.01% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.001% of the total cfDNA. The method may be capable of detecting a percentage of ctDNA that is less than 0.0001% of the total cfDNA. The sample may be a plasma or serum sample. The sample may be a cerebral spinal fluid sample. In some instances, the sample is not a pap smear fluid sample. In some instances, the sample is a cyst fluid sample. In some instances, the sample is a pancreatic fluid sample. The sequence information may comprise information related to at least 10, 20, 30, 40, 100, 200, 300 genomic regions. The genomic regions may comprise genes, exonic regions, intronic regions, untranslated regions, non-coding regions or a combination thereof. The genomic regions may comprise two or more of exonic regions, intronic regions, and untranslated regions. The genomic regions may comprise at least one exonic region and at least one intronic region. At least 5% of the genomic regions may comprise intronic regions. At least about 20% of the genomic regions may comprise exonic regions. The genomic regions may comprise less than 1.5 megabases (Mb) of the genome. The genomic regions may comprise less than 1 Mb of the genome. The genomic regions may comprise less than 500 kilobases (kb) of the genome. The genomic regions may comprise less than 350 kb of the genome. The genomic regions may comprise between 100 kb to 300 kb of the genome. The sequence information may comprise information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to a plurality of genomic regions. The plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. The total size of the genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The total size of the genomic regions of the selector set may be between 100 kb to 300 kb of the genome. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 2. In some instances, the subject is not suffering from a pancreatic cancer. Obtaining sequence information may comprise performing massively parallel sequencing. Massively parallel sequencing may be performed on a subset of a genome of cfDNA from the cfDNA sample. The subset of the genome may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The subset of the genome may comprise between 100 kb to 300 kb of the genome. Obtaining sequence information may comprise using single molecule barcoding. Using single molecule barcoding may comprise attaching barcodes comprising different sequences to nucleic acids from the cfDNA sample. The sequence information may comprise sequence information pertaining to the barcodes. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from the subject. The two or more samples may be the same type of sample. The two or more samples may be two different types of sample. The two or more samples may be obtained from the subject at the same time point. The two or more samples may be obtained from the subject at two or more time points. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more different subjects. The samples from two or more different subjects may be indexed and pooled together prior to obtaining the sequencing information. Using the sequence information may comprise detecting one or more SNVs, indels, fusions, breakpoints, structural variants, variable number of tandem repeats, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, or a combination thereof in selected regions of the subject's genome. Using the sequence information may comprise detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Using the sequence information may comprise detecting at least one SNV, indel, copy number variant, and rearrangement in selected regions of the subject's

genome. In some instances, detecting does not involve performing digital PCR (dPCR). Detecting ctDNA may comprise applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in at least about 60% of cancer subjects from population of cancer subjects. The ctDNA may be derived from a tumor in the subject. Determining the therapy may comprise administering a therapy to the subject. Determining the therapy may comprise modifying a therapeutic regimen. Modifying the therapeutic regimen may comprise terminating a therapeutic regimen. Modifying the therapeutic regimen may comprise adjusting a dosage of the therapy. Modifying the therapeutic regimen may comprise adjusting a frequency of the therapy. The therapeutic regimen may be modified based on a change in the quantity of the ctDNA. The dosage of the therapy may be increased in response to an increase in the quantity of the ctDNA. The dosage of the therapy may be decreased in response to a decrease in the quanitity of the ctDNA. The frequency of the therapy may be increased in response to an increase in the quantity of the ctDNA. The frequency of the therapy may be decreased in response to a decrease in the quanitity of ctDNA.

[0754] Alternatively, the method may comprise (a) obtaining sequence information of cell-free genomic DNA derived from a sample from a subject, wherein the sequence information is derived from regions that are mutated in at least 80% of a population of subjects afflicted with a condition; and (b) determining a therapeutic regimen of a condition in the subject based on the sequence information. The regions that are mutated may comprise a total size of less than 1.5 Mb of the genome. The regions that are mutated may comprise a total size of less than 1 Mb of the genome. The regions that are mutated may comprise a total size of less than 500 kb of the genome. The regions that are mutated may comprise a total size of less than 350 kb of the genome. The regions that are mutated may comprise a total size between 100 kb-300 kb of the genome. The sequence information may be derived from 2 or more regions. The sequence may be derived from 10 or more regions. The sequence may be derived from 50 or more regions. The population of subjects afflicted with the condition may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA). The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 60% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 70% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 80% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 90% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 95% of the population of subjects afflicted with the condition. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 99% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 85% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 90% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 95% of the population of subjects afflicted with the condition. The sequence information may be derived from regions that are mutated in at least 99% of the population of subjects afflicted with the condition. Obtaining sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more 1ncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof. Obtaining sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof. In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. The method may further comprise detecting mutations in the regions based on the sequencing information. Determining the therapeutic regimen may be based on the detection of the mutations. The condition may be a cancer. The cancer may be a solid tumor. The solid tumor may be non-small cell lung cancer (NSCLC). The cancer may be a breast cancer. The breast cancer may be a BRCA1 cancer. The cancer may be a lung cancer, colorectal cancer, prostate cancer, ovarian cancer, esophageal cancer, breast cancer, lymphoma, or leukemia.

[0755] Further disclosed herein are methods for diagnosing, prognosing, or determining a therapeutic regimen for a subject afflicted with or susceptible of having a cancer. The method may comprise (a) obtaining sequence information for selected regions of genomic DNA from a cell-free DNA sample from the subject; (b) using the sequence information to determine the presence or absence of one or more mutations in the selected regions, wherein at least 70% of a population of subjects afflicted with the cancer have mutation(s) in the regions; and (c) providing a report with a diagnosis, prognosis or treatment regimen to the subject, based on the presence or absence of the one or more mutations. The selected regions may comprise a total size of less than 1.5 Mb of the genome. The selected regions may comprise a total size of less than 1 Mb of the genome. The selected regions may comprise a total size of less than 500 kb of the genome. The selected regions mutated may comprise a total size of less than 350 kb of the genome. The selected regions may comprise a total size between 100 kb-300 kb of the genome. The sequence information may be derived from 2 or more selected regions. The sequence may be derived from 10 or more selected regions. The sequence may be derived from 50 or

more selected regions. The population of subjects afflicted with the cancer may be subjects from one or more databases. The one or more databases may comprise The Cancer Genome Atlas (TCGA). The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 60% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 70% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 80% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 90% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 95% of the population of subjects afflicted with the cancer. The sequence information may comprise information pertaining to at least one mutation that may be present in at least about 99% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 85% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 90% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 95% of the population of subjects afflicted with the cancer. The sequence information may be derived from regions that are mutated in at least 99% of the population of subjects afflicted with the cancer. Obtaining sequence information may comprise sequencing noncoding regions. The noncoding regions may comprise one or more lncRNA, snoRNA, siRNA, miRNA, piRNA, tiRNA, PASR, TASR, aTASR, TSSa-RNA, snRNA, RE-RNA, uaRNA, x-ncRNA, hY RNA, usRNA, snaR, vtRNA, T-UCRs, pseudogenes, GRC-RNAs, aRNAs, PALRs, PROMPTs, LSINCTs, or a combination thereof. Obtaining sequence information may comprise sequencing protein coding regions. The protein coding regions may comprise one or more exons, introns, untranslated regions, or a combination thereof. In some instances, at least one of the regions does not comprise KRAS or EGFR. In some instances, at least two of the regions do not comprise KRAS and EGFR. In some instances, at least one of the regions does not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least two of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least three of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. In some instances, at least four of the regions do not comprise KRAS, EGFR, p53, PIK3CA, BRAF, EZH2, or BRCA1. The detection of at least 3 mutations may be indicative of an outcome of the cancer. The detection of one or more mutations in three or more regions may be indicative of an outcome of the cancer. The cancer may be non-small cell lung cancer (NSCLC). The cancer may be a breast cancer. The breast cancer may be a BRCA1 cancer. The cancer may be a lung cancer, colorectal cancer, prostate cancer, ovarian cancer, esophageal cancer, breast cancer, lymphoma, or leukemia. The method of diagnosing or prognosing the cancer has a sensitivity of at least 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method of diagnosing or prognosing the cancer has a specificity of at least 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The method may further comprise administering a therapeutic drug to the subject. The method may further comprise modifying a therapeutic regimen. Modifying the therapeutic regimen may comprise terminating the therapeutic regimen. Modifying the therapeutic regimen may comprise increasing a dosage or frequency of the therapeutic regimen. Modifying the therapeutic regimen may comprise decreasing a dosage or frequency of the therapeutic regimen. Modifying the therapeutic regimen may comprise starting the therapeutic regimen.

[0756] In some embodiment, the method further comprises selecting a therapeutic regimen based on the analysis. In an embodiment, the method further comprises determining a treatment course for the subject based on the analysis. In such embodiments, the presence of tumor cells in an individual, including an estimation of tumor load, provides information to guide clinical decision making, both in terms of institution of and escalation of therapy as well as in the selection of the therapeutic agent to which the patient is most likely to exhibit a robust response.

[0757] The information obtained by CAPP-seq can be used to (a) determine type and level of therapeutic intervention warranted (e.g. more versus less aggressive therapy, monotherapy versus combination therapy, type of combination therapy), and (b) to optimize the selection of therapeutic agents. With this approach, therapeutic regimens can be individualized and tailored according to the specificity data obtained at different times over the course of treatment, thereby providing a regimen that is individually appropriate. In addition, patient samples can be obtained at any point during the treatment process for analysis.

[0758] The therapeutic regimen may be selected based on the specific patient situation. Where CAPP-seq is used as an initial diagnosis, a sample having a positive finding for the presence of ctDNA can indicate the need for additional diagnostic tests to confirm the presence of a tumor, and/or initiation of cytoreductive therapy, e.g. administration of chemotherapeutic drugs, administration of radiation therapy, and/or surgical removal of tumor tissue.

[0759] Further disclosed herein are methods for assessing tumor burden in a subject. The method may comprise (a) obtaining sequence information on cell-free nucleic acids derived from a sample from the subject; (b) using a computer readable medium to determine quantities of circulating tumor DNA (ctDNA) in the sample; (c) assessing tumor burden based on the quantities of ctDNA; and (d) reporting the tumor burden to the subject or a representative of the subject. Determining quantities of ctDNA may comprise determining absolute quantities of ctDNA. Determining quantities of ctDNA may comprise determining relative quantities of ctDNA. Determining quantities of ctDNA may be performed by counting sequence reads pertaining to the ctDNA. Determining quantities of ctDNA may be performed by quantitative PCR. Determining quantities of ctDNA may be performed by digital PCR. Determining quantities of ctDNA may be performed by molecular barcoding of the ctDNA. Molecular barcoding of the ctDNA may comprise attaching barcodes to one or more ends of the ctDNA. The barcode may comprise a random sequence. Two or more barcodes may comprise two or more different random sequences. The barcode may comprise an adaptor sequence. Two or more barcodes may comprise the same adaptor sequence. The barcode may comprise

a primer sequence. Two or more barcodes may comprise the same primer sequence. The primer sequence may be a PCR primer sequence. The primer sequence may be a sequencing primer. Attaching the barcodes to one or more ends of the ctDNA may comprise ligating the barcodes to the one or more ends of the ctDNA. The sequence information may comprise information related to one or more genomic regions. The sequence information may comprise information related to at least 10, 20, 30, 40, 100, 200, 300 genomic regions. The genomic regions may comprise genes, exonic regions, intronic regions, untranslated regions, non-coding regions or a combination thereof. The genomic regions may comprise two or more of exonic regions, intronic regions, and untranslated regions. The genomic regions may comprise at least one exonic region and at least one intronic region. At least 5% of the genomic regions may comprise intronic regions. At least about 20% of the genomic regions may comprise exonic regions. The genomic regions may comprise less than 1.5 megabases (Mb) of the genome. The genomic regions may comprise less than 1 Mb of the genome. The genomic regions may comprise less than 500 kilobases (kb) of the genome. The genomic regions may comprise less than 350 kb of the genome. The genomic regions may comprise between 100 kb to 300 kb of the genome. The sequence information may comprise information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions. The sequence information may comprise information pertaining to a plurality of genomic regions. The plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. At least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions may be based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects. The total size of the genomic regions of the selector set may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The total size of the genomic regions of the selector set may be between 100 kb to 300 kb of the genome. The selector set may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from Table 2. Obtaining sequence information may comprise performing massively parallel sequencing. Massively parallel sequencing may be performed on a subset of a genome of the cell-free nucleic acids from the sample. The subset of the genome may comprise less than 1.5 megabases (Mb), 1 Mb, 500 kilobases (kb), 350 kb, 300 kb, 250 kb, 200 kb, or 150 kb of the genome. The subset of the genome may comprise between 100 kb to 300 kb of the genome. The method may comprise obtaining sequencing information of cell-free DNA samples from two or more samples from the subject. The two or more samples are the same type of sample. The two or more samples are two different types of sample. The two or more samples are obtained from the subject at the same time point. The two or more samples are obtained from the subject at two or more time points. Determining the quantities of ctDNA may comprise detecting one or more SNVs, indels, fusions, breakpoints, structural variants, variable number of tandem repeats, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, or a combination thereof in selected regions of the subject's genome. Determining the quantities of ctDNA may comprise detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Determining the quantities of ctDNA may comprise detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome. Determining the quantities of ctDNA may comprise detecting at least one SNV, indel, copy number variant, and rearrangement in selected regions of the subject's genome. Determining the quantities of ctDNA does not involve performing digital PCR (dPCR). Determining the quantities of ctDNA may comprise applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in one or more cancer subjects from a population of cancer subjects. The selector set may comprise a plurality of genomic regions comprising one or more mutations present in at least about 60% of cancer subjects from population of cancer subjects. The representative of the subject may be a healthcare provider. The healthcare provider may be a nurse, physician, medical technician, or hospital personnel. The representative of the subject may be a family member of the subject. The representative of the subject may be a legal guardian of the subject.

[0760] Further disclosed herein are methods for determining a disease state of a cancer in a subject. The method may comprise (a) obtaining a quantity of circulating tumor DNA (ctDNA) in a sample from the subject; (b) obtaining a volume of a tumor in the subject; and (c) determining a disease state of a cancer in the subject based on a ratio of the quantity of ctDNA to the volume of the tumor. A high ctDNA to volume ratio may be indicative of radiographically occult disease. A low ctDNA to volume ratio may be indicative of non-malignant state. Obtaining the volume of the tumor may comprise obtaining an image of the tumor. Obtaining the volume of the tumor may comprise obtaining a CT scan of the tumor. Obtaining the quantity of ctDNA may comprise digital PCR. Obtaining the quantity of ctDNA may comprise obtaining sequencing information on the ctDNA. The sequencing information may comprise information relating to one or more genomic regions based on a selector set. Obtaining the quantity of ctDNA may comprise hybridization of the ctDNA to an array. The array may comprise a plurality of probes for selective hybridization of one or more genomic regions based on a selector set. The selector set may comprise one or more genomic regions from Table 2. The selector set may comprise one or more genomic regions comprising one or more mutations, wherein the one or more mutations are present in a population of subjects suffering from a cancer. The selector set may comprise a plurality of genomic regions comprising a plurality of mutations, wherein the plurality of mutations are present in at least 60% of a population of subjects suffering from a cancer.

[0761] In some embodiments, the ctDNA content in an individual's blood, or blood derivative, sample is determined at one or more time points, optionally in conjunction with a therapeutic regimen. The presence of the ctDNA correlates with tumor burden, and is useful in monitoring response to therapy, monitoring residual disease, monitoring for the presence of metastases, monitoring total tumor burden, and the

like. Although not required, for some methods CAPP-Seq may be performed in conjunction with tumor imaging methods, e.g. PET/CT scans and the like. Where CAPP-seq is used to estimate tumor burden or residual disease, increased presence of tumor cells over time indicates a need to increase the therapy by escalating dose, selection of agent, etc. Correspondingly, where CAPP-seq shows no evidence of residual disease, a patient may be taken off therapy, or put on a lowered dose.

[0762] CAPP-seq can also be used in clinical trials for new drugs, to determine the efficacy of treatment for a cancer of interest, where a decrease in tumor burden is indicative of efficacy and increased tumor burden is indicative of a lack of efficacy.

[0763] The cancer of interest may be specific for a cancer, for example non-small cell carcinoma, endometrioid uterine carcinoma, etc.; or may be generic for a class of cancers, e.g. epithelial cancers (carcinomas); sarcomas; lymphomas; melanomas; gliomas; teratomas; etc.; or subgenus, e.g. adenocarcinoma; squamous cell carcinoma; and the like.

[0764] The term "diagnosis" may refer to the identification of a molecular or pathological state, disease or condition, such as the identification of a molecular subtype of breast cancer, prostate cancer, or other type of cancer.

[0765] The term "prognosis" may refer to the prediction of the likelihood of cancer-attributable death or progression, including recurrence, metastatic spread, and drug resistance, of a neoplastic disease, such as ovarian cancer. The term "prediction" may refer to the act of foretelling or estimating, based on observation, experience, or scientific reasoning. In one example, a physician may predict the likelihood that a patient will survive, following surgical removal of a primary tumor and/or chemotherapy for a certain period of time without cancer recurrence.

[0766] The terms "treatment," "treating," and the like, may refer to administering an agent, or carrying out a procedure, for the purposes of obtaining an effect. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof and/or may be therapeutic in terms of effecting a partial or complete cure for a disease and/or symptoms of the disease. "Treatment," as used herein, may include treatment of a tumor in a mammal, particularly in a human, and includes: (a) preventing the disease or a symptom of a disease from occurring in a subject which may be predisposed to the disease but has not yet been diagnosed as having it (e.g., including diseases that may be associated with or caused by a primary disease; (b) inhibiting the disease, e.g., arresting its development; and (c) relieving the disease, e.g., causing regression of the disease.

DEFINITIONS

[0767] A number of terms conventionally used in the field of cell culture are used throughout the disclosure. In order to provide a clear and consistent understanding of the specification and claims, and the scope to be given to such terms, the following definitions are provided.

[0768] It is to be understood that this invention is not limited to the particular methodology, protocols, cell lines, animal species or genera, and reagents described, as such may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which will be limited only by the appended claims.

[0769] As used herein the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" may include a plurality of such cells and reference to "the culture" may include reference to one or more cultures and equivalents thereof known to those skilled in the art, and so forth. All technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs unless clearly indicated otherwise.

[0770] "Measuring" or "measurement" in the context of the present teachings may refer to determining the presence, absence, quantity, amount, or effective amount of a substance in a clinical or subject-derived sample, including the presence, absence, or concentration levels of such substances, and/or evaluating the values or categorization of a subject's clinical parameters based on a control.

[0771] Unless otherwise apparent from the context, all elements, steps or features of the invention can be used in any combination with other elements, steps or features.

[0772] General methods in molecular and cellular biochemistry can be found in such standard textbooks as Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001); Short Protocols in Molecular Biology, 4th Ed. (Ausubel et al. eds., John Wiley & Sons 1999); Protein Methods (Bollag et al., John Wiley & Sons 1996); Nonviral Vectors for Gene Therapy (Wagner et al. eds., Academic Press 1999); Viral Vectors (Kaplift & Loewy eds., Academic Press 1995); Immunology Methods Manual (I. Lefkovits ed., Academic Press 1997); and Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998). Reagents, cloning vectors, and kits for genetic manipulation referred to in this disclosure may be available from commercial vendors such as BioRad, Stratagene, Invitrogen, Sigma-Aldrich, and ClonTech.

[0773] The invention has been described in terms of particular embodiments found or proposed by the present inventor to comprise preferred modes for the practice of the invention. It will be appreciated by those of skill in the art that, in light of the present disclosure, numerous modifications and changes can be made in the particular embodiments exemplified without departing from the intended scope of the invention. Due to biological functional equivalency considerations, changes can be made in protein structure without affecting the biological action in kind or amount. All such modifications are intended to be included within the scope of the appended claims.

[0774] The terms "subject," "individual," and "patient" are used interchangeably herein and may refer to a mammal being assessed for treatment and/or being treated. In an embodiment, the mammal is a human. The terms "subject," "individual," and "patient" may encompass, without limitation, individuals having cancer or suspected of having cancer. Subjects may be human, but also include other mammals, particularly those mammals useful as laboratory models for human disease, e.g. mouse, rat, etc. Also included are mammals such as domestic and other species of canines, felines, and the like.

[0775] The terms "cancer," "neoplasm," and "tumor" are used interchangeably herein and may refer to cells which exhibit autonomous, unregulated growth, such that they exhibit an aberrant growth phenotype characterized by a significant loss of control over cell proliferation. Cells of interest

for detection, analysis, or treatment in the present application may include, but are not limited to, precancerous (e.g., benign), malignant, pre-metastatic, metastatic, and non-metastatic cells. Cancers of virtually every tissue are known. The phrase "cancer burden" may refer to the quantum of cancer cells or cancer volume in a subject. Reducing cancer burden accordingly may refer to reducing the number of cancer cells or the cancer volume in a subject. The term "cancer cell" as used herein may refer to any cell that is a cancer cell or is derived from a cancer cell, e.g. clone of a cancer cell. Many types of cancers are known to those of skill in the art, including solid tumors such as carcinomas, sarcomas, glioblastomas, melanomas, lymphomas, myelomas, etc., and circulating cancers such as leukemias. Examples of cancer include, but are not limited to, ovarian cancer, breast cancer, colon cancer, lung cancer, prostate cancer, hepatocellular cancer, gastric cancer, pancreatic cancer, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, head and neck cancer, and brain cancer.

[0776] The "pathology" of cancer may include, but it not limited to, all phenomena that compromise the well-being of the patient. This includes, without limitation, abnormal or uncontrollable cell growth, metastasis, interference with the normal functioning of neighboring cells, release of cytokines or other secretory products at abnormal levels, suppression or aggravation of inflammatory or immunological response, neoplasia, premalignancy, malignancy, invasion of surrounding or distant tissues or organs, such as lymph nodes, etc.

[0777] As used herein, the terms "cancer recurrence" and "tumor recurrence," and grammatical variants thereof, may refer to further growth of neoplastic or cancerous cells after diagnosis of cancer. Particularly, recurrence may occur when further cancerous cell growth occurs in the cancerous tissue. "Tumor spread," similarly, may occur when the cells of a tumor disseminate into local or distant tissues and organs; therefore tumor spread may encompass tumor metastasis. "Tumor invasion" may occur when the tumor growth spreads out locally to compromise the function of involved tissues by compression, destruction, and/or prevention of normal organ function.

[0778] As used herein, the term "metastasis" may refer to the growth of a cancerous tumor in an organ or body part, which is not directly connected to the organ of the original cancerous tumor. Metastasis may include micrometastasis, which is the presence of an undetectable amount of cancerous cells in an organ or body part which is not directly connected to the organ of the original cancerous tumor. Metastasis can also be defined as several steps of a process, such as the departure of cancer cells from an original tumor site, and migration and/or invasion of cancer cells to other parts of the body.

[0779] As used herein, DNA, RNA, nucleic acids, nucleotides, oligonucleotides, polynucleotides may be used interchangeably. Unless explicitly stated otherwise, the term DNA encompasses any type of nucleic acid (e.g., DNA, RNA, DNA/RNA hybrids, and analogues thereof). In instances in which RNA is used in the methods disclosed herein, the methods may further comprise reverse transcription of the RNA to produce a complementary DNA (cDNA) or DNA copy.

[0780] All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[0781] The present invention has been described in terms of particular embodiments found or proposed by the present inventor to comprise preferred modes for the practice of the invention. It will be appreciated by those of skill in the art that, in light of the present disclosure, numerous modifications and changes can be made in the particular embodiments exemplified without departing from the intended scope of the invention. For example, due to codon redundancy, changes can be made in the underlying DNA sequence without affecting the protein sequence. In another example, due to similarities in DNA and RNA, the methods, compositions, and systems may be equally applicable to all types of nucleic acids (e.g., DNA, RNA, DNA/RNA hybrids, and analogues thereof). Moreover, due to biological functional equivalency considerations, changes can be made in protein structure without affecting the biological action in kind or amount. All such modifications are intended to be included within the scope of the appended claims.

[0782] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

EXAMPLES

Example 1

An Ultrasensitive Method for Quantitating Circulating Tumor DNA with Broad Patient Coverage

[0783] Circulating tumor DNA (ctDNA) represents a promising biomarker for noninvasive detection of disease burden and monitoring of recurrence. However, existing ctDNA detection methods are limited by sensitivity, a focus on small numbers of mutations, and/or the need for patient-specific optimization. To address these shortcomings, CAncer Personalized Profiling by Deep Sequencing (CAPP-Seq) was developed, an economical and highly sensitive method for quantifying ctDNA in plasma in nearly every patient. We implemented CAPP-Seq for non-small cell lung cancer (NSCLC) with a design that identified mutations in >95% of tumors, simultaneously detecting point mutations, insertions/deletions, copy number variants, and rearrangements. When tumor mutation profiles were known, we detected ctDNA in 100% of pre-treatment plasma samples from stages II-IV NSCLC and 50% of samples from stage I NSCLC, with a specificity of 95% for mutant allele fractions down to ~0.02%. Absolute quantities of ctDNA were significantly correlated with tumor volume. Furthermore, ctDNA levels in post-treatment samples helped distinguish between residual disease and treatment-related imaging changes and provided earlier response assessment than radiographic approaches. Finally, we explored the utility of this method for biopsy-free

tumor genotyping and cancer screening. CAPP-Seq can be routinely applied clinically to detect and monitor diverse malignancies, thus facilitating personalized cancer therapy. Here we demonstrate the technical performance and explore the clinical utility of CAPP-Seq in patients with early and advanced stage NSCLC.

[0784] Design of a CAPP-Seq Selector for NSCLC.

[0785] For the initial implementation of CAPP-Seq we focused on NSCLC, although our approach can be used for any cancer for which recurrent mutations have been identified. We employed a multi-phase approach to design an NSCLC-specific selector, aiming to identify genomic regions recurrently mutated in this disease (FIG. **1***b*, Table 1). We began by including exons covering recurrent mutations in potential driver genes from the Catalogue of Somatic Mutations in Cancer (COSMIC) database as well as other sources (e.g. KRAS, EGFR, TP53). Next, using whole exome sequencing (WES) data from 407 NSCLC patients profiled by The Cancer Genome Atlas (TCGA), we applied an iterative algorithm to maximize the number of missense mutations per patient while minimizing selector size. Our approach relied on a recurrence index that identified known driver mutations as well as uncharacterized genes that are frequently mutated and are therefore likely to be involved in NSCLC pathogenesis (FIG. **7** and Table 2).

[0786] Approximately 8% of NSCLCs harbor clinically actionable rearrangements involving the receptor tyrosine kinases, ALK, ROS1 and RET. These structural aberrations, which are clinically actionable because they are targets of pharmacologic inhibitors, tend to disproportionately occur in younger patients with significantly less smoking history and whose tumors harbor fewer somatic alterations than most other patients with NSCLC. To utilize the personalized nature and lower false detection rate inherent in the unique junctional sequences of structural rearrangements, we included the introns and exons spanning recurrent fusion breakpoints in these genes in the final design phase (FIG. **1***b*). To detect fusions in tumor and plasma DNA, we developed a breakpoint-mapping algorithm called FACTERA (FIG. **8**). Application of FACTERA to next generation sequencing (NGS) data from 2 NSCLC cell lines known to harbor fusions with previously uncharacterized breakpoints readily identified the breakpoints at nucleotide resolution and these were independently confirmed in both cases (FIG. **9**).

[0787] Collectively, the NSCLC selector design targets 521 exons and 13 introns from 139 recurrently mutated genes, in total covering ~125 kb (FIG. **1***b*). Within this small target (0.004% of the human genome), the selector identifies a median of 4 point mutations and covers 96% of patients with lung adenocarcinoma or squamous cell carcinoma. To validate the number of mutations covered per tumor, we examined the selector region in WES data from an independent cohort of 183 lung adenocarcinoma patients. The selector covered 88% of patients with a median of 4 SNVs per patient, thus validating our selector design algorithm ($P<1.0\times10^{-6}$; FIG. **1***c*). When compared to randomly sampling the exome, regions targeted by the NSCLC selector captured ~4-fold more mutations per patient (at the median, FIG. **1***c*). Due to similarities in key oncogenic machinery across cancers, the NSCLC selector performs favorably on other carcinomas. Indeed, the selector successfully captured 99% of colon, 98% of rectal, and 97% of endometrioid uterine carcinomas, with a median of 12, 7, and 3 mutations per patient, respectively (FIG. **1***d*). This demonstrates the value of targeting hundreds of recurrently mutated genomic regions and shows that a single selector can be designed to simultaneously cover recurrent mutations for multiple malignancies.

[0788] Methodological Optimization and Performance Assessment.

[0789] We performed deep sequencing with the NSCLC selector to achieve ~10,000× coverage (pre-duplication removal, ~10-12 samples per lane), and profiled a total of 90 samples, including 2 NSCLC cell lines, 17 primary tumor biopsies and matched peripheral blood leukocyte (PBL) specimens, and 40 plasma samples from 18 human subjects, including 5 healthy adults and 13 patients with NSCLC before and after various cancer therapies (Tables 3, 20 and 21). To assess and optimize selector performance, we first applied it to cfDNA purified from healthy control plasma, observing efficient and uniform capture of genomic DNA (Tables 3, 20 and 21). Sequenced cfDNA fragments had a median length of ~170 bp (FIG. **2***a*), closely corresponding to the length of DNA contained within a chromatosome. To optimize library preparation from small quantities of cfDNA we explored a variety of modifications to the ligation and post-ligation amplification steps including temperature, incubation time, DNA polymerase, and PCR purification. The optimized protocol increased recovery efficiency by >300% and decreased bias for libraries constructed from as little as 4 ng of cfDNA (FIGS. **10**, **11**, and **12**). Consequently, fluctuations in sequencing depth were minimal (FIG. **2***b,c*).

[0790] The detection limit of CAPP-Seq is affected by (i) the input number and recovery rate of cfDNA molecules, (ii) sample cross-contamination, (iii) potential allelic bias in the capture reagent, and (iv) PCR or sequencing errors (e.g., "technical" background). We examined each of these elements in turn to better understand their potential impact on CAPP-Seq sensitivity. First, by comparing the number of input DNA molecules per sample with estimates of library complexity (FIG. **13***a*), we calculated a cfDNA molecule recovery rate of ≥49% (Tables 3, 20 and 21). This was in agreement with molecule recovery efficiencies calculated using post-PCR mass yields (FIG. **13***b*). Second, by analyzing patient-specific homozygous SNPs across samples, we found cross-contamination of ~0.06% in multiplexed cfDNA (FIG. **14**). While too low to affect ctDNA detection in most applications, we excluded any tumor-derived SNV from further analysis if found as a germline SNP in another profiled patient. To analyze possible capture bias, we next evaluated the allelic skew in heterozygous SNPs (single nucleotide polymorphism) within patient PBL (peripheral blood lymphocyte) samples. We observed a median heterozygous allele fraction of 51% (FIG. **15**), indicating minimal bias toward capture of reference alleles. Finally, we analyzed the distribution of non-reference alleles across the selector for the 40 cfDNA samples, excluding tumor-derived SNVs and germline SNPs (FIG. **2***d*). We found mean and median technical background rates of 0.006% and 0.0003%, respectively (FIG. **2***d*), both considerably lower than previously reported NGS-based methods for ctDNA analysis.

[0791] In addition to technical background, mutant cfDNA could be present in the absence of cancer due to contributions from pre-neoplastic cells from diverse tissues, and such "biological" background may impact sensitivity. We hypothesized that biological background, if present, would be particularly high for recurrently mutated positions in known cancer driver genes and therefore analyzed mutation rates of 107 selected cancer-associated SNVs in all 40 plasma

samples, excluding somatic mutations found in a patient's tumor. Though the median fractional abundance was comparable to the global selector background (~0%), the mean was marginally higher at ~0.01% (FIG. **2***e*). Strikingly, one mutation (TP53 R175H) was detected at a median frequency of ~0.18% across all cfDNA samples, including patients and healthy subjects (FIG. **2***f*). Since this allele is significantly above global background (P<0.01; FIG. **2***f*), we hypothesize that it reflects true biological background and thus excluded it as a potential reporter. To address background more generally, we also normalized for allele-specific differences in background rate when assessing the significance of ctDNA detection. As a result, we found that biological background is not a significant factor for ctDNA quantitation at detection limits above ~0.01%.

[0792] Next, we empirically benchmarked the allele frequency detection limit and linearity of CAPP-Seq by spiking defined concentrations of fragmented genomic DNA from a NSCLC cell line into cfDNA from a healthy individual (FIG. **2***g*) or into genomic DNA from a second NSCLC line (FIG. **16***a*). Defined inputs of NSCLC DNA were accurately detected at fractional abundances between 0.025% and 10% with high linearity ($R^2 \geq 0.994$). Analyses of the influence of the number of SNP reporters on error metrics showed only marginal improvements above a threshold of 4 reporters (FIG. **2***h,i*, FIG. **16***b,c*), equivalent to the median number of SNVs per NSCLC tumor identified by the selector. We also tested whether fusion breakpoints, indels, and CNVs could serve as linear reporters and found that the fractional abundance of these mutation types correlated highly with expected concentrations ($R^2 \geq 0.97$; FIG. **16***d*).

[0793] Identification of somatic mutations in NSCLC patients. Having designed, optimized, and assessed the technical performance of CAPP-Seq, we applied it to the discovery of somatic mutations in tumors collected from a diverse group of 17 NSCLC patients (Table 1 and Table 19). To test the utility of CAPP-Seq for identifying structural rearrangements, which are more frequently seen in tumors from non-smokers, we included 6 patients with clinically confirmed fusions. These translocations served as positive controls, along with SNVs in other tumors previously identified by clinical assays (Table 19). Tumor samples included formalin fixed surgical or biopsy specimens and pleural fluid containing malignant cells. At a mean sequencing depth of ~5,000× (pre-duplicate removal) in tumor and paired germline samples (Tables 3, 20 and 21), we detected 100% of previously identified SNVs and fusions (7 and 8, respectively) and discovered many additional somatic variants (Table 1 and Table 19). Moreover, partner genes and base-pair resolution breakpoints were characterized for each of the 8 rearrangements (FIG. **17**). Tumors containing fusions were almost exclusively from never smokers and, as expected, contained fewer SNVs than those lacking fusions (FIG. **18**). Excluding patients with fusions (<10% of the TCGA design cohort), we identified a median of 6 SNVs (3 missense) per patient (Table 1), in line with our selector design-stage predictions (FIG. **1***b-c*).

[0794] Sensitivity and Specificity.

[0795] Next, we assessed the sensitivity and specificity of CAPP-Seq for disease monitoring and minimal residual disease detection, using plasma samples from 5 healthy controls and 35 serial samples collected from 13 NSCLC patients, all but one of whom had pre- and post-treatment samples available (Table 1; Table 5). CAPP-Seq was used to measure tumor burden across the entire grid of plasma cfDNA samples (13 patient-specific sets of somatic reporters across 40 plasma samples, or 520 pairs), with an approach that integrates information content across multiple instances and classes of somatic mutations to increase sensitivity and specificity. Using ROC analysis, we achieved a maximal sensitivity and specificity of 85% and 95% (AUC=0.95), respectively, for all pre-treated tumors and healthy controls. Sensitivity among stage I tumors was 50% and among stage II-IV patients was 100% with a specificity of 96% (FIG. **3***a,b*). Moreover, when considering both pre and post-treatment samples in an ROC analysis, CAPP-Seq exhibited robust performance, with AUC values of 0.89 for all stages and 0.91 for stages II-IV (P<0.0001; FIG. **19**). Furthermore, by adjusting the ctDNA detection index, we could increase specificity up to 98% while still capturing ⅔ of all cancer-positive samples and ¾ of stage II-IV cancer-positive samples (FIG. **20**). This indicates that our approach could be tuned to deliver a desired sensitivity and specificity depending on the application in question and that CAPP-Seq can achieve robust assessment of tumor burden in NSCLC patients.

[0796] Monitoring of NSCLC Tumor Burden in Plasma Samples.

[0797] We next asked whether significantly detectable levels of ctDNA correlate with radiographically measured tumor volume and clinical response to therapy. Fractions of tumor-derived DNA detected in plasma by SNV and/or indel reporters ranged from ~0.02% to 3.2% (Table 1), with a median of ~0.1% in pre-treatment samples. Moreover, absolute levels of ctDNA in pre-treatment plasma were significantly correlated with tumor volume as measured by computed tomography (CT) and positron emission tomography (PET) imaging (R^2=0.89, P=0.0002; FIG. **3***c*).

[0798] To determine whether ctDNA concentrations reflect disease burden in longitudinal samples, we analyzed plasma cfDNA from three patients with high disease burden who underwent several rounds of therapy for metastatic NSCLC, including surgery, radiotherapy, chemotherapy, and tyrosine kinase inhibitors (FIG. **4***a-c*). As in pre-treatment samples, ctDNA levels were highly correlated with tumor volumes during therapy (R^2=0.95 for P15; R^2=0.85 for P9). In a never-smoker (P6), we detected 3 SNVs and a KIF5B-ALK fusion, and both mutation types were simultaneously detectable in plasma cfDNA and behaved comparably in response to Crizotinib therapy (FIG. **4***c*). In all 3 patients, this behavior was observed whether the mutation type measured was a collection of SNVs and an indel (P15, FIG. **4***a*), multiple fusions (P9, FIG. **4***b*), or SNVs and a fusion (P6, FIG. **4***c*), validating the utility of diverse tumor-derived somatic lesions. Of note, in one patient (P9) we identified both a classic EML4-ALK fusion and two previously unreported fusions involving ROS1: FYN-ROS1 and ROS1-MKX (FIG. **17**). All fusions were confirmed by qPCR amplification of genomic DNA and were independently recovered in plasma samples (Table 5). While the potential function of these novel ROS1 fusions is unknown, to the best of our knowledge this is the first observation of ROS1 and ALK fusions in the same NSCLC patient.

[0799] The NSCLC selector was designed to detect multiple SNVs per tumor and if present, more than 1 type of mutation per tumor. In one patient's tumor (P5), this design allowed us to identify a dominant clone with an activating EGFR mutation as well as a subclone with an EGFR T790M "gatekeeper" mutation. The ratio between clones was identical in a tumor biopsy and simultaneously sampled plasma

(FIG. **4***d*), demonstrating that by detecting multiple reporters per tumor, our method is useful for detecting and quantifying clinically relevant subclones.

[0800] Having validated the performance of CAPP-Seq on advanced stage patients, we next examined other clinical scenarios in which ctDNA biomarkers could be useful. Stage II-III NSCLC patients who undergo definitive radiotherapy with curative intent often have surveillance CT and/or PET/CT scans that are difficult to interpret due to radiation-induced inflammatory and fibrotic changes in the lung and surrounding tissues. These can delay diagnosis of recurrence or lead to unnecessary biopsies and patient anxiety. To compare the results of ctDNA quantitation to routine surveillance imaging, we analyzed pre- and post-radiotherapy plasma cfDNA in 2 patients. For patient P13, who was treated with radiotherapy alone for stage IIB NSCLC, follow-up imaging showed a large mass that was felt to represent residual disease. However, ctDNA at the same time point was undetectable (FIG. **4***e*) and the patient remained disease free 22 months later, supporting the ctDNA result. The second patient (P14) was treated with concurrent chemoradiotherapy for stage IIIB NSCLC and follow-up imaging revealed a near complete response in the thorax (FIG. **4***f*). However, the ctDNA concentration slightly increased compared to pre-treatment, suggesting progression of occult microscopic disease. Indeed, progression was detected clinically 7 months later and the patient ultimately succumbed to NSCLC. These data highlight the use of cfDNA analysis as a complementary modality to imaging studies and as a method for early diagnosis of recurrence.

[0801] We next asked whether the low detection limit of CAPP-Seq would allow monitoring of response to treatment in early stage NSCLC. Approximately 60-70% of stage I NSCLCs are curable with surgery or stereotactic ablative radiotherapy (SABR). Patients P1 (FIG. **4***g*) and P16 (FIG. **4***h*) underwent surgery and SABR, respectively, for stage IB NSCLC. We detected tumor-derived cfDNA in pre-treatment plasma of P1 but not at 3 or 32 months following surgery, suggesting this patient was free of disease and likely cured. For patient P16, the initial surveillance PET-CT scan following SABR showed a residual mass that was interpreted as representing either residual tumor or post-radiotherapy inflammation. We detected no evidence of residual disease by ctDNA, supporting the latter, and the patient remained free of disease at last follow-up 21 months after therapy. Taken together, these results demonstrate the utility of CAPP-Seq as a noninvasive clinical assay for measuring tumor burden in early and advanced stage NSCLC and for monitoring ctDNA during distinct types of therapy.

[0802] Noninvasive Tumor Genotyping and Cancer Screening.

[0803] Finally, we explored whether CAPP-Seq analysis of cfDNA could potentially be used for non-invasive tumor genotyping and cancer screening (e.g., without prior knowledge of tumor mutations). We blinded ourselves to the mutations present in each patient's tumor and applied a novel statistical method to test for the presence of cancer DNA in each plasma sample in our cohort (FIG. **21**). This method identified mutant alleles in all plasma samples containing ctDNA above fractional abundances of 0.4%, with no false positives (FIG. **4***i*). Thus, this approach has utility for non-invasive tumor genotyping in locally advanced or metastatic patients. Since ~95% of nodules identified in patients at high risk for developing NSCLC by low-dose CT are false positives, CAPP-Seq can also serve as a complementary noninvasive screening test.

[0804] In this study, we present CAPP-Seq as a new method for ctDNA quantitation. Key features of our approach include high sensitivity and specificity, coverage of nearly all patients with NSCLC, lack of patient-specific optimization, and low cost. By incorporating optimized library construction and bioinformatics methods, CAPP-Seq achieves the lowest background error rate and lowest detection limit of any NGS-based method used for ctDNA analysis to date. Our approach also reduces the potential impact of stochastic noise and biological variability (e.g., mutations near the detection limit or subclonal tumor evolution) on tumor burden quantitation by integrating information content across multiple instances and classes of somatic mutations. These features facilitated the detection of minimal residual disease and the first report of ctDNA quantitation from stage I NSCLC tumors using deep sequencing. Although we focused on NSCLC, our method can be applied to any malignancy for which recurrent mutation data are available.

[0805] In many patients, levels of ctDNA are considerably lower than the detection thresholds of previously described sequencing-based methods. For example, pre-treatment ctDNA concentration is <0.5% in the majority of patients with lung and colorectal carcinomas (and likely others), and <0.1% in most early and many advanced stage patients. Following therapy, ctDNA concentrations typically drop, rendering highly sensitive methods, like CAPP-Seq, even more critical. Recently, amplicon-based deep sequencing methods were implemented to detect up to 6 recurrently mutated genes per assay. Such approaches are limited by the number and types of mutations that can be simultaneously interrogated, and the reported allele detection limit of ~2% in plasma precludes ctDNA detection in most NSCLC patients. Several studies have reported application of whole exome or genome sequencing to cfDNA for analysis of somatic SNVs (single nucleotide variant) and CNVs (copy number variant). The sensitivity of SNV detection with these approaches is significantly limited by cost of sequencing, and even with 10-fold greater sequencing depth than we used for CAPP-Seq, would be insufficient to detect ctDNA in most NSCLC patients (FIG. **5***a*). Likewise, quantitation of CNVs in plasma via WGS has a reported detection limit of ~1%, limiting this approach to patients with high tumor burden.

[0806] Additional gains in the detection threshold are desirable. Approaches to achieve these gains include using barcoding strategies that suppress PCR errors resulting from library preparation, increasing the amount of plasma used for ctDNA analysis above the average of ~1.5 mL used in this study, further improving ligation and capture efficiency during library preparation, and increasing the size of the selector to increase the number of tumor-specific mutations per patient. A second limitation is the potential for inefficient capture of fusions, which could lead to underestimates of tumor burden (e.g., P9). However, this bias can be analytically addressed when other reporter types are present (e.g., P6; Table 4). Finally, while we found that CAPP-Seq could quantitate CNVs, our current selector design did not prioritize these types of aberrations. Adding coverage for certain CNVs can be useful for monitoring various types of cancers.

[0807] In summary, targeted hybrid capture and high-throughput sequencing of cfDNA allows for highly sensitive and non-invasive detection of ctDNA in cancer patients, at

low cost. CAPP-Seq can be routinely applied clinically for accelerating the personalized detection, therapy, and monitoring of cancer. CAPP-Seq is valuable in a variety of clinical settings, including the assessment of cancer DNA in alternative biological fluids and specimens with low cancer cell content.

[0808] Patient Selection.

[0809] Between April 2010 and June 2012, patients undergoing treatment for newly diagnosed or recurrent NSCLC were enrolled in a study approved by the Stanford University Institutional Review Board and provided informed consent. Enrolled patients had not received blood transfusions within 3 months of blood collection. Patient characteristics are in Tables 3, 20 and 21. All treatments and radiographic examinations were performed as part of standard clinical care. Volumetric measurements of tumor burden were based on visible tumor on CT and calculated according to the ellipsoid formula: (length/2)*(width^2).

[0810] Sample Collection and Processing.

[0811] Peripheral blood from patients was collected in EDTA Vacutainer tubes (BD). Blood samples were processed within 3 hours of collection. Plasma was separated by centrifugation at 2,500×g for 10 min, transferred to microcentrifuge tubes, and centrifuged at 16,000×g for 10 min to remove cell debris. The cell pellet from the initial spin was used for isolation of germline genomic DNA from PBLs (peripheral blood leukocytes) with the DNeasy Blood & Tissue Kit (Qiagen). Matched tumor DNA was isolated from FFPE specimens or from the cell pellet of pleural effusions. Genomic DNA was quantified by Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen).

[0812] Cell-Free DNA Purification and Quantification.

[0813] Cell-free DNA (cfDNA) was isolated from 1-5 mL plasma with the QIAamp Circulating Nucleic Acid Kit (Qiagen). The concentration of purified cfDNA was determined by quantitative PCR (qPCR) using an 81 bp amplicon on chromosome 1 and a dilution series of intact male human genomic DNA (Promega) as a standard curve. Power SYBR Green was used for qPCR on a HT7900 Real Time PCR machine (Applied Biosystems), using standard PCR thermal cycling parameters.

[0814] Illumina NGS Library Construction.

[0815] Indexed Illumina NGS libraries were prepared from cfDNA and shorn tumor, germline, and cell line genomic DNA. For patient cfDNA, 7-32 ng DNA were used for library construction without additional fragmentation. For tumor, germline, and cell line genomic DNA, 69-1000 ng DNA was sheared prior to library construction with a Covaris S2 instrument using the recommended settings for 200 bp fragments. See Table 2 for details.

[0816] The NGS libraries were constructed using the KAPA Library Preparation Kit (Kapa Biosystems) employing a DNA Polymerase possessing strong 3'-5' exonuclease (or proofreading) activity and displaying the lowest published error rate (e.g. highest fidelity) of all commercially available B-family DNA polymerases. The manufacturer's protocol was modified to incorporate with-bead enzymatic and cleanup steps using Agencourt AMPure XP beads (Beckman-Coulter). Ligation was performed for 16 hours at 16° C. using 100-fold molar excess of indexed Illumina TruSeq adapters. Single-step size selection was performed by adding 400 μL (0.8×) of PEG buffer to enrich for ligated DNA fragments. The ligated fragments were then amplified using 500 nM Illumina backbone oligonucleotides and 4-9 PCR cycles, depending on input DNA mass. Library purity and concentration was assessed by spectrophotometer (NanoDrop 2000) and qPCR (KAPA Biosystems), respectively. Fragment length was determined on a 2100 Bioanalyzer using the DNA 1000 Kit (Agilent).

[0817] Design of Library for Hybrid Selection.

[0818] Hybrid selection was performed with a custom SeqCap EZ Choice Library (Roche NimbleGen). This library was designed through the NimbleDesign portal (v1.2.R1) using genome build HG19 NCBI Build 37.1/GRCh37 and with Maximum Close Matches set to 1. Input genomic regions were selected according to the most frequently mutated genes and exons in NSCLC. These regions were identified from the COSMIC database, TCGA, and other published sources. Final selector coordinates are provided in Table 1.

[0819] Hybrid Selection and High Throughput Sequencing.

[0820] NimbleGen SeqCap EZ Choice was used according to the manufacturer's protocol with modifications. Between 9 and 12 indexed Illumina libraries were included in a single capture reaction. Following hybrid selection, the captured DNA fragments were amplified with 12 to 14 cycles of PCR using 1× KAPA HiFi Hot Start Ready Mix and 2 μM Illumina backbone oligonucleotides in 4 to 6 separate 50 μL reactions. The reactions were then pooled and processed with the QIAquick PCR Purification Kit (Qiagen). Multiplexed libraries were sequenced using 2×100 bp pared-end runs on an Illumina HiSeq 2000.

[0821] Mapping and Quality Control of NGS Data.

[0822] Paired-end reads were mapped to the hg19 reference genome with BWA 0.6.2 (default parameters), and sorted/indexed with SAMtools. QC was assessed using a custom Perl script to collect a variety of statistics, including mapping characteristics, read quality, and selector on-target rate (e.g., number of unique reads that intersect the selector space divided by all aligned reads), generated respectively by SAMtools flagstat, FastQC, and BEDTools coverageBed, modified to count each read at most once. Plots of fragment length distribution and sequence depth/coverage were automatically generated for visual QC assessment. To mitigate the impact of sequencing errors, analyses not involving fusions were restricted to properly paired reads, and only bases with a Phred quality score ≥30 (≤0.1% probability of a sequencing error) were further analyzed.

[0823] Analysis of Detection Thresholds by CAPP-Seq.

[0824] Two dilution series were performed to assess the linearity and accuracy of CAPP-Seq for quantitating tumor-derived cfDNA. In one experiment, shorn genomic DNA from a NSCLC cell line (HCC78) was spiked into cfDNA from a healthy individual, while in a second experiment, shorn genomic DNA from one NSCLC cell line (NCI-H3122) was spiked into shorn genomic DNA from a second NSCLC line (HCC78). A total of 32 ng DNA was used for library construction. Following mapping and quality control, homozygous reporters were identified as alleles unique to each sample with at least 20× sequencing depth and an allelic fraction >80%. Fourteen such reporters were identified between HCC78 genomic DNA and plasma cfDNA (FIG. **2***g-h*), whereas 24 reporters were found between NCI-H3122 and HCC78 genomic DNA (FIG. **16**).

[0825] Statistical Analysis.

[0826] The NSCLC selector was validated in silico using an independent cohort of lung adenocarcinomas (FIG. **1***c*). To assess statistical significance, we analyzed the same cohort

using 10,000 random selectors sampled from the exome, each with an identical size distribution to the CAPP-Seq NSCLC selector. The performance of random selectors had a normal distribution, and p-values were calculated accordingly. Note that all identified somatic lesions were considered in this analysis.

[0827] To evaluate the impact of reporter number on tumor burden estimates, we performed Monte Carlo sampling (1,000×), varying the number of reporters available {1, 2, . . . , max n} in two spiking experiments (FIG. **2***g-i*; FIG. **13***b-d*).

[0828] To assess the significance of tumor burden estimates in plasma cfDNA, we compared patient-specific SNV frequencies to the null distribution of selector-wide background alleles. Indels were separately analyzed using mutation-specific background rates and Z statistics. Fusion breakpoints were considered significant when present with >0 read support due to their ultra-low false detection rate. p-values from distinct reporter types were integrated into a single ctDNA detection index, and this was considered significant if the metric was ≤0.05 (≈FPR≤5%), the threshold that maximized CAPP-Seq sensitivity and specificity in ROC analyses (determined by Euclidean distance to a perfect classifier; e.g., TPR=1 and FPR=0; FIG. **3**, FIG. **4**, Table 1, Table 4).

[0829] Related to FIG. **5**, the probability P of recovering at least 2 reads of a single mutant allele in plasma for a given depth and detection limit was modeled by a binomial distribution. Given P, the probability of detecting all identified tumor mutations in plasma (e.g., median of 4 for CAPP-Seq) was modeled by a geometric distribution. Estimates in FIG. **5***a* are based on 250 million 100 bp reads per lane (e.g., using an Illumina HiSeq 2000 platform). Moreover, an on-target rate of 60% was assumed for CAPP-Seq and WES (FIG. **5**).

[0830] Molecular Biology Methods

[0831] Cell Lines.

[0832] The lung adenocarcinoma cell lines NCI-H3122 and HCC78 were obtained from ATCC and DSMZ, respectively, and grown in RPMI 1640 with L-glutamine (Gibco) supplemented with 10% fetal bovine serum (Gembio) and 1% penicillin/streptomycin cocktail. Cells were maintained in mid-log-phase growth in a 37° C. incubator with 5% CO_2. Genomic DNA was purified from freshly harvested cells with the DNeasy Blood & Tissue Kit (Qiagen).

[0833] Pleural Fluid Processing and Flow Cytometry, and Cell Sorting.

[0834] Cells from pleural fluid from patients P9 and P6 were harvested by centrifugation at 300×g for 5 min at 4° C. and washed in FACS staining buffer (HBSS+2% heat-inactivated calf serum [HICS]). Red blood cells were lysed with ACK Lysing Buffer (Invitrogen), and clumps were removed by passing through a 100 μm nylon filter. Filtered cells were spun down and resuspended in staining buffer. While on ice, the cell suspension was blocked for 20 min with 10 μg/mL rat IgG and then stained for 20 min with APC-conjugated mouse anti-human EpCAM (BioLegend, clone 9C4), PerCP-Cy5.5-conjugated mouse anti-human CD45 (eBioscience, clone 2D1), and PerCP-eFluor710-conjugated mouse anti-human CD31 (eBioscience, clone WM59). After staining, cells were washed and resuspended with staining buffer containing 1 μg/mL DAPI, analyzed, and sorted with a FACSAria II cell sorter (BD Biosciences). Cell doublets and DAPI-positive cells were excluded from analysis and sorting. $CD31^{-}CD45^{-}$ $EpCAM^{+}$ cells were sorted into staining buffer, spun down, and flash frozen in liquid nitrogen. DNA was isolated with the QIAamp DNA Micro Kit (Qiagen).

[0835] Optimization of NGS Library Preparation from Low Input cfDNA.

[0836] Protocols for Illumina library construction were compared in a step-wise manner with the goal of (1) optimizing adapter ligation efficiency, (2) reducing the necessary number of PCR cycles following adapter ligation, (3) preserving the naturally occurring size distribution of cfDNA fragments, and (4) minimizing variability in depth of sequencing coverage across all captured genomic regions. Initial optimization was done with NEBNext DNA Library Prep Reagent Set for Illumina (New England BioLabs), which includes reagents for end-repair of the cfDNA fragments, A-tailing, adapter ligation, and amplification of ligated fragments with Phusion High-Fidelity PCR Master Mix. Input was 4 ng cfDNA (obtained from plasma of the same healthy volunteer) for all conditions. Relative allelic abundance in the constructed libraries was assessed by qPCR of 4 genomic loci (Roche NimbleGen: NSC-0237, NSC-0247, NSC-0268, and NSC-0272) and compared by the $2^{-\Delta Ct}$ method.

[0837] Ligations were performed at 20° C. for 15 min (as per the manufacturer's protocol), at 16° C. for 16 hours, or with temperature cycling for 16 hours as previously described. Ligation volumes were varied from the standard (50 μL) down to 10 μL while maintaining a constant concentration of DNA ligase, cfDNA fragments, and Illumina adapters. Subsequent optimizations incorporated ligation at 16° C. for 16 hours in 50 μL reaction volumes.

[0838] Next, we compared standard SPRI bead processing procedures, in which new AMPure XP beads are added after each enzymatic reaction and DNA is eluted from the beads for the next reaction, to with-bead protocol modifications as previously described[3]. We compared 2 concentrations of Illumina adapters in the ligation reaction: 12 nM (10-fold molar excess to cfDNA fragments) and 120 nM (100-fold molar excess).

[0839] Using the optimized library preparation procedures, we next compared the NEBNext DNA Library Prep Reagent Set (with Phusion DNA Polymerase) to the KAPA Library Preparation Kit (with KAPA HiFi DNA Polymerase). The KAPA Library Preparation Kit with our modifications was also compared to the NuGEN SP Ovation Ultralow Library System with automation on Mondrian SP Workstation.

[0840] Evaluation of Library Preparation Modifications on CAPP-Seq Performance.

[0841] We performed CAPP-Seq on 32 ng cfDNA using standard library preparation procedures with the NEBNext kit, or with optimized procedures using either the NEBNext kit or the KAPA Library Preparation Kit. In parallel we performed CAPP-Seq on 4 ng and 128 ng cfDNA using the KAPA kit with our optimized procedures. Indexed libraries were constructed, and hybrid selection was performed in multiplex. The post-capture multiplexed libraries were amplified with Illumina backbone primers for 14 cycles of PCR and then sequenced on a paired-end 100 bp lane of an Illumina HiSeq 2000.

[0842] We also evaluated CAPP-Seq on ultralow input following whole genome amplification (WGA). We used Seq-Plex DNA Amplification Kit (Sigma-Aldrich), which employs degenerate oligonucleotide primer PCR. Briefly, 1 ng cfDNA was amplified with real-time monitoring with SYBR Green I (Sigma-Aldrich) on a HT7900 Real Time PCR machine (Applied Biosystems). Amplification was terminated after 17 cycles yielding 2.8 μg DNA. The primer

removal step yielded ~600 ng DNA, and this entire amount was used for library preparation using the NEBNext kit with optimized procedures as described herein.

[0843] Validation of Variants Detected by CAPP-Seq.

[0844] All structural rearrangements and a subset of tumoral SNVs detected by CAPP-Seq were independently confirmed by qPCR and/or Sanger sequencing of amplified fragments. For HCC78, a 120 bp fragment containing the SLC34A2-ROS1 breakpoint was amplified from genomic DNA using the primers: 5'-AGACGGGAGAAAATAG-CACC-3' (SEQ ID NO: 23) and 5'-ACCAAGGGTTGCA-GAAATCC-3' (SEQ ID NO: 24). For NCI-H3122, a 143 bp fragment containing the EML4-ALK breakpoint was amplified using the primers: 5'-GAGATGGAGTTTCACTCTTGT-TGC-3' (SEQ ID NO: 25) and 5'-GAACCTTTCCATCAT-ACTTAGAAATAC-3' (SEQ ID NO: 26). 5 ng genomic DNA was used as template with 250 nM oligos and 1× Phusion PCR Master Mix (NEB) in 50 μL reactions. Products were resolved on 2.5% agarose gel and bands of the expected size were removed. The amplified DNA fragments were purified using the Qiaquick Gel Extraction Kit (Qiagen) and submitted for Sanger sequencing (Elim Biopharm). For P9, genomic DNA breakpoints were confirmed by qPCR using the primers: 5'-TCCATGGAAGCCAGAAC-3' (SEQ ID NO: 27) and 5'-ATGCTAAGATGTGTCTGTCA-3' (SEQ ID NO: 28) for EML4-ALK; 5'-CCTTAACACAGATGGCTCTTGATGC-3' (SEQ ID NO: 29) and 5'-TCCTCTTTCCACCTTG-GCTTTCC-3' (SEQ ID NO: 30) for ROS1-MKX; and 5'-GGTTCAGAACTACCAATAACAAG-3' (SEQ ID NO: 31) and 5'-ACCTGATGTGTGACCTGATTGATG-3' (SEQ ID NO: 32) for FYN-ROS1. For qPCR, 10 ng of pre-amplified genomic DNA was used as template with 250 nM oligos and 1× Power SyberGreen Master Mix in 10 μL reactions performed in triplicate on a HT7900 Real Time PCR machine (Applied Biosystems). Standard PCR thermal cycling parameters were used. Amplification of amplicons spanning all 3 breakpoints detected in P9 were confirmed in tumor genomic DNA as well as plasma cfDNA, and PBL genomic DNA was used as a negative control.

[0845] CAPP-Seq confirmed somatic tumor mutations (SNVs and rearrangements) that were detected by clinical assays as a part of standard clinical care (Tables 3, 20 and 21). Clinical mutation assays were performed on formalin-fixed paraffin-embedded tissues. SNVs were detected by the SNaPshot assay[4]. Rearrangements were detected by fluorescence in situ hybridization (FISH) using separation probes targeting the ALK locus (Abbott) or ROS1 locus (Cytocell).

[0846] Bioinformatics and Statistical Methods

[0847] CAPP-Seq Detection Threshold Metrics.

[0848] Selector base-level background. We assessed the base-level background distribution of the NSCLC selector (FIG. **2***d*) using all 40 plasma cfDNA samples collected from NSCLC and healthy individuals analyzed in this work (Table 2). Specifically, for each background base in selector positions having ≥500× overall sequencing depth, the outlier-corrected mean across all cfDNA samples was calculated. Although we tested dedicated outlier detection methods, such as iterative Grubbs' method and ROUT, our empirical analyses indicated that simple removal of the minimum and maximum values worked best. Importantly, to restrict our analysis to background bases, each patient sample was pre-filtered to remove germline, loss of heterozygosity (LOH), and/or somatic variant calls made by VarScan 2[6] (somatic p-value=0.01; otherwise, default parameters).

[0849] Significance of SNVs as reporters. To evaluate the significance of tumor-derived SNVs in plasma, we implemented a strategy that integrates cfDNA fractions across all somatic SNVs, performs a position-specific background adjustment, and evaluates statistical significance by Monte Carlo sampling of background alleles across the selector. We note that this approach differs fundamentally from previous methods, where mutations are interrogated individually. Unlike these methods, our strategy dampens the impact of stochastic noise and biological variables (e.g., mutations near the detection limit, or tumor evolution) on tumor burden quantitation, permitting a more robust statistical assessment. In particular, this allows CAPP-Seq to quantitate low levels of ctDNA with potentially high rates of allelic drop out.

[0850] For a given plasma cfDNA sample θ, we begin by adjusting the allelic fraction f for each of n SNVs from patient P in order to minimize the influence of selector technical/biological background on significance estimates. Specifically, for each allele, we perform the following simple operation, $f^*=\max\{0,f-(e-\mu)\}$, where f is the raw allelic fraction in plasma cfDNA, e is the position-specific error rate for the given allele across all cfDNA samples (see above), and μ denotes the mean selector-wide background rate (=0.006% in this study, see section B1.1 and FIG. **2***d*). In effect, this adjustment nudges the mean of all n SNVs closer to the global selector mean μ, mitigating the confounding impact of technical/biological background. Using Monte Carlo simulation, we compare the adjusted mean SNV fraction $F^*(=(\Sigma f^*)/n)$ against the null distribution of background alleles across the selector. Specifically, for each of i iterations (=10,000 in this work), n background alleles are randomly sampled from θ, after which their fractions are adjusted using the above formula and averaged. A SNV p-value for patient P is determined as the percentile of F* with respect to the null distribution of background alleles in θ. Thus, a panel of SNVs from patient P would be assigned a detection p-value of 0.04 if F* ranks in the 96th percentile of adjusted background alleles in θ. We note that background adjustment always improved CAPP-Seq specificity in our ROC analyses.

[0851] Significance of Indels as Reporters.

[0852] We implemented an approach based on population statistics to assess the significance of indels separately from SNVs. For each indel in patient P, we use the Z-test to compare its fraction in a given plasma cfDNA sample θ against its fraction in every cfDNA sample in our cohort (excluding cfDNA samples from the same patient P). To increase statistical robustness, each read strand (positive or negative orientation) is assessed separately, yielding two Z-scores for each indel. These are combined into a single Z-score by Stouffer's method, an unweighted approach for integrative Z statistics. Finally, if patient P has more than 1 indel, all indel-specific Z-scores are combined by Stouffer's method into a final Z statistic, which is trivially converted to a p-value.

[0853] Significance of Fusions as Reporters.

[0854] Given the exceedingly low false positive rate associated with the detection of the same NSCLC fusion breakpoint in independent libraries, the recovery of a tumor-derived genomic fusion in plasma cfDNA by CAPP-Seq was (arbitrarily) assigned a p-value of ~0.

[0855] Integration of Distinct Mutation Types to Estimate Significance of Tumor Burden Quantitation.

[0856] For each patient, we calculate a ctDNA detection index (akin to a false positive rate) based on p-value integration from his or her array of reporters (Table 1 and Table 19).

For cases where only a single reporter type is present in a patient's tumor, the corresponding p-value is used. If SNV and indel reporters are detected, and if each independently has a p-value <0.1, we combine their respective p-values by Fisher's method (Fisher, 1925), and the resulting p-value is used. Otherwise, given the prioritization of SNVs in the selector design, the SNV p-value is used. If a fusion breakpoint identified in a tumor sample (e.g., involving ROS1, ALK, or RET) is recovered in plasma cfDNA from the same patient, it trumps all other mutation types, and its p-value (~0) is used. If a fusion detected in the tumor is not found in corresponding plasma (potentially due to hybridization inefficiency; see section C4), the p-value for any remaining mutation type(s) is used. Importantly, as new patients are processed, we cross check reporter types across the growing sample database to improve specificity (described in section B1.6, below) and identify potential red flags.

[0857] Indel/Fusion Correction for Sensitivity and Specificity Assessment.

[0858] Related to FIG. **3**, after calculating a ctDNA detection index for every set of reporters across all cfDNA samples using the methods described herein, we applied an additional step to increase specificity. Namely, to exploit the lower technical background of indels and fusion breakpoints as compared to SNVs, we applied an "indel/fusion correction". Specifically, if indel/fusion reporters found in patient X's tumor could be uniquely detected in patient X's plasma cfDNA (e.g., not detected in any other patient or control cfDNA sample), then the ctDNA detection index corresponding to patient X was set to 1 (e.g., ctDNA not detectable) in every unmatched cfDNA sample. In other words, patient X's reporters would not be called a false positive in another patient. Although we have not yet encountered two patients with the same indel/fusion reporter(s), if this was the case, the correction would not be applied from one patient to the other.

[0859] To perform this correction in a blinded manner, as shown for FIG. **3** (panels a and b), we identified germline SNPs in each cfDNA and PBL sample, and assigned each cfDNA sample to the tumor/normal pair with highest SNP concordance (after un-blinding, all cfDNA samples were found to be correctly matched to their corresponding tumor/normal pairs). As shown in FIG. **19**, this correction consistently increased CAPP-Seq specificity. Germline SNPs were identified using VarScan 2, with a p-value threshold of 0.01, minimum sequence coverage of 100×, a minimum average quality score of 30 (Phred), and otherwise default parameters.

[0860] Sensitivity and Specificity Analysis.

[0861] We tested CAPP-Seq performance in a blinded fashion by masking all patient identifying information, including disease stage, cfDNA time point, treatment, etc. We then tested our detection metrics described herein for correctly calling tumor burden across the entire grid of de-identified plasma cfDNA samples (13 patient-specific sets of somatic reporters across 40 plasma samples, or 520 pairs). To calculate sensitivity and specificity, we "un-blinded" ourselves and grouped patient samples into cancer-positive (e.g. cancer was present in the patient's body), cancer-negative (e.g. patient was cured), or cancer-unknown (e.g. insufficient data to determine true classification) categories. We considered every time point of patients with radiographic evidence of recurrence and all stage IV patients as cancer-positive, regardless of clinical evaluation at the time point in question. The post-treatment time point of patient 13 (P13; stage IIB NSCLC) was considered cancer-unknown due to "No Evidence of Disease (NED)" status at last follow-up, nearly 2 years from their treatment (FIG. **4***e*). Patient 2 (P2; stage IIIB NSCLC), was classified as NED following complete surgical resections, and was also considered cancer-unknown. All post-treatment stage I NSCLC patient samples were conservatively considered "cancer unknown" rather than true negatives due to limited follow-up.

[0862] Analysis of Library Complexity

[0863] Library Complexity Estimation.

[0864] We estimated the number of haploid genome equivalents per library using 330 genome equivalents per ing of input DNA (Table 2), and calculated overall 'molecule recovery' as the median depth after duplicate removal divided by the smaller of (i) the median depth before duplicate removal and (ii) the estimated number of haploid genome equivalents. Molecule recovery at a given sequencing depth was estimated to be 38% for cfDNA, 37% for tumor DNA, and 48% for PBLs (highest DNA input mass among all samples).

[0865] In contrast to genomic DNA, plasma cfDNA is naturally fragmented and has a highly stereotyped size distribution related to nucleosome spacing, with a median length of ~170 bp and very low dispersion (FIG. **2***a*, Tables 3, 20 and 21). As such, we hypothesized that independent input molecules with identical start/end coordinates may inflate the duplication rate of cfDNA, leading to an underestimated molecule recovery rate.

[0866] We tested this hypothesis by analyzing heterozygous germline SNPs, reasoning that DNA fragments (e.g., paired end reads) with identical start/end coordinates and differing by a single a priori defined germline variant are more likely to represent independent starting molecules than technical artifacts (e.g., PCR duplicates). Heterozygous SNPs were identified in all ninety samples (Table 2) using VarScan 2 (as described herein), and filtered for variants with an allele frequency between 40% and 60% that are present in the Common SNPs subset of dbSNP (version 137.0). For each heterozygous common SNP, A/B, we counted all fragments with unique start/end coordinates that support A, B, or AB. Among molecules with a given A/B SNP, there is a 50% chance of getting A and B together when randomly sampling two molecules (AB or BA), and there is a combined 50% chance of getting either AA or BB. Since the number of unique start/end positions for AB (denoted N) represents at least twice as many molecules (≥2N), and a combined ≥2N molecules can be assumed missing from unique start/end coordinates that support A or B, a lower bound on total missing library complexity is determined by the formula, 3N/S, where S denotes the sum of unique start/end coordinates covering A, B, and AB. Across SNPs in each input sample, we calculated an average of 30% missing library complexity in cfDNA samples, and 4% and 6% missing library complexity in tumor and PBL genomic DNA, respectively (FIG. **13***a*). Molecule recovery rates adjusted for estimated loss of complexity are provided in Table 2, and indicate a mean molecule recovery of at least 49% in cfDNA, 37% in tumor genomic DNA (mostly FFPE) and 51% in PBL genomic DNA.

[0867] Duplication Rate.

[0868] Common deduping tools, such as SAMtools rmdup and Picard tools MarkDuplicates (http://picard.sourceforge.net), identify and/or collapse reads based on sequence coordinates and quality, not sequence composition. This can result in the removal of tumor-derived reads (representing distinct molecules) that happen to share sequence coordinates with

germline reads. This is particularly problematic for cfDNA since for a large fraction of molecules there are other unique molecules with the same start and end (see above). To address this issue, we developed a custom Perl script that ignores bases with low quality (here, Phred Q<30), and collapses only those fragments (read pairs) with 100% sequence identity that also share genomic coordinates. The resulting post-duplicate reads are provided alongside corresponding non-deduped data in Tables 2 and 4, which respectively cover sequencing statistics and cfDNA monitoring results.

[0869] Library Complexity Measured Via PCR and Mass Input.

[0870] As a separate estimation of library complexity, for each Illumina NGS library constructed from cfDNA, we calculated the fraction of expected library yield from the actual yield and the expected (ideal) yield (FIG. **13***b*). The actual library yield was determined from the molarity and volume of the constructed libraries (prior to hybrid selection). The expected library yield was calculated from the mass of cfDNA used for library preparation and the number of PCR cycles performed, with the assumption that ligation was 100% efficient and PCR was 95% efficient at each cycle. A PCR efficiency of 95% was observed from qPCR performed on serial dilutions of Illumina TruSeq libraries (average of R^2>0.999 from 4 independent experiments).

[0871] CAPP-Seq Selector Design.

[0872] Most human cancers are relatively heterogeneous for somatic mutations in individual genes. Specifically, in most human tumors, recurrent somatic alterations of single genes account for a minority of patients, and only a minority of tumor types can be defined using a small number of recurrent mutations (<5-10) at predefined positions. Therefore, the design of the selector is vital to the CAPP-Seq method because (1) it dictates which mutations can be detected in with high probability for a patient with a given cancer, and (2) the selector size (in kb) directly impacts the cost and depth of sequence coverage. For example, the hybrid selection libraries available in current whole exome capture kits range from 51-71 Mb, providing ~40-60 fold maximum theoretical enrichment versus whole genome sequencing. The degree of potential enrichment is inversely proportional to the selector size such that for a ~100 kb selector, >10,000 fold enrichment should be achievable.

[0873] We employed a six-phase design strategy to identify and prioritize genomic regions for the CAPP-Seq NSCLC selector as detailed below. Three phases were used to incorporate known and suspected NSCLC driver genes, as well as genomic regions known to participate in clinically actionable fusions (phases 1, 5, 6), while another three phases employed an algorithmic approach to maximize both the number of patients covered and SNVs per patient (phases 2-4). The latter relied upon a metric that we termed "Recurrence Index" (RI), defined for this example as the number of NSCLC patients with SNVs that occur within a given kilobase of exonic sequence (e.g., No. of patients with mutations/exon length in kb). RI thus serves to measure patient-level recurrence frequency at the exon level, while simultaneously normalizing for gene/exon size. As a source of somatic mutation data uniformly genotyped across a large cohort of patients, in phases 2-4, we analyzed non-silent SNVs identified in TCGA whole exome sequencing data from 178 patients in the Lung Squamous Cell Carcinoma dataset (SCC) and from 229 patients in the Lung Adenocarcinoma (LUAD) datasets (TCGA query date was Mar. 13, 2012). Thresholds for each metric (e.g. RI and patients per exon) were selected to statistically enrich for known/suspected drivers in SCC and LUAD data (FIG. **7**). RefSeq exon coordinates (hg19) were obtained via the UCSC Table Browser (query date was Apr. 11, 2012).

[0874] The following algorithm was used to design the CAPP-Seq selector (parenthetical descriptions match design phases noted in FIG. **1***b*).

[0875] Phase 1 (Known Drivers)

Initial seed genes were chosen based on their frequency of mutation in NSCLCs. Analysis of COSMIC (v57) identified known driver genes that are recurrently mutated in ≥9% of NSCLC (denominator ≥500 cases). Specific exons from these genes were selected based on the pattern of SNVs previously identified in NSCLC. The seed list also included single exons from genes with recurrent mutations that occurred at low frequency but had strong evidence for being driver mutations, such as BRAF exon 15, which harbors V600E mutations in <2% of NSCLC.

[0876] Phase 2 (Max. Coverage)

For each exon with SNVs covering ≥5 patients in LUAD and SCC, we selected the exon with highest RI that identified at least 1 new patient when compared to the prior phase. Among exons with equally high RI, we added the exon with minimum overlap among patients already captured by the selector. This was repeated until no further exons met these criteria.

[0877] Phase 3 (RI≥30)

For each remaining exon with an RI≥30 and with SNVs covering ≥3 patients in LUAD and SCC, we identified the exon that would result in the largest reduction in patients with only 1 SNV. To break ties among equally best exons, the exon with highest RI was chosen. This was repeated until no additional exons satisfied these criteria.

[0878] Phase 4 (RI≥20)

Same procedure as phase 3, but using RI≥20.

[0879] Phase 5 (Predicted Drivers)

We included all exons from additional genes previously predicted to harbor driver mutations in NSCLC.

[0880] Phase 6 (Add Fusions)

For recurrent rearrangements in NSCLC involving the receptor tyrosine kinases ALK, ROS1, and RET, the introns most frequently implicated in the fusion event and the flanking exons were included.

[0881] All exons included in the selector, along with their corresponding HUGO gene symbols and genomic coordinates, as well as patient statistics for NSCLC and a variety of other cancers, are provided in Table 1, organized by selector design phase.

[0882] CAPP-Seq Computational Pipeline

[0883] Mutation Discovery: SNVs/Indels.

[0884] For detection of somatic SNV and insertion/deletion events, we employed VarScan 2 (somatic p-value=0.01, minimum variant frequency=5%, strand filter=true, and otherwise default parameters). Somatic variant calls (SNV or indel) present at less than 0.5% mutant allelic frequency in the paired normal sample (PBLs), but in a position with at least 1000× overall depth in PBLs and 100× depth in the tumor, and with at least 1× read depth on each strand, were retained (Tables 3, 20 and 21). While the selector was designed to predominantly capture exons, in practice, it also captures limited sequence content flanking each targeted region. For instance, this phenomenon is the basis for the (thus far) uniformly successful recovery by CAPP-Seq of fusion partners (which are not included within the selector) for kinase genes such as ALK and ROS1 recurrently rearranged in NSCLC. As

such, we also considered variant calls detected within 500 bps of defined selector coordinates. These calls were eliminated if present in non-coding repeat regions, since repeats may confound mapping accuracy. Repeat sequence coordinates were obtained using the RepeatMasker track in the UCSC table browser (hg19). Given a low, but measurable cross-contamination rate of ~0.06% in multiplexed cfDNA samples, (FIG. **14**) we also excluded any SNVs found as germline SNPs in samples from the same lane. Additionally, we excluded SNVs in the top 99.9^{th} percentile of global selector background (>0.27% sample-wide background rate; see FIG. **2***d* and section B1.1 above). Finally, we excluded any SNVs not present at a depth of at least 500× in at least 1 cfDNA sample. Variant annotation was automatically downloaded from the SeattleSeq Annotation 137 web server. Complete details for all identified SNVs and indels are provided in Tables 3, 20 and 21. Of note, all depth thresholds refer to pre-duplication removal reads.

[0885] Mutation Discovery: Fusions.

[0886] For practical and robust de novo enumeration of genomic fusion events and breakpoints from paired-end next-generation sequencing data, we developed a novel heuristic approach, termed FACTERA (FACile Translocation Enumeration and Recovery Algorithm). FACTERA has minimal external dependencies, works directly on a preexisting .bam alignment file, and produces easily interpretable output. Major steps of the algorithm are summarized below, and are complemented by a graphical schematic to illustrate key elements of the breakpoint identification process (FIG. **8**). FACTERA is coded in Perl and freely available upon request.

[0887] As input, FACTERA requires a .bam alignment file of paired-end reads produced by BWA, exon coordinates in .bed format (e.g., hg19 RefSeq coordinates), and a 0.2 bit reference genome to enable fast sequence retrieval (e.g., hg19). In addition, the analysis can be optionally restricted to reads that overlap particular genomic regions (.bed file), such as the CAPP-Seq selector used in this work.

[0888] FACTERA processes the input in three sequential phases: identification of discordant reads, detection of breakpoints at base pair-resolution, and in silico validation of candidate fusions. Each phase is described in detail below.

[0889] Identification of Discordant Reads.

[0890] To iteratively reduce the sequence space for gene fusion identification, FACTERA, like other algorithms (e.g. BreakDancer), identifies and classifies discordant read pairs. Such reads indicate a nearby fusion event since they either map to different chromosomes or are separated by an unexpectedly large insert size (e.g. total fragment length), as determined by the BWA mapping algorithm. The bitwise flag accompanying each aligned read encodes a variety of mapping characteristics (e.g., improperly paired, unmapped, wrong orientation, etc.) and is leveraged to rapidly filter the input for discordant pairs. The closest exon of each discordant read is subsequently identified, and used to cluster discordant pairs into distinct gene-gene groups, yielding a list of genomic regions R adjacent to candidate fusion sites. For each member gene of a discordant gene pair, the genomic region R_i is defined by taking the minimum of all 3' exon/read coordinates in the cluster, and the maximum of all 5' exon/read coordinates in the cluster. These regions are used to prioritize the search for breakpoints in the next phase (FIG. **8***a*).

[0891] Detection of Breakpoints at Base Pair-Resolution.

[0892] Discordant read pairs may be introduced by NGS library preparation and/or sequencing artifacts (e.g., jumping PCR). However, they are also likely to flank the breakpoints of bona fide fusion events. As such, all discordant gene pairs identified in the preceding phase are ranked in decreasing order of discordant read depth (duplicate fragments are eliminated to correct for possible PCR bias), and genomic regions with a depth of at least 2× (by default) are further evaluated for potential breakpoints. Within each region, FACTERA analyzes all properly paired reads in which one of the two reads is "soft-clipped," or truncated (see FIG. **8***a*). Soft-clipped reads allow for precise breakpoint determination, and are easily identified by parsing the CIGAR string associated with each mapped read, which compactly specifies the alignment operation used on each base (e.g. My=y contiguous bases were mapped, Sx=x bases were skipped). To simplify this step, only soft-clipped reads with the following two patterns are considered, SxMy and MySx, and the number of skipped bases x is required to be at least 16 (≤1 in 4.3B by random chance) to reduce the impact of non-specific sequence alignments.

[0893] To validate potential genomic breakpoints, defined as the edges of soft-clipped reads, FACTERA executes the following routine, depicted in FIG. **8**. For each discordant gene pair (e.g. genes w and v in FIG. **8***a*), all candidate breakpoints are tabulated, and the support (e.g. read frequency) for each is determined Breakpoints supported by less than 2 reads (by default) are excluded from further analysis. Starting with the two breakpoints with highest support, FACTERA selects a representative soft-clipped read for each breakpoint, such that the length of the clipped sequence is closest to half of the read length (FIG. **8***b*). If the mapped region of one read matches the soft-clipped region of the other, FACTERA records a putative fusion event. To assess inter-read concordance (e.g. see reads 1 and 2 in FIG. **8***c*), FACTERA employs the following algorithm. The mapped region of read 1 is parsed into all possible subsequences of length k (e.g., k-mers) using a sliding window (k=10, by default). Each k-mer, along with its lowest sequence index in read 1, is stored in a hash table data structure, allowing k-mer membership to be assessed in constant time (FIG. **8***c*, left panel). Subsequently, the soft clipped sequence of read 2 is parsed into subsequences of length k, and the hash table is interrogated for matching k-mers (FIG. **8***c*, right panel). If a minimum matching threshold is achieved (=0.5×the minimum length of the two compared subsequences), then the two reads are considered concordant. FACTERA will process at most 1000 (by default) putative breakpoint pairs for each discordant gene pair. Moreover, for each gene pair, FACTERA will only compare reads whose orientations are compatible with valid fusions. Such reads have soft-clipped sequences facing opposite directions (FIG. **8***d*, top panel). When this condition is not satisfied, FACTERA uses the reverse complement of read 1 for k-mer analysis (FIG. **8***d*, bottom panel).

[0894] In some instances, genomic subsequences flanking the true breakpoint may be nearly or completely identical, causing the aligned portions of soft-clipped reads to overlap. Unfortunately, this prevents an unambiguous determination of the breakpoint. As such, FACTERA incorporates a simple algorithm to arbitrarily adjust the breakpoint in one read (e.g., read 2) to match the other (e.g., read 1). Depending upon read orientation, there are two ways this can occur, both of which are illustrated in FIG. **8***e*. For each read, FACTERA calculates

the distance between the breakpoint and the read coordinate corresponding to the first k-mer match between reads. For example, as anecdotally illustrated in FIG. **8***e*, x is defined as the distance between the breakpoint coordinate of read 1 and the index of the first matching k-mer, j, whereas y denotes the corresponding distance for read 2. The offset is estimated as the difference in distances (x, y) between the two reads (see FIG. **8***e*).

[0895] In Silico Validation of Candidate Fusions.

[0896] To confirm each candidate breakpoint in silico, FACTERA performs a local realignment of reads against a template fusion sequence (±500 bp around the putative breakpoint) extracted from the 0.2 bit reference genome. BLAST is currently employed for this purpose, although BLAT or other fast aligners could be substituted. A BLAST database is constructed by collecting all reads that map to each candidate fusion sequence, including discordant reads and soft-clipped reads, as well as all unmapped reads in the original input .bam file. All reads that map to a given fusion candidate with at least 95% identity and a minimum length of 90% of the input read length (by default) are retained, and reads that span or flank the breakpoint are counted. As a final step, output redundancies are minimized by removing fusion sequences within a 20 bp interval of any fusion sequence with greater read support and with the same sequence orientation (to avoid removing reciprocal fusions).

[0897] FACTERA produces a simple output text file, which includes for each fusion sequence, the gene pair, the chromosomal sequence coordinates of the breakpoint, the fusion orientation (e.g., forward-forward or forward-reverse), the genomic sequences within 50 bp of the breakpoint, and depth statistics for reads spanning and flanking the breakpoint. Fusions identified in patients analyzed in this work are provided in Tables 3, 20 and 21.

[0898] Experimental Validation of FACTERA.

[0899] To experimentally evaluate the performance of FACTERA, we generated NGS data from two NSCLC cell lines, HCC78 (21.5M×100 bp paired-end reads) and NCI-H3122 (19.4M×100 bp paired-end reads), each of which has a known rearrangement (ROS1 and ALK, respectively) with a breakpoint that has, to the best of our knowledge, not been previously published. FACTERA readily revealed evidence for a reciprocal SLC34A2-ROS1 translocation in the former and an EML4-ALK fusion in the latter. Precise breakpoints predicted by FACTERA were experimentally validated by PCR amplification and Sanger sequencing (FIG. **9**; see also Validation of Variants Detected by CAPP-Seq). Importantly, FACTERA completed each run in practical time (~90 sec), using only a single thread on a hexa-core 3.4 GHz Intel Xeon E5690 chip. These initial results illustrate the utility of FACTERA as part of the CAPP-Seq analysis pipeline.

[0900] Templated Fusion Discovery.

[0901] We implemented a user-directed option to "hunt" for fusions within expected candidate genes. A fusion could be missed by FACTERA if the fusion detection criteria employed by FACTERA are incompletely satisfied—such as if discordant reads, but not soft-clipped reads, are identified—and will most likely occur when fusion allele frequency in the tumor is extremely low. As input, the method is supplied with candidate fusion gene sequences as "baits". All unmapped and soft-clipped reads in the input .bam file are subsequently aligned to these templates (using blastn) to identify reads that have sufficient similarity to both (for each read, 95% identity, e-value<1.0e-5, and at least 30% of the read length must map to the template, by default). Such reads are output as a list to the user for manual analysis.

[0902] We tested this simple approach on a low purity tumor sample found to harbor an ALK fusion by FISH, but not FACTERA (e.g., case P9). Using templates for ALK and its common fusion partner, ELM4, we identified 4 reads that mapped to both, in a region with an overall depth of ~1900×. The estimated allele frequency of 0.21% is strikingly similar to the 0.22% tumor purity measured by FACS (FIG. **17**), confirming the utility of the templated fusion discovery method. We subsequently FACS-depleted CD45+ immune populations and re-sequenced this patient's tumor. In the enriched tumor sample, FACTERA identified the EML4-ALK fusion, along with two novel ROS1 fusions (FIG. **4***b*, Tables 3, 20 and 21).

[0903] Mutation Recovery:SNVs/Indels.

[0904] Using a custom Perl script, previously identified reporter alleles were intersected with a SAMtools mpileup file generated for each plasma cfDNA sample, and the number and frequency of supporting reads was calculated for each reporter allele. Only reporters in properly paired reads at positions with at least 500× overall depth (pre-duplication removal) were considered (Table 4).

[0905] Mutation Recovery: Fusions.

[0906] For enumeration of fusion frequency in sequenced plasma DNA, FACTERA executes the last step of the discovery phase (e.g., in silico validation of candidate fusions, above) using the set of previously identified fusion templates. The fusion allele frequency is calculated as α/β, where α is the number of breakpoint-spanning reads, and β is the mean overall depth within a genomic region ±5 bps around the breakpoint. Regarding the NSCLC selector described in this work, the latter calculation was always performed on the single gene contained in the NSCLC selector library. If both fusion genes are targeted within a selector library, overall depth is estimated by taking the mean depth calculated for both genes.

[0907] Notably, in some cases we observed lower fusion allele frequencies than would be expected for heterozygous alleles (e.g., see cell line fusions in Tables 3, 20 and 21). This was seen in cell lines, in an empirical spiking experiment, and in one patient's tumor and plasma samples (e.g., P6), and could potentially result from inefficient "pull-down" of fusions whose partners are not represented in the selector. Regardless, fusions are useful reporters—they possess virtually no background signal and show linear behavior over defined concentrations in a spiking experiment (FIG. **16***d*). Moreover, allelic frequencies in plasma are easily adjusted for such inefficiencies by dividing the measured frequency in plasma by the corresponding frequency in the tumor. In cases where sequenced tumor tissue is impure, tumor content can be estimated using the frequencies of SNVs (or indels) as a reference frame, allowing the fusion fraction to be normalized accordingly (Table 4).

[0908] Screening Plasma cfDNA without Knowledge of Tumor DNA.

[0909] We devised the following statistical algorithm as an initial step toward non-invasive tumor genotyping and cancer screening with CAPP-Seq. The method identifies candidate SNVs using iterative models of (i) background noise in paired germline DNA (in this work, PBLs), (ii) base-pair resolution background frequencies in plasma cfDNA across the selector,

and (iii) sequencing error in cfDNA. Examples are provided in FIG. **21**. The algorithm works in four main steps, detailed below.

[0910] As input, the algorithm takes allele frequencies from a single plasma cfDNA sample and analyzes high quality background alleles, defined in a first step for each genomic position as the non-dominant base with highest fractional abundance. Only alleles with depth of at least 500× and strand bias <90% (conservative, by default) are analyzed. For consistency with variant calling, we allowed the screening approach to interrogate selector regions within 500 bp of defined coordinates, expanding the effective sequence space from ~125 kb to ~600 kb.

[0911] Second, the binomial distribution is used to test whether a given input cfDNA allele is significantly different from the corresponding paired germline allele (FIG. **21***a-b*). Here the probability of success is taken to be the frequency of the background allele in PBLs, and the number of trials is the allele's corresponding depth in plasma cfDNA. To avoid contributions from alleles in rare circulating tumor cells that might contaminate PBLs, input alleles with a fractional abundance greater than 0.5% in paired PBLs (by default) or a Bonferroni-adjusted binomial probability greater than 2.08×10^{-8} are not further considered (alpha of 0.05/[~600 kb*4 alleles per position]).

[0912] Third, a database of cfDNA background allele frequencies is assembled. Here, we used samples analyzed in the present study (e.g., pre-treatment NSCLC samples and 1 sample from a healthy volunteer), except the input sample is left out to avoid bias. Based on the assumption that all background allele fractions follow a normal distribution, a Z-test is employed to test whether a given input allele differs significantly from typical cfDNA background at the same position (FIG. **21***a-b*). All alleles within the selector are evaluated, and those with an average background frequency of 5% or greater (by default) or a Bonferroni-adjusted single-tailed Z-score <5.6 are not further considered (alpha of 0.05, adjusted as above).

[0913] Finally, candidate alleles are tested for remaining possible sequencing errors. This step leverages the observation that non-tumor variants (e.g., "errors") in plasma cfDNA tend to have a higher duplication rate than bona fide variants detectable in the patient's tumor (data not shown). As such, the number of supporting reads is compared for each input allele between nondeduped (all fragments meeting QC criteria) and deduped data (only unique fragments meeting QC criteria). An outlier analysis is then used to distinguish candidate tumor-derived SNVs from remaining background noise (FIG. **21***a-c*). Specifically, to reveal outlier tendency in the data, the square root of the robust distance Rd (Mahalanobis distance) is compared against the square root of the quantiles of a chi-squared distribution Cs. This transformation reveals natural separation between true SNVs and false positives in cancer patients (FIG. **21***a, c*), and notably, reveals an absence of outlier structure in patient samples lacking tumor-derived SNVs (FIG. **21***b, c*). To automatically call SNVs without prior knowledge, the screening approach iterates through data points by decreasing Rb and recalculating the Pearson's correlation coefficient Rho between Rd and Cs for points 1 to i, where Rd_i is the current maximum Rd. The algorithm iteratively reports outliers (e.g., candidate SNVs) until it terminates when Rho≥0.85

Example 2

Designing a Personalized Selector Set

[0914] In certain circumstances, monitoring tumor burden in a patient known to have cancer is likely to be impractical using an 'off-the-shelf' strategy applying knowledge from a cohort of patients with the same tumor type, to selectively capture genomic regions that are recurrently mutated in that tumor type using CAPP-Seq. These situations include, but are not limited to, cases where (1) the tumor is of an unknown primary histology (e.g., CUP); (2) the histology is known, but is too rare to have a sufficient number of patients with that tumor type previously profiled to define the average patient's tumor somatic genetic landscape (e.g., soft tissue sarcoma subtyped); (3) the histology is known but the average/median number of recurrent somatic lesions in that tumor type are too low to achieve desired sensitivity levels (e.g., pediatric tumors, etc.); or (4) the histology is known and the average/median number of recurrent somatic lesions is reasonable, but the average burden of tumor volume is so small that additional sensitivity can be achieved using more mutations per tumor (e.g., early stages of malignant melanoma). In such cases, a personalized strategy for monitoring tumor burden is likely to overcome these hurdles for disease monitoring.

[0915] Here, tumor(s) from a patient known to have cancer are genotyped by profiling the tumor genome, exome, or targeted region expected to be enriched for somatic aberrations. The genotype of the cancer may be compared to a genotype of the germline of the same patient. The resulting lesions are then catalogued and used to build a custom, personalized selector comprising a set of biotinylated oligonucleotides for selective hybrid affinity capture of corresponding circulating tumor DNA (ctDNA) molecules. Cell-free DNA circulating in blood or body fluids and harboring such ctDNA molecules would be isolated, and used to build shotgun genomic libraries that include ligation of molecular tags ('barcodes') that distinguish such sequences from others, allowing for suppression of spurious errors introduced during the amplification of cfDNA using thermostable DNA polymerases as part of polymerase chain reaction. The personalized selector would then be applied for capture of the fragments of interest, sequenced and analyzed in the same manner as the 'off-the-shelf' CAPP-Seq workflow, allowing the tracking and quantitation of those mutations originally discovered in the primary tumor within the corresponding cfDNA. As an alternative to affinity based hybrid capture of ctDNA/cfDNA, amplicons specific to the corresponding region could be interrogated by PCR, with such fragments selectively indexed using molecular barcodes that similarly allow distinction of sequencing errors introduced during PCR.

Example 3

Use of a Selector Set to Diagnose a Cancer

[0916] A plasma sample is obtained from a female subject with an abnormal lump in her breast. Cell-free DNA (cfDNA) is extracted from the plasma sample. An end repair reaction is performed on the cfDNA by mixing the components in a sterile microfuge tube (or other suitable sterile container) as follows:

Component	Volume (μL)
cfDNA	1-75
Phosphorylation Reaction Buffer (10X)	10
T4 DNA polymerase	5
T4 Polynucleotide kinase	5
dNTPs	4
DNA Polymerase I, Large (Klenow)	1
Sterile H_2O	-bring total volume up to 100 μL

[0917] The end repair reaction mixture is incubated in a thermal cycler for 30 minutes at 20° C.

[0918] Clean-up of the end repaired cfDNA is performed by adding 160 μL (1.6×) of resuspended AMPure XP beads to the end repair reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is incubated for 5 minutes at room temperature. The reaction is placed on a magnetic stand to separate the beads from the supernatant. After the solution is clear (approximately 5 minutes), the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 10 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by adding 40 μL of sterile water and vortexing or pipetting the water up and down. The reaction is placed back on the magnetic stand. Once the solution is clear, 32 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0919] dA-tailing of the end repaired cfDNA is performed by mixing the following components in the sterile microfuge tube as follows:

Component	Volume (μL)
End repaired cfDNA	32
NEBuffer 2 (10X)	5
Deoxyadenosine 5'-Triphosphate	10
Klenow Fragment (3'→5' exo-)	3

[0920] The dA-tailing reaction is incubated in a thermal cycle for 30 minutes at 37° C.

[0921] Clean-up of the dA-tailed cfDNA is performed by adding 90 μL (1.8×) of resuspended AMPure XP beads to the dA-tailing reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is incubated for 5 minutes at room temperature. The reaction is placed on a magnetic stand to separate the beads from the supernatant. After the solution is clear (approximately 5 minutes), the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 10 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by adding 15 μL of sterile water and vortexing or pipetting the water up and down. The reaction is placed back on the magnetic stand. Once the solution is clear, 10 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0922] Adaptor ligation of the dA-tailed cfDNA is performed by mixing the following components in the sterile microfuge tube as follows:

Component	Volume (μL)
dA-tailed cfDNA	10
Quick Ligation Reaction Buffer (2X)	25
Illumina Adaptor	10
Quick T4 DNA Ligase	5

[0923] The adaptor ligation reaction is incubated at 16° C. for 16 hours. The adaptor ligation reaction is terminated by adding 3 μL of USER™ enzyme mix by pipetting up and down and incubation at 37° C.

[0924] Clean-up of the adaptor-ligated cfDNA is performed by adding 90 μL (1.8×) of resuspended AMPure XP beads to the adaptor ligation reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is incubated for 5 minutes at room temperature. The reaction is placed on a magnetic stand to separate the beads from the supernatant. After the solution is clear (approximately 5 minutes), the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 10 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by adding 105 μL of sterile water and vortexing or pipetting the water up and down. The reaction is placed back on the magnetic stand. Once the solution is clear, 100 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0925] Universal PCR amplification is performed on the adaptor-ligated cfDNA using primers targeting the adaptors. The PCR amplification is conducted using 14 amplification cycles. Selector set probes are used to selectively capture a subset of the amplified products of the adaptor ligated cfDNA. Sequencing reactions are performed on the captured amplified products. The captured amplified cfDNA is sequenced on a paired-end 100 bp lane of an Illumina HiSeq 2000.

[0926] The sequencing information is analyzed by detecting mutations in one or more genomic regions based on a selector set. The selector set contains information pertaining to mutations occurring in one or more genomic regions, wherein the mutations are present in at least about 70% of a population of subjects suffering from a breast cancer. In order to determine the statistical significance of the mutations detected in the sample, p-values for the different classes of mutations are calculated. A ctDNA detection index is used to evaluate the statistical significance of detecting two or more classes of mutations.

[0927] A report of the mutations detected in the sample and the statistical significance of the detection of the mutations is provided to a physician. Based on the detection of at least three mutations in three genomic regions, the physician diagnoses a breast cancer in the subject.

Example 4

Use of a Selector Set to Determine a Status or Outcome of a Cancer

[0928] Cell-free DNA (cfDNA) is purified from a sample from a subject diagnosed with a prostate cancer. An end repair

reaction is performed on the cfDNA by mixing the components in a sterile microfuge tube (or other suitable sterile container) as follows:

Component	Volume (μL)
1-5 μg cfDNA	1-85
10X End Repair Buffer	10
End Repair Enzyme Mix	5
Sterile H_2O	-bring total volume up to 100 μL

[0929] The end repair reaction mixture is incubated in a thermal cycler for 30 minutes at 20° C.

[0930] Clean-up of the end repaired cfDNA is performed by adding 160 μL (1.6×) of resuspended AMPure XP beads to the end repair reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 15 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by resuspending the beads thoroughly in 32.5 μL of elution buffer and incubating at room temperature for 2 minutes. The reaction is placed back on the magnetic stand at room temperature for 15 minutes or until the solution is clear. 30 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0931] dA-tailing of the end repaired cfDNA is performed by mixing the following components in the sterile microfuge tube as follows:

Component	Volume (μL)
End repaired cfDNA	30
10X A-tailing buffer	5
A-tailing enzyme	3
Sterile water	12

[0932] The dA-tailing reaction is incubated in a thermal cycle for 30 minutes at 30° C.

[0933] Clean-up of the dA-tailed cfDNA is performed by adding 90 μL (1.8×) of resuspended AMPure XP beads to the dA-tailing reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 15 minutes or until the reaction is clear. After the solution is clear (approximately 5 minutes), the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by resuspending the beads thoroughly in 32.5 μL of elution buffer and incubating at room temperature for 2 minutes. The reaction is placed back on the magnetic stand for 15 minutes at room temperature or until the solution is clear. 30 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0934] Adaptor ligation of the dA-tailed cfDNA is performed by mixing the following components in the sterile microfuge tube as follows:

Component	Volume (μL)
dA-tailed cfDNA	30
5X Ligation Buffer	10
Illumina Adaptor	5
DNA Ligase	5

[0935] The adaptor ligation reaction is incubated at 16° C. for 16 hours.

[0936] Clean-up of the adaptor-ligated cfDNA is performed by adding 50 μL of resuspended AMPure XP beads to the adaptor ligation reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 15 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. The beads are resuspended in 52.5 μL of elution buffer. The reaction is placed back on the magnetic stand and incubated at room temperature for 15 minutes or until the solution is clear. 50 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0937] A second clean-up of the adaptor-ligated cfDNA is performed by adding 50 μL of resuspended AMPure XP beads to the adaptor ligation reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 15 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. The beads are resuspended in 32.5 μL of elution buffer and incubated at room temperature for 2 minutes. The reaction is placed back on the magnetic stand and incubated at room temperature for 15 minutes or until the solution is clear. 30 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0938] Universal PCR amplification is performed on the adaptor-ligated cfDNA using primers targeting the adaptors. The PCR amplification is conducted using 16 amplification cycles. Selector set probes are used to selectively capture a subset of the amplified adaptor ligated cfDNA. The amplified cfDNA is sequenced on a paired-end 100 bp lane of an Illumina HiSeq 2000.

[0939] The sequencing information is analyzed by detecting mutations in one or more genomic regions based on a selector set. The selector set contains information pertaining to mutations occurring in one or more genomic regions, wherein the mutations are present in at least about 70% of a population of subjects suffering from a breast cancer. A quantity of circulating tumor-DNA (ctDNA) is determined based on the sequencing reads.

[0940] A report comprising the quantity of the ctDNA is provided to a physician. Based on the quantity of the ctDNA, the physician provides a prognosis of the prostate cancer in the subject.

Example 5

Use of a Selector Set to Determine a Therapeutic Regimen for the Treatment of a Cancer

[0941] Cell-free DNA (cfDNA) is purified from a sample from a subject diagnosed with a thyroid cancer. An end repair reaction is performed on the cfDNA by mixing the components in a sterile microfuge tube (or other suitable sterile container) as follows:

Component	Volume (μL)
1-5 μg cfDNA	1-85
10X End Repair Buffer	10
End Repair Enzyme Mix	5
Sterile H_2O	-bring total volume up to 100 μL

[0942] The end repair reaction mixture is incubated in a thermal cycler for 30 minutes at 20° C.

[0943] Clean-up of the end repaired cfDNA is performed by adding 160 μL (1.6×) of resuspended AMPure XP beads to the end repair reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 15 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by resuspending the beads thoroughly in 32.5 μL of elution buffer and incubating at room temperature for 2 minutes. The reaction is placed back on the magnetic stand at room temperature for 15 minutes or until the solution is clear. 30 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0944] dA-tailing of the end repaired cfDNA is performed by mixing the following components in the sterile microfuge tube as follows:

Component	Volume (μL)
End repaired cfDNA	30
10X A-tailing buffer	5
A-tailing enzyme	3
Sterile water	12

[0945] The dA-tailing reaction is incubated in a thermal cycle for 30 minutes at 30° C.

[0946] Clean-up of the dA-tailed cfDNA is performed by adding 90 μL (1.8×) of resuspended AMPure XP beads to the dA-tailing reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 15 minutes or until the reaction is clear. After the solution is clear (approximately 5 minutes), the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. cfDNA is eluted from the beads by resuspending the beads thoroughly in 32.5 μL of elution buffer and incubating at room temperature for 2 minutes. The reaction is placed back on the magnetic stand for 15 minutes at room temperature or until the solution is clear. 30 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0947] Adaptor ligation of the dA-tailed cfDNA is performed by mixing the following components in the sterile microfuge tube as follows:

Component	Volume (μL)
dA-tailed cfDNA	30
5X Ligation Buffer	10
Adaptor	5
DNA Ligase	5

[0948] The adaptor ligation reaction is incubated at 16° C. for 16 hours. The concentration of the adaptor is increased through the duration of the incubation. The adaptor is a Y-shaped adaptor. The 5' strand of the split portion of the Y-shaped contains a molecular barcode and a sample index. The double stranded portion of the Y-shaped adaptor contains a universal sequence. The universal sequence is used for PCR enrichment and sequencing.

[0949] Clean-up of the adaptor-ligated cfDNA is performed by adding 50 μL of resuspended AMPure XP beads to the adaptor ligation reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 5 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 15 minutes while the reaction is on the magnetic stand. The beads are resuspended in 52.5 μL of elution buffer. The reaction is placed back on the magnetic stand and incubated at room temperature for 5 minutes or until the solution is clear. 50 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0950] A second clean-up of the adaptor-ligated cfDNA is performed by adding 50 μL of resuspended AMPure XP beads to the adaptor ligation reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 5 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 10 minutes while the reaction is on the magnetic stand. The beads are resuspended in 105 μL of elution buffer and incubated at room temperature for 2 minutes. The reaction is placed back on the magnetic stand and incubated at room temperature until the solution is clear. 100 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube).

[0951] Bead based size selection of the adaptor ligated cfDNA is performed by adding 80 μL of AMPure XP beads to the adaptor ligated cfDNA. The reaction is mixed by vortexing the reaction or pipetting the solution up and down at least 10 times. The reaction is incubated at room temperature for 5 minutes. The reaction is placed on a magnetic stand for 5 minutes or until the solution is clear. Once the solution is clear, the supernatant is transferred to a new tube. 20 μL of AMPure XP beads are added to the supernatant (vortex or pipet up and down to mix) and incubated at room temperature for 5 minutes. The reaction is placed on the magnetic stand for 5 minutes or until the solution is clear. Once the solution is clear, the supernatant is removed and discarded. While on the magnetic stand, the beads are washed twice using 200 μL of freshly prepared 80% ethanol. The ethanol washes are incubated at room temperature for 30 seconds and removed and discarded. The beads are air dried at room temperature for 10 minutes. cfDNA is eluted from the beads by resuspending the beads in 25 μL of sterile water or 0.1× TE Buffer. The reaction is placed back on the magnetic stand. Once the solution is clear, 20 μL of the supernatant is transferred to a new microfuge tube.

[0952] PCR enrichment of the adaptor ligated cfDNA is by mixing the following components:

Component	Volume (μL)
Adaptor ligated cfDNA	20
Universal PCR Primer (25 μM)	2.5
Index Primer (25 μM)	2.5
Phusion High-Fidelity PCR Master Mix	25

[0953] The PCR enrichment is performed using the cycling conditions of 1 cycle at 98° C. for 30 seconds, 17 cycles of 98° C. for 10 seconds, 65° C. for 30 seconds, and 72° C. for 30 seconds, followed by 1 cycle of 72° C. for 5 minutes and a hold at 4° C.

[0954] Clean-up of the PCR enriched cfDNA is performed by adding 50 μL (1×) of resuspended AMPure XP beads to the PCR enriched cfDNA reaction mixture. The AMPure beads are mixed into the solution on a vortex mixer or by pipetting up and down (e.g., 10 times or more). The reaction is placed on a magnetic stand and incubated at room temperature for 5 minutes or until the solution is clear. After the solution is clear, the supernatant is removed and discarded. The beads are washed twice by adding 200 μL of 80% freshly prepared ethanol to the reaction while in the magnetic stand. For each wash, the ethanol solution is added at room temperature for 30 seconds. The supernatant is removed and discarded. The beads are air dried for 10 minutes while the reaction is on the magnetic stand. The beads are resuspended in 30 μL of 0.1× TE. The reaction is placed back on the magnetic stand and incubated at room temperature for until the solution is clear. 25 μL of the supernatant is transferred to a fresh, sterile container (e.g., microfuge tube). The enriched cfDNA is diluted 20-fold with the addition of nuclease free water

[0955] The enriched cfDNA is hybridized to an array comprising selector set probes. The quantity of the circulating tumor DNA (ctDNA) is determined using array-based hybridization. An image of the array is obtained and the quantity of the ctDNA is calculated based on the intensity signals on the array.

[0956] A report comprising the quantity of the ctDNA, the mutations found, and a list of anti-cancer therapies is provided to a physician. Based on the quantity of the ctDNA, the types of mutations found, and the list of anti-cancer therapies, the physician provides a therapeutic regimen for treating of the thyroid cancer in the subject.

TABLE 1

									Fusion		Pre-treatment		
Case	Age	Sex	Histology	Stage	TNM	Smoking history	No. of SNVs (non-silent)	Indels	ALK/ ROS1	Partner	ctDNA (%)	ctDNA (pg/mL)	Tumor (cc)
P12	86	F	SCC	IA	T1bN0M0	Heavy	6 (3)	1			ND	ND	5.5
P1	66	M	Adeno	IB	T2aN0M0	Heavy	12 (3)	4			0.025	1.9	23.1
P16	82	F	Adeno	IB	T2aN0M0	Heavy	26 (5)	2			0.019	2.5	22.5
P17	85	F	Adeno	IB	T2aN0M0	Heavy	2 (2)	0			ND	ND	10.2
P13	90	F	SCC	IIB	T3aN0M0	Heavy	5 (4)	0			1.78	269.8	339.3
P2	61	M	Large Cell	IIIA	T3aN1M0	Heavy	12 (3)	1			0.896	64.7	23.1
P3	67	F	Adeno	IIIB	T1bN3M0	Light	1 (1)	0			0.095	16.2	7.9
P14	55	M	Adeno	IIIB	T1aN3M0	Heavy	8 (5)	0			0.05	10.2	5.2
P15	41	M	Adeno	IIIB	T3N3M0	Light	25 (10)	1			0.58	108.1	121.8
P4	47	F	Adeno	IV	T2aN2M1b	Heavy	3 (2)	0			0.039	2.1	12.4
P5	49	F	Adeno	IV	T1bN0M1a	None	4 (3)	0			3.2	143.8	82.1
P6	54	M	Adeno	IV	T3N2M1b	None	3 (2)	0	ALK	KIF5B	1.0	350.2	NA
P9	47	F	Adeno	IV	T4N3M1a	None	0	0	ALK ROS1	EML4 MKX, FYN	0.04	3.8	66.2
P10	35	F	Adeno	IIIA	T4N0M0	None	0	0	ROS1	SLC34 A2	—	—	—
P11	38	F	Adeno	IIIA	T3N2M0	None	2 (1)	0	ROS1	CD74	—	—	—
P7	50	M	Adeno	IV	T1aN2M1b	Light	0	0	ALK	EML4	—	—	—
P8	48	F	Adeno	IV	T4N0M1b	None	1 (0)	0	ALK	EML4	—	—	—

Patient characteristics and pre-treatment CAPP-Seq monitoring results.

ND, mutant DNA was not detected above background.

NA, tumor volume could not be reliably assessed.

Dashes, plasma sample not available.

Smoking history, ≥20 pack years (Heavy), >0 and <20 pack years (Light).

Additional details are provided in Tables 3, 4, 20 and 21.

TABLE 2

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
AKT1	chr14	105246424	105246553	7.7	0.25
BRAF	chr7	140453074	140453193	66.7	2.21
BRAF	chr7	140481375	140481493	58.8	3.93
CDKN2A	chr9	21970900	21971207	97.4	11.30
CDKN2A	chr9	21974475	21974826	19.9	13.02
CTNNB1	chr3	41266016	41266244	26.2	14.00
EGFR	chr7	55241613	55241736	24.2	14.25
EGFR	chr7	55242414	55242513	80.0	15.97
EGFR	chr7	55248985	55249171	26.7	16.95
EGFR	chr7	55259411	55259567	89.2	19.90
ERBB2	chr17	37880164	37880263	0.0	19.90
ERBB2	chr17	37880978	37881164	21.4	20.88
HRAS	chr11	533765	533944	16.7	21.38
HRAS	chr11	534211	534322	26.8	22.11
KEAP1	chr19	10599867	10600044	16.9	22.85
KEAP1	chr19	10600323	10600529	72.5	26.54
KEAP1	chr19	10602252	10602938	36.4	31.45
KEAP1	chr19	10610070	10610709	28.1	34.64
KEAP1	chr19	10597327	10597494	11.9	35.14
KRAS	chr12	25380167	25380346	22.2	36.12
KRAS	chr12	25398207	25398318	500.0	46.93
MEK1	chr15	66727364	66727575	0.0	46.93
MET	chr7	116411902	116412043	14.1	47.42
NFE2L2	chr2	178098732	178098999	115.7	52.09
NOTCH1	chr9	139396723	139396940	4.6	52.09
NOTCH1	chr9	139399124	139399556	0.0	52.09
NOTCH1	chr9	139390522	139392010	2.0	52.58
NOTCH1	chr9	139397633	139397782	0.0	52.58
NRAS	chr1	115256420	115256599	27.8	53.32
NRAS	chr1	115258670	115258781	0.0	53.32
PIK3CA	chr3	178935997	178936122	150.8	55.28
PIK3CA	chr3	178951881	178952152	14.7	56.02
PTEN	chr10	89624226	89624305	12.5	56.27
PTEN	chr10	89653781	89653866	0.0	56.27
PTEN	chr10	89685269	89685314	65.2	56.76
PTEN	chr10	89690802	89690846	0.0	56.76
PTEN	chr10	89692769	89693008	20.8	57.49
PTEN	chr10	89711874	89712016	21.0	57.74
PTEN	chr10	89717609	89717776	35.7	58.48
PTEN	chr10	89720650	89720875	13.3	58.72
STK11	chr19	1206912	1207202	13.7	58.97
STK11	chr19	1218415	1218499	23.5	59.21
STK11	chr19	1219322	1219412	11.0	59.46
STK11	chr19	1220371	1220504	29.9	59.46
STK11	chr19	1220579	1220716	29.0	59.46
STK11	chr19	1221211	1221339	31.0	59.46
STK11	chr19	1221947	1222005	0.0	59.46
STK11	chr19	1222983	1223171	0.0	59.46
STK11	chr19	1226452	1226646	0.0	59.46
TP53	chr17	7577018	7577155	405.8	64.86
TP53	chr17	7577498	7577608	450.5	70.27
TP53	chr17	7578176	7578289	342.1	73.71
TP53	chr17	7579311	7579590	110.7	76.66
TP53	chr17	7578370	7578554	367.6	83.54
REG1B	chr2	79313937	79314056	83.3	83.78
TPTE	chr21	10970008	10970062	72.7	84.28
CSMD3	chr8	113246593	113246706	70.2	84.77
TP53	chr17	7573926	7574033	83.3	85.50
FAM135B	chr8	139151228	139151339	71.4	86.00
U2AF1	chr21	44524424	44524512	56.2	86.24
THSD7A	chr7	11501637	11501770	67.2	86.49
MLL3	chr7	151962122	151962294	63.6	86.73
EYA4	chr6	133849862	133849943	61.0	86.98
HCN1	chr5	45267190	45267355	54.2	87.22
AKR1B10	chr7	134222945	134223029	58.8	87.71
SLC6A5	chr11	20668379	20668480	49.0	87.96
DPP10	chr2	116525872	116525980	55.0	88.45
SCN7A	chr2	167327124	167327216	43.0	88.70
SNTG1	chr8	51621445	51621538	53.2	88.94

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
VPS13A	chr9	79946925	79947029	47.6	89.19
IL1RAPL1	chrX	29938065	29938211	47.6	89.43
CTNNA2	chr2	80085138	80085305	47.6	89.68
CSMD3	chr8	113323206	113323395	47.4	89.93
FAM5C	chr1	190203501	190203607	46.7	90.17
CACNA1E	chr1	181708282	181708389	37.0	90.42
KRTAP5-5	chr11	1651070	1651784	43.4	91.15
PDE1C	chr7	31864480	31864601	41.0	91.40
RYR2	chr1	237806626	237806747	41.0	91.65
NRXN1	chr2	50733632	50733755	40.3	91.89
COL19A1	chr6	70637800	70637924	40.0	92.14
CSMD3	chr8	113697634	113697961	39.6	92.38
LRP1B	chr2	141665445	141665646	34.7	92.63
GKN2	chr2	69173435	69173592	38.0	92.87
CD5L	chr1	157805624	157805945	37.3	93.12
SPTA1	chr1	158627266	158627484	36.5	93.37
DHX9	chr1	182812428	182812569	35.2	93.61
ADAMTS20	chr12	43858393	43858535	35.0	93.86
NLRP4	chr19	56382192	56382363	34.9	93.86
CDH18	chr5	19473334	19473825	34.6	94.35
MYH2	chr17	10450791	10450935	34.5	94.84
OR5L2	chr11	55594694	55595630	32.0	94.84
OR4A15	chr11	55135359	55136394	30.9	94.84
OR6F1	chr1	247875130	247876057	28.0	94.84
OR4C6	chr11	55432642	55433572	29.0	95.09
OR2T4	chr1	248524882	248525929	31.5	95.09
FAM5C	chr1	190067147	190068264	31.3	95.09
PSG2	chr19	43575851	43576106	35.2	95.09
ITM2A	chrX	78618438	78618636	30.2	95.09
TNN	chr1	175092535	175092799	45.3	95.09
GATA3	chr10	8105958	8106101	20.8	95.09
HCN1	chr5	45461947	45462109	30.7	95.09
OCA2	chr15	28211835	28211968	44.8	95.09
CTNNA2	chr2	80816428	80816610	27.3	95.09
CNTN5	chr11	99715818	99715994	33.9	95.09
POM121L12	chr7	53103364	53104255	31.4	95.09
LRRC7	chr1	70225887	70226076	26.3	95.09
CNTNAP5	chr2	125530375	125530594	36.4	95.09
SLC4A10	chr2	162751188	162751335	33.8	95.09
SETD2	chr3	47142947	47143045	30.3	95.09
GFRAL	chr6	55216050	55216381	30.1	95.09
SORCS3	chr10	106927015	106927107	32.3	95.33
POTEG	chr14	19553416	19553937	32.6	95.33
F9	chrX	138630521	138630650	30.8	95.58
SLC26A3	chr7	107416896	107416989	21.3	95.58
UNC5D	chr8	35606044	35606213	29.4	95.58
PDE4DIP	chr1	144882775	144882881	37.4	95.58
MRPL1	chr4	78870950	78871032	48.2	95.58
COL25A1	chr4	109784474	109784543	42.9	95.58
SPTA1	chr1	158650372	158650519	33.8	95.58
TNR	chr1	175331798	175331945	33.8	95.58
GALNT13	chr2	155157921	155158102	33.0	95.58
EIF3E	chr8	109241298	109241424	39.4	95.58
SLC5A1	chr22	32445929	32446001	54.8	95.58
COASY	chr17	40717000	40717065	45.5	95.58
TBX15	chr1	119467268	119467440	40.5	95.58
PYHIN1	chr1	158908869	158909037	35.5	95.58
PSG5	chr19	43690493	43690557	46.2	95.58
BTRC	chr10	103290993	103291090	20.4	95.58
MDGA2	chr14	47324226	47324357	30.3	95.58
GUCY1A3	chr4	156629387	156629446	33.3	95.58
HGF	chr7	81386504	81386619	34.5	95.58
TIMD4	chr5	156346467	156346552	34.9	95.58
AK5	chr1	77752625	77752812	31.9	95.58
ODZ3	chr4	183245173	183245405	30.0	95.58
COL5A2	chr2	189927897	189927996	30.0	95.58
NTM	chr11	132180005	132180126	32.8	95.58
LTBP1	chr2	33500031	33500157	39.4	95.58
PRSS1	chr7	142458405	142458565	31.1	95.58

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
CDKN2A	chr9	21971001	21971207	125.6	95.58
CNGB3	chr8	87738758	87738885	31.3	95.58
SI	chr3	164777689	164777815	31.5	95.58
SI	chr3	164767578	164767663	46.5	95.58
TMEM132D	chr12	129822178	129822362	32.4	95.58
ASTN1	chr1	176998769	176998877	27.5	95.58
SAGE1	chrX	134987410	134987551	42.3	95.58
THSD7A	chr7	11464322	11464459	36.2	95.58
ADAMTS12	chr5	33683963	33684160	30.3	95.58
NRXN1	chr2	50463926	50464108	43.7	95.58
CSMD3	chr8	113562899	113563102	34.3	95.58
CSMD3	chr8	113364644	113364763	41.7	95.58
EPB41L4B	chr9	112018415	112018504	22.2	95.58
POLR3B	chr12	106820974	106821136	24.5	95.58
ATP10B	chr5	160097469	160097674	34.0	95.58
CSMD1	chr8	3165216	3165343	31.3	95.58
FBN2	chr5	127648325	127648487	30.7	95.58
EXOC5	chr14	57684699	57684786	22.7	95.58
ANKRD304	chr10	37440987	37441049	47.6	95.58
TRIML1	chr4	189065189	189065287	40.4	95.58
SPTA1	chr1	158631076	158631199	32.3	95.58
POLDIP2	chr17	26684313	26684473	31.1	95.58
KLHL1	chr13	70314525	70314688	30.5	95.58
TRIM58	chr1	248039201	248039791	23.7	95.58
GRIA3	chrX	122537262	122537370	27.5	95.58
CNOT4	chr7	135048605	135048818	23.4	95.58
NAV3	chr12	78582388	78582557	23.5	95.58
NAV3	chr12	78400198	78401225	21.4	95.58
TRPC5	chrX	111195270	111195648	21.1	95.58
LRRC2	chr3	46592956	46593081	23.8	95.58
ADAMTS16	chr5	5239793	5240038	24.4	95.58
ACER2	chr9	19424697	19424839	21.0	95.58
AMOT	chrX	112024113	112024346	21.4	95.58
OBP2A	chr9	138439716	138439827	26.8	95.58
INHBA	chr7	41729247	41730140	19.0	95.58
INHBA	chr7	41739584	41739972	7.7	95.58
EPHA5	chr4	66189831	66189937	28.0	95.58
EPHA5	chr4	66197690	66197846	12.7	95.58
EPHA5	chr4	66201649	66201843	10.3	95.58
EPHA5	chr4	66213771	66213921	19.9	95.58
EPHA5	chr4	66217106	66217316	19.0	95.58
EPHA5	chr4	66218740	66218840	19.8	95.58
EPHA5	chr4	66230734	66230920	16.0	95.58
EPHA5	chr4	66231649	66231775	23.6	95.58
EPHA5	chr4	66233058	66233158	19.8	95.58
EPHA5	chr4	66242698	66242798	0.0	95.58
EPHA5	chr4	66270091	66270194	19.2	95.58
EPHA5	chr4	66280001	66280161	6.2	95.58
EPHA5	chr4	66286158	66286283	0.0	95.58
EPHA5	chr4	66356094	66356430	14.8	95.58
EPHA5	chr4	66361105	66361261	6.4	95.58
EPHA5	chr4	66467358	66468022	9.0	95.58
EPHA5	chr4	66509062	66509163	0.0	95.58
EPHA5	chr4	66535279	66535460	5.5	95.58
EPHA3	chr3	89156892	89156992	0.0	95.58
EPHA3	chr3	89176340	89176441	19.6	95.58
EPHA3	chr3	89259009	89259670	9.1	95.58
EPHA3	chr3	89390065	89390221	25.5	95.58
EPHA3	chr3	89390904	89391240	8.9	95.58
EPHA3	chr3	89444986	89445111	15.9	95.58
EPHA3	chr3	89448467	89448656	5.3	95.58
EPHA3	chr3	89456418	89456521	0.0	95.58
EPHA3	chr3	89457198	89457299	0.0	95.58
EPHA3	chr3	89462290	89462416	23.6	95.58
EPHA3	chr3	89468354	89468540	5.3	95.58
EPHA3	chr3	89478236	89478336	0.0	95.58
EPHA3	chr3	89480299	89480509	19.0	95.58
EPHA3	chr3	89498374	89498524	6.6	95.58
EPHA3	chr3	89499326	89499520	10.3	95.58

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
EPHA3	chr3	89521613	89521769	19.1	95.58
EPHA3	chr3	89528546	89528652	9.3	95.58
PTPRD	chr9	8317857	8317958	19.6	95.58
PTPRD	chr9	8319830	8319966	0.0	95.58
PTPRD	chr9	8331581	8331736	6.4	95.58
PTPRD	chr9	8338921	8339047	15.7	95.58
PTPRD	chr9	8340342	8340469	7.8	95.58
PTPRD	chr9	8341089	8341268	0.0	95.58
PTPRD	chr9	8341692	8341978	7.0	95.58
PTPRD	chr9	8375935	8376090	6.4	95.58
PTPRD	chr9	8376606	8376726	8.3	95.58
PTPRD	chr9	8389231	8389407	0.0	95.58
PTPRD	chr9	8404536	8404660	0.0	95.58
PTPRD	chr9	8436590	8436690	9.9	95.58
PTPRD	chr9	8437168	8437268	0.0	95.58
PTPRD	chr9	8449724	8449837	26.3	95.58
PTPRD	chr9	8454536	8454637	0.0	95.58
PTPRD	chr9	8460410	8460571	18.5	95.58
PTPRD	chr9	8465465	8465675	28.4	95.58
PTPRD	chr9	8470989	8471090	9.8	95.58
PTPRD	chr9	8484118	8484378	19.2	95.58
PTPRD	chr9	8485226	8485327	0.0	95.58
PTPRD	chr9	8485761	8486349	6.8	95.58
PTPRD	chr9	8492861	8492979	8.4	95.58
PTPRD	chr9	8497204	8497305	9.8	95.58
PTPRD	chr9	8499646	8499840	10.3	95.58
PTPRD	chr9	8500753	8501059	9.8	95.58
PTPRD	chr9	8504260	8504405	6.8	95.58
PTPRD	chr9	8507300	8507434	7.4	95.58
PTPRD	chr9	8517847	8518429	15.4	95.58
PTPRD	chr9	8521276	8521546	18.5	95.58
PTPRD	chr9	8523468	8523568	9.9	95.58
PTPRD	chr9	8524924	8525035	8.9	95.58
PTPRD	chr9	8526585	8526685	0.0	95.58
PTPRD	chr9	8527298	8527399	19.6	95.58
PTPRD	chr9	8528590	8528779	21.1	95.58
PTPRD	chr9	8633316	8633458	21.0	95.58
PTPRD	chr9	8636698	8636844	13.6	95.58
PTPRD	chr9	8733761	8733861	0.0	95.58
KDR	chr4	55946107	55946330	4.5	95.58
KDR	chr4	55948115	55948215	0.0	95.58
KDR	chr4	55948702	55948802	19.8	95.58
KDR	chr4	55953773	55953925	19.6	95.58
KDR	chr4	55955034	55955140	18.7	95.58
KDR	chr4	55955540	55955640	0.0	95.58
KDR	chr4	55955857	55955969	8.8	95.58
KDR	chr4	55956122	55956245	0.0	95.58
KDR	chr4	55958782	55958882	19.8	95.58
KDR	chr4	55960968	55961122	12.9	95.58
KDR	chr4	55961737	55961838	19.6	95.58
KDR	chr4	55962395	55962509	8.7	95.58
KDR	chr4	55963828	55963933	28.3	95.58
KDR	chr4	55964303	55964439	0.0	95.58
KDR	chr4	55964863	55964970	18.5	95.58
KDR	chr4	55968063	55968195	7.5	95.58
KDR	chr4	55968528	55968675	13.5	95.58
KDR	chr4	55970809	55971151	14.6	95.58
KDR	chr4	55971998	55972107	18.2	95.58
KDR	chr4	55972853	55972977	8.0	95.58
KDR	chr4	55973903	55974060	12.7	95.58
KDR	chr4	55976569	55976733	12.1	95.58
KDR	chr4	55976820	55976935	8.6	95.58
KDR	chr4	55979470	55979648	11.2	95.58
KDR	chr4	55980292	55980432	0.0	95.58
KDR	chr4	55981040	55981209	5.9	95.58
KDR	chr4	55981447	55981578	30.3	95.58
KDR	chr4	55984770	55984967	0.0	95.58
KDR	chr4	55987260	55987360	9.9	95.58
KDR	chr4	55991376	55991477	0.0	95.58

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
NTRK3	chr15	88420165	88420351	0.0	95.58
NTRK3	chr15	88423500	88423659	6.3	95.58
NTRK3	chr15	88428895	88428995	0.0	95.58
NTRK3	chr15	88472421	88472665	4.1	95.58
NTRK3	chr15	88476242	88476415	23.0	95.58
NTRK3	chr15	88483853	88483984	7.6	95.58
NTRK3	chr15	88522575	88522694	0.0	95.58
NTRK3	chr15	88524456	88524591	0.0	95.58
NTRK3	chr15	88576087	88576276	10.5	95.58
NTRK3	chr15	88669501	88669604	28.8	95.58
NTRK3	chr15	88670374	88670475	0.0	95.58
NTRK3	chr15	88671903	88672003	0.0	95.58
NTRK3	chr15	88678331	88678628	23.5	95.58
NTRK3	chr15	88679129	88679271	7.0	95.58
NTRK3	chr15	88679697	88679840	13.9	95.58
NTRK3	chr15	88680634	88680792	0.0	95.58
NTRK3	chr15	88690549	88690650	0.0	95.58
NTRK3	chr15	88726634	88726734	9.9	95.58
NTRK3	chr15	88727442	88727543	9.8	95.58
RB1	chr13	48878048	48878185	0.0	95.58
RB1	chr13	48881415	48881542	23.4	95.58
RB1	chr13	48916734	48916850	8.5	95.58
RB1	chr13	48919215	48919335	8.3	95.58
RB1	chr13	48921929	48922030	0.0	95.58
RB1	chr13	48923075	48923175	0.0	95.58
RB1	chr13	48934152	48934263	17.9	95.58
RB1	chr13	48936950	48937093	0.0	95.58
RB1	chr13	48939018	48939118	0.0	95.58
RB1	chr13	48941629	48941739	27.0	95.58
RB1	chr13	48942651	48942751	0.0	95.58
RB1	chr13	48947534	48947634	19.8	95.58
RB1	chr13	48951053	48951170	0.0	95.58
RB1	chr13	48953707	48953808	19.6	95.58
RB1	chr13	48954154	48954254	0.0	95.58
RB1	chr13	48954288	48954389	9.8	95.58
RB1	chr13	48955382	48955579	0.0	95.58
RB1	chr13	49027128	49027247	0.0	95.58
RB1	chr13	49030339	49030485	20.4	95.58
RB1	chr13	49033823	49033969	6.8	95.58
RB1	chr13	49037866	49037971	0.0	95.58
RB1	chr13	49039133	49039247	8.7	95.58
RB1	chr13	49039340	49039504	12.1	95.58
RB1	chr13	49047460	49047561	0.0	95.58
RB1	chr13	49050836	49050979	0.0	95.58
RB1	chr13	49051465	49051565	0.0	95.58
RB1	chr13	49054120	49054220	0.0	95.58
ERBB4	chr2	212248339	212248785	6.7	95.58
ERBB4	chr2	212251577	212251875	10.0	95.58
ERBB4	chr2	212252643	212252743	0.0	95.58
ERBB4	chr2	212285165	212285336	11.6	95.58
ERBB4	chr2	212286730	212286830	9.9	95.58
ERBB4	chr2	212288879	212289026	6.8	95.58
ERBB4	chr2	212293120	212293220	0.0	95.58
ERBB4	chr2	212295669	212295825	12.7	95.58
ERBB4	chr2	212426627	212426813	5.3	95.58
ERBB4	chr2	212483901	212484000	0.0	95.58
ERBB4	chr2	212488646	212488769	0.0	95.58
ERBB4	chr2	212495186	212495319	0.0	95.58
ERBB4	chr2	212522465	212522566	19.6	95.58
ERBB4	chr2	212530047	212530202	6.4	95.58
ERBB4	chr2	212537885	212537985	9.9	95.58
ERBB4	chr2	212543776	212543909	7.5	95.58
ERBB4	chr2	212566691	212566891	10.0	95.58
ERBB4	chr2	212568823	212568924	0.0	95.58
ERBB4	chr2	212570029	212570129	9.9	95.58
ERBB4	chr2	212576774	212576901	7.8	95.58
ERBB4	chr2	212578259	212578373	8.7	95.58
ERBB4	chr2	212587117	212587259	0.0	95.58
ERBB4	chr2	212589800	212589919	16.7	95.58

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
ERBB4	chr2	212615346	212615446	0.0	95.58
ERBB4	chr2	212652749	212652884	7.4	95.58
ERBB4	chr2	212812154	212812341	21.3	95.82
ERBB4	chr2	212989476	212989628	13.1	95.82
ERBB4	chr2	213403163	213403263	0.0	95.82
NTRK1	chr1	156785575	156785676	0.0	95.82
NTRK1	chr1	156811872	156811985	0.0	95.82
NTRK1	chr1	156830726	156830938	0.0	95.82
NTRK1	chr1	156834132	156834233	9.8	95.82
NTRK1	chr1	156834505	156834605	0.0	95.82
NTRK1	chr1	156836685	156836786	0.0	95.82
NTRK1	chr1	156837895	156838041	6.8	95.82
NTRK1	chr1	156838296	156838439	0.0	95.82
NTRK1	chr1	156841414	156841547	0.0	95.82
NTRK1	chr1	156843424	156843751	3.0	95.82
NTRK1	chr1	156844133	156844233	0.0	95.82
NTRK1	chr1	156844340	156844440	0.0	95.82
NTRK1	chr1	156844697	156844800	0.0	95.82
NTRK1	chr1	156845311	156845458	13.5	95.82
NTRK1	chr1	156845871	156846002	22.7	95.82
NTRK1	chr1	156846191	156846364	11.5	95.82
NTRK1	chr1	156848913	156849154	16.5	95.82
NTRK1	chr1	156849790	156849949	0.0	95.82
NTRK1	chr1	156851248	156851434	0.0	95.82
NF1	chr17	29422307	29422407	0.0	95.82
NF1	chr17	29483000	29483144	0.0	95.82
NF1	chr17	29486019	29486119	9.9	95.82
NF1	chr17	29490203	29490394	5.2	95.82
NF1	chr17	29496908	29497015	9.3	95.82
NF1	chr17	29508423	29508523	0.0	95.82
NF1	chr17	29508715	29508815	0.0	95.82
NF1	chr17	29509525	29509683	6.3	95.82
NF1	chr17	29527439	29527613	17.1	95.82
NF1	chr17	29528054	29528177	0.0	95.82
NF1	chr17	29528415	29528516	0.0	95.82
NF1	chr17	29533257	29533389	0.0	95.82
NF1	chr17	29541468	29541603	7.4	95.82
NF1	chr17	29546022	29546136	8.7	95.82
NF1	chr17	29548867	29549008	7.0	95.82
NF1	chr17	29550461	29550585	0.0	95.82
NF1	chr17	29552112	29552268	0.0	95.82
NF1	chr17	29553452	29553702	4.0	95.82
NF1	chr17	29554222	29554322	0.0	95.82
NF1	chr17	29554532	29554632	9.9	95.82
NF1	chr17	29556042	29556483	4.5	95.82
NF1	chr17	29556852	29556992	7.1	95.82
NF1	chr17	29557277	29557400	8.1	95.82
NF1	chr17	29557851	29557951	0.0	95.82
NF1	chr17	29559090	29559207	0.0	95.82
NF1	chr17	29559717	29559899	10.9	95.82
NF1	chr17	29560019	29560231	4.7	95.82
NF1	chr17	29562628	29562790	12.3	95.82
NF1	chr17	29562935	29563039	0.0	95.82
NF1	chr17	29576001	29576137	0.0	95.82
NF1	chr17	29579936	29580037	0.0	95.82
NF1	chr17	29585361	29585520	0.0	95.82
NF1	chr17	29586048	29586148	9.9	95.82
NF1	chr17	29587386	29587533	13.5	95.82
NF1	chr17	29588728	29588875	0.0	95.82
NF1	chr17	29592246	29592357	0.0	95.82
NF1	chr17	29652837	29653270	4.6	95.82
NF1	chr17	29654516	29654857	8.8	95.82
NF1	chr17	29657313	29657516	9.8	95.82
NF1	chr17	29661855	29662049	15.4	95.82
NF1	chr17	29663350	29663491	14.1	95.82
NF1	chr17	29663652	29663932	0.0	95.82

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
NF1	chr17	29664385	29664600	4.6	95.82
NF1	chr17	29664817	29664917	9.9	95.82
NF1	chr17	29665042	29665157	0.0	95.82
NF1	chr17	29665721	29665823	19.4	95.82
NF1	chr17	29667522	29667663	7.0	95.82
NF1	chr17	29670026	29670153	15.6	95.82
NF1	chr17	29676137	29676269	15.0	95.82
NF1	chr17	29677200	29677336	0.0	95.82
NF1	chr17	29679274	29679432	12.6	95.82
NF1	chr17	29683477	29683600	0.0	95.82
NF1	chr17	29683977	29684108	7.6	95.82
NF1	chr17	29684286	29684387	9.8	95.82
NF1	chr17	29685497	29685640	6.9	95.82
NF1	chr17	29685959	29686060	0.0	95.82
NF1	chr17	29687504	29687721	0.0	95.82
NF1	chr17	29701030	29701173	6.9	95.82
APC	chr5	112043414	112043579	0.0	95.82
APC	chr5	112090587	112090722	0.0	95.82
APC	chr5	112102014	112102115	9.8	95.82
APC	chr5	112102885	112103087	9.9	95.82
APC	chr5	112111325	112111434	9.1	95.82
APC	chr5	112116486	112116600	0.0	95.82
APC	chr5	112128134	112128234	0.0	95.82
APC	chr5	112136975	112137080	0.0	95.82
APC	chr5	112151191	112151290	0.0	95.82
APC	chr5	112154662	112155041	2.6	95.82
APC	chr5	112157590	112157690	0.0	95.82
APC	chr5	112162804	112162944	0.0	95.82
APC	chr5	112163614	112163714	0.0	95.82
APC	chr5	112164552	112164669	16.9	95.82
APC	chr5	112170647	112170862	0.0	95.82
APC	chr5	112173249	112179823	3.5	96.07
ATM	chr11	108098337	108098437	0.0	96.07
ATM	chr11	108098502	108098615	8.8	96.07
ATM	chr11	108099904	108100050	0.0	96.07
ATM	chr11	108106396	108106561	0.0	96.07
ATM	chr11	108114679	108114845	0.0	96.07
ATM	chr11	108115514	108115753	4.2	96.07
ATM	chr11	108117690	108117854	0.0	96.07
ATM	chr11	108119659	108119829	5.8	96.07
ATM	chr11	108121427	108121799	0.0	96.07
ATM	chr11	108122563	108122758	0.0	96.07
ATM	chr11	108123541	108123641	9.9	96.07
ATM	chr11	108124540	108124766	0.0	96.07
ATM	chr11	108126941	108127067	7.9	96.07
ATM	chr11	108128207	108128333	0.0	96.07
ATM	chr11	108129707	108129807	0.0	96.07
ATM	chr11	108137897	108138069	5.8	96.07
ATM	chr11	108139136	108139336	0.0	96.07
ATM	chr11	108141781	108141882	0.0	96.07
ATM	chr11	108141977	108142133	0.0	96.07
ATM	chr11	108143246	108143346	0.0	96.07
ATM	chr11	108143448	108143579	7.6	96.07
ATM	chr11	108150217	108150335	0.0	96.07
ATM	chr11	108151721	108151895	0.0	96.07
ATM	chr11	108153436	108153606	11.7	96.07
ATM	chr11	108154953	108155200	4.0	96.07
ATM	chr11	108158326	108158442	0.0	96.07
ATM	chr11	108159703	108159830	7.8	96.07
ATM	chr11	108160328	108160528	5.0	96.07
ATM	chr11	108163345	108163520	0.0	96.07
ATM	chr11	108164039	108164204	0.0	96.07
ATM	chr11	108165653	108165786	0.0	96.07
ATM	chr11	108168011	108168111	9.9	96.07
ATM	chr11	108170440	108170612	5.8	96.07
ATM	chr11	108172374	108172516	0.0	96.07
ATM	chr11	108173579	108173756	0.0	96.07
ATM	chr11	108175401	108175579	11.2	96.07
ATM	chr11	108178617	108178717	0.0	96.07

TABLE 2-continued

Genomic Region				Coverage (unique LUAD & SCC patients; n = 407) % patients ≥1	
Gene	Chr	Start (bp)	End (bp)	RI	SNV
ATM	chr11	108180886	108181042	0.0	96.07
ATM	chr11	108183131	108183231	9.9	96.07
ATM	chr11	108186543	108186644	0.0	96.07
ATM	chr11	108186737	108186840	9.6	96.07
ATM	chr11	108188099	108188248	0.0	96.07
ATM	chr11	108190680	108190785	0.0	96.07
ATM	chr11	108192027	108192147	0.0	96.07
ATM	chr11	108196036	108196271	4.2	96.07
ATM	chr11	108196784	108196952	0.0	96.07
ATM	chr11	108198371	108198485	0.0	96.07
ATM	chr11	108199747	108199965	4.6	96.07
ATM	chr11	108200940	108201148	0.0	96.07
ATM	chr11	108202170	108202284	0.0	96.07
ATM	chr11	108202605	108202764	0.0	96.07
ATM	chr11	108203488	108203627	0.0	96.07
ATM	chr11	108204603	108204704	9.8	96.07
ATM	chr11	108205695	108205836	21.1	96.07
ATM	chr11	108206571	108206688	8.5	96.07
ATM	chr11	108213948	108214098	0.0	96.07
ATM	chr11	108216469	108216635	0.0	96.07
ATM	chr11	108217998	108218099	9.8	96.07
ATM	chr11	108224492	108224607	8.6	96.07
ATM	chr11	108225519	108225619	0.0	96.07
ATM	chr11	108235808	108235945	7.2	96.07
ATM	chr11	108236051	108236235	10.8	96.07
FGFR4	chr5	176516598	176516699	0.0	96.07
FGFR4	chr5	176517390	176517654	3.8	96.07
FGFR4	chr5	176517735	176517836	9.8	96.07
FGFR4	chr5	176517938	176518105	0.0	96.07
FGFR4	chr5	176518685	176518809	0.0	96.07
FGFR4	chr5	176519321	176519512	0.0	96.07
FGFR4	chr5	176519646	176519785	0.0	96.07
FGFR4	chr5	176520138	176520552	4.8	96.07
FGFR4	chr5	176520654	176520776	0.0	96.07
FGFR4	chr5	176522330	176522441	8.9	96.07
FGFR4	chr5	176522533	176522724	0.0	96.07
FGFR4	chr5	176523057	176523180	0.0	96.07
FGFR4	chr5	176523272	176523373	0.0	96.07
FGFR4	chr5	176523604	176523742	0.0	96.07
FGFR4	chr5	176524292	176524398	0.0	96.07
FGFR4	chr5	176524527	176524677	0.0	96.07
ALK	chr2	29446207	29448431	—	—
ROS1	chr6	117641031	117658503	—	—
RET	chr10	43606655	43612179	—	—
PDGFRA	chr4	55140698	55141140	—	—
FGFR1	chr8	38275746	38277253	—	—

TABLE 3

Sample count	Sample description/ patient (P#)/ healthy control (C#)	Sample source	DNA mass used for library (ng)	Expected haploid genome copies (330 × per ng)	Total cfDNA concen-tration in plasma (ng/mL)	Volume of plasma used for library (mL)	Library mass used for capture (ng)
1	H3122 0.1% into HCC78	Cell line	128	42240			111
2	H3122 1% into HCC78	Cell line	128	42240			111
3	H3122 10% into HCC78	Cell line	128	42240			111
4	H3122100%	Cell line	128	42240			111
5	HCC78 100%	Cell line	128	42240			111
6	HCC78 10% into C1 plasma DNA 4 cycles	Cell line/ plasma DNA	128	42240			83.3
7	HCC78 10% into C1 plasma DNA 8 cycles SigmaWGA	Cell line/ plasma DNA	1	330			83.3
8	HCC78 10% into C1 plasma DNA 6 cycles	Cell line/ plasma DNA	32	10560			83.3
9	HCC78 10% into C1 plasma DNA 8 cycles NEBNextOvernightBead	Cell line/ plasma DNA	32	10560			83.3
10	HCC78 10% into C1 plasma DNA 8 cycles OrigNEBNext 15 min Lig	Cell line/ plasma DNA	32	10560			83.3
11	HCC78 10% into C1 plasma DNA 4 ng 9 cycles	Cell line/ plasma DNA	4	1320			83.3
12	HCC78 0.025% into C1 plasma DNA	Cell line/ plasma DNA	32	10560			83.3
13	HCC78 0.05% into C1 plasma DNA	Cell line/ plasma DNA	32	10560			83.3
14	HCC78 0.1% into C1 plasma DNA	Cell line/ plasma DNA	32	10560			83.3
15	HCC78 0.5% into C1 plasma DNA	Cell line/ plasma DNA	32	10560			83.3
16	HCC78 1% into C1 plasma DNA	Cell line/ plasma DNA	32	10560			83.3
17	P1	PBL	500	165000			83.3
18	P2	PBL	500	165000			83.3
19	P3	PBL	500	165000			83.3
20	P4	PBL	500	165000			83.3
21	P5	PBL	500	165000			83.3
22	P6	PBL	500	165000			83.3
23	P7	PBL	500	165000			83.3
24	P8	PBL	500	165000			83.3
25	P9	PBL	500	165000			83.3
26	P10	PBL	400	132000			83.3
27	P11	PBL	500	165000			83.3
28	P12	PBL	200	66000			83.3
29	P13	PBL	200	66000			83.3
30	P14	PBL	200	66000			83.3
31	P15	PBL	200	66000			83.3
32	P16	PBL	200	66000			83.3
33	P17	PBL	200	66000			83.3
34	P1	Tumor	500	165000			83.3
35	P2	Tumor	500	165000			83.3
36	P3	Tumor	500	165000			83.3
37	P4	Tumor	200	66000			83.3
38	P5	Tumor	100	33000			83.3
39	P6	Tumor	1000	330000			83.3
40	P7	Tumor	500	165000			83.3
41	P8	Tumor	500	165000			83.3
42	P9	Tumor	69	22770			83.3
43	P10	Tumor	500	165000			83.3
44	P11	Tumor	500	165000			83.3
45	P12	Tumor	125	41196			83.3
46	P13	Tumor	5	1516			83.3

TABLE 3-continued

Sample count	Sample description/ patient (P#)/ healthy control (C#)	Sample source	DNA mass used for library (ng)	Expected haploid genome copies (330 × per ng)	Total cfDNA concen-tration in plasma (ng/mL)	Volume of plasma used for library (mL)	Library mass used for capture (ng)
47	P14	Tumor	125	41197			83.3
48	P15	Tumor	4	1427			83.3
49	P16	Tumor	12	3872			83.3
50	P17	Tumor	97	31904			83.3
51	C1	Plasma DNA	32	10560	12.49	2.56	83.3
52	C2	Plasma DNA	2	793	14.24	0.17	83.3
53	C3	Plasma DNA	37	12218	7.82	4.73	83.3
54	C4	Plasma DNA	1	375	6.64	0.17	83.3
55	C5	Plasma DNA	21	6834	14.44	1.43	83.3
56	P1 time point 1	Plasma DNA	13	4290	7.33	1.77	83.3
57	P1 time point 2	Plasma DNA	7	2310	7.52	0.93	83.3
58	P1 time point 3	Plasma DNA	36	11755	19.87	1.79	83.3
59	P2 time point 1	Plasma DNA	13	4290	7.22	1.80	83.3
60	P2 time point 2	Plasma DNA	16	5280	10.93	1.46	83.3
61	P2 time point 3	Plasma DNA	35	11462	13.12	2.65	83.3
62	P3 time point 1	Plasma DNA	15	4950	17.17	0.87	83.3
63	P3 time point 2	Plasma DNA	16	5280	12.84	1.25	83.3
64	P4 time point 1	Plasma DNA	10	3300	5.40	1.85	83.3
65	P4 time point 2	Plasma DNA	16	5280	18.19	0.88	83.3
66	P5 time point 1	Plasma DNA	9	2970	4.49	2.00	83.3
68	P5 time point 2	Plasma DNA	29	9549	37.06	0.78	83.3
67	P5 time point 3	Plasma DNA	15	4950	5.37	2.79	83.3
69	P6 time point 1	Plasma DNA	17	5610	35.11	0.48	83.3
70	P6 time point 2	Plasma DNA	20	6600	85.87	0.23	83.3
71	P9 time point 1	Plasma DNA	12	3960	9.22	1.30	83.3
72	P9 time point 2	Plasma DNA	17	5610	11.38	1.49	83.3
73	P9 time point 3	Plasma DNA	16	5280	10.41	1.54	83.3
74	P9 time point 4	Plasma DNA	35	11622	19.42	1.81	83.3
75	P9 time point 5	Plasma DNA	36	11775	33.70	1.06	83.3
76	P12 time point 1	Plasma DNA	17	5507	11.03	1.51	83.3
77	P12 time point 2	Plasma DNA	28	9230	15.57	1.80	83.3
78	P13 time point 1	Plasma DNA	25	8291	15.18	1.65	83.3
79	P13 time point 2	Plasma DNA	15	5043	9.24	1.65	83.3
80	P14 time point 1	Plasma DNA	17	5716	20.36	0.85	83.3
81	P14 time point 2	Plasma DNA	35	11596	27.49	1.28	83.3
82	P15 time point 1	Plasma DNA	25	8111	18.57	1.32	83.3
83	P15 time point 2	Plasma DNA	31	10308	17.86	1.75	83.3
84	P15 time point 3	Plasma DNA	7	2305	5.28	1.32	83.3
85	P15 time point 4	Plasma DNA	23	7525	5.74	3.97	83.3
86	P15 time point 5	Plasma DNA	8	2517	2.88	2.65	83.3
87	P16 time point 1	Plasma DNA	17	5688	13.02	1.32	83.3
88	P16 time point 2	Plasma DNA	6	2089	10.14	0.62	83.3
89	P16 time point 3	Plasma DNA	32	10579	17.49	1.83	83.3
90	P17 time point 1	Plasma DNA	12	4056	9.28	1.32	83.3

TABLE 4

Patient number	Plasma time point	% ctDNA[a]	Total plasma DNA (ng/mL)	ctDNA (pg/mL)	ctDNA detection index[b]
P12	1	ND	11.032	ND	NS
P12	2	ND	15.571	ND	NS
P1	1	0.025	7.326	1.854	0.005
P1	2	ND	7.520	ND	NS
P1	3	ND	19.869	ND	NS
P16	1	0.019	13.023	2.474	0.05
P16	2	ND	10.140	ND	NS
P16	3	ND	17.492	ND	NS
P17	1	ND	9.285	ND	NS
P13	1	1.777	15.184	269.821	<0.0001
P13	2	ND	9.237	ND	NS
P2	1	0.896	7.221	64.698	<0.0001
P2	2	0.038	10.927	4.152	0.03
P2	3	ND	13.120	ND	NS
P3	1	0.095	17.171	16.237	0.009

TABLE 4-continued

Patient number	Plasma time point	% ctDNA[a]	Total plasma DNA (ng/mL)	ctDNA (pg/mL)	ctDNA detection index[b]
P3	2	ND	12.841	ND	NS
P14	1	0.050	20.356	10.179	0.02
P14	2	0.042	27.491	11.416	0.02
P15	1	0.582	18.568	108.117	3.2E−05
P15	2	ND	17.859	ND	NS
P15	3	ND	5.276	ND	NS
P15	4	0.421	5.742	24.201	1.7E−06
P15	5	0.855	2.881	24.639	0.0001
P4	1	0.039	5.400	2.125	0.04
P4	2	ND	18.191	ND	NS
P5	1	3.201	4.491	143.781	<0.0001
P5	2	0.074	37.064	27.557	0.02
P5	3	0.351	5.372	18.861	0.0006
P6	1	0.998	35.100	350.190	~0
P6	2	0.230	85.900	197.951	~0

TABLE 4-continued

Patient number	Plasma time point	% ctDNA[a]	Total plasma DNA (ng/mL)	ctDNA (pg/mL)	ctDNA detection index[b]
P9	1	0.042	9.221	3.828	~0
P9	2	0.005	11.383	0.585	~0
P9	3	0.050	10.406	5.184	~0
P9	4	0.019	19.419	3.615	~0
P9	5	ND	33.697	ND	NS

[a]Mean fraction across all SNV/indel reporters if present, or fusions if no other reporter types present. The subclonal T790M reporter identified in P5 was excluded.
[b]Analogous to false positive rate

TABLE 5

Case	Mutant allele	Ref. allele	Chr	Position	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)
					Time point 1			Time point 2			Time point 3		
P12	T	C	chr 4	55973786	0	1508	0.000	2	2104	0.095	—	—	—
P12	T	G	chr 6	117650296	1	5165	0.019	2	6148	0.033	—	—	—
P12	G	T	chr 7	41729291	0	3773	0.000	0	4634	0.000	—	—	—
P12	T	A	chr 9	8471102	0	3637	0.000	0	4246	0.000	—	—	—
P12	G	T	chr 12	25380276	0	3633	0.000	0	4186	0.000	—	—	—
P12	A	C	chr 19	10602473	0	1873	0.000	0	2399	0.000	—	—	—
P12	−C	T	chr 17	7577057	0	3451	0.000	0	3779	0.000	—	—	—
P1	A	G	chr 1	156785560	0	4572	0.000	3	6202	0.048	0	5220	0.000
P1	T	G	chr 1	157806043	0	1838	0.000	0	2266	0.000	0	1902	0.000
P1	G	C	chr 1	248525206	0	2828	0.000	0	4529	0.000	0	3327	0.000
P1	C	T	chr 2	33500291	1	943	0.106	0	943	0.000	0	935	0.000
P1	A	C	chr 4	55946307	0	6856	0.000	0	8817	0.000	0	6279	0.000
P1	G	A	chr 4	55963949	0	5742	0.000	0	7335	0.000	2	5766	0.035
P1	A	C	chr 4	55968672	0	5856	0.000	0	7431	0.000	0	6376	0.000
P1	C	T	chr 6	117642146	0	5266	0.000	4	6849	0.058	0	5407	0.000
P1	T	G	chr 9	8376700	3	5535	0.054	0	7322	0.00	0	6196	0.000
P1	T	C	chr 9	8733625	1	827	0.121	0	1398	0.000	0	1110	0.000
P1	T	G	chr 10	43611663	0	3722	0.000	0	4565	0.000	0	6741	0.000
P1	T	G	chr 15	88522525	1	4919	0.020	4	6736	0.059	12	5693	0.211
P1	+G	C	chr 17	7578474	0	1762	0.000	0	2373	0.000	5	4578	0.109
P1	−A	G	chr 17	29552244	1	4484	0.022	0	6485	0.000	0	4640	0.000
P1	+T	C	chr 17	29553484	0	3657	0.000	0	4713	0.000	0	3618	0.000
P1	−T	C	chr 17	29592185	3	3694	0.081	0	3247	0.000	0	3692	0.000
P16	A	G	chr 1	156843429	7	3107	0.225	1	3492	0.029	0	3602	0.000
P16	T	C	chr 1	181708291	0	5009	0.000	0	6962	0.000	4	6865	0.058
P16	A	C	chr 1	24852532	0	4484	0.000	0	5927	0.000	0	5948	0.000
P16	A	C	chr 2	125530343	0	5051	0.000	0	6591	0.000	2	6029	0.033
P16	A	C	chr 2	212530083	0	5112	0.000	1	5986	0.017	0	6462	0.000
P16	C	T	chr 2	212587119	0	5929	0.000	1	7481	0.013	0	7205	0.000
P16	T	G	chr 4	55958900	0	4585	0.000	6	5818	0.103	0	5664	0.000
P16	C	T	chr 4	55962358	1	4558	0.022	0	6406	0.000	1	6077	0.016
P16	A	C	chr 4	55968588	0	6084	0.000	0	8376	0.000	1	8537	0.012
P16	G	A	chr 4	55970963	0	5646	0.000	0	7604	0.000	0	7359	0.000
P16	A	C	chr 4	55971241	0	1562	0.000	0	2209	0.000	0	1952	0.000
P16	T	G	chr 5	19473838	0	3180	0.000	0	4028	0.000	1	4127	0.024
P16	A	G	chr 5	112176654	9	5308	0.170	0	6481	0.000	0	5211	0.000
P16	T	G	chr 5	176520134	4	4790	0.084	1	5207	0.019	0	5946	0.000
P16	T	G	chr 7	11501543	0	2141	0.000	0	2950	0.000	0	3026	0.000
P16	A	C	chr 7	53103357	0	2252	0.000	0	2737	0.000	1	2816	0.036
P16	T	C	chr 7	116411990	0	5193	0.000	0	7080	0.000	0	6466	0.000
P16	A	C	chr 10	43606641	0	3519	0.000	0	4261	0.000	0	4521	0.000
P16	A	G	chr 11	534195	1	2729	0.037	0	3262	0.000	0	3629	0.000
P16	G	C	chr 11	108143456	0	5308	0.000	0	6992	0.000	0	6833	0.000
P16	A	C	chr 12	25398284	0	4346	0.000	3	5866	0.051	0	5458	0.000
P16	A	C	chr 13	48947619	0	4639	0.000	1	6236	0.016	0	4765	0.000
P16	T	C	chr 13	70314492	0	2414	0.000	0	2752	0.000	0	2261	0.000
P16	A	T	chr 13	70314809	0	731	0.000	0	610	0.000	0	564	0.000
P16	C	G	chr 15	88472337	0	2467	0.000	0	3236	0.000	0	3274	0.000
P16	A	C	chr 17	7578132	0	2568	0.000	0	3369	0.000	0	3492	0.000
P16	+T	A	chr 2	212295977	0	483	0.000	0	356	0.000	0	302	0.000
P16	−C	T	chr 19	1220638	0	2848	0.000	0	3186	0.000	0	4066	0.000
P17	T	G	chr 7	81386606	1	4524	0.022	—	—	—	—	—	—
P17	A	C	chr 12	25398285	0	4165	0.000	—	—	—	—	—	—
P13	T	C	chr 1	190067540	202	5609	3.601	7	6967	0.100	—	—	—
P13	T	C	chr 5	45461969	147	5251	2.799	0	6568	0.000	—	—	—
P13	G	C	chr 8	38276015	7	5937	0.118	0	8357	0.000	—	—	—
P13	T	C	chr 15	88483904	1	5854	0.017	4	7528	0.053	—	—	—
P13	T	C	chr 17	7577538	93	3962	2.347	3	5035	0.060	—	—	—

TABLE 5-continued

Case	Mutant allele	Ref. allele	Chr	Position	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)
P2	A	C	chr 2	50463926	49	6724	0.729	0	4981	0.000	0	5636	0.000
P2	G	A	chr 3	89457148	40	4838	0.827	0	4311	0.000	0	4114	0.000
P2	T	G	chr 3	89468286	5	4667	0.107	2	3625	0.055	6	3411	0.176
P2	T	A	chr 3	89480240	15	5073	0.296	0	4321	0.000	0	3984	0.000
P2	T	A	chr 4	66189669	4	950	0.421	5	1436	0.348	0	1237	0.000
P2	T	G	chr 4	66242868	16	2107	0.759	0	1655	0.000	0	1879	0.000
P2	A	C	chr 5	176522747	46	2220	2.072	0	1377	0.000	0	3196	0.000
P2	C	T	chr 6	117648229	70	7819	0.895	0	5985	0.000	0	5951	0.000
P2	A	C	chr 12	78400637	35	7907	0.443	1	6326	0.016	1	6402	0.016
P2	T	G	chr 12	78400910	106	8211	1.291	1	6289	0.016	2	6260	0.032
P2	T	C	chr 17	7577551	112	5629	1.990	2	3814	0.052	2	4934	0.041
P2	T	G	chr 19	1207247	15	1124	1.335	0	747	0.000	0	1214	0.000
P2	+A	C	chr 2	79314100	16	3280	0.488	0	2390	0.000	0	2299	0.000
P3	A	C	chr 17	7578253	6	6345	0.095	0	8583	0.000	—	—	—
P14	C	A	chr 1	156841521	0	7377	0.000	0	5043	0.000	—	—	—
P14	T	G	chr 3	89176334	0	4981	0.000	0	4471	0.000	—	—	—
P14	A	G	chr 7	55249159	6	9223	0.065	1	6567	0.015	—	—	—
P14	G	T	chr 7	55259515	1	7207	0.014	0	5418	0.000	—	—	—
P14	T	C	chr 10	43607789	0	7552	0.000	1	5382	0.019	—	—	—
P14	C	T	chr 17	7577545	0	6379	0.000	4	4773	0.084	—	—	—
P14	T	C	chr 17	29553484	16	4983	0.321	8	3728	0.215	—	—	—
P14	G	C	chr 19	1223125	0	5804	0.000	0	3984	0.000	—	—	—
P15	T	G	chr 1	70226008	53	5317	0.997	0	6580	0.000	3	7204	0.042
P15	A	C	chr 1	144882833	30	11651	0.257	0	12602	0.000	0	15616	0.000
P15	A	C	chr 1	190203515	0	3976	0.000	0	5011	0.000	0	5485	0.000
P15	A	C	chr 1	248525334	32	5748	0.557	5	5359	0.093	1	7423	0.013
P15	A	C	chr 2	155157911	0	4151	0.000	0	5195	0.000	0	6295	0.000
P15	A	G	chr 2	212495103	10	1941	0.515	0	2439	0.000	1	2469	0.041
P15	T	G	chr 3	89528742	16	2224	0.719	0	2523	0.000	1	2776	0.036
P15	T	G	chr 4	55979517	197	7397	2.663	1	7458	0.013	0	10521	0.000
P15	A	C	chr 4	66189751	0	2556	0.000	0	3448	0.000	0	4235	0.000
P15	A	C	chr 4	66233002	6	981	0.612	0	1212	0.000	0	1542	0.000
P15	A	C	chr 4	66233003	6	1027	0.584	0	1258	0.000	0	1579	0.000
P15	T	G	chr 4	66233146	59	4970	1.187	0	5644	0.000	0	5923	0.000
P15	A	C	chr 5	176523126	59	5192	1.136	0	4356	0.000	4	7533	0.053
P15	A	C	chr 5	176524647	1	6308	0.016	0	5473	0.000	0	7179	0.000
P15	A	C	chr 7	41729339	33	5544	0.595	0	5817	0.000	0	8610	0.000
P15	A	C	chr 8	87738607	0	744	0.000	0	1531	0.000	0	1094	0.000
P15	A	C	chr 8	113563115	34	4123	0.825	0	4571	0.000	0	4569	0.000
P15	A	C	chr 9	8528716	56	6479	0.864	6	6339	0.095	0	8990	0.000
P15	A	T	chr 9	138439735	56	5497	1.019	0	5288	0.000	0	7310	0.000
P15	A	C	chr 10	43608292	21	5832	0.360	0	4912	0.000	0	7629	0.000
P15	T	C	chr 10	43608755	5	6687	0.075	1	6772	0.015	0	10118	0.000
P15	A	C	chr 11	55135855	63	5692	1.107	0	5984	0.000	0	9570	0.000
P15	T	C	chr 12	25398284	27	3573	0.756	1	4691	0.021	0	5193	0.000
P15	T	C	chr 13	48954333	0	2498	0.000	0	3674	0.000	0	3696	0.000
P15	T	G	chr 13	48954451	1	2233	0.045	0	3214	0.000	0	3319	0.000
P15	+T	G	chr 17	29533514	4	1758	0.228	0	2705	0.000	0	2333	0.000
P4	T	C	chr 2	212248555	6	7623	0.079	5	10563	0.047	—	—	—
P4	T	C	chr 12	25398281	0	5359	0.000	0	9389	0.000	—	—	—
P5	T	C	chr 7	55249071	42	4736	0.887	0	5978	0.000	10	5597	0.179
P5	G	T	chr 7	55259515	503	11349	4.432	12	5955	0.202	58	12222	0.475
P5	A	G	chr 11	55135338	86	4063	2.117	0	2802	0.000	10	4798	0.208
P5	T	C	chr 17	7577097	227	7429	3.056	1	4643	0.022	36	9723	0.370
P6	A	G	chr 12	78400791	84	13970	0.601	28	10128	0.276	—	—	—
P6	T	G	chr 12	129822187	78	8680	0.899	9	6604	0.136	—	—	—
P6	A	G	chr 17	7578275	140	9376	1.493	22	7897	0.279	—	—	—
P6*	KIF 5B-ALK	—	chr 10/ chr 2	—	28	15006	0.187/ 1.56	2	9989	0.02/ 0.167	—	—	—
P9	EML 4-ALK	—	chr 2/ chr 2	—	0	10688	0.000	0	13647	0.000	0	13521	0.000
P9	FYN-ROS 1	—	chr 6/ chr 6	—	0	9261	0.000	0	6826	0.000	2	10693	0.019
P9	ROS-1-MKX	—	chr 6/ chr 10	—	10	8029	0.125	1	6485	0.015	13	9943	0.131
P12	T	C	chr 4	55973786	—	—	—	—	—	—			
P12	T	G	chr 6	117650296	—	—	—	—	—	—			
P12	G	T	chr 7	41729291	—	—	—	—	—	—			
P12	T	A	chr 9	8471102	—	—	—	—	—	—			

TABLE 5-continued

Case	Mutant allele	Ref. allele	Chr	Position	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)
P12	G	T	chr 12	25380276	—	—	—	—	—	—			
P12	A	C	chr 19	10602473	—	—	—	—	—	—			
P12	−C	T	chr 17	7577057	—	—	—	—	—	—			
P1	A	G	chr 1	156785560	—	—	—	—	—	—			
P1	T	G	chr 1	157806043	—	—	—	—	—	—			
P1	G	C	chr 1	248525206	—	—	—	—	—	—			
P1	C	T	chr 2	33500291	—	—	—	—	—	—			
P1	A	C	chr 4	55946307	—	—	—	—	—	—			
P1	G	A	chr 4	55963949	—	—	—	—	—	—			
P1	A	C	chr 4	55968672	—	—	—	—	—	—			
P1	C	T	chr 6	117642146	—	—	—	—	—	—			
P1	T	G	chr 9	8376700	—	—	—	—	—	—			
P1	T	C	chr 9	8733625	—	—	—	—	—	—			
P1	T	G	chr 10	43611663	—	—	—	—	—	—			
P1	T	G	chr 15	88522525	—	—	—	—	—	—			
P1	+G	C	chr 17	7578474	—	—	—	—	—	—			
P1	−A	G	chr 17	29552244	—	—	—	—	—	—			
P1	+T	C	chr 17	29553484	—	—	—	—	—	—			
P1	−T	C	chr 17	29592185	—	—	—	—	—	—			
P16	A	G	chr 1	156843429	—	—	—	—	—	—			
P16	T	C	chr 1	181708291	—	—	—	—	—	—			
P16	A	C	chr 1	248525326	—	—	—	—	—	—			
P16	A	C	chr 2	125530343	—	—	—	—	—	—			
P16	A	C	chr 2	212530083	—	—	—	—	—	—			
P16	C	T	chr 2	212587119	—	—	—	—	—	—			
P16	T	G	chr 4	55958900	—	—	—	—	—	—			
P16	C	T	chr 4	55962358	—	—	—	—	—	—			
P16	A	C	chr 4	55968588	—	—	—	—	—	—			
P16	G	A	chr 4	55970963	—	—	—	—	—	—			
P16	A	C	chr 4	55971241	—	—	—	—	—	—			
P16	T	G	chr 5	19473838	—	—	—	—	—	—			
P16	A	G	chr 5	112176654	—	—	—	—	—	—			
P16	T	G	chr 5	176520134	—	—	—	—	—	—			
P16	T	G	chr 7	11501543	—	—	—	—	—	—			
P16	A	C	chr 7	53103357	—	—	—	—	—	—			
P16	T	C	chr 7	116411990	—	—	—	—	—	—			
P16	A	C	chr 10	43606641	—	—	—	—	—	—			
P16	A	G	chr 11	534195	—	—	—	—	—	—			
P16	G	C	chr 11	108143456	—	—	—	—	—	—			
P16	A	C	chr 12	25398284	—	—	—	—	—	—			
P16	A	C	chr 13	48947619	—	—	—	—	—	—			
P16	T	C	chr 13	70314492	—	—	—	—	—	—			
P16	A	T	chr 13	70314809	—	—	—	—	—	—			
P16	C	G	chr 15	88472337	—	—	—	—	—	—			
P16	A	C	chr 17	7578132	—	—	—	—	—	—			
P16	+T	C	chr 2	212295977	—	—	—	—	—	—			
P16	−C	T	chr 19	1220638	—	—	—	—	—	—			
P17	T	G	chr 7	81386606	—	—	—	—	—	—			
P17	A	C	chr 12	25398285	—	—	—	—	—	—			
P13	T	C	chr 1	190067540	—	—	—	—	—	—			
P13	T	C	chr 5	45461969	—	—	—	—	—	—			
P13	G	C	chr 8	38276015	—	—	—	—	—	—			
P13	T	C	chr 15	88483904	—	—	—	—	—	—			
P13	T	C	chr 17	7577538	—	—	—	—	—	—			
P2	A	C	chr 2	50463926	—	—	—	—	—	—			
P2	G	A	chr 3	89457148	—	—	—	—	—	—			
P2	T	G	chr 3	89468286	—	—	—	—	—	—			
P2	T	A	chr 3	89480240	—	—	—	—	—	—			
P2	T	A	chr 4	66189669	—	—	—	—	—	—			
P2	T	G	chr 4	66242868	—	—	—	—	—	—			
P2	A	C	chr 5	176522747	—	—	—	—	—	—			
P2	C	T	chr 6	117648229	—	—	—	—	—	—			
P2	A	C	chr 12	78400637	—	—	—	—	—	—			
P2	T	G	chr 12	78400910	—	—	—	—	—	—			
P2	T	C	chr 17	7577551	—	—	—	—	—	—			
P2	T	G	chr 19	1207247	—	—	—	—	—	—			
P2	+A	C	chr 2	79314100	—	—	—	—	—	—			
P3	A	C	chr 17	7578253	—	—	—	—	—	—			
P14	C	A	chr 1	156841521	—	—	—	—	—	—			
P14	T	G	chr 3	89176334	—	—	—	—	—	—			
P14	A	G	chr 7	55249159	—	—	—	—	—	—			
P14	G	T	chr 7	55259515	—	—	—	—	—	—			
P14	T	C	chr 10	43607789	—	—	—	—	—	—			

TABLE 5-continued

Case	Mutant allele	Ref. allele	Chr	Position	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)	Mutant allele depth	Total depth	ctDNA (%)
P14	C	T	chr 17	7577545	—	—	—	—	—	—			
P14	T	C	chr 17	29553484	—	—	—	—	—	—			
P14	G	C	chr 19	1223125	—	—	—	—	—	—			
P15	T	G	chr 1	70226008	33	5346	0.617	124	6200	2.000			
P15	A	C	chr 1	144882833	23	9807	0.235	117	12719	0.920			
P15	A	C	chr 1	190203515	0	3870	0.000	0	3965	0.000			
P15	A	C	chr 1	248525334	27	4232	0.638	56	5397	1.038			
P15	A	C	chr 2	155157911	0	4146	0.000	0	4508	0.000			
P15	A	G	chr 2	212495103	6	1796	0.334	0	2025	0.00			
P15	T	G	chr 3	89528742	21	1741	1.206	25	2118	1.180			
P15	T	G	chr 4	55979517	84	6351	1.323	219	7158	3.060			
P15	A	C	chr 4	66189751	0	2590	0.000	0	2706	0.000			
P15	A	C	chr 4	66233002	10	759	1.318	0	852	0.000			
P15	A	C	chr 4	66233003	10	791	1.264	0	868	0.000			
P15	T	G	chr 4	66233146	24	4571	0.525	45	4578	0.983			
P15	A	C	chr 5	176523126	27	3904	0.692	111	4798	2.313			
P15	A	C	chr 5	176524647	0	4637	0.000	0	5864	0.000			
P15	A	C	chr 7	41729339	16	4749	0.337	27	5865	0.460			
P15	A	C	chr 8	87738607	1	847	0.118	12	1098	1.093			
P15	A	C	chr 8	113563115	4	3404	0.118	91	3470	2.622			
P15	A	C	chr 9	8528716	17	5373	0.316	85	6082	1.398			
P15	A	T	chr 9	138439735	1	4332	0.023	19	5349	0.355			
P15	A	C	chr 10	43608292	2	3998	0.050	52	4959	1.049			
P15	T	C	chr 10	43608755	4	5518	0.072	24	6410	0.374			
P15	A	C	chr 11	55135855	13	4702	0.276	105	5200	2.019			
P15	T	C	chr 12	25398284	2	3896	0.051	42	3951	1.063			
P15	T	C	chr 13	48954333	0	2591	0.000	0	2786	0.000			
P15	T	G	chr 13	48954451	24	2204	1.089	7	2633	0.266			
P15	+T	G	chr 17	29533514	5	1618	0.309	0	1481	0.000			
P4	T	C	chr 2	212248555	—	—	—	—	—	—			
P4	T	C	chr 12	25398281	—	—	—	—	—	—			
P5	T	C	chr 7	55249071	—	—	—	—	—	—			
P5	G	T	chr 7	55259515	—	—	—	—	—	—			
P5	A	G	chr 11	55135338	—	—	—	—	—	—			
P5	T	C	chr 17	7577097	—	—	—	—	—	—			
P6	A	G	chr 12	78400791	—	—	—	—	—	—			
P6	T	G	chr 12	129822187	—	—	—	—	—	—			
P6	A	G	chr 17	7578275	—	—	—	—	—	—			
P6*	KIF 5B-ALK	—	chr 10/ chr 2	—	—	—	—	—	—	—			
P9	EML 4-ALK	—	chr 2/ chr 2	—	0	9837	0.000	0	8667	0.000			
P9	FYN-ROS 1	—	chr 6/ chr 6	—	2	7700	0.026	0	7483	0.000			
P9	ROS 1-MKX	—	chr 6/ chr 10	—	2	6695	0.030	0	6186	0.000			

*By comparing to the mean fraction of SNV reporters in this tumor, the capture efficiency of this fusion was estimated to be 12%.
The % ctDNA of this fusion was therefore normalized by dividing it by 0.12.
ctDNA concentrations pre- and post-adjustment are shown separated by a forward slash.

TABLE 6

Chromosome	Start (bp)	End (bp)	Gene
chr3	178935997	178936122	P1K3CA
chr3	178951909	178952140	P1K3CA
chr17	7578369	7578555	TP53
chr17	7578176	7578289	TP53
chr17	7577018	7577155	TP53
chr17	7577498	7577608	TP53
chr10	8115700	8115987	GATA3
chr6	170871003	170871217	TBP
chr17	7579310	7579537	TP53
chr3	178921508	178921607	PIK3CA
chr10	8111435	8111561	GATA3
chr9	141107487	141107586	FAM157B

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr21	36252853	36253010	RUNX1
chr16	68862076	68862207	CDH1
chr3	178927973	178928126	PIK3CA
chr17	7573926	7574033	TP53
chr3	178916822	178916947	PIK3CA
chr16	68845586	68845763	CDH1
chr16	67070541	67070658	CBFB
chr16	68835595	68835782	CDH1
chr16	68844099	68844244	CDH1
chr2	46707798	46707897	TMEM247
chr10	89692779	89693004	PTEN
chr17	37880164	37880263	ERBB2

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chrX	135960073	135960245	RBMX
chr3	178938773	178938945	PIK3CA
chr5	56160564	56160762	MAP3K1
chr6	26031882	26032137	HIST1H3B
chr2	198266708	198266854	SF3B1
chr2	129075869	129075968	HS6ST1
chr12	115118686	115118896	TBX3
chr5	56176937	56177100	MAP3K1
chr2	110301729	110301828	SEPT10
chr19	49458943	49459090	BAX
chr16	68772199	68772314	CDH1
chr10	51853598	51853697	FAM21A
chr16	68855913	68856123	CDH1
chr16	68849435	68849649	CDH1
chr20	29632610	29632721	FRG1B
chr16	68842595	68842751	CDH1
chr14	23523714	23523882	CDH24
chr9	116187612	116187711	C9orf43
chr1	120611934	120612033	NOTCH2
chr17	7576839	7576938	TP53
chr16	68842326	68842470	CDH1
chr19	8130917	8131065	FBN3
chr3	152554179	152554351	P2RY1
chrX	48887845	48887953	TFE3
chr11	89774235	89774445	TRIM49C
chr6	74227784	74227974	EEF1A1
chr6	32551969	32552138	HLA-DRB1
chr10	105484023	105484122	SH3PXD2A
chr2	27717413	27717546	FNDC4
chr1	203154334	203154457	CHI3L1
chr12	970255	970354	WNK1
chr19	11577555	11577654	ELAVL3
chr10	123258008	123258119	FGFR2
chr5	56183204	56183347	MAP3K1
chr2	9989495	9989594	TAF1B
chr17	12011106	12011226	MAP2K4
chr7	100245061	100245160	ACTL6B
chr12	4479577	4479740	FGF23
chrX	49040087	49040186	PRICKLE3
chr1	7890009	7890108	PER3
chr19	55325382	55325489	KIR2DL4
chr6	26406256	26406425	BTN3A1
chr10	8105955	8106101	GATA3
chr12	112600810	112600909	HECTD4
chr6	32489735	32489834	HLA-DRB5
chr17	4875688	4875787	CAMTA2
chr12	6702256	6702394	CHD4
chr16	67645853	67646024	CTCF
chr2	70905836	70906015	ADD2
chr19	40383694	40383994	FCGBP
chr6	26123758	26124001	HIST1H2BC
chr9	78789961	78790208	PCSK5
chr8	12285063	12285251	FAM86B2
chr19	54745495	54745683	LILRA6
chr12	115120615	115120804	TBX3
chr7	150884014	150884269	ASB10
chr1	12979941	12980233	PRAMEF7
chr14	38060746	38061532	FOXA1
chr6	32549361	32549564	HLA-DRB1
chr16	15696435	15696534	KIAA0430
chrX	154133100	154133269	F8
chr11	2434050	2434149	TRPM5
chr20	61444567	61444666	OGFR
chr17	12016549	12016677	MAP2K4
chr5	67591053	67591152	PIK3R1
chr15	22077530	22077704	POTEB
chr7	100773700	100773848	SERPINE1
chrX	119292961	119293092	RHOXF2
chrX	31525397	31525570	DMD
chr1	27057726	27057935	ARID1A
chr7	37960219	37960318	EPDR1
chr1	230979428	230979527	C1orf198
chr19	55399501	55399643	FCAR
chr19	48945465	48945576	GRIN2D
chr21	36171597	36171759	RUNX1

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr21	36206715	36206875	RUNX1
chr16	15802633	15802732	MYH11
chr3	49412866	49413022	RHOA
chr1	170521253	170521403	GORAB
chr17	12032455	12032604	MAP2K4
chr11	64457862	64457961	NRXN2
chr13	27664011	27664110	USP12
chr19	41596308	41596469	CYP2A13
chr17	61619619	61619778	KCNH6
chr2	27356067	27356187	PREB
chr3	178917469	178917568	PIK3CA
chr2	36994265	36994429	VIT
chr16	1291453	1291623	TPSAB1
chr16	68857310	68857497	CDH1
chr12	12870830	12871245	CDKN1B
chr20	20269279	20269470	C20orf26
chr19	40376674	40377035	FCGBP
chr22	35661297	35661545	HMGXB4
chr17	77768442	77768692	CBX8
chr9	136131208	136131417	ABO
chr11	46563793	46564008	AMBRA1
chr20	48604441	48604540	SNAI1
chr7	31735082	31735235	PPP1R17
chr3	52402776	52402875	DNAH1
chr7	150718274	150718416	ATG9B
chr11	66335454	66335553	CTSF
chr1	245582880	245583047	KIF26B
chr16	68846037	68846166	CDH1
chr16	75690147	75690321	TERF2IP
chr6	25600807	25600906	LRRC16A
chr6	42897308	42897459	CNPY3
chr1	145323629	145323728	NBPF10
chr20	20033035	20033189	CRNKL1
chr4	126328145	126328244	FAT4
chr9	377062	377200	DOCK8
chr1	242253179	242253347	PLD5
chr4	144545278	144545443	FREM3
chr6	26250514	26250662	HIST1H3F
chr10	96058144	96058294	PLCE1
chr19	6183141	6183251	ACSBG2
chr1	159683726	159683856	CRP
chr19	4292605	4292733	TMIGD2
chr11	49204697	49204796	FOLH1
chrX	18822012	18822167	PPEF1
chr1	33099237	33099336	ZBTB8OS
chr19	1084236	1084345	HMHA1
chr12	133198140	133198306	P2RX2
chr8	70981417	70981516	PRDM14
chrX	114082632	114082731	HTR2C
chr12	130827125	130827224	PIWIL1
chrX	49071851	49071973	CACNA1F
chr2	98165868	98165967	ANKRD36B
chr6	25776821	25776982	SLC17A4
chr12	130648737	130648882	FZD10
chr20	58547016	58547178	CDH26
chr16	57760038	57760137	CCDC135
chr6	102124572	102124671	GRIK2
chr19	14748955	14749064	EMR3
chr10	49388900	49389051	FRMPD2
chr14	24035486	24035628	AP1G2
chrX	11162125	11162224	ARHGAP6
chr11	121000378	121000477	TECTA
chr19	16841999	16842098	NWD1
chr13	33638165	33638278	KL
chr22	46712077	46712236	GTSE1
chrX	119005874	119005976	NDUFA1
chr17	73490976	73491115	KIAA0195
chr13	36886455	36886614	SPG20
chr20	45867566	45867733	ZMYND8
chr9	2643252	2643404	VLDLR
chr1	29189399	29189498	OPRD1
chr5	175782615	175782744	KIAA1191
chr7	107823275	107823374	NRCAM
chr22	26868797	26868905	HPS4
chr1	54417811	54417910	LRRC42

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr3	41268698	41268843	CTNNB1
chr2	241631330	241631462	AQP12A
chr10	89653774	89653873	PTEN
chr6	167790033	167790192	TCP10
chr3	86010665	86010764	CADM2
chr2	85051080	85051179	TRABD2A
chr6	48035985	48036157	PTCHD4
chr19	40392454	40392803	FCGBP
chr11	57569454	57569631	CTNND1
chr21	41719609	41719831	DSCAM
chrX	91873396	91873743	PCDH11X
chr3	130116501	130116761	COL6A5
chr1	154841539	154842331	KCNN3
chr1	12939484	12939864	PRAMEF4
chr9	105767464	105767685	CYLC2
chr7	65705508	65705729	TPST1
chr7	151945139	151945695	MLL3
chr17	58260549	58260772	USP32
chr3	121825055	121825335	CD86
chr15	23684996	23686765	GOLGA6L2
chr3	105421039	105421268	CBLB
chr9	84202595	84202743	TLE1
chr4	151223776	151223944	LRBA
chr18	30260131	30260290	KLHL14
chr19	52521620	52521747	ZNF614
chr19	40741890	40741989	AKT2
chr20	45188680	45188779	SLC13A3
chr1	115222989	115223088	AMPD1
chr4	151829485	151829619	LRBA
chr6	79577309	79577408	IRAK1BP1
chr6	37619915	37620077	MDGA1
chr2	197709223	197709322	PGAP1
chr8	28595035	28595180	EXTL3
chr5	195143	195277	LRRC14B
chr5	56161166	56161283	MAP3K1
chr16	29821397	29821552	MAZ
chr2	138000026	138000142	THSD7B
chr1	171154852	171154984	FMO2
chr17	10312638	10312804	MYH8
chr7	103051866	103052033	SLC26A5
chrX	107396857	107396956	ATG4A
chr22	19127367	19127537	DGCR14
chr12	99074054	99074180	APAF1
chr7	106526579	106526737	PIK3CG
chr7	128658008	128658107	TNPO3
chr16	20575995	20576156	ACSM2B
chr12	4554415	4554554	FGF6
chr19	36584932	36585066	WDR62
chr2	25978885	25978984	ASXL2
chr1	234565965	234566064	TARBP1
chr1	17266447	17266546	CROCC
chr12	6091076	6091175	VWF
chr13	31903637	31903805	B3GALTL
chr19	48342911	48343010	CRX
chr19	41743890	41743989	AXL
chr5	67589536	67589662	PIK3R1
chr22	17978441	17978581	CECR2
chrX	128631870	128632008	SMARCA1
chr12	7343047	7343152	PEX5
chr4	17585105	17585265	LAP3
chr20	3209483	3209657	SLC4A11
chr7	100187263	100187416	FBXO24
chr16	29994054	29994222	TAOK2
chrX	29935580	29935713	IL1RAPL1
chr20	42965923	42966022	R3HDML
chr2	27601435	27601551	ZNF513
chr1	155265227	155265359	PKLR
chr6	32634275	32634384	HLA-DQB1
chr6	27805734	27806085	HIST1H2AK
chr7	21639511	21639694	DNAH11
chr21	37444932	37445118	CBR1
chr5	56177409	56178674	MAP3K1
chr17	63221254	63221455	RGS9
chrX	8434191	8434393	VCX3B
chr2	89277989	89278194	IGKV3-7

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chrX	131762516	131762943	HS6ST2
chr12	57389411	57389645	GPR182
chr2	130831830	130833037	POTEF
chr2	132021022	132022032	POTEE
chr17	80159566	80159706	CCDC57
chr4	154191481	154191580	TRIM2
chr15	91424577	91424680	FURIN
chr9	34655582	34655681	IL11RA
chr9	16552595	16552754	BNC2
chr3	53217129	53217228	PRKCD
chr17	61570811	61570910	ACE
chr17	29483000	29483144	NF1
chr6	74072452	74072621	KHDC3L
chr14	24040434	24040654	JPH4
chr2	27248450	27248605	MAPRE3
chr16	67100584	67100701	CBFB
chr16	28506484	28509154	APOBR
chr2	129025859	129026228	HS6ST1
chr13	39587206	39587683	PROSER1
chr2	90078036	90078271	IGKV3D-20
chr2	234680925	234681080	UGT1A4
chrX	70472830	70472964	ZMYM3
chr1	68624805	68624930	WLS
chr2	25967089	25967232	ASXL2
chr16	68847224	68847371	CDH1
chrX	21450737	21450903	CNKSR2
chr12	122242643	122242817	SETD1B
chr19	51628888	51629053	SIGLEC9
chr22	37962637	37962797	CDC42EP1
chr17	56056512	56056674	VEZF1
chr4	1388323	1389290	CRIPAK
chr2	21236080	21236261	APOB
chrX	38144854	38146403	RPGR
chrX	119004943	119005377	RNF113A
chr12	11244066	11244726	TAS2R43
chr5	26881368	26881733	CDH9
chr7	19184660	19184944	FERD3L
chr5	56170879	56171089	MAP3K1
chrX	151899871	151900721	MAGEA12
chr1	240370193	240371705	FMN2
chrX	99661924	99663462	PCDH19
chr6	26043522	26043738	HIST1H2BB
chr19	55450429	55451645	NLRP7
chr12	125396334	125398031	UBC
chr14	69256737	69257170	ZFP36L1
chr14	74042018	74042190	ACOT2
chr9	69421915	69422014	ANKRD20A4
chr19	49913009	49913132	CCDC155
chr5	78181482	78181581	ARSB
chr7	45120238	45120361	NACAD
chrX	65819448	65819550	EDA2R
chr1	8421429	8421528	RERE
chr15	43701211	43701310	TP53BP1
chr5	56155571	56155727	MAP3K1
chr6	27776208	27776370	HIST1H2AI
chr1	233136089	233136234	PCNXL2
chr17	56386382	56386481	BZRAP1
chr17	53398052	53398151	HLF
chrX	12627839	12628000	FRMPD4
chr12	7527906	7528005	CD163L1
chr6	86332253	86332355	SYNCRIP
chr6	33263911	33264010	RGL2
chr13	37569561	37569733	ALG5
chr1	67242935	67243067	TCTEX1D1
chr10	72604229	72604395	SGPL1
chr16	68867202	68867354	CDH1
chr10	99379269	99379410	MORN4
chr1	150917574	150917673	SETDB1
chr5	139192984	139193083	PSD2
chr20	31041506	31041605	C20orf112
chr16	4862086	4862249	GLYR1
chr6	32551966	32552065	HLA-DRB1
chr6	28093419	28093518	ZSCAN16
chr1	26608848	26608947	UBXN11
chr9	19096669	19096768	HAUS6

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr7	128491508	128491682	FLNC
chr12	25398207	25398318	KRAS
chr1	247320232	247320337	ZNF124
chr13	25021153	25021324	PARP4
chr2	159481539	159481710	PKP4
chr3	178922283	178922382	PIK3CA
chr1	170695376	170695531	PRRX1
chr2	169996960	169997059	LRP2
chr10	89624216	89624315	PTEN
chr10	30653807	30653906	MTPAP
chr9	95277169	95277330	ECM2
chr2	166245337	166246185	SCN2A
chr6	26056018	26056591	HIST1H1C
chr16	26147093	26147546	HS3ST4
chr6	13977497	13978084	RNF182
chr6	27115003	27115268	HIST1H2AH
chrX	34960975	34962645	FAM47B
chr6	26216531	26216717	HIST1H2BG
chr12	52699842	52700028	KRT86
chr14	60938268	60938455	C14orf39
chr2	234652186	234652467	DNAJB3
chrX	134427657	134427940	ZNF75D
chr7	86415633	86415917	GRM3
chr11	118770650	118770853	BCL9L
chr1	190067487	190068154	FAM5C
chr17	16335334	16335540	TRPV2
chr9	27948860	27950484	LINGO2
chr6	29454651	29455669	MAS1L
chr17	40714738	40715280	COASY
chr10	46998994	47000218	GPRIN2
chr2	240982115	240982328	PRR21
chr15	20739674	20740539	GOLGA6L6
chr1	149858553	149858867	HIST2H2AC
chr12	46245782	46246211	ARID2
chr20	54961338	54961557	AURKA
chr2	165550902	165551967	COBLL1
chr19	38102565	38104062	ZNF540
chr12	81111100	81111321	MYF5
chrX	48418496	48419227	TBC1D25
chr1	149857850	149858181	HIST2H2BE
chr6	31238879	31239110	HLA-C
chr15	83926260	83926491	BNC1
chr3	130095170	130095628	COL6A5
chr17	39637092	39637327	KRT35
chr7	72412439	72414041	POM121
chr16	67644735	67645508	CTCF
chr1	235345090	235345864	ARID4B
chr17	262972	263748	C17orf97
chrX	149638545	149639506	MAMLD1
chr6	33169118	33169361	SLC39A7
chr21	47664836	47665081	MCM3AP
chr6	167754301	167754657	TTLL2
chr1	176863700	176863947	ASTN1
chr13	92345526	92346013	GPC5
chr21	40570750	40571559	BRWD1
chr9	138395416	138395776	MRPS2
chrX	148037135	148037947	AFF2
chrX	111155633	111155994	TRPC5
chr19	37879509	37880727	ZNF527
chr11	46419037	46419290	AMBRA1
chr12	125834061	125834885	TMEM132B
chr11	58919837	58920661	FAM111A
chr7	26224165	26225183	NFE2L3
chr1	156843433	156843687	NTRK1
chr2	80529541	80530775	LRRTM1
chrX	30260295	30261122	MAGEB4
chr6	114378676	114379176	HS3ST5
chr12	11546089	11546743	PRB2
chr2	160035346	160035601	TANC1
chr15	86807529	86808033	AGBL1
chr17	21318697	21319944	KCNJ12
chr1	176564440	176564697	PAPPA2
chr3	56667119	56667625	FAM208A
chrX	48681326	48681991	HDAC6
chr19	36673393	36674068	ZNF565

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr12	5020628	5021917	KCNA1
chr17	38643341	38643607	TNS4
chr9	121929580	121930447	DBC1
chr20	50768781	50769650	ZFP64
chr2	145147115	145147384	ZEB2
chr14	107012976	107013246	IGHV3-49
chrX	99551395	99551785	PCDH19
chr15	33954548	33954820	RYR3
chr5	90086904	90087074	GPR98
chr11	61093081	61093180	DDB1
chrX	10176284	10176394	CLCN4
chr1	151261057	151261156	ZNF687
chr16	61689465	61689593	CDH8
chr1	78478781	78478899	DNAJB4
chr7	132937850	132937949	EXOC4
chr17	16029395	16029520	NCOR1
chr19	54677851	54678003	MBOAT7
chr16	56782198	56782316	NUP93
chr1	181726104	181726203	CACNA1E
chr1	186014822	186014958	HMCN1
chr14	74967586	74967732	LTBP2
chr19	55598889	55598988	EPS8L1
chr16	22337427	22337526	POLR3E
chr3	180685863	180686032	FXR1
chrX	51238892	51238991	NUDT11
chr21	41710061	41710186	DSCAM
chr6	161071369	161071529	LPA
chr5	140960310	140960450	DIAPH1
chr3	51417547	51417646	DOCK3
chr1	200973893	200974061	KIF21B
chr21	28315698	28315866	ADAMTS5
chr10	105207115	105207214	CALHM2
chr10	28905130	28905247	WAC
chr20	1961171	1961313	PDYN
chr19	36303266	36303419	PRODH2
chr11	60183853	60184022	MS4A14
chr10	100503639	100503813	HPSE2
chr17	18205577	18205748	TOP3A
chr2	66798407	66798506	MEIS1
chr1	169578746	169578897	SELF
chr20	47601265	47601377	ARFGEF2
chr1	33745911	33746010	ZNF362
chrX	2779578	2779693	GYG2
chr19	18652622	18652721	FKBP8
chr17	8158786	8158885	PFAS
chr12	120634575	120634737	RPLP0
chr9	139357442	139357555	SEC16A
chr2	233785115	233785250	NGEF
chr4	190874201	190874300	FRG1
chr17	8224159	8224311	ARHGEF15
chr11	63670086	63670185	MARK2
chr4	147724635	147724766	TTC29
chrX	135592240	135592378	HTATSF1
chr22	31487660	31487833	SMTN
chr19	11287290	11287450	KANK2
chr6	28403762	28403873	ZSCAN23
chr19	33470934	33471065	RHPN2
chr2	204820384	204820527	ICOS
chr12	72338080	72338179	TPH2
chr7	100284271	100284442	GIGYF1
chr16	2134563	2134662	TSC2
chr17	74943920	74944019	MGAT5B
chr7	27148011	27148110	HOXA3
chr1	6257711	6257816	RPL22
chr8	2975909	2976008	CSMD1
chr5	56167736	56167858	MAP3K1
chr5	56168458	56168557	MAP3K1
chr5	56174806	56174928	MAP3K1
chr5	56181758	56181890	MAP3K1
chr19	13370377	13370515	CACNA1A
chr7	142124195	142124360	TRBV6-8
chr2	90139365	90139530	IGKV1D-16
chr1	13474848	13474984	PRAMEF18
chr4	46967043	46967142	GABRA4
chr11	116827663	116827780	SIK3

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	28890297	28890396	TBC1D29
chr2	80773031	80773188	CTNNA2
chr15	41099875	41100006	ZFYVE19
chr7	91779843	91780009	LRRD1
chr1	155290250	155290349	FDPS
chr9	3346665	3346764	RFX3
chr9	97873812	97873911	FANCC
chr11	49053333	49053432	TRIM49B
chr1	43296635	43296768	ERMAP
chr5	32233878	32234040	MTMR12
chr19	14876469	14876616	EMR2
chr9	2729501	2729632	KCNV2
chr1	120484228	120484368	NOTCH2
chr3	108723918	108724024	MORC1
chr12	41419004	41419118	CNTN1
chr12	115115373	115115472	TBX3
chr8	48887307	48887473	MCM4
chr22	19121811	19121973	DGCR14
chr11	68305213	68305349	PPP6R3
chr1	176926813	176926964	ASTN1
chr21	40834342	40834441	SH3BGR
chr12	130184678	130184777	TMEM132D
chr1	19464531	19464665	UBR4
chr6	127648146	127648289	ECHDC1
chr1	160279965	160280064	COPA
chr12	28114824	28114930	PTHLH
chr11	119210188	119210296	C1QTNF5
chr12	130833861	130833960	PIWIL1
chrX	10535214	10535386	MID1
chr12	21168622	21168721	SLCO1B7
chr5	154173388	154173559	LARP1
chr12	6344636	6344735	CD9
chr17	61557129	61557273	ACE
chr7	130023231	130023333	CPA1
chr6	39507793	39507967	KIF6
chr2	198267360	198267494	SF3B1
chr17	11597184	11597315	DNAH9
chr2	74763874	74763973	LOXL3
chr11	62381006	62381105	ROM1
chr19	33098607	33098732	ANKRD27
chr11	6452419	6452518	HPX
chr12	54645831	54645967	CBX5
chr1	149871794	149871947	BOLA1
chr12	1740510	1740609	WNT5B
chr9	113637768	113637891	LPAR1
chr7	128852210	128852309	SMO
chr17	73774671	73774804	H3F3B
chr8	48811029	48811129	PRKDC
chr12	111923516	111923669	ATXN2
chr12	130648737	130648882	FZD10
chr17	67190035	67190134	ABCA10
chr12	18852727	18852884	PLCZ1
chr17	60023828	60023961	MED13
chr1	26885297	26885428	RPS6KA1
chr19	1783032	1783131	ATP8B3
chr12	111321893	111322028	CCDC63
chr2	15374720	15374819	NBAS
chr2	220494023	220494122	SLC4A3
chr1	2303919	2304030	MORN1
chr1	16891301	16891413	NBPF1
chrX	114242494	114242639	IL13RA2
chr1	212911795	212911894	NSL1
chr20	58559693	58559860	CDH26
chrX	122528818	122528980	GRIA3
chr7	97833266	97833437	LMTK2
chr12	66531836	66531938	TMBIM4
chr22	41752346	41752480	ZC3H7B
chr11	46917434	46917569	LRP4
chr1	151509205	151509369	CGN
chr7	143055976	143056091	FAM131B
chr1	45288144	45288243	PTCH2
chr10	94653105	94653277	EXOC6
chr11	74880240	74880339	SLCO2B1
chr1	153043147	153043246	SPRR2B
chr18	66354903	66355002	TMX3

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	37868180	37868300	ERBB2
chr3	176769248	176769347	TBL1XR1
chr19	55107146	55107252	LILRA1
chrX	117570664	117570787	WDR44
chr8	80677449	80677555	HEY1
chr5	67589148	67589270	PIK3R1
chr1	160769621	160769720	LY9
chr12	100660697	100660854	DEPDC4
chr17	74623496	74623665	ST6GALNAC1
chr6	135511265	135511400	MYB
chr6	44224078	44224233	SLC35B2
chr20	30534289	30534388	PDRG1
chr17	66871754	66871874	ABCA8
chr8	103284778	103284938	UBR5
chr17	59557505	59557604	TBX4
chrX	47500669	47500827	ELK1
chr17	62892221	62892320	LRRC37A3
chr19	51323154	51323291	KLK1
chr15	71952870	71952969	THSD4
chr1	116280844	116280956	CASQ2
chr1	113616169	113616268	LRIG2
chr19	40368617	40368716	FCGBP
chr20	18429620	18429719	DZANK1
chr3	31725366	31725492	OSBPL10
chr3	31871578	31871702	OSBPL10
chr3	101572096	101572247	NFKBIZ
chr9	15489983	15490122	PSIP1
chr3	115395121	115395258	GAP43
chr12	20806921	20807085	PDE3A
chr1	107691295	107691450	NTNG1
chr11	126136657	126136817	SRPR
chr16	70595532	70595687	SF3B3
chr6	4943855	4943954	CDYL
chr16	29472706	29472854	SULT1A4
chr4	71500187	71500286	ENAM
chr4	100521721	100521890	MTTP
chr11	289843	289955	ATHL1
chr16	28913577	28913676	ATP2A1
chr15	38614441	38614610	SPRED1
chr1	16265790	16265922	SPEN
chrX	39922947	39923046	BCOR
chr1	12405430	12405566	VPS13D
chr12	53041956	53042121	KRT2
chr2	108479164	108479276	RGPD4
chr6	35108523	35108661	TCP11
chr12	108603943	108604056	WSCD2
chr8	104709325	104709424	RIMS2
chr5	129243892	129243991	CHSY3
chr13	24860362	24860472	SPATA13
chrX	48672846	48672973	HDAC6
chr5	37169183	37169282	C5orf42
chrX	74296356	74296489	ABCB7
chr17	26101296	26101431	NOS2
chr10	90537855	90537957	LIPN
chr2	198363398	198363572	HSPD1
chr17	73100131	73100285	SLC16A5
chr20	25755848	25755947	FAM182B
chr15	25966885	25966984	ATP10A
chr9	12702270	12702442	TYRP1
chr9	35616075	35616246	CD72
chr1	44134854	44134953	KDM4A
chr2	1926144	1926291	MYT1L
chr12	91371888	91371987	EPYC
chr15	43668295	43668424	TUBGCP4
chr3	151107766	151107923	MED12L
chr12	13529164	13529263	C12orf36
chr19	47492800	47492932	ARHGAP35
chrX	134185955	134186116	FAM127B
chr5	137289941	137290040	FAM13B
chr20	61907831	61908003	ARFGAP1
chr5	14358286	14358456	TRIO
chr4	1838155	1838299	LETM1
chr2	99634662	99634812	TSGA10
chr10	43597800	43597900	RET
chr3	148871280	148871435	HPS3

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chrX	114524321	114524420	LUZP4
chr12	57498952	57499095	STAT6
chr3	112710096	112710195	GTPBP8
chr3	178937358	178937523	PIK3CA
chr1	149939345	149939444	OTUD7B
chr6	76640678	76640798	IMPG1
chr2	71839770	71839936	DYSF
chr15	75111492	75111633	LMAN1L
chr1	170695408	170695542	PRRX1
chr7	120496734	120496833	TSPAN12
chr1	51767863	51767962	TTC39A
chr15	101447325	101447483	ALDH1A3
chr1	29609284	29609432	PTPRU
chr15	28769084	28769183	GOLGA8G
chr14	64580037	64580136	SYNE2
chr6	26217292	26217391	HIST1H2AE
chr19	49982165	49982304	F1T3LG
chrX	130409472	130409571	IGSF1
chr1	11317096	11317206	MTOR
chr1	206611313	206611448	SRGAP2
chr17	41931250	41931349	CD300LG
chr19	10781687	10781835	ILF3
chr6	131925317	131925460	MED23
chr3	184035081	184035180	EIF4G1
chrX	85403969	85404068	DACH2
chr1	215408279	215408415	KCNK2
chr15	83523395	83523552	HOMER2
chr18	14850212	14850381	ANKRD30B
chr4	173961083	173961251	GALNTL6
chr9	123888015	123888114	CNTRL
chr1	175067599	175067698	TNN
chr7	73279501	73279649	WBSCR28
chr7	100170019	100170193	SAP25
chr12	89818981	89819119	POC1B
chr8	53038606	53038705	ST18
chr13	67205357	67205532	PCDH9
chr16	1129032	1129207	SSTR5
chr20	50400809	50400984	SALL4
chr12	69656160	69656335	CPSF6
chr2	43452473	43452871	ZFP36L2
chr17	66246372	66246549	AMZ2
chr12	56478825	56479002	ERBB3
chr17	15964870	15965148	NCOR1
chr12	76424349	76425063	PHLDA1
chr20	2774880	2775058	CPXM1
chr12	112460033	112460211	ERP29
chrX	107018375	107018553	TSC22D3
chrX	23397728	23398007	PTCHD1
chr16	28884770	28885050	SH2B1
chr15	42052535	42052714	MGA
chr19	12154700	12154982	ZNF878
chr6	90660210	90661582	BACH2
chr22	17450867	17451048	GAB4
chr3	36484913	36485095	STAC
chr21	40794924	40795106	LCA5L
chr14	52186773	52187058	FRMD6
chr14	21215830	21216115	EDDM3A
chr1	197479778	197480064	DENND1B
chr6	75892983	75893167	COL12A1
chr1	240656325	240656741	GREM2
chr19	53793013	53793430	BIRC8
chr3	38991613	38991798	SCN11A
chr17	16326826	16327011	TRPV2
chrX	17750085	17750270	NHS
chr19	814467	814653	LPPR3
chrX	118284279	118284465	KIAA1210
chr8	88885139	88886088	DCAF4L2
chrX	125685469	125686221	DCAF12L1
chr22	22730662	22730850	IGLV5-45
chr11	125325767	125325955	FEZ1
chr16	3293330	3293518	MEFV
chr2	202149564	202149752	CASP8
chr5	153149726	153149915	GRIA1
chrX	147743619	147744201	AFF2
chr4	16504297	16504487	LDB2

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr20	41419913	41420104	PTPRT
chr4	122853548	122853848	TRPC3
chr19	51165615	51165807	SHANK1
chr7	100349542	100350751	ZAN
chr1	114225697	114226132	MAGI3
chr17	68171418	68172398	KCNJ2
chr11	120352006	120352199	ARHGEF12
chr20	31671212	31671649	BPIFB4
chr4	139980482	139980676	ELF2
chr16	62055070	62055265	CDH8
chr6	26188993	26189188	HIST1H4D
chr2	209025575	209025770	CRYGA
chr14	95053767	95053963	SERPINA5
chr5	140589609	140590840	PCDHB12
chr1	120458146	120458943	NOTCH2
chr2	166201097	166201297	SCN2A
chr12	10978186	10978500	TAS2R10
chr8	109796470	109797276	TMEM74
chr6	11190311	11191332	NEDD9
chr2	56144945	56145147	EFEMP1
chr1	160920835	160921038	ITLN2
chr5	118835029	118835233	HSD17B4
chr3	3189136	3189340	TRNT1
chr2	132288158	132288363	CCDC74A
chr3	48694416	48694739	CELSR3
chr12	53775932	53776139	SP1
chr17	76799655	76799862	USP36
chr12	5153646	5155078	KCNA5
chr3	196434455	196434663	CEP19
chr7	77789381	77789589	MAGI2
chr7	37780042	37780878	GPR141
chr6	154412131	154412458	OPRM1
chr19	52537524	52538587	ZNF432
chr16	396353	396826	AXIN1
chr14	72139080	72139290	SIPA1L1
chr16	9857874	9858522	GRIN2A
chr6	26199107	26199319	HIST1H2AD
chr2	90025216	90025428	IGKV2D-26
chr3	129389466	129389678	TMCC1
chr20	23016242	23017093	SSTR4
chr1	89448781	89449435	RBMXL1
chr20	896597	896810	ANGPT4
chr17	39645669	39645882	KRT36
chr16	15702157	15702370	KIAA0430
chr21	38884439	38884773	DYRK1A
chr7	128119301	128119515	METTL2B
chr20	5903618	5904478	CHGB
chr11	64627437	64627774	EHD1
chr19	58370284	58371379	ZNF587
chr1	19439144	19439360	UBR4
chr5	140580561	140581432	PCDHB11
chr19	51021545	51022418	LRRC4B
chr16	22926373	22926864	HS3ST2
chr14	95921719	95921937	SYNE3
chr17	46629395	46629738	HOXB3
chr9	5300143	5300363	RLN2
chr13	36049384	36050060	MAB21L1
chr14	94087992	94089111	UNC79
chr1	248039225	248039570	TRIM58
chr8	124195350	124195571	FAM83A
chr1	28920327	28920548	RAB42
chr12	129558460	129559468	TMEM132D
chrX	30872291	30873432	TAB3
chrX	5811008	5811361	NLGN4X
chr15	32929243	32929936	ARHGAP11A
chr6	78172232	78172742	HTR1B
chr3	121206757	121207555	POLQ
chrX	78216026	78216941	P2RY10
chr12	7045007	7046169	ATN1
chr6	26271218	26271576	HIST1H3G
chr19	8807879	8808586	ACTL9
chr1	206224465	206224827	AVPR1B
chr2	182542798	182543322	NEUROD1
chrX	17768049	17768340	SCML1
chr6	17637545	17637837	NUP153

TABLE 6-continued

Chromosome	Start (bp)	End (bp)	Gene
chr21	39086564	39087179	KCNJ6
chr14	106173557	106173791	IGHA1
chr17	38253388	38253622	NR1D1
chr11	96117311	96117840	CCDC82
chr12	16430302	16430537	SLC15A5
chr1	214170479	214171545	PROX1
chr19	15511969	15512206	AKAP8L
chr2	131414336	131414574	POTEJ
chr12	71977910	71978453	LGR5
chr7	82763886	82764264	PCLO
chr5	76028287	76029029	F2R
chr6	155450748	155451384	TIAM2
chr14	24845638	24845883	NFATC4
chr15	53907841	53908086	WDR72
chr13	108518035	108518788	FAM155A
chr12	47629615	47630076	PCED1B
chr19	51645627	51646011	SIGLEC7
chrX	77244949	77245411	ATP7A
chr7	126173022	126173578	GRM8
chr19	19906155	19906464	ZNF506
chrX	102931121	102931368	MORF4L2
chr4	25005572	25005819	LGI2
chr2	227872736	227872983	COL4A4
chrX	75003458	75004574	MAGEE2
chr7	108204867	108205255	THAP5
chr19	52217058	52217307	HAS1
chr9	139390622	139390871	NOTCH1
chr19	52888047	52888439	ZNF880
chr1	237947086	237948219	RYR2
chrX	30268638	30269645	MAGEB1
chrX	64721695	64722832	ZC3H12B
chr1	221912293	221913068	DUSP10
chr7	39503849	39504102	POU6F2
chr19	51273961	51274852	GPR32
chrX	12735732	12736884	FRMPD4
chrX	152225667	152226243	PNMA3
chr3	88039974	88040230	HTR1F
chr8	56435861	56436761	XKR4
chrX	155003546	155004222	SPRY3
chr17	26861800	26862057	FOXN1
chrX	68382801	68383058	PJA1
chr5	137680988	137681245	FAM53C
chr1	12942943	12943201	PRAMEF4
chr1	231344719	231344977	TRIM67
chr2	99013186	99013590	CNGA3
chr1	171251125	171251384	FMO1
chr7	96635419	96635681	DLX6
chr6	139487509	139487771	HECA
chr7	88423579	88424170	C7orf62
chr7	99956434	99956697	PILRB
chr2	133402800	133402997	GPR39
chr1	183511386	183511584	SMG7
chr12	56397549	56397814	SUOX
chr19	35232114	35232613	ZNF181
chr7	150171134	150171635	GIMAP8
chr7	75028333	75028600	TRIM73
chr1	25572974	25573241	C1orf63
chr22	39909830	39910166	SMCR7L
chr10	91198587	91198856	SLC16A12
chr20	61542180	61542889	DIDO1
chr20	50701236	50701661	ZFP64
chr3	13860452	13860792	WNT7A
chr9	111625372	111625798	ACTL7A
chr19	7676699	7677125	CAMSAP3
chrX	103080349	103080690	RAB9B
chrX	135593185	135594143	HTATSF1
chrX	112058601	112058874	AMOT
chr14	20019841	20020114	POTEM
chr2	239164300	239164505	PER2
chr6	153043014	153043357	MYCT1
chr11	209436	209711	RIC8A
chr2	51254719	51255150	NRXN1
chrX	118971733	118971941	UPF3B

TABLE 7

Chromosome	Start (bp)	End (bp)	Gene
chr12	25398207	25398318	KRAS
chr6	170870990	170871089	TBP
chr7	128587317	128587416	IRF5
chr9	96438892	96439020	PHF2
chr11	117789286	117789385	TMPRSS13
chr17	7577018	7577155	TP53
chr17	7578369	7578551	TP53
chr17	7577498	7577608	TP53
chr17	56833438	56833614	PPM1E
chr3	178935997	178936122	PIK3CA
chr17	7578176	7578289	TP53
chr12	132547047	132547146	EP400
chr7	140453074	140453193	BRAF
chr9	140918128	140918227	CACNA1B
chr21	46924329	46924470	COL18A1
chr18	48591824	48591932	SMAD4
chr5	112116486	112116600	APC
chr1	154841790	154842346	KCNN3
chr19	58549260	58549532	ZSCAN1
chr17	72350401	72350579	KIF19
chr19	39330909	39331008	HNRNPL
chr22	29885015	29886640	NEFH
chr3	41266058	41266157	CTNNB1
chr19	54754649	54754796	LILRB5
chr2	1271163	1271319	SNTG2
chr12	133219467	133219580	POLE
chr1	27100070	27100208	ARID1A
chr5	112173345	112179738	APC
chr9	12775812	12775911	LURAP1L
chr19	56599373	56599472	ZNF787
chr13	46170598	46171110	FAM194B
chr1	29138925	29139024	OPRD1
chr10	17659090	17659189	PTPLA
chr2	11810043	11810142	NTSR2
chr20	32664822	32664921	RALY
chr12	53068987	53069344	KRT1
chr14	93154359	93154541	RIN3
chr19	17932137	17932290	INSL3
chr6	16326657	16328230	ATXN1
chr20	46279801	46279900	NCOA3
chr1	85039985	85040084	CTBS
chr19	1064981	1065080	ABCA7
chr1	21044068	21044167	KIF17
chr2	187558955	187559054	FAM171B
chr17	6899436	6899571	ALOX12
chr7	130418475	130418574	KLF14
chr9	124855210	124855332	TTLL11
chr7	1586652	1586812	TMEM184A
chr8	143808950	143809194	THEM6
chr4	88535232	88537514	DSPP
chr1	228504471	228504671	OBSCN
chr11	320605	320806	IFITM3
chr20	44420643	44420748	DNTTIP1
chr17	74381511	74381610	SPHK1
chr19	2226674	2226773	DOT1L
chr15	66274640	66274739	MEGF11
chr16	84224917	84225016	ADAD2
chr16	31154139	31154238	PRSS36
chr7	6566298	6566397	GRID2IP
chr3	121351263	121351362	HCLS1
chr1	200880977	200881173	C1orf106
chr3	178916650	178916958	PIK3CA
chr2	98611944	98612043	TMEM131
chr19	17393464	17393570	ANKLE1
chr5	112128134	112128233	APC
chr20	60887455	60887588	LAMA5
chr16	602312	602512	SOLH
chr1	152487916	152488147	CRCT1
chr8	145001587	145001785	PLEC
chr13	28367011	28367110	GSX1
chr12	124824644	124824743	NCOR2
chr11	76751523	76751622	B3GNT6
chr17	40706742	40706907	HSD17B1
chr18	56887497	56887636	GRP
chr3	178951963	178952087	PIK3CA

TABLE 7-continued

Chromosome	Start (bp)	End (bp)	Gene
chr10	104159146	104159245	NFKB2
chr15	78441709	78441808	IDH3A
chr2	42275814	42275913	PKDCC
chr11	95825253	95826577	MAML2
chr19	56041254	56041623	SBK2
chrX	66765031	66766111	AR
chr19	58384471	58386127	ZNF814
chr1	26608827	26609017	UBXN11
chr8	144775907	144776528	ZNF707
chr16	24788422	24788646	TNRC6A
chr19	2732780	2733356	SLC39A3
chr17	36508384	36508582	SOCS7
chr3	51417547	51417646	DOCK3
chr19	15284978	15285087	NOTCH3
chr8	120220760	120220859	MAL2
chr15	60690041	60690140	ANXA2
chr16	15122734	15122889	PDXDC1
chr11	61658750	61658849	FADS3
chr19	4499590	4499689	HDGFRP2
chr19	17392865	17393018	ANKLE1
chr16	3304157	3304672	MEFV
chr20	43348541	43348751	WISP2
chr5	140214076	140216118	PCDHA7
chr13	111367954	111368317	ING1
chr13	32885653	32885906	ZAR1L
chr6	44243153	44243560	TMEM151B
chr17	4693053	4693343	GLTPD2
chr20	3732264	3732634	HSPA12B
chr17	39684144	39684438	KRT19
chr19	6737467	6737587	GPR108
chr19	49611231	49611330	SNRNP70
chr12	124829233	124829400	NCOR2
chr4	153249359	153249520	FBXW7
chr19	17448911	17449010	GTPBP3
chr8	145742795	145742894	RECQL4
chr20	590521	590620	TCF15
chr12	122242643	122242817	SETD1B
chr7	150037524	150037698	RARRES2
chr1	227922917	227923082	JMJD4
chr7	44924577	44924676	PURB
chr10	105110691	105110790	PCGF6
chr19	45867243	45867377	ERCC2
chr12	57619208	57619447	NXPH4
chr20	37377138	37377455	ACTR5
chr6	29910532	29910744	HLA-A
chr2	239049467	239050143	KLHL30
chr9	25677697	25677954	TUSC1
chr13	21562370	21563346	LATS2
chr2	39187172	39187520	ARHGEF33
chr18	3188779	3188977	MYOM1
chr22	20780023	20780297	SCARF2
chr6	53516875	53517036	KLHL31
chr19	36002347	36002446	DMKN
chr2	36825104	36825203	FEZ2
chr1	153907243	153907342	DENND4B
chr10	29760066	29760172	SVIL
chr22	29091697	29091861	CHEK2
chr3	150421508	150421607	FAM194A
chr20	44520189	44520288	CTSA
chr12	113376370	113376469	OAS3
chr12	122359394	122359516	WDR66
chr19	47768029	47768203	CCDC9
chr19	17337506	17337605	OCEL1
chr10	102988328	102988427	LBX1
chr2	148683599	148683730	ACVR2A
chr11	17035660	17035759	PLEKHA7
chrX	295101	295252	PPP2R3B
chr17	17119693	17119817	FLCN
chr5	112162804	112162944	APC
chr8	8860573	8860681	ERI1
chr10	85996984	85997269	LRIT1
chr7	2577780	2578372	BRAT1
chr6	29911106	29911320	HLA-A
chr19	41173536	41174022	NUMBL
chr19	40023093	40023309	EID2B

TABLE 7-continued

Chromosome	Start (bp)	End (bp)	Gene
chr19	48305145	48306174	TPRX1
chr16	20359830	20360505	UMOD
chr17	56435046	56435862	RNF43
chr1	155178610	155179012	MTX1
chr10	46998897	47000240	GPRIN2
chr19	1004686	1005532	GRIN3B
chr10	71905568	71906151	TYSND1
chr1	206680982	206681265	RASSF5
chr17	18918361	18918512	SLC5A10
chr7	139167933	139168064	KLRG2
chr19	49850446	49850620	TEAD2
chr4	3257543	3257642	MSANTD1
chr10	135186743	135186842	ECHS1
chr7	5372281	5372407	TNRC18
chr12	6777069	6777203	ZNF384
chr8	113240984	113241120	CSMD3
chr19	10679188	10679329	CDKN2D
chr19	984406	984555	WDR18
chr16	2059524	2059623	ZNF598
chr16	2059622	2059736	ZNF598
chr19	1789555	1789722	ATP8B3
chr1	175129889	175129988	KIAA0040
chr22	50920999	50921167	ADM2
chr7	1022847	1023021	CYP2W1
chr19	10431749	10431848	RAVER1
chr15	79092746	79092845	ADAMTS7
chr1	248020555	248020715	TRIM58
chr17	48433882	48433981	XYLT2
chr22	24121377	24121516	MMP11
chr12	25378547	25378707	KRAS
chr1	22149808	22149981	HSPG2
chr3	114057954	114058053	ZBTB20
chr15	102264303	102264477	TARSL2
chr6	160769761	160769860	SLC22A3
chr6	137113136	137113249	MAP3K5
chr16	88691009	88691153	ZC3H18
chr4	170678954	170679053	C4orf27
chr14	105267578	105268105	ZBTB42
chr4	1388323	1389466	CRIPAK
chr17	70119682	70120347	SOX9
chr15	100252709	100252893	MEF2A
chr11	44331308	44331531	ALX4
chr17	7579311	7579537	TP53
chr3	150127941	150128485	TSC22D2
chr2	95537567	95537796	TEKT4
chrX	54209386	54209576	FAM120C
chr19	58879172	58880386	ZNF837
chr22	19968871	19969107	ARVCF
chr20	48808010	48808450	CEBPB
chr12	7045137	7045925	ATN1
chr22	50615457	50616807	PANX2
chr5	140248963	140250986	PCDHA11
chr11	65810208	65811054	GAL3ST3
chr17	63533584	63533941	AXIN2
chr21	46929314	46929468	COL18A1
chr17	56448271	56448394	RNF43
chr8	144874504	144874603	SCRIB
chr8	145689544	145689660	CYHR1
chr3	56591226	56591325	CCDC66
chr12	124886949	124887107	NCOR2
chr1	204120808	204120953	ETNK2
chr9	138903634	138903747	NACC2
chr19	17622601	17622700	PGLS
chr18	34205515	34205642	FHOD3
chr19	50249868	50249967	TSKS
chr22	50921108	50921207	ADM2
chr17	48619220	48619319	EPN3
chr11	76751512	76751611	B3GNT6
chr16	84229435	84229581	ADAD2
chr19	49965140	49965293	ALDH16A1
chr19	51015392	51015547	ASPDH
chr2	241696750	241696849	KIF1A
chrX	153657038	153657199	ATP6AP1
chr20	49411648	49411747	BCAS4
chr8	145692341	145692493	KIFC2

TABLE 7-continued

Chromosome	Start (bp)	End (bp)	Gene
chr7	150498638	150498812	TMEM176A
chr5	112164552	112164669	APC
chr1	204228390	204228489	PLEKHA6
chr1	115258670	115258781	NRAS
chr4	113436024	113436123	NEUROG2
chr16	1820881	1820994	NME3
chr6	82461335	82461758	FAM46A
chr22	29837536	29837753	RFPL1
chr16	1270027	1270898	CACNA1H
chr3	126260607	126261395	CHST13
chr2	239009072	239009337	ESPNL
chr4	4228254	4228473	OTOP1
chr15	90320120	90320492	MESP2
chr2	56411816	56411994	CCDC85A
chr6	102503254	102503433	GRIK2
chr7	42003929	42006215	GLI3
chr22	20130457	20131117	ZDHHC8
chr19	7747292	7747622	TRAPPC5
chr1	17266400	17266587	CROCC
chr1	41976327	41976661	HIVEP3
chr17	59489706	59489894	C17orf82
chr19	17836780	17838754	MAP1S
chr14	77491801	77493810	IRF2BPL
chr10	134999542	135000160	KNDC1
chr5	24487851	24488260	CDH10
chr15	93588263	93588738	RGMA
chr3	122631701	122631897	SEMA5B
chr9	96051072	96051774	WNK2
chr2	171572939	171573733	SP5
chr11	44286426	44286625	ALX4
chr14	24040237	24040437	JPH4
chr6	74161445	74161693	MB21D1
chr9	4117863	4118590	GLIS3
chr5	53813829	53815535	SNX18
chr7	20824042	20824957	SP8
chrX	153688537	153688790	PLXNA3
chr8	88885042	88886058	DCAF4L2
chr12	5153619	5154540	KCNA5
chr19	31767495	31770449	TSHZ3
chr8	143694521	143695458	ARC
chr16	88599613	88601371	ZFPM1
chr8	144378009	144378869	ZNF696
chr15	65369394	65370354	KBTBD13
chr11	76750643	76751605	B3GNT6
chr12	53045562	53045778	KRT2
chr5	140228182	140230609	PCDHA9
chr16	87677885	87678577	JPH3
chr3	126733052	126733175	PLXNA1
chr19	622286	622385	POLRMT
chr22	38483130	38483271	BAIAP2L2
chr9	136918393	136918563	BRD3
chr1	8421091	8421204	RERE
chr1	6257711	6257816	RPL22
chr2	208633363	208633462	FZD5
chr7	75677461	75677560	MDH2
chr11	379584	379683	B4GALNT4
chr13	39425847	39425976	FREM2
chr19	44031239	44031338	ETHE1
chr2	202344754	202344898	STRADB
chr5	38407050	38407204	EGFLAM
chr2	211179634	211179766	MYL1
chr1	52306003	52306102	NRD1
chr19	14083711	14083810	RFX1
chr18	48604661	48604790	SMAD4
chr14	105070741	105070840	TMEM179
chr10	89692825	89692999	PTEN
chr10	89720678	89720824	PTEN
chr6	166571879	166572046	T
chr5	140174693	140176839	PCDHA2
chr11	63767113	63767235	MACROD1
chr6	110746108	110746285	SLC22A16
chr4	7043077	7044601	CCDC96
chr4	147560303	147560536	POU4F2
chr17	70118880	70119113	SOX9
chr8	77616518	77618658	ZFHX4

TABLE 7-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	79898713	79899611	MYADML2
chrX	50350756	50350945	SHROOM4
chrX	82763440	82764401	POU3F4
chr20	61443686	61444940	OGFR
chr4	24801299	24801573	SOD3
chr3	142840198	142841090	CHST2
chr12	53207441	53207638	KRT4
chr5	140262267	140264211	PCDHA13
chr9	139943392	139943527	ENTPD2
chr3	183951001	183951136	VWA5B2
chr2	46707801	46707900	TMEM247
chr1	152659327	152659480	LCE2B
chr2	87088917	87089016	CD8B
chr22	38051312	38051481	SH3BP1
chr11	6411896	6411995	SMPD1
chr17	260141	260300	C17orf97
chrX	110987946	110988045	ALG13
chr16	58549882	58549981	SETD6
chr19	51843758	51843857	VSIG10L
chr2	176957772	176957871	HOXD13
chr18	3452173	3452272	TGIF1
chrX	30326562	30327361	NR0B1
chr13	58298909	58299163	PCDH17
chr2	51254720	51255173	NRXN1
chr20	57766218	57769660	ZNF831
chr13	19751124	19751658	TUBA3C
chr19	48182629	48183772	GLTSCR1
chr1	237947095	237947554	RYR2
chr8	142367086	142368005	GPR20
chr10	124895626	124895884	HMX3
chr13	58206825	58209075	PCDH17
chr19	10224348	10224527	PPAN-P2RY11
chr5	176025233	176026162	GPRIN1
chr5	140515026	140517383	PCDHB5
chr5	140480344	140482621	PCDHB3
chr14	104641320	104644148	KIF26A
chr2	96780973	96781614	ADRA2B
chr2	226446656	226447604	NYAP2
chr20	43932953	43933349	MATN4
chrX	120008789	120009265	CT47B1
chr5	140207725	140209879	PCDHA6
chr8	77763206	77768391	ZFHX4
chr7	96653655	96653869	DLX5
chr12	108985546	108986113	TMEM119
chr8	98289177	98289987	TSPYL5
chr13	46287373	46288410	SPERT
chr5	140255094	140257222	PCDHA12
chr18	76753062	76754866	SALL3
chr2	1481012	1481232	TPO
chr16	30666088	30666368	PRR14
chr4	8582726	8583313	GPR78
chr22	38476923	38477343	SLC16A8
chr4	134071393	134073880	PCDH10
chr18	76753062	76755371	SALL3
chr7	53103553	53104199	POM121L12
chr12	110019199	110019355	MVK
chr1	117086970	117087119	CD58
chr4	140811098	140811206	MAML3
chr8	120429023	120429177	NOV
chr5	36035806	36035971	UGT3A2
chr2	74687408	74687551	WBP1
chr13	38320291	38320455	TRPC4
chr16	12009241	12009340	GSPT1
chr16	77246457	77246556	SYCE1L
chr20	6032926	6033034	LRRN4
chr1	55081692	55081845	FAM151A
chr12	122685078	122685207	LRRC43
chr11	108117690	108117854	ATM
chr17	5037181	5037291	USP6
chr7	102112900	102113056	LRWD1
chr3	139258468	139258567	RBP1
chr12	95044117	95044216	TMCC3
chr5	5239832	5239994	ADAMTS16
chr6	33263902	33264001	RGL2
chr1	17265510	17265609	CROCC

TABLE 7-continued

Chromosome	Start (bp)	End (bp)	Gene
chr19	1912910	1913009	ADAT3
chr8	11831510	11831609	DEFB136
chr16	230483	230582	HBQ1
chr6	166826249	166826375	RPS6KA2
chr10	126480292	126480402	METTL10
chr12	121432052	121432151	HNF1A
chr10	26446311	26446444	MYO3A
chr1	45671916	45672015	ZSWIM5
chr1	150530472	150530571	ADAMTSL4
chr4	8594554	8594653	CPZ
chr4	8603026	8603125	CPZ
chr3	129293178	129293333	PLXND1
chr4	5862760	5862884	CRMP1
chr1	15850563	15850695	CASP9
chr12	25380212	25380311	KRAS
chr19	54754728	54754827	LILRB5
chr15	26026180	26026312	ATP10A
chr15	42371702	42371801	PLA2G4D
chr14	29261265	29261364	C14orf23
chr7	87564340	87564501	ADAM22
chr16	2070132	2070231	NPW
chr9	135947042	135947141	CEL
chr9	133884777	133884876	LAMC3
chr19	41858871	41858970	TGFB1
chr12	53183933	53184032	KRT3
chr4	126237800	126242717	FAT4
chr4	57843294	57843729	NOA1
chr19	47548478	47548679	NPAS1
chr1	160062149	160062473	IGSF8
chr18	3456402	3456579	TGIF1
chr18	3456402	3456579	TGIF1
chr17	1359313	1359412	CRK
chr20	44642762	44642913	MMP9
chr19	47878845	47878944	DHX34
chr17	41133021	41133120	RUNDC1
chr1	47685454	47685632	TALI
chr19	48197450	48197892	GLTSCR1
chr10	27702255	27703028	PTCHD3
chr3	189526071	189526306	TP63
chr8	52320849	52322051	PXDNL
chr1	99470032	99470213	LPPR5
chr8	144997022	144999732	PLEC
chr15	69325531	69325630	NOX5
chr14	86087944	86089826	FLRT2
chr16	614762	615096	C16orf11
chr17	35300116	35300417	LHX1
chr2	220283206	220283444	DES
chr5	140572180	140574513	PCDHB10
chr2	1651970	1653391	PXDN
chr16	1840641	1842408	IGFALS
chr12	54379054	54379706	HOXC10
chr7	154862696	154863298	HTR5A
chr2	177036378	177036844	HOXD3
chr10	135012167	135012731	KNDC1
chr7	86415677	86416247	GRM3
chr7	43484121	43485149	HECW1
chr5	140557678	140559997	PCDHB8
chr5	140220991	140223330	PCDHA8
chr5	140753704	140756051	PCDHGA6
chr1	213031947	213032350	FLVCR1
chr8	10583340	10584034	SOX7
chr2	43451492	43452683	ZFP36L2
chr12	4479530	4479942	FGF23
chr17	3627472	3628884	GSG2
chr22	37964298	37964746	CDC42EP1
chr4	57180524	57182759	KIAA1211
chr1	117078658	117078762	CD58
chr11	124750401	124750500	ROBO3
chr11	64026609	64026708	PLCB3
chr16	88105674	88105818	BANP
chr19	5110698	5110797	KDM4B
chr11	76751543	76751642	B3GNT6
chr19	10407123	10407222	ICAM5
chr1	27621004	27621120	WDTC1
chr5	158630536	158630642	RNF145

TABLE 7-continued

Chromosome	Start (bp)	End (bp)	Gene
chr19	55815034	55815194	BRSK1
chr5	112769461	112770529	TSSK1B
chr22	18300931	18301135	MICAL3
chr17	21318662	21319944	KCNJ12
chr1	117122056	117122290	IGSF3
chr13	29598831	29600873	MTUS2
chr15	45007619	45007892	B2M
chr1	87045630	87045903	CLCA4
chr16	10788328	10788537	TEKT5

TABLE 8

Chromosome	Start (bp)	End (bp)	Gene
chr7	148508714	148508813	EZH2
chr6	134495648	134495770	SGK1
chr19	19260043	19260165	MEF2B
chr6	37138900	37139211	PIM1
chr7	2985452	2985590	CARD11
chr6	26234721	26234922	HIST1H1D
chr3	38182243	38182342	MYD88
chr6	26031980	26032147	HIST1H3B
chr6	27834958	27835057	HIST1H1B
chr19	10335365	10335542	S1PR2
chr6	26056101	26056498	HIST1H1C
chr18	60985340	60985897	BCL2
chr17	63049621	63049729	GNA13
chr12	49426498	49426597	MLL2
chr6	37138732	37138831	PIM1
chr6	37138342	37138441	PIM1
chr3	38182622	38182777	MYD88
chr15	45003728	45003827	B2M
chr6	26124500	26124827	HIST1H2AC
chr6	26156732	26157169	HIST1H1E
chr6	26250484	26250639	HIST1H3F
chr19	6586219	6586366	CD70
chr15	45007783	45007882	B2M
chr2	242066173	242066272	PASK
chr2	96809958	96810091	DUSP2
chr17	63052507	63052611	GNA13
chr17	7577018	7577155	TP53
chrX	113965789	113965942	HTR2C
chr1	120458108	120458207	NOTCH2
chr3	176750758	176750924	TBL1XR1
chr17	62006764	62006863	CD79B
chr14	80328148	80328247	NRXN3
chr5	89923411	89923541	GPR98
chr17	40951085	40951254	CNTD1
chr4	153249338	153249437	FBXW7
chr7	2963866	2963999	CARD11
chr12	92539163	92539311	BTG1
chr6	26158538	26158769	HIST1H2BD
chr6	27860546	27860875	HIST1H2AM
chr1	2489781	2489907	TNFRSF14
chr16	85936621	85936795	IRF8
chr6	26123760	26124023	HIST1H2BC
chr6	27100943	27101241	HIST1H2AG
chr6	27114217	27114519	HIST1H2BK
chr6	26045793	26046018	HIST1H3C
chr3	183273160	183273402	KLHL6
chr1	85733326	85733577	BCL10
chr17	63010422	63010942	GNA13
chr6	27100151	27100263	HIST1H2BJ
chr7	5569165	5569288	ACTB
chr3	187443286	187443417	BCL6
chr19	42599939	42600081	POU2F2
chr1	2488088	2488187	TNFRSF14
chr17	7578401	7578530	TP53
chr12	113496043	113496165	DTX1
chr11	128391798	128391897	ETS1
chr7	34724163	34724296	NPSR1
chr12	92537876	92538195	BTG1

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr8	122626696	122627104	HAS2
chr16	11348700	11349138	SOCS1
chrX	1584584	1585235	P2RY8
chr15	39544367	39544819	C15orf54
chr6	27861294	27861585	HIST1H2BO
chr8	114185958	114186078	CSMD3
chr8	57228764	57228900	SDR16C5
chr6	14118180	14118296	CD83
chr19	19261467	19261566	MEF2B
chr10	98781006	98781170	SLIT1
chr5	32090982	32091118	PDZD2
chr2	125555706	125555805	CNTNAP5
chr5	7414684	7414783	ADCY2
chr11	17482173	17482272	ABCC8
chr5	88119528	88119627	MEF2C
chr1	173819464	173819617	DARS2
chr1	181727082	181727247	CACNA1E
chr7	148506392	148506491	EZH2
chr1	117078701	117078800	CD58
chr1	117086988	117087131	CD58
chr7	82763869	82763975	PCLO
chr12	13769407	13769569	GRIN2B
chr5	145393394	145393518	SH3RF2
chr19	43766018	43766117	PSG9
chr20	25003575	25003728	ACSS1
chr11	60229847	60230006	MS4A1
chr11	89531416	89531515	TRIM49
chr8	101730371	101730470	PABPC1
chr15	66729083	66729230	MAP2K1
chr4	24544556	24544655	DHX15
chr16	3786650	3786816	CREBBP
chr6	134493799	134493912	SGK1
chr3	60522592	60522695	FHIT
chr1	9784333	9784479	PIK3CD
chr19	10934463	10934575	DNM2
chr15	26806082	26806181	GABRB3
chr17	7577498	7577608	TP53
chr5	112176808	112176907	APC
chr1	82408728	82408842	LPHN2
chr1	190195307	190195406	FAM5C
chr7	2977540	2977666	CARD11
chr11	118343087	118343186	MLL
chr3	16419284	16419420	RFTN1
chr6	27839714	27839833	HIST1H3I
chr11	49208195	49208321	FOLH1
chr11	18194889	18195049	MRGPRX4
chrX	102931279	102931380	MORF4L2
chr8	3141777	3141876	CSMD1
chr5	149677048	149677147	ARSI
chrX	70784450	70784603	OGT
chr3	38181907	38182033	MYD88
chr9	35800705	35800838	NPR2
chr19	21476425	21476524	ZNF708
chr16	85954792	85954891	IRF8
chr4	158257566	158257665	GRIA2
chr11	14899653	14899752	CYP2R1
chr18	30349821	30350141	KLHL14
chr22	23523625	23524360	BCR
chr9	4118465	4118649	GLIS3
chr5	124079896	124080638	ZNF608
chrX	92927664	92928269	NAP1L3
chr1	167096068	167096479	DUSP27
chr4	115997524	115997764	NDST4
chr6	27777852	27778102	HIST1H3H
chrX	86773014	86773267	KLHL4
chr7	138601542	138601795	KIAA1549
chr1	179562711	179562985	TDRD5
chr8	128750609	128751108	MYC
chr4	154624731	154625043	TLR2
chr1	149857823	149858147	HIST2H2BE
chr17	51900491	51900825	KIF2B
chr8	116616196	116616816	TRPS1
chr4	88583982	88584348	DMP1
chrX	41586526	41586894	GPR82
chr14	55241653	55241762	SAMD4A

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr8	85774530	85774688	RALYL
chr5	89949226	89949325	GPR98
chr7	91503615	91503714	MTERF
chr2	136872579	136872678	CXCR4
chr5	80643592	80643749	ACOT12
chr14	21897075	21897174	CHD8
chr22	41525893	41526007	EP300
chr4	126319938	126320070	FAT4
chr17	6012926	6013086	WSCD1
chr9	95085704	95085803	NOL8
chr2	11354943	11355042	ROCK2
chr1	59844415	59844514	FGGY
chr13	37401779	37401890	RFXAP
chr12	48190799	48190925	HDAC7
chr2	198353036	198353135	HSPD1
chr10	48428818	48428917	GDF10
chr17	26961625	26961724	KIAA0100
chr1	150915478	150915577	SETDB1
chr7	1527451	1527550	INTS1
chr3	93755496	93755595	ARL13B
chr1	7700459	7700613	CAMTA1
chr11	130784481	130784580	SNX19
chr2	1687837	1687936	PXDN
chrX	138886629	138886758	ATP11C
chr10	121677458	121677557	SEC23IP
chr16	58562378	58562552	CNOT1
chr2	75425942	75426041	TACR1
chr6	102337597	102337696	GRIK2
chr9	35376114	35376213	UNC13B
chr15	52529678	52529843	MYO5C
chr4	100784919	100785018	DAPP1
chrX	135288683	135288782	FHL1
chr3	50005082	50005181	RBM6
chr19	15366097	15366196	BRD4
chr3	183209816	183209915	KLHL6
chr3	183210322	183210468	KLHL6
chr21	35169764	35169863	ITSN1
chr12	66923602	66923701	GRIP1
chr8	68931783	68931906	PREX2
chr9	119202908	119203007	ASTN2
chr9	23701450	23701549	ELAVL2
chr5	121758997	121759096	SNCAIP
chr8	113303749	113303869	CSMD3
chr12	6439763	6439877	TNFRSF1A
chr2	141245185	141245308	LRP1B
chr2	141291589	141291709	LRP1B
chr2	142004794	142004923	LRP1B
chr10	22653781	22653948	SPAG6
chr12	119942896	119942995	CCDC60
chr10	115365936	115366041	NRAP
chr4	159634270	159634412	PPID
chr1	160319338	160319460	NCSTN
chr12	132683716	132683815	GALNT9
chr11	111715323	111715446	ALG9
chr18	28714581	28714715	DSC1
chr22	36661645	36661744	APOL1
chrX	125955246	125955345	CXorf64
chr18	21526107	21526248	LAMA3
chr7	21550781	21550880	SP4
chr8	124975517	124975638	FER1L6
chr8	124195469	124195568	FAM83A
chr1	91740276	91740375	HFM1
chr1	229772414	229772513	URB2
chr12	49943314	49943413	KCNH3
chr6	72006181	72006280	OGFRL1
chr13	32907254	32907353	BRCA2
chr17	41847111	41847210	DUSP3
chr8	99441265	99441364	KCNS2
chr4	85626546	85626664	WDFY3
chr4	85687017	85687116	WDFY3
chr4	85717707	85717806	WDFY3
chr12	57598409	57598532	LRP1
chr2	149528565	149528664	EPC2
chr2	122204912	122205083	CLASP1
chr11	66008985	66009084	PACS1

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	155458538	155458637	TIAM2
chr8	124664138	124664237	KLHL38
chr2	202264099	202264216	TRAK2
chr21	37833351	37833450	CLDN14
chr17	74276367	74276532	QRICH2
chr17	1563134	1563295	PRPF8
chr1	92470012	92470111	BRDT
chr16	14334156	14334255	MKL2
chr12	115120815	115120932	TBX3
chr12	108013890	108013989	BTBD11
chr6	152697629	152697728	SYNE1
chr8	110463284	110463383	PKHD1L1
chr5	32074455	32074554	PDZD2
chr15	65917821	65917920	SLC24A1
chr14	32615457	32615556	ARHGAP5
chr2	103148789	103148888	SLC9A4
chr5	79733658	79733757	ZFYVE16
chr14	92088143	92088242	CATSPERB
chr15	89056238	89056337	DET1
chr1	35857812	35857953	ZMYM4
chr6	38743648	38743747	DNAH8
chr2	125204372	125204471	CNTNAP5
chr2	125669029	125669128	CNTNAP5
chr5	36671219	36671318	SLC1A3
chr4	3419114	3419268	RGS12
chr8	110984837	110984940	KCNV1
chr11	64645602	64645701	EHD1
chr7	31378618	31378717	NEUROD6
chr8	35544062	35544227	UNC5D
chr17	33288569	33288668	ZNF830
chr19	37210027	37210126	ZNF567
chr4	187524795	187524894	FAT1
chr20	3321138	3321237	C20orf194
chr1	109795535	109795634	CELSR2
chr11	100863129	100863281	TMEM133
chr5	67591036	67591135	PIK3R1
chr9	37740424	37740523	FRMPD1
chrX	32663134	32663233	DMD
chr2	169781166	169781313	ABCB11
chr18	64239223	64239322	CDH19
chr8	623942	624041	ERICH1
chr9	82319697	82319817	TLE4
chr20	35812674	35812773	RPN2
chr14	35873721	35873820	NFKBIA
chr6	83838787	83838886	DOPEY1
chr2	73675936	73676035	ALMS1
chr11	73715528	73715630	UCP3
chr6	126210203	126210302	NCOA7
chr20	36963988	36964087	BPI
chr6	26252135	26252245	HIST1H2BH
chr2	69627576	69627675	NFU1
chr20	480476	480578	CSNK2A1
chr7	140453074	140453193	BRAF
chr11	7021848	7021947	ZNF214
chr18	32428253	32428352	DTNA
chr11	70271422	70271521	CTTN
chr15	50784917	50785016	USP8
chr3	164730749	164730848	SI
chr1	27105515	27105614	ARID1A
chr17	18001578	18001677	DRG2
chr11	125472697	125472843	STT3A
chr18	56390352	56390451	MALT1
chr4	186380380	186380479	CCDC110
chr1	160850943	160851102	ITLN1
chr5	131825077	131825176	IRF1
chr10	129868564	129868714	PTPRE
chr10	54527901	54528035	MBL2
chr2	171071238	171071338	MYO3B
chr18	3193815	3193956	MYOM1
chr1	1290159	1290330	MXRA8
chr3	2924828	2924931	CNTN4
chr5	52096606	52096705	PELO
chr10	90773913	90774012	FAS
chr13	25352432	25352544	RNF17
chr7	80285855	80286016	CD36

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr5	132084038	132084167	CCNI2
chr7	64439781	64439907	ZNF117
chr16	84089609	84089740	MBTPS1
chr19	39959395	39959501	SUPT5H
chr19	19576162	19576261	GATAD2A
chr4	155490817	155490916	FGB
chr4	66231649	66231775	EPHA5
chr1	111957552	111957666	OVGP1
chr6	105243453	105243560	HACE1
chr11	118770652	118770757	BCL9L
chr2	55756013	55756128	CCDC104
chr2	27319583	27319682	KHK
chr14	81743377	81743476	STON2
chr7	82784463	82784562	PCLO
chrX	53573396	53573553	HUWE1
chr2	200233327	200233430	SATB2
chr8	77762480	77762598	ZFHX4
chr1	37945890	37946030	ZC3H12A
chr1	37948734	37948833	ZC3H12A
chr5	19571723	19571822	CDH18
chr9	134497232	134497374	RAPGEF1
chr10	30747012	30747165	MAP3K8
chr2	27247017	27247116	MAPRE3
chr6	76385659	76385758	SENP6
chr2	79313492	79313630	REG1B
chr7	129806268	129806367	TMEM209
chr12	39688222	39688321	KIF21A
chr10	101841194	101841293	CPN1
chr17	40475021	40475161	STAT3
chr8	75898251	75898352	CRISPLD1
chr10	131640350	131640449	EBF3
chr7	14758166	14758310	DGKB
chr9	101904817	101904985	TGFBR1
chr3	100557041	100557140	ABI3BP
chr3	100604998	100605097	ABI3BP
chr19	58131754	58131853	ZNF134
chr3	146311834	146311933	PLSCR5
chr16	53503855	53503954	RBL2
chr1	154931304	154931403	PYGO2
chr6	80223202	80223301	LCA5
chr1	24840825	24840924	RCAN3
chr6	27277340	27277439	POM121L2
chr14	102822104	102822234	CINP
chr12	57496608	57496707	STAT6
chrX	153997441	153997585	DKC1
chr12	26553060	26553191	ITPR2
chr12	26755302	26755428	ITPR2
chr5	35047903	35048002	AGXT2
chr14	50889816	50889915	MAP4K5
chrX	154159899	154159998	F8
chr9	34635668	34635767	SIGMAR1
chr7	113558284	113558383	PPP1R3A
chr6	27799221	27799320	HIST1H4K
chr2	152518644	152518743	NEB
chr1	236718595	236718764	HEATR1
chr17	78343567	78343667	RNF213
chr7	122634962	122635061	TAS2R16
chr6	394881	394980	IRF4
chr5	137599979	137600078	GFRA3
chr2	189849534	189849633	COL3A1
chr1	185269130	185269262	IVNS1ABP
chr5	83259023	83259179	EDIL3
chr12	53900804	53900926	NPFF
chr1	231334792	231334915	TRIM67
chr17	5037181	5037291	USP6
chr3	151165871	151165970	IGSF10
chr19	55143390	55143489	LILRB1
chr6	26216767	26216866	HIST1H2BG
chr1	12785590	12785705	AADACL3
chrX	70612724	70612844	TAF1
chr15	91019894	91020050	IQGAP1
chr3	112324459	112324558	CCDC80
chr5	149631536	149631635	CAMK2A
chr17	50235066	50235165	CA10
chr4	36075310	36075445	ARAP2

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr15	99250974	99251073	IGF1R
chr14	65259812	65259911	SPTB
chr7	47944073	47944172	PKD1L1
chr21	34166539	34166638	C21orf62
chr3	173322726	173322825	NLGN1
chr10	25313261	25313360	THNSL1
chr1	201038599	201038729	CACNA1S
chr8	144990426	144990525	PLEC
chr13	28197172	28197271	POLR1D
chr12	41900352	41900451	PDZRN4
chr20	139395	139494	DEFB127
chr7	146997232	146997382	CNTNAP2
chr6	26443795	26443894	BTN3A3
chr16	30093780	30093879	PPP4C
chr10	22030840	22030939	MLLT10
chr15	44120405	44120504	WDR76
chr16	11076734	11076848	CLEC16A
chr6	49937259	49937358	DEFB113
chr7	127014541	127014640	ZNF800
chr3	37514844	37514951	ITGA9
chr5	140221244	140221343	PCDHA8
chr19	1055059	1055158	ABCA7
chr2	238275682	238275781	COL6A3
chr2	238280539	238280638	COL6A3
chr6	27782778	27782877	HIST1H2BM
chr16	72833925	72834028	ZFHX3
chr9	78686641	78686814	PCSK5
chr13	26620899	26620998	SHISA2
chr15	66727404	66727503	MAP2K1
chr5	21783466	21783603	CDH12
chr7	73950496	73950605	GTF2IRD1
chr7	92733518	92733617	SAMD9
chr20	57581376	57581540	CTSZ
chr1	116283348	116283449	CASQ2
chr22	50471719	50471818	TTLL8
chr7	75192479	75192578	HIP1
chr19	58965614	58965713	ZNF324B
chr11	31392295	31392406	DNAJC24
chr5	80369181	80369280	RASGRF2
chr8	116426513	116426636	TRPS1
chr8	116599420	116599519	TRPS1
chr20	32341030	32341129	ZNF341
chr21	28338441	28338573	ADAMTS5
chr10	105209455	105209554	CALHM2
chr16	29824386	29824485	PRRT2
chr14	54886703	54886802	CDKN3
chr2	116534779	116534878	DPP10
chr12	56397541	56397640	SUOX
chr1	151339198	151339297	SELENBP1
chr21	18981289	18981462	BTG3
chr3	196529887	196530035	PAK2
chrX	118540596	118540695	SLC25A43
chr20	48127564	48127716	PTGIS
chr20	3543855	3544010	ATRN
chr5	35709125	35709224	SPEF2
chr5	35807232	35807355	SPEF2
chr6	26199865	26199964	HIST1H2BF
chr2	160136377	160136476	WDSUB1
chr10	96014649	96014806	PLCE1
chr10	123987351	123987523	TACC2
chr6	41899465	41899568	BYSL
chr10	16996387	16996547	CUBN
chr7	122809280	122809379	SLC13A1
chr6	84925034	84925133	KIAA1009
chr12	15813547	15813674	EPS8
chr16	5041881	5041980	SEC14L5
chr2	48028009	48028108	MSH6
chr2	170735009	170735108	UBR3
chr2	234545387	234545561	UGT1A10
chr2	9770341	9770440	YWHAQ
chr1	12726644	12726743	AADACL4
chrX	119509339	119509438	ATP1B4
chr7	94740570	94740703	PPP1R9A
chr5	39138726	39138825	FYB
chr17	4007975	4008074	ZZEF1

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr12	111089106	111089205	HVCN1
chr22	32193585	32193689	DEPDC5
chr19	38996930	38997029	RYR1
chr1	1421489	1421615	ATAD3B
chr14	37154076	37154175	SLC25A21
chr3	140281652	140281798	CLSTN2
chr17	38447286	38447385	CDC6
chr6	51617998	51618151	PKHD1
chr10	21076130	21076237	NEBL
chr11	65108869	65109033	DPF2
chr18	52899739	52899902	TCF4
chrX	151819978	151820077	GABRQ
chrX	70347869	70347968	MED12
chr19	52537324	52537423	ZNF432
chr21	32638490	32638633	TIAM1
chr2	230861466	230861639	FBXO36
chr1	236966822	236966921	MTR
chrX	84526133	84526234	ZNF711
chr20	55966758	55966857	RBM38
chr4	7728506	7728630	SORCS2
chrX	153628143	153628282	RPL10
chr20	30681665	30681819	HCK
chr2	9514894	9514993	ASAP2
chr15	50223389	50223488	ATP8B4
chrX	140996390	140996491	MAGEC1
chr16	3788559	3788673	CREBBP
chr16	3808854	3808973	CREBBP
chr6	134491958	134492057	SGK1
chr6	134494403	134494502	SGK1
chr6	134494599	134494704	SGK1
chr6	134495130	134495229	SGK1
chr4	151817527	151817626	LRBA
chr3	23934688	23934787	NKIRAS1
chrX	13680790	13680889	TCEANC
chr19	15164540	15164639	CASP14
chr8	24813192	24813291	NEFL
chr12	122658390	122658539	IL31
chr6	70859719	70859818	COL19A1
chrX	119059299	119059398	NKAP
chr12	18800809	18800962	PIK3C2G
chr8	48777075	48777174	PRKDC
chr7	100172827	100172926	LRCH4
chr9	133948158	133948257	LAMC3
chr17	62006585	62006684	CD79B
chr13	114009637	114009796	GRTP1
chr6	73043453	73043552	RIMS1
chr3	187447106	187447205	BCL6
chr5	176522495	176522594	FGFR4
chr18	6311538	6311637	L3MBTL4
chr15	95001365	95001475	MCTP2
chr15	75798216	75798316	PTPN9
chr2	215843515	215843614	ABCA12
chr2	32865336	32865477	TTC27
chr3	27216096	27216195	NEK10
chr4	62813853	62813952	LPHN3
chr11	9597421	9597520	WEE1
chr6	106552825	106552924	PRDM1
chr3	107517429	107517528	BBX
chr10	128923737	128923865	DOCK1
chr13	111109686	111109785	COL4A2
chr3	122338609	122338708	PARP15
chr22	17690369	17690468	CECR1
chr4	83279811	83279973	HNRNPD
chr4	76572212	76572341	G3BP2
chr5	179201689	179201788	MAML1
chr3	123385065	123385193	MYLK
chr11	5529961	5530060	UBQLN3
chr11	57156049	57156181	PRG2
chr6	151673552	151673651	AKAP12
chr18	54547185	54547284	WDR7
chr8	15519664	15519805	TUSC3
chr3	196288280	196288379	WDR53
chr18	47101791	47101899	LIPG
chr19	56300172	56300343	NLRP11
chr9	86530434	86530533	KIF27

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr8	25715787	25715886	EBF2
chr22	41320365	41320486	XPNPEP3
chr2	170042198	170042297	LRP2
chr12	18891329	18891491	CAPZA3
chr1	223465866	223465965	SUSD4
chr1	2491261	2491417	TNFRSF14
chr6	17856257	17856356	KIF13A
chr8	86354301	86354420	CA3
chr1	94341859	94341958	DNTTIP2
chr2	177033872	177033971	HOXD3
chr2	128409047	128409146	GPR17
chr14	21269809	21269908	RNASE1
chr17	7579314	7579413	TP53
chr4	160274689	160274788	RAPGEF2
chr1	183498026	183498177	SMG7
chr7	105738160	105738259	SYPL1
chr10	118220477	118220597	PNLIPRP3
chr6	32943160	32943298	BRD2
chr19	8028461	8028560	ELAVL1
chr2	211542610	211542709	CPS1
chr10	103870285	103870458	LDB1
chrX	18528907	18529006	CDKL5
chr15	73067306	73067405	ADPGK
chr11	124524550	124524689	SIAE
chr14	47120706	47120805	RPL10L
chr12	32875343	32875442	DNM1L
chr15	41797166	41797265	LTK
chr18	44139410	44139565	LOXHD1
chr11	68480737	68480875	MTL5
chr1	62327222	62327339	INADL
chr14	73576049	73576200	RBM25
chr15	41384224	41384380	INO80
chrX	105152975	105153074	NRK
chr17	79478986	79479112	ACTG1
chr6	55659076	55659225	BMP5
chr19	1376496	1376595	MUM1
chr19	54377264	54377408	MYADM
chr12	83289884	83289983	TMTC2
chr2	165557109	165557208	COBLL1
chr17	29314961	29315124	RNF135
chr16	77326994	77327093	ADAMTS18
chr6	41877064	41877163	MED20
chr5	11236802	11236935	CTNND2
chr5	11364764	11364863	CTNND2
chr4	88011129	88011228	AFF1
chr8	139601454	139601553	COL22A1
chr17	28530189	28530357	SLC6A4
chr19	16594755	16594854	CALR3
chr9	74597635	74597734	C9orf85
chr3	49060488	49060605	NDUFAF3
chr14	64628861	64628990	SYNE2
chr1	154076518	154076617	NUP210L
chr1	115829207	115829306	NGF
chr12	21032377	21032476	SLCO1B3
chr3	50289828	50289970	GNAI2
chr6	101100600	101100765	ASCC3
chrX	82763773	82763872	POU3F4
chr14	21792809	21792927	RPGRIP1
chr15	91454076	91454191	MAN2A2
chr1	212792672	212792771	ATF3
chr7	2976714	2976813	CARD11
chr7	2983982	2984143	CARD11
chr9	101797295	101797436	COL15A1
chr6	26217266	26217365	HIST1H2AE
chr1	180257497	180257652	ACBD6
chr3	183474315	183474477	YEATS2
chr7	82997199	82997298	SEMA3E
chr19	964872	964971	ARID3A
chr18	47379858	47379957	MYO5B
chr2	190561034	190561133	ANKAR
chr4	38830591	38830690	TLR6
chr17	5366848	5367009	DHX33
chr4	52894133	52894265	SGCB
chr7	57529173	57529272	ZNF716
chr1	196715017	196715116	CFH

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr12	25398207	25398318	KRAS
chrX	77245260	77245359	ATP7A
chr4	144797907	144798008	GYPE
chr11	111613246	111613389	PPP2R1B
chr20	10622189	10622288	JAG1
chr6	27833334	27833433	HIST1H2AL
chr10	75037936	75038095	TTC18
chr4	41748177	41748324	PHOX2B
chr7	154790359	154790494	PAXIP1
chr12	59276650	59276814	LRIG3
chr10	91514274	91514430	KIF20B
chrX	19702095	19702194	SH3KBP1
chr1	33134339	33134455	RBBP4
chr16	84050214	84050313	SLC38A8
chr13	33329979	33330094	PDS5B
chr6	40360213	40360338	LRFN2
chr15	42178024	42178123	SPTBN5
chr15	42182286	42182403	SPTBN5
chr15	75705265	75705364	SIN3A
chr8	43211901	43212038	POTEA
chr15	45059892	45059991	TRIM69
chr1	145663185	145663284	RNF115
chr13	107822916	107823015	FAM155A
chr12	64062062	64062165	DPY19L2
chr1	207133970	207134069	FCAMR
chr18	28934566	28934665	DSG1
chr16	89986545	89986644	TUBB3
chr19	4219587	4219755	ANKRD24
chr4	110221723	110221822	COL25A1
chr9	79829223	79829322	VPS13A
chr14	60470335	60470434	LRRC9
chr5	141059826	141059925	ARAP3
chr7	34097670	34097775	BMPER
chr7	34118612	34118757	BMPER
chr16	67645458	67645557	CTCF
chr4	71024052	71024151	C4orf40
chr1	183085901	183086038	LAMC1
chr6	41903669	41903768	CCND3
chr5	137733914	137734032	KDM3B
chr19	12976129	12976295	MAST1
chr19	18547782	18547915	ISYNA1
chr18	28980843	28980983	DSG4
chr18	28989414	28989554	DSG4
chr1	215408276	215408375	KCNK2
chr8	17447205	17447304	PDGFRL
chr15	76726408	76726507	SCAPER
chr17	38935753	38935879	KRT27
chr4	53773623	53773758	SCFD2
chr9	8517993	8518092	PTPRD
chr18	44470498	44470597	PIAS2
chr1	115142824	115142973	DENND2C
chr1	204956545	204956668	NFASC
chr12	112321438	112321537	MAPKAPK5
chr4	39505494	39505605	UGDH
chr20	8637830	8637931	PLCB1
chr8	56986618	56986718	RPS20
chr15	101586185	101586357	LRRK1
chr21	28213316	28213484	ADAMTS1
chr21	28216821	28216939	ADAMTS1
chr13	99361820	99361919	SLC15A1
chr11	47738969	47739068	FNBP4
chr3	51929063	51929162	IQCF1
chr11	108385066	108385165	EXPH5
chrX	83129052	83129151	CYLC1
chr19	12902639	12902790	JUNB
chr15	31324879	31324978	TRPM1
chr4	106157669	106157768	TET2
chr4	106157669	106157768	TET2
chr14	30093357	30093464	PRKD1
chr10	29162152	29162251	C10orf126
chr14	23887408	23887507	MYH7
chr1	237777351	237777450	RYR2
chr1	237872333	237872432	RYR2
chr1	237955374	237955473	RYR2
chr14	90650530	90650629	KCNK13

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	56401576	56401738	DST
chr6	56506744	56506899	DST
chr5	86703814	86703913	CCNH
chr20	50408497	50408596	SALL4
chr2	62729571	62729685	TMEM17
chr1	94485168	94485267	ABCA4
chr9	13122077	13122176	MPDZ
chr9	13125254	13125353	MPDZ
chr9	13222236	13222335	MPDZ
chr6	66205085	66205184	EYS
chrX	79947321	79947477	BRWD3
chr6	43153193	43153348	CUL9
chr22	16287258	16287357	POTEH
chr16	30777747	30777859	RNF40
chr6	56880036	56880135	BEND6
chr10	73337660	73337759	CDH23
chr6	75965903	75966002	TMEM30A
chr6	75969062	75969206	TMEM30A
chr3	39942307	39942417	MYRIP
chr10	103920213	103920312	NOLC1
chr14	103438375	103438474	CDC42BPB
chr19	40884019	40884118	PLD3
chr5	137520200	137520365	KIF20A
chr12	34179714	34179813	ALG10
chr8	1513979	1514078	DLGAP2
chr1	151508712	151508821	CGN
chr12	7087502	7087669	LPCAT3
chr12	107144432	107144571	RFX4
chr2	237032525	237032624	AGAP1
chr7	33035844	33035943	FKBP9
chr18	50936909	50937008	DCC
chr1	206239399	206239498	C1orf186
chr6	107780193	107780292	PDSS2
chr2	80801287	80801439	CTNNA2
chr6	26020776	26020886	HIST1H3A
chr3	160960295	160960441	NMD3
chr13	111372024	111372140	ING1
chr12	12037378	12037521	ETV6
chr2	168074675	168074810	XIRP2
chr10	34985245	34985347	PARD3
chr5	135382023	135382184	TGFBI
chr1	35472551	35472699	ZMYM6
chr5	101627159	101627258	SLCO4C1
chr5	13777310	13777464	DNAH5
chr3	38592168	38592289	SCN5A
chr4	157688996	157689095	PDGFC
chr2	178481432	178481531	TTC30A
chr5	16453121	16453265	ZNF622
chr9	33385768	33385867	AQP7
chrX	26157157	26157552	MAGEB18
chr13	51915293	51915474	SERPINE3
chr18	13825985	13826401	MC5R
chr10	15138569	15138755	C10orf111
chr1	215848722	215848909	USH2A
chr18	64176264	64176451	CDH19
chr11	118764907	118765342	CXCR5
chr19	13264454	13264647	IER2
chr6	167753817	167754016	TTLL2
chr8	105509842	105510291	LRP12
chr14	44974728	44975179	FSCB
chr5	137801551	137801752	EGR1
chr14	26917682	26917884	NOVA1
chrX	91133711	91133913	PCDH11X
chr2	129025756	129025960	HS6ST1
chr11	65623475	65623681	CFL1
chr4	126411276	126411748	FAT4
chrX	102529117	102529327	TCEAL5
chr15	56390324	56390539	RFX7
chr2	155711425	155711641	KCNJ3
chr11	110451414	110451631	ARHGAP20
chr18	74728958	74729176	MBP
chr3	168834000	168834219	MECOM
chr12	49723932	49724157	TROAP
chrX	125686292	125686517	DCAF12L1
chr16	2165393	2165622	PKD1

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr16	2049881	2050111	ZNF598
chr18	24496280	24496517	CHST9
chr4	52861378	52861618	LRRC66
chr5	140346836	140347078	PCDHAC2
chr4	156135335	156135577	NPY2R
chr20	49626535	49626782	KCNG1
chr5	5182162	5182410	ADAMTS16
chr8	13357238	13357493	DLC1
chr2	77746652	77746909	LRRTM4
chr1	114680207	114680472	SYT6
chr3	52521566	52521836	NISCH
chrX	72667220	72667491	CDX4
chr7	89856395	89856678	STEAP2
chr6	139694759	139695043	CITED2
chr5	139908231	139908521	ANKHD1-EIF4EBP3
chr7	119915452	119915743	KCND2
chr19	53013964	53014256	ZNF578
chr1	28800091	28800385	PHACTR4
chr19	53384711	53385007	ZNF320
chr10	123970892	123971189	TACC2
chr5	140482099	140482396	PCDHB3
chr11	100998623	100998923	PGR
chr8	107719046	107719353	OXR1
chr9	27950200	27950510	LINGO2
chrX	151935296	151935608	MAGEA3
chr3	156763153	156763466	LEKR1
chr18	65179922	65181766	DSEL
chr7	110762993	110764937	LRRN3
chr4	30725160	30725981	PCDH7
chr1	226923743	226925140	ITPKB
chr4	188924172	188924867	ZFP42
chr9	16435555	16436253	BNC2
chr13	84453589	84455218	SLITRK1
chr5	140207820	140209113	PCDHA6
chr13	58207180	58209076	PCDH17
chrX	73962177	73963052	KIAA2022
chrX	27998915	27999442	DCAF8L1
chr13	46357646	46358180	SIAH3
chrX	109694664	109695215	RGAG1
chrX	35820634	35821206	MAGEB16
chr3	7620283	7620915	GRM7
chr19	22362809	22363934	ZNF676
chr5	75913775	75914411	F2RL2
chr4	80327835	80328489	GK2
chr1	227842666	227843353	ZNF678
chr2	1652069	1652771	PXDN
chr4	38775463	38775787	TLR10
chr6	26197079	26197411	HIST1H3D
chr8	98289660	98289998	TSPYL5
chr8	104897619	104898393	RIMS2
chr18	64172177	64172523	CDH19
chr12	86198768	86199549	RASSF9
chr19	44610962	44611310	ZNF224
chr15	23931598	23931947	NDN
chr17	61432393	61432746	TANC2
chr3	165548257	165548615	BCHE
chr10	55581999	55582810	PCDH15
chr1	86591496	86591856	COL24A1
chr19	56423331	56423694	NLRP13
chr17	2202994	2203359	SMG6
chrX	91090526	91090897	PCDH11X
chr14	23344753	23345125	LRP10
chr6	107390142	107390514	BEND3
chr20	23028428	23028801	THBD
chr19	21366346	21366722	ZNF431
chr15	86312120	86312500	KLHL25
chr15	70961395	70961785	UACA
chr3	39227655	39228049	XIRP1
chr2	108626767	108627163	SLC5A7
chrX	141291116	141291519	MAGEC2
chr6	94120276	94120685	EPHA7
chr4	187509930	187510340	FAT1
chr6	28213078	28213491	ZKSCAN4
chr8	18729496	18729912	PSD3

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr1	190067191	190068138	FAM5C
chr2	198950504	198950925	PLCL1
chr3	150127293	150127719	TSC22D2
chr1	61553848	61554286	NFIA
chr19	58639971	58640431	ZNF329
chr5	140181824	140182291	PCDHA3
chr16	30456075	30456549	SEPHS2
chr22	20819371	20819850	KLHL22
chr13	32912282	32912764	BRCA2
chr17	21318946	21319434	KCNJ12
chr5	138208750	138209240	LRRTM2
chr5	129520740	129521232	CHSY3
chr8	8748736	8749231	MFHAS1
chr2	186653719	186654217	FSIP2
chr19	42752828	42753332	ERF
chr5	140553945	140554450	PCDHB7
chr8	103663550	103664076	KLF10
chr5	140516575	140517107	PCDHB5
chr15	23811063	23811606	MKRN3
chr19	35232198	35232754	ZNF181
chr1	29069584	29070145	YTHDF2
chr7	106508558	106509120	PIK3CG
chr17	18022175	18022740	MYO15A
chr16	2812143	2812722	SRRM2
chr11	129739436	129740022	NFRKB
chr1	151377896	151378497	POGZ
chr1	14108395	14109018	PRDM2
chr1	75038484	75039109	C1orf173
chrX	26212155	26212785	MAGEB6
chr7	82387890	82388031	PCLO
chr7	82453578	82453677	PCLO
chr6	37138548	37138655	PIM1
chr6	37140805	37140904	PIM1
chr4	126328000	126328099	FAT4
chr4	126336747	126336846	FAT4
chr4	126337678	126337777	FAT4
chr4	126389661	126389760	FAT4
chr8	113266468	113266567	CSMD3
chr8	113308061	113308235	CSMD3
chr8	113314021	113314195	CSMD3
chr8	113332120	113332219	CSMD3
chr8	113347557	113347703	CSMD3
chr8	113348910	113349009	CSMD3
chr8	113353773	113353872	CSMD3
chr8	113364644	113364763	CSMD3
chr8	113569046	113569145	CSMD3
chr8	113585729	113585886	CSMD3
chr8	113599294	113599464	CSMD3
chr8	113668445	113668544	CSMD3
chr8	113702216	113702315	CSMD3
chr8	113812390	113812503	CSMD3
chr8	113871373	113871495	CSMD3
chr8	114448912	114449011	CSMD3
chr12	49416049	49416148	MLL2
chr12	49418360	49418491	MLL2
chr12	49420593	49420692	MLL2
chr12	49427948	49428047	MLL2
chr12	49433338	49433437	MLL2
chr12	49437982	49438087	MLL2
chr12	49438185	49438305	MLL2
chr12	49444450	49444549	MLL2
chr12	49447258	49447424	MLL2
chr7	2978312	2978465	CARD11
chr7	2979449	2979548	CARD11
chr7	2987232	2987331	CARD11
chr7	148523590	148523689	EZH2
chr1	2489164	2489273	TNFRSF14
chr1	2493111	2493254	TNFRSF14
chr17	7578176	7578289	TP53
chr6	56327843	56327954	DST
chr6	56330875	56330993	DST
chr6	56368794	56368896	DST
chr6	56458548	56458647	DST
chr6	56466227	56466326	DST
chr6	56499259	56499414	DST

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	56499598	56499751	DST
chr6	56501352	56501451	DST
chr6	56515723	56515830	DST
chr2	141072503	141072668	LRP1B
chr2	141259305	141259404	LRP1B
chr2	141299447	141299546	LRP1B
chr2	141356244	141356343	LRP1B
chr2	141459289	141459414	LRP1B
chr2	141680580	141680679	LRP1B
chr2	141819709	141819808	LRP1B
chr1	215799117	215799216	USH2A
chr1	215813913	215814012	USH2A
chr1	215844296	215844395	USH2A
chr1	215901422	215901521	USH2A
chr1	215953269	215953368	USH2A
chr1	215955383	215955538	USH2A
chr1	215960043	215960142	USH2A
chr1	216052082	216052181	USH2A
chr1	216108069	216108168	USH2A
chr1	216262354	216262481	USH2A
chr1	216270424	216270555	USH2A
chr1	216462621	216462752	USH2A
chr1	216497541	216497640	USH2A
chr19	19257550	19257684	MEF2B
chr4	187530336	187530474	FAT1
chr4	187534231	187534330	FAT1
chr4	187549398	187549497	FAT1
chr4	187557842	187557941	FAT1
chr6	134492772	134492871	SGK1
chr1	185833601	185833760	HMCN1
chr1	185969270	185969369	HMCN1
chr1	185972849	185972976	HMCN1
chr1	186039743	186039889	HMCN1
chr1	186062637	186062736	HMCN1
chr1	186083110	186083255	HMCN1
chr1	186135939	186136074	HMCN1
chr1	186143645	186143774	HMCN1
chr1	186158943	186159042	HMCN1
chr8	116631744	116631843	TRPS1
chr16	3789578	3789725	CREBBP
chr16	3823772	3823871	CREBBP
chr16	3900300	3900399	CREBBP
chr16	85945170	85945269	IRF8
chr5	89943517	89943616	GPR98
chr5	89971896	89972026	GPR98
chr5	90040945	90041044	GPR98
chr5	90049479	90049578	GPR98
chr5	90087039	90087138	GPR98
chr5	90106831	90106930	GPR98
chr18	6943213	6943312	LAMA1
chr18	6947161	6947295	LAMA1
chr18	6955351	6955464	LAMA1
chr18	6980519	6980636	LAMA1
chr18	6983097	6983233	LAMA1
chr18	6985525	6985642	LAMA1
chr18	7032064	7032175	LAMA1
chr18	7080285	7080456	LAMA1
chr4	85642561	85642725	WDFY3
chr4	85695972	85696134	WDFY3
chr8	77775531	77775630	ZFHX4
chr10	16873250	16873416	CUBN
chr10	16930415	16930565	CUBN
chr10	16957870	16957969	CUBN
chr10	16979723	16979822	CUBN
chr10	17087038	17087137	CUBN
chr10	17130189	17130288	CUBN
chr3	187440245	187440389	BCL6
chr3	187442728	187442866	BCL6
chr3	187449498	187449597	BCL6
chr5	123982951	123983050	ZNF608
chr5	123985296	123985395	ZNF608
chr8	2824183	2824282	CSMD1
chr8	2857479	2857653	CSMD1
chr8	3038631	3038736	CSMD1
chr8	3165895	3165994	CSMD1

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr8	4494995	4495094	CSMD1
chr9	13221370	13221499	MPDZ
chr9	13224501	13224600	MPDZ
chr5	13716704	13716803	DNAH5
chr5	13737466	13737565	DNAH5
chr5	13766039	13766138	DNAH5
chr5	13770878	13770977	DNAH5
chr5	13864534	13864633	DNAH5
chr5	13883032	13883131	DNAH5
chr5	13894758	13894930	DNAH5
chr5	13920588	13920726	DNAH5
chr22	23610594	23610702	BCR
chr6	152443540	152443639	SYNE1
chr6	152651704	152651803	SYNE1
chr6	152683304	152683458	SYNE1
chr6	152702294	152702393	SYNE1
chr6	152730693	152730844	SYNE1
chr5	82789317	82789416	VCAN
chr5	82843789	82843903	VCAN
chr5	82875823	82875922	VCAN
chrX	32509457	32509556	DMD
chrX	32583801	32583900	DMD
chrX	32613873	32613993	DMD
chrX	32662259	32662358	DMD
chrX	32717291	32717390	DMD
chrX	33146223	33146322	DMD
chr3	164700030	164700198	SI
chr3	164700764	164700863	SI
chr3	164735548	164735661	SI
chr3	164776750	164776870	SI
chr3	164786865	164786983	SI
chr7	48337962	48338084	ABCA13
chr7	48547445	48547544	ABCA13
chr7	48550679	48550795	ABCA13
chr7	48559750	48559849	ABCA13
chr7	48682883	48682989	ABCA13
chr1	181452980	181453079	CACNA1E
chr1	181724372	181724533	CACNA1E
chr1	181745236	181745364	CACNA1E
chr1	181759580	181759692	CACNA1E
chr17	40469171	40469270	STAT3
chr17	40474377	40474476	STAT3
chr17	40478125	40478224	STAT3
chr17	40485908	40486067	STAT3
chr17	40491329	40491428	STAT3
chr12	18534699	18534814	PIK3C2G
chr12	18544055	18544186	PIK3C2G
chr12	18573871	18573970	PIK3C2G
chr12	18699256	18699366	PIK3C2G
chr12	18747415	18747514	PIK3C2G
chr2	169995086	169995216	LRP2
chr2	170010969	170011113	LRP2
chr2	170012783	170012915	LRP2
chr2	170025048	170025186	LRP2
chr2	170101242	170101341	LRP2
chr2	170115542	170115641	LRP2
chr19	6590080	6590179	CD70
chr19	6590851	6591013	CD70
chr17	74011104	74011203	EVPL
chr5	11110994	11111093	CTNND2
chr5	11397142	11397315	CTNND2
chr5	11411647	11411764	CTNND2
chr3	64547253	64547427	ADAMTS9
chr3	64579949	64580048	ADAMTS9
chr18	60793423	60793599	BCL2-NA
chr18	60774470	60774594	BCL2-NA
chr8	128764154	128764209	MYC-IGH
chr14	106329109	106330460	IGH@-MYC
chr3	187461513	187463197	BCL6-NA
chr11	69346747	69346916	CCND1-NA
chr18	60763905	60763963	BCL2-NA
chr14	106323422	106328049	IGH@-MYC
chr18	60764357	60764467	BCL2-NA
chr14	106239409	106242027	IGH@-BCL6
chr14	106329407	106329468	IGHJ6

TABLE 8-continued

Chromosome	Start (bp)	End (bp)	Gene
chr14	106330023	106330072	IGHJ5
chr14	106330424	106330470	IGHJ4
chr14	106330796	106330845	IGHJ3
chr14	106331408	106331460	IGHJ2
chr14	106331616	106331668	IGHJ1
chr14	106494134	106494445	IGHV2-5.1
chr14	106494531	106494597	IGHV2-5.2
chr14	106518399	106518704	IGHV3-7.1
chr14	106518807	106518932	IGHV3-7.2
chr14	106725200	106725505	IGHV3-23.1
chr14	106725608	106725733	IGHV3-23.2
chr14	106815721	106816026	IGHV3-33.1
chr14	106816127	106816253	IGHV3-33.2
chr14	106829593	106829895	IGHV4-34.1
chr14	106829978	106830076	IGHV4-34.2
chr14	106877618	106877926	IGHV4-39.1
chr14	106878009	106878126	IGHV4-39.2
chr14	106993813	106994118	IGHV3-48.1
chr14	106994221	106994346	IGHV3-48.2
chr14	107034728	107035033	IGHV5-51.1
chr14	107035116	107035221	IGHV5-51.2
chr14	107169930	107170235	IGHV1-69.1
chr14	107170321	107170428	IGHV1-69.2
chrX	100611039	100611256	BTK

TABLE 9

Chromosome	Start (bp)	End (bp)
chr17	7572917	7573017
chr17	7573926	7574033
chr17	7576510	7576691
chr17	7576839	7576939
chr17	7577018	7577155
chr17	7577498	7577608
chr17	7578176	7578289
chr17	7578361	7578554
chr17	7579310	7579590
chr17	7579660	7579760
chr17	7579825	7579925
chr17	8926070	8926201
chr17	10402290	10402409
chr17	10416183	10416283
chr17	20799111	20799211
chr17	21319121	21319799
chr17	26874644	26874744
chr17	26962100	26962200
chr17	27248705	27248826
chr17	37879790	37879913
chr17	37880164	37880264
chr17	37880978	37881164
chr17	37881301	37881457
chr17	37881567	37881667
chr17	37881959	37882106
chr17	37882813	37882913
chr17	40556937	40557366
chr17	40837331	40837431
chr17	44845791	44846006
chr17	51900603	51902313
chr17	56344761	56344861
chr17	66938077	66938177
chr17	75208099	75208231
chr17	79414082	79414310
chr9	1056621	1056916
chr9	4118776	4118876
chr9	8404536	8404660
chr9	8485225	8485325
chr9	16419232	16419555
chr9	17761421	17761521
chr9	18776929	18777149
chr9	21968184	21968284
chr9	21968697	21968797

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr9	21970900	21971207
chr9	21974475	21974826
chr9	21994137	21994330
chr9	34976546	34976646
chr9	78789901	78790045
chr9	94486757	94487239
chr9	104385599	104385715
chr9	111617085	111618027
chr9	113538082	113538182
chr9	115806435	115806535
chr9	121929789	121930084
chr9	126135805	126136023
chr9	129957369	129957496
chr9	131479014	131479114
chr3	1424636	1424791
chr3	9989136	9989306
chr3	10320050	10320150
chr3	18390797	18390897
chr3	26751364	26751574
chr3	36484921	36485092
chr3	38748769	38748874
chr3	38802762	38802862
chr3	38891989	38892089
chr3	41266057	41266157
chr3	48691721	48691885
chr3	49698932	49699032
chr3	52473987	52474098
chr3	54921982	54922082
chr3	55504230	55504574
chr3	64132763	64132863
chr3	65415276	65415406
chr3	66023701	66023853
chr3	73453378	73453543
chr3	74315631	74315800
chr3	77623656	77623874
chr3	78649348	78649459
chr3	79174597	79174697
chr3	81586069	81586169
chr3	89259079	89259477
chr3	89468390	89468530
chr3	93615419	93615535
chr3	96706192	96706776
chr3	102196391	102196491
chr3	112991330	112991514
chr3	114069871	114070512
chr3	119886475	119886856
chr3	120366691	120366791
chr3	126071038	126071314
chr3	126736297	126736397
chr3	134920319	134920485
chr3	142681457	142681817
chr3	147127985	147128783
chr3	154146591	154147111
chr3	164712044	164712193
chr3	178935997	178936122
chr3	178951881	178952152
chr3	180359871	180359971
chr3	186760514	186760878
chr1	4771972	4772524
chr1	10384028	10384128
chr1	10386319	10386419
chr1	11591621	11591767
chr1	12266840	12266983
chr1	12785336	12785494
chr1	16133895	16133995
chr1	16474992	16475531
chr1	17668793	17668897
chr1	23234504	23234604
chr1	27087376	27087504
chr1	27100070	27100208
chr1	27224052	27224180
chr1	27332657	27332876
chr1	33957160	33957271
chr1	36937065	36937198
chr1	46826375	46826500

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr1	58946674	58946836
chr1	67147569	67147934
chr1	70446049	70446149
chr1	70493913	70494013
chr1	86591234	86591334
chr1	89734411	89734539
chr1	103345310	103345410
chr1	103364473	103364573
chr1	103477947	103478047
chr1	103491355	103491508
chr1	111216542	111216792
chr1	114340236	114340481
chr1	152286493	152287124
chr1	154988884	154989109
chr1	155629489	155629589
chr1	156823777	156823877
chr1	157514663	157514763
chr1	157804283	157804383
chr1	158063128	158063236
chr1	158064475	158064857
chr1	158151970	158152070
chr1	158224988	158225088
chr1	158324296	158324396
chr1	158325168	158325273
chr1	158590011	158590111
chr1	158592793	158592957
chr1	158609659	158609797
chr1	158626350	158626450
chr1	158637683	158637802
chr1	159002313	159002481
chr1	159021500	159021857
chr1	161721457	161721571
chr1	175087761	175087884
chr1	181731701	181731801
chr1	183849784	183849884
chr1	185891510	185891631
chr1	185902883	185902983
chr1	185958619	185958779
chr1	186008859	186009011
chr1	186043872	186044023
chr1	186105987	186106087
chr1	190067507	190067651
chr1	193028302	193028402
chr1	196227370	196227470
chr1	196642119	196642271
chr1	204518477	204518600
chr1	210977341	210977489
chr1	211093048	211093383
chr1	215972269	215972459
chr1	216419942	216420152
chr1	231935853	231935956
chr1	232649895	232650139
chr1	237791187	237791287
chr1	240555786	240555886
chr1	244640829	244640929
chr1	249212294	249212442
chr18	6896500	6896625
chr18	9887970	9888070
chr18	31537401	31537501
chr18	44560449	44560926
chr18	48591822	48591922
chr18	48593388	48593557
chr18	55247286	55247431
chr18	59195225	59195330
chr18	60642639	60642792
chr18	74963050	74963150
chr19	5244043	5244327
chr19	7687216	7687330
chr19	8609180	8609348
chr19	8808080	8808405
chr19	11134193	11134307
chr19	21992321	21992491
chr19	22157365	22157560
chr19	22364208	22364308
chr19	22942330	22942459

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr19	31039134	31040222
chr19	35842909	35843009
chr19	37210102	37210288
chr19	37440440	37440631
chr19	37975081	37975181
chr19	40408485	40408619
chr19	43698598	43698698
chr19	46443799	46443899
chr19	47935612	47935712
chr19	49385288	49385460
chr19	51189493	51189612
chr19	51330100	51330200
chr19	51645679	51645779
chr19	52272233	52272763
chr19	52327729	52328003
chr19	52538241	52538341
chr19	53612569	53612917
chr19	54310748	54310919
chr19	55593821	55593970
chr19	55815034	55815194
chr19	57293320	57293489
chr19	58048608	58048914
chr19	58601318	58601479
chr13	19751147	19751388
chr13	28014250	28014404
chr13	32745317	32745417
chr13	33590917	33591017
chr13	36046536	36046673
chr13	36686054	36686248
chr13	36748858	36749006
chr13	36886455	36886614
chr13	36888368	36888468
chr13	37427676	37427776
chr13	38237609	38237780
chr13	48985655	48985755
chr13	58206729	58208431
chr13	58298776	58299424
chr13	66878762	66878862
chr13	73636053	73636410
chr13	74518112	74518212
chr13	78477298	78477398
chr13	92345718	92345965
chr13	94482408	94482634
chr13	102047547	102047710
chr13	114083266	114083407
chr16	9857863	9858061
chr16	14041834	14041934
chr16	31926456	31926578
chr16	49430407	49430537
chr16	51171054	51171343
chr16	55362612	55363185
chr16	55690614	55690714
chr16	61760997	61761119
chr16	61851396	61851498
chr16	61891022	61891142
chr16	61935287	61935387
chr16	65022116	65022246
chr16	66956016	66956116
chr16	67333329	67333429
chr16	77228294	77228416
chr16	77353745	77353960
chr16	77769724	77769883
chr16	80638284	80638443
chr16	80654659	80654832
chr16	81969862	81969962
chr16	86544658	86544806
chr5	1221945	1222045
chr5	5239910	5240010
chr5	5319135	5319251
chr5	9190413	9190523
chr5	19473404	19473753
chr5	23510028	23510136
chr5	23521131	23521288
chr5	24487934	24488034
chr5	24491769	24491869

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr5	24505220	24505357
chr5	24537581	24537715
chr5	26915768	26915868
chr5	32712169	32712504
chr5	35876342	35876442
chr5	42719339	42719496
chr5	63256293	63257273
chr5	67576744	67576844
chr5	67591054	67591154
chr5	83402530	83402630
chr5	100222166	100222266
chr5	101592839	101592939
chr5	101593646	101593791
chr5	101834362	101834478
chr5	109190864	109190964
chr5	128983457	128983588
chr5	135692537	135692637
chr5	136324145	136324264
chr5	140182123	140182957
chr5	140222588	140222721
chr5	140562893	140563039
chr5	140811004	140811172
chr5	148407162	148407434
chr5	156346460	156346560
chr5	161116251	161116351
chr5	169423079	169423179
chr5	172659657	172659843
chr5	178413398	178413499
chr5	180661254	180661354
chr12	939168	939326
chr12	4735912	4736043
chr12	6632065	6632165
chr12	7635238	7635358
chr12	9162021	9162133
chr12	9754096	9754196
chr12	9833518	9833629
chr12	23696144	23696318
chr12	23728695	23728795
chr12	23893800	23893973
chr12	25378548	25378707
chr12	25380167	25380346
chr12	25398207	25398318
chr12	29614786	29614941
chr12	41900282	41900382
chr12	43944891	43944991
chr12	45410193	45410293
chr12	52910577	52910677
chr12	55420590	55421030
chr12	56647957	56648057
chr12	57553698	57553798
chr12	62786832	62786982
chr12	63544118	63544218
chr12	75601408	75601619
chr12	85266927	85267027
chr12	85450573	85450673
chr12	85517871	85517971
chr12	85531627	85531747
chr12	86373752	86374125
chr12	89744605	89744705
chr12	94975667	94975767
chr12	99640240	99640522
chr12	100704821	100704971
chr12	103352293	103352639
chr12	104476511	104476611
chr12	106460715	106460815
chr12	108169109	108169494
chr12	108985454	108985675
chr12	113704047	113704147
chr12	117768680	117768780
chr12	118198821	118199091
chr12	118599728	118599828
chr12	121972398	121972498
chr12	128899734	128899839
chr12	130184676	130185148
chr2	271852	271952

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr2	1241659	1241789
chr2	1643073	1643194
chr2	16736322	16736422
chr2	17830731	17830863
chr2	23985078	23985216
chr2	27668162	27668316
chr2	31588841	31588975
chr2	31805700	31805800
chr2	37873276	37873554
chr2	48808233	48809281
chr2	49217690	49217790
chr2	50463971	50464075
chr2	51254718	51255389
chr2	60687822	60688673
chr2	65540875	65541009
chr2	70188461	70188561
chr2	70903839	70903939
chr2	70910754	70910854
chr2	71791187	71791343
chr2	80136815	80136915
chr2	85622655	85622755
chr2	96689056	96689188
chr2	96780826	96781620
chr2	100209964	100210093
chr2	100623671	100623845
chr2	107459958	107460195
chr2	113310221	113310393
chr2	116497314	116497469
chr2	116599786	116599921
chr2	125232315	125232456
chr2	125521553	125521724
chr2	125555813	125555913
chr2	138169196	138169428
chr2	141031997	141032097
chr2	141093237	141093339
chr2	141135748	141135855
chr2	141356206	141356306
chr2	141533666	141533807
chr2	141680553	141680680
chr2	141773363	141773463
chr2	141819634	141819770
chr2	160194147	160194247
chr2	163291905	163292035
chr2	166901565	166901776
chr2	167129083	167129183
chr2	171687559	171687659
chr2	176958128	176958397
chr2	177036729	177036833
chr2	178098900	178099000
chr2	189868982	189869090
chr2	196729066	196729624
chr2	202211264	202211398
chr2	207619755	207620090
chr2	212248422	212248600
chr2	212285165	212285336
chr2	212426709	212426809
chr2	212488646	212488769
chr2	216245639	216245756
chr2	222301117	222301217
chr2	229890703	229890803
chr2	230910820	230911186
chr2	232262574	232262674
chr2	232602722	232602848
chr8	2037925	2038025
chr8	3216750	3216854
chr8	3265606	3265706
chr8	19810820	19810932
chr8	21903621	21903721
chr8	22020086	22020245
chr8	26722104	26722204
chr8	32617813	32617913
chr8	40554763	40554920
chr8	52323809	52323967
chr8	53853280	53853437
chr8	65493779	65494351

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr8	65528512	65528749
chr8	66539499	66539599
chr8	68930066	68930166
chr8	68965331	68965481
chr8	68968064	68968209
chr8	69011916	69012078
chr8	69046323	69046428
chr8	69445217	69445384
chr8	70476210	70476382
chr8	70512837	70512988
chr8	72755917	72756277
chr8	73480060	73480226
chr8	75756215	75756334
chr8	75898308	75898408
chr8	88365863	88366014
chr8	88885072	88886105
chr8	89179945	89180045
chr8	89198704	89198827
chr8	90937332	90937538
chr8	103289299	103289399
chr8	104337212	104337312
chr8	104427349	104427573
chr8	104897532	104898129
chr8	105263251	105263391
chr8	105436474	105436617
chr8	105463473	105463632
chr8	105502960	105503590
chr8	107718772	107718872
chr8	110980399	110980499
chr8	110984663	110984987
chr8	113267482	113267656
chr8	113657337	113657454
chr8	113933855	113933980
chr8	116616215	116616713
chr8	131916133	131916287
chr8	132051787	132052129
chr8	133141743	133141878
chr8	133187699	133187855
chr8	133899105	133899443
chr8	143694946	143695283
chr8	145637921	145638050
chr7	1275537	1275639
chr7	1481833	1481981
chr7	13935471	13935571
chr7	19156555	19156655
chr7	19184867	19184967
chr7	19765216	19765326
chr7	21641095	21641195
chr7	21675492	21675604
chr7	23292925	23293078
chr7	27194678	27194789
chr7	30795196	30795296
chr7	30961680	30961845
chr7	37252939	37253062
chr7	37780836	37780936
chr7	37955698	37956077
chr7	37988438	37988538
chr7	41739625	41739896
chr7	42017156	42017321
chr7	43635647	43635747
chr7	53103572	53104215
chr7	55241613	55241736
chr7	55242414	55242514
chr7	55248985	55249171
chr7	55259411	55259567
chr7	55260446	55260546
chr7	55266409	55266556
chr7	55268007	55268107
chr7	56126054	56126214
chr7	64166897	64166997
chr7	70885934	70886081
chr7	71175745	71175913
chr7	77797233	77797333
chr7	82544292	82545776
chr7	82764240	82764705

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr7	86394604	86394909
chr7	86415649	86416412
chr7	87144624	87144724
chr7	87150091	87150192
chr7	90355906	90356006
chr7	90546915	90547039
chr7	90741856	90741996
chr7	91709224	91709379
chr7	96650097	96650278
chr7	103969418	103969616
chr7	106508066	106508622
chr7	107315389	107315554
chr7	112461883	112462205
chr7	113558285	113559041
chr7	116411902	116412043
chr7	116417433	116417533
chr7	116418829	116419011
chr7	116422041	116422151
chr7	116423357	116423523
chr7	116435708	116435845
chr7	116435940	116436178
chr7	126086324	126086424
chr7	136699666	136700924
chr7	136938210	136938384
chr7	137075965	137076082
chr7	137206611	137206712
chr7	149461895	149461995
chr7	154585788	154585912
chr7	157926664	157926764
chr10	7325922	7326022
chr10	7657944	7658061
chr10	7679343	7679443
chr10	16932369	16932526
chr10	16970155	16970302
chr10	16982093	16982193
chr10	18242214	18242444
chr10	37507960	37508353
chr10	46967493	46967638
chr10	50121608	50121845
chr10	50374913	50375054
chr10	55912911	55913011
chr10	60558158	60558320
chr10	60936613	60936719
chr10	76857469	76857643
chr10	83635326	83635572
chr10	84118494	84118624
chr10	84498319	84498419
chr10	84733543	84733671
chr10	87362203	87362368
chr10	89692769	89693008
chr10	89711874	89712016
chr10	89717609	89717776
chr10	101816769	101816909
chr10	108458993	108459093
chr10	108536291	108536450
chr10	117221448	117221548
chr10	118030462	118030562
chr10	118236163	118236331
chr10	125804147	125804313
chr10	134015449	134015620
chr6	13365730	13365859
chr6	26189081	26189229
chr6	26204900	26205079
chr6	26251883	26251983
chr6	27100313	27100466
chr6	27775630	27775730
chr6	27858505	27858605
chr6	28053312	28053476
chr6	28097349	28097508
chr6	28543183	28543603
chr6	32411532	32411687
chr6	43323452	43323552
chr6	43612899	43612999
chr6	50682827	50683088
chr6	50696568	50696734

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr6	57472350	57472450
chr6	62604499	62604599
chr6	66115088	66115196
chr6	66200486	66200600
chr6	66204686	66205195
chr6	69653721	69653886
chr6	69666536	69666701
chr6	70070763	70071313
chr6	72011366	72011466
chr6	85446427	85446549
chr6	88218758	88218893
chr6	90432687	90432787
chr6	90459323	90459423
chr6	100841453	100841709
chr6	102250205	102250313
chr6	102337630	102337731
chr6	106552782	106553406
chr6	110714075	110714175
chr6	110763684	110763784
chr6	126080648	126080811
chr6	144263522	144263622
chr6	146720219	146720760
chr6	152476026	152476177
chr6	152497528	152497695
chr6	152674397	152674568
chr6	152690106	152690262
chr6	152694241	152694341
chr6	152712452	152712552
chr6	152737526	152737735
chr6	152768592	152768757
chr6	152786396	152786527
chr6	153073312	153073419
chr6	154412217	154412459
chr6	159653069	159654492
chr6	161011975	161012133
chr6	161152118	161152218
chr6	162683696	162683796
chr6	165715082	165715681
chr6	165792682	165792855
chr6	169053730	169053846
chr11	688330	688460
chr11	5536660	5537043
chr11	6261558	6261799
chr11	6265195	6265586
chr11	7324310	7324582
chr11	16007815	16007946
chr11	17757872	17757972
chr11	20959318	20959418
chr11	27679627	27680104
chr11	35640969	35641155
chr11	36520039	36520190
chr11	61607848	61607948
chr11	63137777	63137877
chr11	63398570	63398670
chr11	64419506	64419626
chr11	64559643	64559747
chr11	65349852	65349952
chr11	88911836	88911936
chr11	89916080	89916180
chr11	92087059	92087237
chr11	92430530	92430630
chr11	92616214	92616314
chr11	103907637	103908702
chr11	106680787	106681050
chr11	110450613	110450997
chr11	113853841	113854011
chr11	117374596	117374696
chr11	120996278	120996554
chr11	120998709	120999023
chr11	121000544	121000880
chr11	121016607	121016766
chr11	124794764	124794965
chr11	125255446	125255618
chr4	10445956	10446241
chr4	15780039	15780139

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr4	20535194	20535338
chr4	26719547	26719647
chr4	42964907	42965096
chr4	44176852	44177137
chr4	44624430	44624593
chr4	46252419	46252519
chr4	47427826	47427960
chr4	47514565	47514798
chr4	47560186	47560286
chr4	56428686	56428786
chr4	65188406	65188510
chr4	66356148	66356269
chr4	66467764	66467879
chr4	69885946	69886046
chr4	85611658	85611817
chr4	89653282	89653382
chr4	96761626	96762304
chr4	107956653	107956753
chr4	126355420	126355575
chr4	140432823	140432923
chr4	140640548	140640648
chr4	155719085	155719401
chr4	157771436	157771536
chr4	162307004	162307531
chr4	162459317	162459452
chr4	162697105	162697221
chr4	164272378	164272478
chr4	166981179	166981340
chr4	177052722	177052880
chr4	177608374	177608562
chr4	186380565	186380685
chr4	188924037	188924757
chr4	189012603	189012986
chr20	1961041	1961141
chr20	2375833	2375933
chr20	9525015	9525141
chr20	11904073	11904280
chr20	21494122	21494286
chr20	21687128	21687381
chr20	21695384	21695484
chr20	23016302	23016579
chr20	23807160	23807271
chr20	25462603	25462736
chr20	30584567	30584911
chr20	34022210	34022502
chr20	35060182	35060992
chr20	39831690	39831790
chr20	41419851	41420050
chr20	44680388	44680488
chr20	44839096	44839196
chr20	44845464	44845632
chr20	54578942	54579092
chr20	57036020	57036317
chr20	57042377	57042477
chr20	57829250	57829609
chr20	60448783	60448956
chr20	60887687	60887836
chr20	61542441	61542593
chr20	62045440	62045546
chr20	62121861	62122041
chr14	24534219	24534319
chr14	26917394	26917983
chr14	29237047	29237609
chr14	42356533	42356867
chr14	45644655	45644819
chr14	52507374	52507519
chr14	52534642	52534742
chr14	59930652	59930752
chr14	60193838	60193938
chr14	65007951	65008051
chr14	65241975	65242075
chr14	68052667	68052814
chr14	77275824	77275975
chr14	88945797	88945916
chr14	90650479	90650701

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr14	94964338	94964727
chr14	100317245	100317345
chr15	23811018	23812264
chr15	23931441	23932345
chr15	25222023	25222176
chr15	25959023	25959376
chr15	27572031	27572143
chr15	28358245	28358373
chr15	28474819	28474993
chr15	31851286	31851386
chr15	33954415	33954834
chr15	48062706	48063822
chr15	49217061	49217161
chr15	49327722	49327822
chr15	51350414	51350603
chr15	54586092	54586262
chr15	59368237	59368337
chr15	70960099	70960400
chr15	83226521	83226621
chr15	83332553	83332716
chr15	89861806	89861906
chr15	91548919	91549019
chr15	92459483	92459583
chr21	10906904	10907040
chr21	15872949	15873049
chr21	22849679	22849779
chr21	28296534	28296889
chr21	28327057	28327190
chr21	31538304	31538861
chr21	32638628	32639147
chr21	36206723	36206823
chr21	38302607	38302707
chr21	41414457	41414593
chr22	17072731	17073103
chr22	18028302	18028573
chr22	22892498	22892655
chr22	24584017	24584117
chr22	26423369	26423517
chr22	38121538	38121797
chr22	39626088	39626233
chr22	40140117	40140217
chr22	41076963	41077877
chr22	45281730	45281830
chr22	46318839	46318939
chr22	46327001	46327101
chr22	50302891	50302995
chrX	62875367	62875467
track name = 169383__2__EAC__ESCC__P1__tiled__region description = "169383__2__EAC__ESCC__P1__tiled__region"		
chr1	4771948	4772532
chr1	10384007	10384142
chr1	10386292	10386433
chr1	11591595	11591812
chr1	12266815	12267022
chr1	12785310	12785530
chr1	16133874	16134022
chr1	16474969	16475570
chr1	17668768	17668903
chr1	23234563	23234640
chr1	27087341	27087520
chr1	27100036	27100196
chr1	27224031	27224205
chr1	27332631	27332916
chr1	33957137	33957306
chr1	36937037	36937218
chr1	46826340	46826530
chr1	58946640	58946856
chr1	67147542	67147961
chr1	70446024	70446162
chr1	70493889	70494028
chr1	86591212	86591366
chr1	89734388	89734570

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr1	103345277	103345441
chr1	103364482	103364585
chr1	103477912	103478067
chr1	103491322	103491539
chr1	111216507	111216827
chr1	114340208	114340487
chr1	152286458	152287142
chr1	154988850	154989140
chr1	155629459	155629598
chr1	156823749	156823893
chr1	157514639	157514778
chr1	157804259	157804398
chr1	158063104	158063274
chr1	158064444	158064872
chr1	158151949	158152095
chr1	158224954	158225107
chr1	158324264	158324418
chr1	158325139	158325290
chr1	158589989	158590127
chr1	158592764	158592986
chr1	158609634	158609806
chr1	158626324	158626478
chr1	158637649	158637836
chr1	159002322	159002462
chr1	159021467	159021888
chr1	161721433	161721604
chr1	175087761	175087837
chr1	181731679	181731823
chr1	183849761	183849902
chr1	185891485	185891654
chr1	185902860	185903000
chr1	185958590	185958811
chr1	186008830	186009036
chr1	186043845	186044057
chr1	186105955	186106104
chr1	190067484	190067688
chr1	193028273	193028424
chr1	196227345	196227477
chr1	196642095	196642300
chr1	204518442	204518513
chr1	204518532	204518636
chr1	210977314	210977532
chr1	211093019	211093407
chr1	215972244	215972489
chr1	216419909	216420166
chr1	231935832	231935967
chr1	232649860	232650150
chr1	237791159	237791305
chr1	240555759	240555899
chr1	244640802	244640949
chr1	249212268	249212484
chr2	271830	271973
chr2	1241635	1241815
chr2	1643045	1643218
chr2	16736289	16736435
chr2	17830709	17830893
chr2	23985053	23985228
chr2	27668133	27668348
chr2	31588817	31588986
chr2	31805670	31805812
chr2	37873269	37873554
chr2	48808203	48809302
chr2	49217658	49217805
chr2	50463940	50464095
chr2	51254686	51255431
chr2	60687799	60688541
chr2	60688589	60688702
chr2	65540840	65541032
chr2	70188433	70188573
chr2	70903818	70903950
chr2	70910723	70910874
chr2	71791163	71791373
chr2	80136788	80136945
chr2	85622632	85622777
chr2	96689034	96689210

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr2	96780794	96780980
chr2	96780989	96781583
chr2	100209938	100210120
chr2	100623643	100623852
chr2	107459926	107460218
chr2	113310213	113310419
chr2	116497291	116497491
chr2	116599761	116599934
chr2	125232291	125232499
chr2	125521532	125521738
chr2	125555782	125555935
chr2	138169170	138169453
chr2	141031962	141032115
chr2	141093202	141093352
chr2	141135727	141135867
chr2	141356182	141356332
chr2	141533637	141533838
chr2	141680522	141680713
chr2	141773336	141773490
chr2	141819611	141819787
chr2	160194115	160194274
chr2	163291875	163292068
chr2	166901530	166901827
chr2	167129060	167129201
chr2	171687525	171687684
chr2	176958102	176958419
chr2	177036702	177036838
chr2	178098877	178099014
chr2	189868949	189869125
chr2	196729031	196729646
chr2	202211243	202211416
chr2	207619734	207620118
chr2	212248394	212248638
chr2	212285144	212285364
chr2	212426674	212426825
chr2	212488614	212488802
chr2	216245604	216245790
chr2	222301096	222301253
chr2	229890670	229890831
chr2	230910788	230910863
chr2	230910893	230911219
chr2	232262549	232262683
chr2	232602697	232602862
chr3	1424605	1424830
chr3	9989104	9989331
chr3	10320029	10320167
chr3	18390764	18390911
chr3	26751332	26751593
chr3	36484887	36485121
chr3	38748748	38748901
chr3	38802738	38802890
chr3	38891968	38892111
chr3	41266031	41266180
chr3	48691698	48691896
chr3	49698908	49699050
chr3	52473962	52474135
chr3	54921956	54922107
chr3	55504206	55504592
chr3	64132739	64132871
chr3	65415249	65415431
chr3	66023669	66023883
chr3	73453346	73453564
chr3	74315603	74315810
chr3	77623631	77623903
chr3	78649326	78649496
chr3	79174566	79174720
chr3	81586036	81586187
chr3	89259047	89259516
chr3	89468362	89468540
chr3	93615429	93615497
chr3	96706159	96706811
chr3	102196370	102196513
chr3	112991299	112991548
chr3	114069844	114070552
chr3	119886440	119886877

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr3	120366660	120366816
chr3	126071010	126071337
chr3	126736266	126736420
chr3	134920286	134920512
chr3	142681424	142681850
chr3	147127961	147128744
chr3	154146561	154147122
chr3	164712020	164712229
chr3	178935989	178936159
chr3	178951854	178952172
chr3	180359849	180359989
chr3	186760479	186760905
chr4	10445924	10446277
chr4	15780006	15780164
chr4	20535166	20535374
chr4	26719517	26719656
chr4	42964885	42965135
chr4	44176820	44177180
chr4	44624405	44624623
chr4	46252385	46252533
chr4	47427794	47427983
chr4	47514535	47514821
chr4	47560165	47560305
chr4	56428653	56428795
chr4	65188382	65188519
chr4	66356138	66356310
chr4	66467733	66467920
chr4	69885942	69886018
chr4	85611637	85611856
chr4	89653250	89653391
chr4	96761594	96762345
chr4	107956626	107956767
chr4	126355386	126355613
chr4	140432793	140432937
chr4	140640513	140640670
chr4	155719061	155719263
chr4	155719266	155719440
chr4	157771411	157771560
chr4	162306983	162307544
chr4	162459288	162459469
chr4	162697083	162697258
chr4	164272346	164272500
chr4	166981167	166981370
chr4	177052690	177052909
chr4	177608339	177608589
chr4	186380544	186380716
chr4	188924014	188924778
chr4	189012569	189013000
chr5	1221917	1222062
chr5	5239880	5240016
chr5	5319100	5319284
chr5	9190387	9190571
chr5	19473374	19473770
chr5	23510005	23510167
chr5	23521135	23521242
chr5	24487910	24488051
chr5	24491745	24491901
chr5	24505195	24505377
chr5	24537560	24537733
chr5	26915734	26915884
chr5	32712148	32712533
chr5	35876318	35876455
chr5	42719315	42719537
chr5	63256259	63257289
chr5	67576711	67576869
chr5	67591026	67591168
chr5	83402497	83402643
chr5	100222131	100222294
chr5	101592816	101592954
chr5	101593621	101593831
chr5	101834341	101834510
chr5	109190841	109190981
chr5	128983423	128983610
chr5	135692516	135692676
chr5	136324121	136324297

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr5	140182098	140182175
chr5	140182238	140182310
chr5	140182483	140182560
chr5	140182568	140182999
chr5	140222608	140222716
chr5	140562868	140563072
chr5	140810983	140811186
chr5	148407138	148407459
chr5	156346437	156346573
chr5	161116217	161116374
chr5	169423054	169423192
chr5	172659629	172659873
chr5	178413369	178413506
chr5	180661231	180661366
chr6	13365707	13365888
chr6	26189054	26189267
chr6	26204879	26205123
chr6	26251854	26251928
chr6	27100289	27100390
chr6	27775638	27775733
chr6	27858484	27858636
chr6	28053289	28053512
chr6	28097314	28097546
chr6	28543159	28543619
chr6	32411501	32411728
chr6	43323426	43323566
chr6	43612876	43613010
chr6	50682804	50683125
chr6	50696544	50696763
chr6	57472315	57472479
chr6	62604464	62604609
chr6	66115055	66115208
chr6	66200455	66200533
chr6	66204665	66205232
chr6	69653686	69653908
chr6	69666511	69666726
chr6	70070736	70071348
chr6	72011344	72011477
chr6	85446401	85446573
chr6	88218733	88218908
chr6	90432663	90432805
chr6	90459293	90459434
chr6	100841429	100841754
chr6	102250183	102250352
chr6	102337598	102337756
chr6	106552761	106553422
chr6	110714053	110714195
chr6	110763663	110763800
chr6	126080621	126080846
chr6	144263490	144263637
chr6	146720185	146720267
chr6	146720285	146720789
chr6	152475999	152476214
chr6	152497504	152497725
chr6	152674362	152674576
chr6	152690077	152690293
chr6	152694212	152694351
chr6	152712417	152712574
chr6	152737502	152737758
chr6	152768557	152768784
chr6	152786372	152786558
chr6	153073282	153073456
chr6	154412187	154412481
chr6	159653040	159654524
chr6	161011944	161012018
chr6	161012044	161012124
chr6	161152089	161152256
chr6	162683672	162683816
chr6	165715061	165715713
chr6	165792676	165792880
chr6	169053698	169053880
chr7	1275509	1275657
chr7	1481809	1482022
chr7	13935444	13935597
chr7	19156534	19156676

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr7	19184834	19184992
chr7	19765194	19765358
chr7	21641065	21641211
chr7	21675470	21675641
chr7	23292900	23293104
chr7	27194687	27194765
chr7	30795172	30795308
chr7	30961647	30961869
chr7	37252914	37253091
chr7	37780804	37780958
chr7	37955674	37956097
chr7	37988449	37988570
chr7	41739601	41739907
chr7	42017131	42017334
chr7	43635612	43635757
chr7	53103546	53104247
chr7	55241591	55241759
chr7	55242381	55242526
chr7	55248951	55249200
chr7	55259376	55259601
chr7	55260416	55260574
chr7	55266386	55266601
chr7	55267986	55268123
chr7	56126106	56126174
chr7	64166862	64166999
chr7	70885904	70886116
chr7	71175714	71175940
chr7	77797201	77797352
chr7	82544260	82545825
chr7	82764215	82764746
chr7	86394572	86394940
chr7	86415622	86416440
chr7	87144592	87144745
chr7	87150067	87150208
chr7	90355871	90356017
chr7	90546891	90547068
chr7	90741831	90742014
chr7	91709202	91709364
chr7	96650075	96650318
chr7	103969395	103969652
chr7	106508035	106508636
chr7	107315365	107315580
chr7	112461861	112462244
chr7	113558256	113559072
chr7	116411893	116412071
chr7	116417408	116417550
chr7	116418808	116419051
chr7	116422018	116422192
chr7	116423323	116423536
chr7	116435673	116435857
chr7	116435918	116436202
chr7	126086294	126086440
chr7	136699641	136700948
chr7	136938176	136938402
chr7	137075941	137076107
chr7	137206576	137206725
chr7	149461871	149462013
chr7	154585764	154585930
chr7	157926629	157926788
chr8	2037890	2038038
chr8	3216724	3216867
chr8	3265574	3265725
chr8	19810795	19810967
chr8	21903592	21903737
chr8	22020057	22020161
chr8	22020162	22020272
chr8	26722069	26722222
chr8	32617785	32617930
chr8	40554738	40554958
chr8	52323785	52323998
chr8	53853251	53853462
chr8	65493757	65494032
chr8	65494042	65494191
chr8	65494227	65494391
chr8	65528487	65528769

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr8	66539522	66539629
chr8	68930042	68930178
chr8	68965307	68965511
chr8	68968042	68968240
chr8	69011892	69012095
chr8	69046292	69046443
chr8	69445187	69445397
chr8	70476182	70476406
chr8	70512812	70513020
chr8	72755882	72755998
chr8	72756022	72756311
chr8	73480037	73480248
chr8	75756185	75756373
chr8	75898285	75898431
chr8	88365837	88366051
chr8	88885047	88885297
chr8	88885302	88885455
chr8	88885462	88885546
chr8	88885557	88886139
chr8	89179917	89180056
chr8	89198682	89198824
chr8	90937307	90937548
chr8	103289276	103289419
chr8	104337191	104337331
chr8	104427316	104427607
chr8	104897511	104898133
chr8	105263221	105263396
chr8	105436451	105436660
chr8	105463451	105463660
chr8	105502926	105503597
chr8	107718744	107718895
chr8	110980374	110980514
chr8	110984634	110985018
chr8	113267454	113267659
chr8	113657314	113657497
chr8	113933824	113934003
chr8	116616185	116616753
chr8	131916109	131916318
chr8	132051754	132052139
chr8	133141714	133141903
chr8	133187669	133187884
chr8	133899079	133899474
chr8	143694920	143695315
chr8	145637888	145638073
chr9	1056598	1056953
chr9	4118744	4118887
chr9	8404515	8404679
chr9	8485200	8485335
chr9	16419203	16419596
chr9	17761389	17761540
chr9	18776899	18777178
chr9	21968154	21968293
chr9	21968669	21968817
chr9	21970869	21971023
chr9	21971074	21971146
chr9	21974444	21974836
chr9	21994114	21994361
chr9	34976514	34976673
chr9	78789876	78790085
chr9	94486731	94487269
chr9	104385577	104385747
chr9	111617052	111618037
chr9	113538057	113538205
chr9	115806414	115806551
chr9	121929754	121930113
chr9	126135772	126136060
chr9	129957342	129957493
chr9	131478983	131479144
chr10	7325892	7326035
chr10	7657919	7658083
chr10	7679314	7679460
chr10	16932336	16932562
chr10	16970181	16970256
chr10	16970256	16970333
chr10	16982066	16982204

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr10	18242180	18242471
chr10	37507974	37508044
chr10	37508064	37508164
chr10	37508244	37508367
chr10	46967459	46967672
chr10	50121586	50121874
chr10	50374891	50375073
chr10	55912880	55913038
chr10	60558133	60558347
chr10	60936578	60936756
chr10	76857443	76857660
chr10	83635304	83635585
chr10	84118469	84118642
chr10	84498284	84498429
chr10	84733509	84733693
chr10	87362169	87362396
chr10	89692737	89692810
chr10	89692877	89692951
chr10	89692972	89693037
chr10	89711887	89711966
chr10	89717577	89717711
chr10	101816743	101816921
chr10	108458965	108459086
chr10	108536260	108536481
chr10	117221415	117221564
chr10	118030440	118030586
chr10	118236140	118236349
chr10	125804115	125804341
chr10	134015427	134015645
chr11	688305	688471
chr11	5536639	5537070
chr11	6261526	6261826
chr11	6265171	6265623
chr11	7324276	7324605
chr11	16007783	16007975
chr11	17757839	17757982
chr11	20959285	20959435
chr11	27679603	27680129
chr11	35640938	35641197
chr11	36520053	36520129
chr11	61607817	61607889
chr11	61607892	61607971
chr11	63137763	63137898
chr11	63398538	63398691
chr11	64419483	64419650
chr11	64559608	64559755
chr11	65349827	65349958
chr11	88911803	88911964
chr11	89916058	89916193
chr11	92087030	92087280
chr11	92430500	92430642
chr11	92616180	92616319
chr11	103907615	103908737
chr11	106680753	106681085
chr11	110450591	110451018
chr11	113853816	113854026
chr11	117374573	117374714
chr11	120996248	120996577
chr11	120998678	120999035
chr11	121000523	121000916
chr11	121016583	121016805
chr11	124794740	124794986
chr11	125255420	125255637
chr12	939134	939350
chr12	4735883	4736074
chr12	6632038	6632173
chr12	7635216	7635385
chr12	9161987	9162169
chr12	9754072	9754208
chr12	9833497	9833670
chr12	23696109	23696284
chr12	23728719	23728827
chr12	23893769	23893994
chr12	25378518	25378628
chr12	25378668	25378736

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr12	25380138	25380301
chr12	25380308	25380385
chr12	25398223	25398302
chr12	29614758	29614988
chr12	41900251	41900405
chr12	43944864	43945004
chr12	45410164	45410312
chr12	52910577	52910693
chr12	55420557	55421064
chr12	56647927	56648077
chr12	57553677	57553810
chr12	62786802	62787016
chr12	63544092	63544240
chr12	75601385	75601647
chr12	85266893	85267046
chr12	85450548	85450693
chr12	85517848	85517985
chr12	85531603	85531772
chr12	86373723	86374150
chr12	89744581	89744720
chr12	94975641	94975790
chr12	99640216	99640562
chr12	100704786	100704860
chr12	100704936	100705008
chr12	103352259	103352658
chr12	104476489	104476628
chr12	106460689	106460851
chr12	108169076	108169504
chr12	108985426	108985704
chr12	113704019	113704099
chr12	113704104	113704186
chr12	117768645	117768792
chr12	118198800	118199112
chr12	118599695	118599850
chr12	121972376	121972515
chr12	128899739	128899867
chr12	130184644	130184753
chr12	130184809	130184885
chr12	130184909	130185183
chr13	19751265	19751337
chr13	28014224	28014388
chr13	32745291	32745431
chr13	33590892	33591034
chr13	36046510	36046687
chr13	36686030	36686279
chr13	36748835	36749047
chr13	36886430	36886636
chr13	36888345	36888480
chr13	37427655	37427803
chr13	38237585	38237799
chr13	48985620	48985766
chr13	58206696	58208461
chr13	58298751	58299440
chr13	66878731	66878891
chr13	73636031	73636448
chr13	74518091	74518231
chr13	78477276	78477414
chr13	92345686	92345987
chr13	94482380	94482674
chr13	102047515	102047743
chr13	114083234	114083422
chr14	24534187	24534331
chr14	26917367	26917620
chr14	26917622	26918013
chr14	29237015	29237623
chr14	42356512	42356900
chr14	45644623	45644841
chr14	52507344	52507559
chr14	52534609	52534757
chr14	59930630	59930772
chr14	60193815	60193959
chr14	65007927	65008064
chr14	65241942	65242105
chr14	68052644	68052856
chr14	77275791	77276011

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr14	88945764	88945951
chr14	90650455	90650744
chr14	94964307	94964766
chr14	100317221	100317362
chr15	23810994	23812075
chr15	23812104	23812292
chr15	23931449	23932333
chr15	25221988	25222205
chr15	25958993	25959413
chr15	27572003	27572180
chr15	28358221	28358406
chr15	31851263	31851391
chr15	33954389	33954851
chr15	48062677	48063833
chr15	49217039	49217144
chr15	49327689	49327837
chr15	51350391	51350633
chr15	54586066	54586272
chr15	59368205	59368361
chr15	70960066	70960430
chr15	83226495	83226647
chr15	83332530	83332744
chr15	89861776	89861937
chr15	91548896	91549040
chr15	92459456	92459611
chr16	9857831	9858081
chr16	14041800	14041955
chr16	31926427	31926615
chr16	49430376	49430561
chr16	51171023	51171301
chr16	51171303	51171376
chr16	55362580	55362656
chr16	55362685	55363221
chr16	55690580	55690735
chr16	61760969	61761154
chr16	61851374	61851518
chr16	61890990	61891168
chr16	61935255	61935413
chr16	65022085	65022275
chr16	66955989	66956133
chr16	67333294	67333438
chr16	77228269	77228449
chr16	77353719	77354000
chr16	77769694	77769909
chr16	80638256	80638472
chr16	80654631	80654861
chr16	81969830	81969975
chr16	86544628	86544841
chr17	7572889	7573037
chr17	7573894	7574051
chr17	7576519	7576721
chr17	7576809	7576957
chr17	7576984	7577173
chr17	7577469	7577648
chr17	7578154	7578326
chr17	7578339	7578589
chr17	7579289	7579605
chr17	7579639	7579772
chr17	7579804	7579954
chr17	8926044	8926230
chr17	10402265	10402434
chr17	10416150	10416306
chr17	21319098	21319177
chr17	21319178	21319512
chr17	21319628	21319704
chr17	21319713	21319791
chr17	26874611	26874770
chr17	26962076	26962218
chr17	27248681	27248859
chr17	37879768	37879938
chr17	37880143	37880277
chr17	37880943	37881204
chr17	37881268	37881488
chr17	37881538	37881678
chr17	37881938	37882156

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr17	37882788	37882929
chr17	40556914	40557410
chr17	40837296	40837461
chr17	44845763	44846043
chr17	51900576	51902327
chr17	56344734	56344879
chr17	66938048	66938188
chr17	75208118	75208269
chr17	79414058	79414333
chr18	6896474	6896661
chr18	9887936	9888078
chr18	31537375	31537513
chr18	44560419	44560951
chr18	48591799	48591943
chr18	48593364	48593571
chr18	55247265	55247478
chr18	59195203	59195340
chr18	60642613	60642830
chr18	74963029	74963171
chr19	5244021	5244366
chr19	7687191	7687359
chr19	8609156	8609380
chr19	8808048	8808442
chr19	11134169	11134342
chr19	21992290	21992423
chr19	21992445	21992515
chr19	22157475	22157578
chr19	22364180	22364329
chr19	22942295	22942368
chr19	22942385	22942466
chr19	31039111	31040259
chr19	35842887	35843029
chr19	37210077	37210321
chr19	37440412	37440659
chr19	37975047	37975200
chr19	40408494	40408566
chr19	43698620	43698690
chr19	46443765	46443918
chr19	47935615	47935688
chr19	49385265	49385472
chr19	51189464	51189643
chr19	51330069	51330210
chr19	51645656	51645811
chr19	52272206	52272803
chr19	52327701	52328023
chr19	52538216	52538362
chr19	53612546	53612935
chr19	54310713	54310933
chr19	55593798	55593996
chr19	55815008	55815221
chr19	57293289	57293504
chr19	58048575	58048946
chr19	58601285	58601365
chr19	58601390	58601493
chr20	1961010	1961155
chr20	2375810	2375956
chr20	9524993	9525161
chr20	11904050	11904305
chr20	21494090	21494308
chr20	21687095	21687423
chr20	21695360	21695505
chr20	23016277	23016591
chr20	23807162	23807243
chr20	25462581	25462771
chr20	30584543	30584934
chr20	34022178	34022541
chr20	35060150	35061028
chr20	39831656	39831808
chr20	41419816	41420067
chr20	44680361	44680513
chr20	44839061	44839205
chr20	44845441	44845647
chr20	54578919	54579121
chr20	57035998	57036348
chr20	57042353	57042494

TABLE 9-continued

Chromosome	Start (bp)	End (bp)
chr20	57829216	57829641
chr20	60448762	60448971
chr20	60887662	60887869
chr20	61542409	61542627
chr20	62045419	62045559
chr20	62121839	62122023
chr21	10906931	10907002
chr21	15872914	15873066
chr21	22849657	22849793
chr21	28296513	28296924
chr21	28327036	28327210
chr21	31538276	31538876
chr21	32638600	32639172
chr21	36206693	36206849
chr21	38302583	38302718
chr21	41414424	41414618
chr22	17072700	17073132
chr22	18028270	18028588
chr22	22892467	22892686
chr22	24583982	24584137
chr22	26423336	26423552
chr22	38121507	38121820
chr22	39626067	39626279
chr22	40140083	40140243
chr22	41076940	41077923
chr22	45281706	45281848
chr22	46318804	46318959
chr22	46326974	46327113
chr22	50302869	50303006
chrX	62875334	62875489

TABLE 10

chr16	3786650	3786816	CREBBP
chr16	3788559	3788673	CREBBP
chr9	33798483	33798620	PRSS3
chr7	148508714	148508813	EZH2
chr22	23230232	23230432	IGLL5
chr18	60985291	60985897	BCL2
chr12	57496608	57496707	STAT6
chr7	2979473	2979572	CARD11
chr6	27114203	27114486	HIST1H2BK
chr9	33796640	33796800	PRSS3
chr1	39322631	39322730	RRAGC
chr1	2491261	2491417	TNFRSF14
chr17	62006585	62006684	CD79B
chr12	49415825	49415934	MLL2
chrX	150573387	150573530	VMA21
chr1	150727476	150727626	CTSS
chr9	33798014	33798113	PRSS3
chr6	26156786	26157248	HIST1H1E
chr20	17639667	17640053	RRBP1
chr1	2492058	2492157	TNFRSF14
chr12	49424675	49424816	MLL2
chr12	49433246	49433389	MLL2
chrX	153663644	153663743	ATP6AP1
chr8	20074730	20074835	ATP6V1B2
chr18	9887338	9887437	TXNDC2
chr16	57983250	57983349	CNGB1
chr22	41565506	41565620	EP300
chr2	119604215	119604314	EN1
chr3	183273198	183273297	KLHL6
chr7	142131525	142131624	TRBV5-6
chr4	146695657	146695824	ZNF827
chr19	19260016	19260115	MEF2B
chr20	48522108	48522207	SPATA2
chr2	51254638	51254737	NRXN1
chr10	94452434	94452533	HHEX
chr1	150470131	150470230	TARS2
chr19	50861848	50861947	NAPSA
chr19	55903047	55903146	RPL28
chr5	149792187	149792312	CD74

TABLE 10-continued

chr6	26124577	26124741	HIST1H2AC
chr9	1056514	1056613	DMRT2
chr1	2489781	2489907	TNFRSF14
chr1	2493111	2493254	TNFRSF14
chr17	48823117	48823216	LUC7L3
chr1	52933846	52933945	ZCCHC11
chr12	49440435	49440534	MLL2
chr6	26234697	26234796	HIST1H1D
chr3	42787414	42787519	CCDC13
chr7	121653384	121653483	PTPRZ1
chr16	1823023	1823122	MRPS34
chr12	92539163	92539311	BTG1
chr3	141162243	141162342	ZBTB38
chr10	90773888	90774026	FAS
chr8	40011192	40011291	C8orf4
chr6	26123881	26123980	HIST1H2BC
chr12	113496061	113496212	DTX1
chr2	43452587	43452686	ZFP36L2
chr5	140176763	140176862	PCDHA2
chr6	37138342	37138441	PIM1
chr11	86133615	86133757	CCDC81
chr7	87912060	87912159	STEAP4
chr2	182413251	182413350	CERKL
chr6	32906520	32906619	HLA-DMB
chr12	39756899	39757015	KIF21A
chr15	45814435	45814534	SLC30A4
chr15	42147520	42147619	SPTBN5
chr9	33799025	33799178	PRSS3
chr6	132270569	132270668	CTGF
chr2	232660773	232660872	COPS7B
chr10	101147908	101148058	CNNM1
chr17	5036195	5036294	USP6
chr6	160953562	160953681	LPA
chr1	160182899	160183055	PEA15
chrX	119388935	119389034	ZBTB33
chr14	51237122	51237221	NIN
chr1	78401606	78401705	NEXN
chr7	27204662	27204761	HOXA9
chr16	85954780	85954882	IRF8
chr16	19883730	19883829	GPRC5B
chr20	39991558	39991657	EMILIN3
chr9	90260809	90260929	DAPK1
chr9	34658516	34658680	IL11RA
chr12	49425799	49426352	MLL2
chr20	55840785	55840987	BMP7
chr6	27835005	27835210	HIST1H1B
chr2	240981542	240982132	PRR21
chr19	22155090	22155726	ZNF208
chr4	1388358	1388594	CRIPAK
chr12	11461506	11461743	PRB4
chr12	11214190	11214473	TAS2R46
chr3	147108721	147109023	ZIC4
chr7	48312016	48312587	ABCA13
chr12	49430943	49432739	MLL2
chr12	49420059	49420689	MLL2
chr1	203274734	203274876	BTG2
chr22	41566409	41566575	EP300
chr4	126242585	126242703	FAT4
chr11	27384450	27384549	CCDC34
chr19	10335462	10335561	S1PR2
chr4	38799494	38799593	TLR1
chr6	136594219	136594325	BCLAF1
chr22	29885560	29885659	NEFH
chr10	70547910	70548021	CCAR1
chr13	33716428	33716527	STARD13
chrX	142718477	142718576	SLITRK4
chr20	23615890	23616004	CST3
chr7	138969220	138969319	UBN2
chr1	21808089	21808262	NBPF3
chr16	28603746	28603845	SULT1A2
chr2	166872116	166872237	SCN1A
chr1	214170811	214170950	PROX1
chr21	35189750	35189849	ITSN1
chr16	3781275	3781374	CREBBP
chr8	113484819	113484936	CSMD3
chr17	61775911	61776071	LIMD2
chr12	18644384	18644492	PIK3C2G

TABLE 10-continued

chr8	48736418	48736557	PRKDC
chr9	133957445	133957548	LAMC3
chrX	125955251	125955356	CXorf64
chr14	50246930	50247040	KLHDC2
chrX	21450701	21450800	CNKSR2
chr17	45214636	45214735	CDC27
chr4	17706617	17706716	FAM184B
chr3	75787081	75787180	ZNF717
chr9	130742270	130742416	FAM102A
chr1	171123267	171123366	FMO6P
chr1	21031259	21031369	KIF17
chr2	96617076	96617175	ANKRD36C
chr4	148589689	148589796	PRMT10
chr2	160239061	160239160	BAZ2B
chr16	1279269	1279439	TPSB2
chr1	46087006	46087105	CCDC17
chr8	52733109	52733270	PCMTD1
chr6	26045849	26045948	HIST1H3C
chr1	2489164	2489273	TNFRSF14
chr6	168377012	168377111	HGC6.3
chr10	129901079	129901178	MKI67
chr17	7578458	7578557	TP53
chr12	85521621	85521720	LRRIQ1
chr9	139753456	139753584	MAMDC4
chr14	80327738	80327837	NRXN3
chr1	149883474	149883573	SV2A
chrX	32663176	32663275	DMD
chr22	26829629	26829728	ASPHD2
chr19	35828674	35828773	CD22
chr12	49416398	49416497	MLL2
chr12	49427855	49427954	MLL2
chr12	49437417	49437565	MLL2
chr12	49439847	49439957	MLL2
chr12	49444221	49444346	MLL2
chr12	49446989	49447104	MLL2
chr6	110714271	110714393	DDO
chrX	23410992	23411091	PTCHD1
chr7	299761	299860	FAM20C
chr1	85733436	85733535	BCL10
chr6	27861455	27861569	HIST1H2BO
chr7	13935512	13935611	ETV1
chr7	70231146	70231245	AUTS2
chr17	79479257	79479380	ACTG1
chr18	40854102	40854201	SYT4
chr2	114691855	114691963	ACTR3
chr14	47426601	47426752	MDGA2
chr3	50293623	50293752	GNAI2
chr7	2977540	2977666	CARD11
chr11	118343589	118343688	MLL
chr1	10689828	10689937	PEX14
chr11	111249844	111249943	POU2AF1
chr9	91965694	91965793	SECISBP2
chr17	43011718	43011817	KIF18B
chr3	64536567	64536738	ADAMTS9
chr1	111957481	111957580	OVGP1
chr17	18145198	18145313	LLGL1
chr19	13054633	13054732	CALR
chr6	29911909	29912008	HLA-A
chrX	153663458	153663557	ATP6AP1
chr1	158913594	158913693	PYHIN1
chr5	158141107	158141206	EBF1
chr1	228475527	228475626	OBSCN
chr3	9594028	9594127	LHFPL4
chr8	2910008	2910136	CSMD1
chr1	12337499	12337598	VPS13D
chr6	41903681	41903780	CCND3
chr1	150443929	150444028	RPRD2
chr6	74229045	74229144	EEF1A1
chr6	128298067	128298199	PTPRK
chr8	20073915	20074014	ATP6V1B2
chr10	97101320	97101435	SORBS1
chr4	155505491	155505598	FGA
chr12	104379380	104379506	TDG
chr12	11506386	11506485	PRB1
chr19	15132617	15132731	CCDC105
chr8	145024683	145024815	PLEC
chr16	67911411	67911559	EDC4

TABLE 10-continued

chr11	66639494	66639630	PC
chr6	165711465	165711590	C6orf118
chrX	79932311	79932457	BRWD3
chr15	54586092	54586262	UNC13C
chr12	108954825	108954924	SART3
chr20	29631533	29631632	FRG1B
chr12	57905480	57905651	MARS
chr21	43256219	43256318	PRDM15
chr6	170627609	170627708	FAM120B
chr8	8750154	8750253	MFHAS1
chr1	240370922	240371023	FMN2
chr1	214818796	214818895	CENPF
chr22	37425300	37425399	MPST
chr10	51465512	51465691	AGAP7
chr12	46244635	46244816	ARID2
chr1	68512352	68512761	DIRAS3
chrX	7811644	7811830	VCX
chr7	127894552	127894740	LEP
chr4	189012637	189012828	TRIML2
chr20	43726332	43726529	KCNS1
chr5	140605138	140605339	PCDHB14
chr6	78172235	78173066	HTR1B
chr18	30350002	30350211	KLHL14
chrX	152244293	152244510	PNMA6D
chr12	11286598	11286821	TAS2R30
chr1	31194363	31194587	MATN1
chr4	187524361	187524591	FAT1
chr17	63010386	63010623	GNA13
chr19	50549249	50549492	ZNF473
chr14	104643890	104644134	KIF26A
chr16	1306816	1307061	TPSD1
chr7	151945006	151945257	MLL3
chrX	27839561	27839820	MAGEB10
chr22	23523223	23523841	BCR
chr17	57290162	57290449	SMG8
chr6	26056112	26056422	HIST1H1C
chr14	86088478	86088869	FLRT2
chr9	42410027	42410426	ANKRD20A2
chr17	16612377	16612838	CCDC144A
chr14	33292488	33292963	AKAP6
chr6	1390621	1391103	FOXF2
chr11	85436759	85437249	SYTL2
chr1	245027101	245027593	HNRNPU
chr13	41239782	41240281	FOXO1
chr5	150945512	150946027	FAT2
chr1	201178875	201180218	IGFN1
chr12	49433523	49434895	MLL2
chrX	140993858	140995691	MAGEC1
chr11	70332042	70332575	SHANK2
chr2	55252510	55253083	RTN4
chr19	16687146	16687737	MED26
chrX	125298695	125299314	DCAF12L2
chr7	82582905	82583627	PCLO
chr18	65180943	65181675	DSEL
chr5	5461354	5462093	KIAA0947
chr3	40528745	40529496	ZNF619
chr1	249141669	249142463	ZNF672
chr2	136872540	136873336	CXCR4
chr1	24201100	24201996	CNR2
chr11	6567173	6568114	DNHD1
chr16	89350182	89351139	ANKRD11
chr12	49422855	49422954	MLL2
chr12	49428357	49428456	MLL2
chr12	49428594	49428718	MLL2
chr12	49433004	49433141	MLL2
chr12	49435961	49436060	MLL2
chr12	49438185	49438305	MLL2
chr12	49440042	49440207	MLL2
chr12	49441747	49441852	MLL2
chr12	49447258	49447424	MLL2
chr12	49448260	49448359	MLL2
chr16	3786036	3786204	CREBBP
chr16	3808854	3808973	CREBBP
chr16	3817794	3817893	CREBBP
chr16	3819151	3819250	CREBBP
chr16	3828011	3828183	CREBBP
chr16	3830732	3830879	CREBBP

TABLE 10-continued

chr16	3900823	3900922	CREBBP
chr9	33794780	33794879	PRSS3
chr1	2488088	2488187	TNFRSF14
chr22	23235876	23235998	IGLL5
chr22	23237632	23237731	IGLL5
chr1	16893673	16893846	NBPF1
chr1	16910088	16910191	NBPF1
chr1	16918406	16918505	NBPF1
chr2	96614261	96614360	ANKRD36C
chr1	145299788	145299887	NBPF10
chr1	145302725	145302824	NBPF10
chr1	145314191	145314290	NBPF10
chr1	145323629	145323728	NBPF10
chr1	145336256	145336355	NBPF10
chr1	145368413	145368512	NBPF10
chr1	148010883	148011056	NBPF14
chr1	148013295	148013394	NBPF14
chr1	148017501	148017665	NBPF14
chr1	148021552	148021651	NBPF14
chr1	148025746	148025845	NBPF14
chr7	148506392	148506491	EZH2
chr7	151836759	151836876	MLL3
chr7	151859918	151860017	MLL3
chr7	151878655	151878754	MLL3
chr12	57493776	57493875	STAT6
chr12	57498246	57498369	STAT6
chr12	57499029	57499128	STAT6
chr1	146406508	146406607	NBPF12
chr1	146436711	146436810	NBPF12
chr1	146448373	146448546	NBPF12
chr1	146457897	146458070	NBPF12
chr8	3047432	3047531	CSMD1
chr8	3081250	3081389	CSMD1
chr18	60793423	60793599	BCL2-NA
chr18	60774470	60774594	BCL2-NA
chr18	60763905	60763963	BCL2-NA
chr18	60764357	60764467	BCL2-NA
chr14	107169930	107170235	IGHV1-69.1
chr14	107170321	107170428	IGHV1-69.2
chr14	106610312	106610623	IGHV3-15.1
chr14	106610726	106610852	IGHV3-15.2
chr14	106691672	106691977	IGHV3-21.1
chr14	106692078	106692203	IGHV3-21.2
chr14	106725200	106725505	IGHV3-23.1
chr14	106725608	106725733	IGHV3-23.2
chr14	106791004	106791309	IGHV3-30.1
chr14	106791410	106791536	IGHV3-30.2
chr14	106993813	106994118	IGHV3-48.1
chr14	106994221	106994346	IGHV3-48.2
chr14	107218675	107218980	IGHV3-74.1
chr14	107219083	107219365	IGHV3-74.2
chr14	106829593	106829895	IGHV4-34.1
chr14	106829978	106830076	IGHV4-34.2
chr14	106877618	106877926	IGHV4-39.1
chr14	106878009	106878126	IGHV4-39.2
chr14	106329407	106329468	IGHJ6
chr14	106330023	106330072	IGHJ5
chr14	106330424	106330470	IGHJ4
chr14	106330796	106330845	IGHJ3
chr14	106331408	106331460	IGHJ2
chr14	106331616	106331668	IGHJ1

TABLE 11

Chromosome	Start (bp)	End (bp)	Gene
chr17	7578383	7578554	TP53
chr17	7577018	7577155	TP53
chr17	7578176	7578289	TP53
chr9	21971016	21971199	CDKN2A
chr17	7577498	7577608	TP53
chr3	178935997	178936122	PIK3CA
chr9	21970899	21971199	CDKN2A
chr20	29628226	29628331	FRG1B

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	7579311	7579580	TP53
chr2	178098803	178098974	NFE2L2
chr20	29625872	29625984	FRG1B
chr9	139412203	139412382	NOTCH1
chr1	145302645	145302744	NBPF10
chr3	178951963	178952086	PIK3CA
chr9	20414286	20414385	MLLT3
chr4	153247174	153247380	FBXW7
chr11	534211	534322	HRAS
chr17	7576839	7576938	TP53
chr1	145367713	145367822	NBPF10
chr19	40367823	40367922	FCGBP
chr6	29910600	29910699	HLA-A
chr7	86394555	86394735	GRM3
chr5	24511435	24511616	CDH10
chr8	107782022	107782216	ABRA
chr1	27100070	27100208	ARID1A
chr17	26684313	26684473	POLDIP2
chr2	141359045	141359175	LRP1B
chr16	72188111	72188258	PMFBP1
chr9	139402683	139402837	NOTCH1
chr3	157146110	157146277	VEPH1
chr12	124798915	124799014	FAM101A
chrX	79999593	79999692	BRWD3
chr18	14542736	14543021	POTEC
chr16	65032521	65032725	CDH11
chr14	19553544	19553820	POTEG
chr12	81471975	81472120	ACSS3
chr7	55209978	55210130	EGFR
chr6	119337959	119338094	FAM184A
chr6	152763209	152763380	SYNE1
chrX	79281102	79281201	TBX22
chr3	109049450	109049549	DPPA4
chr7	111368415	111368577	DOCK4
chr22	22127161	22127271	MAPK1
chr14	62547692	62547864	SYT16
chr1	16464346	16464479	EPHA2
chr16	20442541	20442643	ACSM5
chr16	10995894	10996041	CIITA
chr16	64984709	64984857	CDH11
chr9	37014992	37015149	PAX5
chr6	31975095	31975194	CYP21A2
chr9	139418202	139418374	NOTCH1
chr7	53103444	53104150	POM121L12
chr6	27839691	27840063	HIST1H3I
chr5	89943366	89943472	GPR98
chr2	125192072	125192237	CNTNAP5
chr14	69701455	69701571	EXD2
chr3	181430808	181430907	SOX2
chr7	6426828	6426927	RAC1
chr22	41652714	41652828	RANGAP1
chr6	123869598	123869757	TRDN
chr12	113515302	113515401	DTX1
chr20	1961099	1961356	PDYN
chr1	217955515	217955664	SPATA17
chr19	24010293	24010549	RPSAP58
chr9	21974671	21974770	CDKN2A
chr2	202131209	202131506	CASP8
chr11	40135933	40137642	LRRC4C
chr2	80529766	80530938	LRRTM1
chr3	178916835	178916947	PIK3CA
chr16	75512868	75513584	CHST6
chr19	22154038	22157389	ZNF208
chr8	139163586	139165359	FAM135B
chr6	26204884	26205157	HIST1H4E
chr12	11545925	11546908	PRB2
chr5	63256300	63257092	HTR1A
chr7	154561126	154561281	DPP6
chr7	95157419	95157521	ASB4
chr6	57512589	57512692	PRIM2
chr8	113585729	113585886	CSMD3
chr4	147560456	147560568	POU4F2
chr1	145368535	145368634	NBPF10
chr7	37955723	37956081	SFRP4
chr13	88327767	88330089	SLITRK5

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chr4	187539226	187542862	FAT1
chr6	27834626	27835171	HIST1H1B
chr5	140261864	140264052	PCDHA13
chr6	66204658	66205125	EYS
chr1	57257787	57258100	C1orf168
chr7	21639523	21639717	DNAH11
chr18	5397092	5397423	EPB41L3
chr2	202149545	202150040	CASP8
chr1	157555965	157556231	FCRL4
chr5	24537560	24537765	CDH10
chr8	73848741	73850116	KCNB2
chr1	197390340	197391060	CRB1
chr18	13884632	13885468	MC2R
chr4	187627773	187630693	FAT1
chr5	26885756	26885968	CDH9
chr7	88962839	88966280	ZNF804B
chr5	140175844	140176834	PCDHA2
chr19	30934638	30936599	ZNF536
chrX	140993367	140996183	MAGEC1
chr19	20727462	20728771	ZNF737
chr8	88885017	88886181	DCAF4L2
chr15	23811026	23812218	MKRN3
chr4	134071331	134073620	PCDH10
chr12	7636016	7636248	CD163
chr7	11675952	11676535	THSD7A
chr6	96651055	96652002	FUT9
chr10	84745206	84745340	NRG3
chr1	248028042	248028156	TRIM58
chr3	30691783	30691948	TGFBR2
chr3	183756306	183756405	HTR3D
chr1	198713182	198713332	PTPRC
chr14	52520338	52520463	NID2
chr15	26806217	26806316	GABRB3
chr8	139601514	139601677	COL22A1
chr1	176738745	176738864	PAPPA2
chr2	138414365	138414539	THSD7B
chr2	209308081	209308255	PTH2R
chr8	113256622	113256747	CSMD3
chr8	114110998	114111145	CSMD3
chr7	11630120	11630219	THSD7A
chr16	20570578	20570738	ACSM2B
chr7	142459655	142459790	PRSS1
chr11	132016188	132016287	NTM
chr5	176709465	176709582	NSD1
chr10	55955479	55955595	PCDH15
chr5	11082807	11082958	CTNND2
chr19	54466452	54466611	CACNG8
chr1	104115728	104115870	AMY2B
chr5	13719087	13719207	DNAH5
chr14	47504313	47504489	MDGA2
chr1	75072309	75072554	C1orf173
chr17	21318730	21319867	KCNJ12
chr5	23522737	23522988	PRDM9
chr7	34118610	34118795	BMPER
chr13	36700036	36700223	DCLK1
chr5	140236299	140237322	PCDHA10
chrX	37026543	37029321	FAM47C
chr4	96761393	96762283	PDHA2
chr3	147113699	147114230	ZIC4
chr18	64172066	64172406	CDH19
chr5	140214168	140216301	PCDHA7
chrX	74494188	74494382	UPRT
chr17	80788943	80790329	ZNF750
chr14	44973722	44976121	FSCB
chr19	57174960	57176561	ZNF835
chr1	240370333	240371753	FMN2
chr1	216850475	216850747	ESRRG
chr2	15415719	15415924	NBAS
chr19	52618923	52620045	ZNF616
chr5	23526394	23527863	PRDM9
chr5	140180883	140182957	PCDHA3
chr19	22940344	22941816	ZNF99
chr12	4479555	4479838	FGF23
chr14	23346446	23346654	LRP10
chr19	43268169	43268378	PSG8

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chr19	54677829	54678114	MBOAT7
chr12	57484985	57485458	NAB2
chr19	22940344	22942465	ZNF99
chr10	135438780	135438991	FRG2B
chr18	63547682	63547974	CDH7
chr3	169540213	169540508	LRRIQ4
chr13	58206801	58208988	PCDH17
chr5	45262057	45262790	HCN1
chr7	121943817	121944308	FEZF1
chr19	10610148	10610643	KEAP1
chr12	11420457	11421069	PRB3
chr13	108518048	108518794	FAM155A
chr22	37603210	37603433	SSTR3
chr9	119976687	119976992	ASTN2
chrX	34960983	34962806	FAM47B
chr6	116937940	116938344	RSPH4A
chr5	140552560	140554406	PCDHB7
chr9	112898500	112900136	PALM2-AKAP2
chr19	5455842	5456254	ZNRF4
chr18	13826242	13826657	MC5R
chr3	155199247	155200710	PLCH1
chr7	63679732	63680528	ZNF735
chr3	148458872	148459825	AGTR1
chr15	23889143	23890841	MAGEL2
chr5	140474515	140476703	PCDHB2
chrX	142795188	142795519	SPANXN2
chr1	190067523	190068200	FAM5C
chr8	145770918	145771163	ARHGAP39
chr1	205038983	205039124	CNTN2
chr2	141081461	141081635	LRP1B
chr3	132435600	132435753	NPHP3
chr3	109026902	109027050	DPPA2
chr6	119341141	119341266	FAM184A
chr1	205779409	205779509	SLC41A1
chr20	33033160	33033259	ITCH
chr18	64178804	64178922	CDH19
chr6	129714206	129714305	LAMA2
chr19	39103250	39103382	MAP4K1
chr2	200173514	200173613	SATB2
chr11	45241141	45241257	PRDM11
chr2	28634819	28634950	FOSL2
chr3	97439104	97439254	EPHA6
chr14	105996001	105996100	TMEM121
chr3	30713558	30713699	TGFBR2
chr15	43927924	43928024	CATSPER2
chr1	149905298	149905423	MTMR11
chr17	36927373	36927506	PIP4K2B
chr3	140167411	140167510	CLSTN2
chr17	10248783	10248933	MYH13
chr10	33199173	33199272	ITGB1
chr19	55401000	55401099	FCAR
chr12	109972412	109972571	UBE3B
chr2	160206240	160206346	BAZ2B
chr3	157820502	157820663	SHOX2
chr19	50370310	50370461	PNKP
chr20	44519119	44519289	NEURL2
chr2	79254932	79255059	REG3G
chr1	196295846	196296019	KCNT2
chr14	30046462	30046561	PRKD1
chr6	30297128	30297276	TRIM39
chr1	240492638	240492768	FMN2
chr19	58991733	58991832	ZNF446
chr4	1957842	1957941	WHSC1
chr15	75641370	75641469	NEIL1
chr6	55113473	55113572	HCRTR2
chr3	157920860	157921034	RSRC1
chr8	113249418	113249577	CSMD3
chr8	113317028	113317139	CSMD3
chr8	113599294	113599464	CSMD3
chr2	50280442	50280583	NRXN1
chr13	32972545	32972675	BRCA2
chr18	909476	909586	ADCYAP1
chr18	14513660	14513784	POTEC
chr8	110509147	110509296	PKHD1L1
chr3	168838843	168839000	MECOM

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chr4	187557302	187557401	FAT1
chr17	10369589	10369733	MYH4
chr3	37048468	37048567	MLH1
chr5	167812223	167812360	WWC1
chr10	131640391	131640490	EBF3
chr20	278639	278738	ZCCHC3
chr10	131565150	131565249	MGMT
chr1	155887289	155887463	KIAA0907
chr6	26216706	26216848	HIST1H2BG
chr12	54396230	54396329	HOXC9
chr19	18084918	18085017	KCNN1
chr20	20177277	20177408	C20orf26
chr9	120470840	120471007	TLR4
chr12	106708135	106708271	TCP11L2
chr15	84581961	84582060	ADAMTSL3
chr9	139405104	139405257	NOTCH1
chr9	139412588	139412744	NOTCH1
chr9	139413069	139413215	NOTCH1
chr4	79455602	79455769	FRAS1
chr5	45396639	45396738	HCN1
chr19	17088177	17088350	CPAMD8
chr9	136234190	136234289	SURF4
chr20	13839907	13840080	SEL1L2
chr11	122849988	122850144	BSX
chr12	50367086	50367244	AQP6
chr6	41165983	41166105	TREML2
chr4	62679517	62679616	LPHN3
chr5	101834370	101834544	SLCO6A1
chr2	79349113	79349251	REG1A
chr17	7573926	7574033	TP53
chr2	217124225	217124379	MARCH4
chr15	89424689	89424834	HAPLN3
chr16	77465304	77465454	ADAMTS18
chr5	11364859	11364958	CTNND2
chr5	11397142	11397315	CTNND2
chr10	103826304	103826403	HPS6
chr7	83610636	83610794	SEMA3A
chr2	202151181	202151317	CASP8
chr19	11470231	11470368	LPPR2
chr19	12951792	12951891	MAST1
chr10	108389024	108389131	SORCS1
chr5	26903771	26903931	CDH9
chr2	1459847	1460008	TPO
chr19	2121161	2121310	AP3D1
chrX	78616825	78616976	ITM2A
chr6	66044911	66045010	EYS
chr5	13901431	13901564	DNAH5
chr19	37733482	37733581	ZNF383
chr22	50654145	50654296	SELO
chrX	12734345	12734914	FRMPD4
chr8	92972553	92972728	RUNX1T1
chr1	161518210	161518385	FCGR3A
chr2	164466119	164467942	FIGN
chr6	46107689	46108044	ENPP4
chr11	22301127	22301308	ANO5
chr19	54313207	54314440	NLRP12
chr3	126707543	126708597	PLXNA1
chr3	73432745	73433987	PDZRN3
chr14	69256532	69257127	ZFP36L1
chr7	72413633	72413897	POM121
chr3	147127953	147128847	ZIC1
chr1	186275518	186277191	PRG4
chr11	30032399	30034074	KCNA4
chr4	9783798	9785062	DRD5
chr1	74506981	74507589	LRRIQ3
chr7	87913171	87913539	STEAP4
chr6	32020543	32020731	TNXB
chr15	23931583	23932342	NDN
chrX	34148022	34150221	FAM47A
chr6	26271336	26271610	HIST1H3G
chr7	146829389	146829579	CNTNAP2
chr1	12887252	12887626	PRAMEF11
chr19	22362785	22364285	ZNF676
chr2	227924129	227924320	COL4A4
chr10	68686714	68688016	LRRTM3

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chrX	90690578	90691202	PABPC5
chr5	11346513	11346705	CTNND2
chr22	32586994	32587271	RFPL2
chr12	49420048	49420673	MLL2
chr12	130184385	130185157	TMEM132D
chr7	57187578	57188705	ZNF479
chr4	164393536	164394866	TKTL2
chr7	86415632	86416017	GRM3
chr8	56015392	56015675	XKR4
chr20	50139751	50140541	NFATC2
chr2	56419678	56420320	CCDC85A
chr15	48500014	48500300	SLC12A1
chr5	140589637	140590605	PCDHB12
chr6	26158464	26158753	HIST1H2BD
chr4	162306901	162307559	FSTL5
chr17	10303757	10304049	MYH8
chr6	26225382	26225675	HIST1H3E
chr6	26056152	26056553	HIST1H1C
chr3	113724487	113724692	KIAA1407
chr19	55106643	55106849	LILRA1
chr4	110667390	110667596	CFI
chr14	42355977	42357172	LRFN5
chr7	57528633	57529305	ZNF716
chr19	53643670	53644867	ZNF347
chr12	78400291	78400964	NAV3
chr6	112671162	112671571	RFPL4B
chr6	87725250	87726087	HTR1E
chr6	31323135	31323344	HLA-B
chr12	125397966	125398268	UBC
chr1	111215825	111217040	KCNA3
chr3	129695648	129695952	TRH
chr1	38227190	38227732	EPHA10
chr1	16475128	16475543	EPHA2
chr9	121929388	121930239	DBC1
chr5	140536956	140537263	PCDHB17
chr12	130921488	130921798	RIMBP2
chr10	25886753	25887455	GPR158
chr2	108626705	108626921	SLC5A7
chr22	40815101	40815317	MKL1
chr1	149784912	149785128	HIST2H3D
chrX	26212040	26212352	MAGEB6
chr21	28338152	28338578	ADAMTS5
chr19	58384571	58386285	ZNF814
chr20	9561037	9561466	PAK7
chr4	52860789	52862055	LRRC66
chr1	148594406	148594625	NBPF15
chr19	36357105	36357326	KIRREL2
chr3	197427592	197427813	KIAA0226
chr19	55450491	55451378	NLRP7
chr17	7751610	7752329	KDM6B
chr19	46056874	46057096	OPA3
chr10	117884807	117885029	GFRA1
chr8	110980373	110980810	KCNV1
chr1	214170118	214171410	PROX1
chr2	99012565	99013653	CNGA3
chr5	140767491	140769260	PCDHGB4
chr9	104448968	104449193	GRIN3A
chrX	139865919	139866497	CDR1
chr12	11506213	11506796	PRB1
chr1	13183334	13183781	LOC440563
chr11	5529367	5530481	UBQLN3
chr1	99771595	99772516	LPPR4
chr10	29821794	29822127	SVIL
chr19	21606157	21607280	ZNF493
chr6	27114237	27114572	HIST1H2BK
chr6	146755039	146755795	GRM1
chr11	3680851	3681449	ART1
chr7	136699800	136700565	CHRM2
chr2	77745480	77746850	LRRTM4
chr16	10273896	10274136	GRIN2A
chr4	44176944	44177184	KCTD8
chr8	77775640	77775987	ZFHX4
chr1	151774039	151774827	LINGO4
chr18	7955121	7955364	PTPRM
chr4	111397595	111398076	ENPEP

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chr15	85405870	85406115	ALPK3
chr15	33954862	33955107	RYR3
chr14	47426599	47426844	MDGA2
chr19	31038870	31040061	ZNF536
chr10	124339093	124339339	DMBT1
chr4	70079788	70080273	UGT2B11
chrX	127185699	127185946	ACTRT1
chr1	215847663	215848873	USH2A
chr7	119914917	119915730	KCND2
chr2	11802095	11802346	NTSR2
chrX	104463740	104464104	TEX13A
chr6	134210543	134210908	TCF21
chr4	187524461	187525111	FAT1
chr4	73012713	73012968	NPFFR2
chr7	31377951	31378781	NEUROD6
chr14	59112195	59113679	DACT1
chr10	50819199	50820233	SLC18A3
chr12	78444554	78444927	NAV3
chr11	55032386	55032645	TRIM48
chr9	116136230	116136606	HDHD3
chr5	38481935	38482197	LIFR
chrX	151869580	151870255	MAGEA6
chr11	18158850	18159706	MRGPRX3
chr19	56952581	56954106	ZNF667
chr1	157557066	157557332	FCRL4
chr8	113697693	113697959	CSMD3
chr2	106497875	106498398	NCK2
chr1	149859079	149859463	HIST2H2AB
chr9	140611226	140611610	EHMT1
chr22	17288659	17288927	XKR3
chr1	176668317	176668585	PAPPA2
chr5	3599604	3600292	IRX1
chrX	125685233	125686313	DCAF12L1
chr7	150325294	150325564	GIMAP6
chr14	70633590	70634903	SLC8A3
chr2	51254660	51255051	NRXN1
chr17	29220392	29220784	ATAD5
chr20	23016731	23017265	SSTR4
chr9	27949284	27950605	LINGO2
chr8	85799838	85800011	RALYL
chr5	90040933	90041032	GPR98
chr2	31598257	31598393	XDH
chr22	41547849	41547948	EP300
chr22	41566409	41566575	EP300
chr1	89616145	89616258	GBP7
chr4	57896470	57896569	POLR2B
chr1	205034916	205035037	CNTN2
chr6	50803913	50804012	TFAP2B
chr16	61687869	61687980	CDH8
chr6	105474260	105474359	LIN28B
chr5	139192995	139193155	PSD2
chr2	141027810	141027915	LRP1B
chr2	141032077	141032176	LRP1B
chr2	141299369	141299475	LRP1B
chr2	141457952	141458107	LRP1B
chr2	141571225	141571375	LRP1B
chr2	141806555	141806673	LRP1B
chr17	6665466	6665565	XAF1
chr15	45003728	45003827	B2M
chr22	37098522	37098621	CACNG2
chr3	186917475	186917629	RTP1
chr1	89448464	89448606	RBMXL1
chr14	95033316	95033446	SERPINA4
chr7	154172023	154172122	DPP6
chr10	87487702	87487801	GRID1
chr3	109028016	109028177	DPPA2
chr9	104499619	104499718	GRIN3A
chr12	81503338	81503483	ACSS3
chr6	26027282	26027418	HIST1H4B
chr7	38530647	38530746	AMPH
chr3	125879695	125879845	ALDH1L1
chr12	101510456	101510576	ANO4
chr7	55221703	55221845	EGFR
chr19	55106218	55106361	LILRA1
chr19	55107219	55107318	LILRA1

TABLE 11-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	152674397	152674568	SYNE1
chr6	152786397	152786535	SYNE1
chr13	112722099	112722213	SOX1
chr19	22951999	22952126	ZNF99
chr9	73164457	73164590	TRPM3
chr2	125281879	125282029	CNTNAP5
chr2	125284861	125285033	CNTNAP5
chr5	36686233	36686332	SLC1A3

TABLE 12

Chromosome	Start (bp)	End (bp)
chr12	22068691	22068802
chr12	25378548	25378707
chr12	25380167	25380346
chr12	25398207	25398318
chr12	78400210	78401193
chr17	7573926	7574033
chr17	7577018	7577155
chr17	7577498	7577608
chr17	7578176	7578289
chr17	7578361	7578461
chr17	7579311	7579546
chr17	26684313	26684473
chr17	37880164	37880264
chr17	37880978	37881164
chr17	37881567	37881667
chr17	50008349	50008492
chr17	51900393	51902406
chr2	21228063	21235357
chr2	29416090	29416788
chr2	29419631	29419731
chr2	29420408	29420542
chr2	29430037	29430138
chr2	29432648	29432748
chr2	29436849	29436949
chr2	29443572	29443701
chr2	29445192	29445292
chr2	29445378	29445478
chr2	29446207	29448431
chr2	40655642	40657411
chr2	77745476	77746913
chr2	79384697	79384824
chr2	79385451	79385589
chr2	80085138	80085305
chr2	80101224	80101466
chr2	80136739	80136923
chr2	80529444	80530834
chr2	107423137	107423369
chr2	125261887	125262126
chr2	141242924	141243077
chr2	141665445	141665615
chr2	155711287	155711820
chr2	167760006	167760383
chr2	168099081	168108352
chr2	178098798	178098974
chr2	185800516	185803687
chr2	228881145	228884872
chr2	237172849	237172949
chr3	41266080	41266180
chr3	73432629	73433920
chr3	147108740	147109014
chr3	147113642	147114236
chr3	147127972	147128848
chr3	158983022	158983162
chr3	164905717	164908582
chr3	178935997	178936122
chr3	178951881	178952152
chr7	18705889	18706013
chr7	53103417	53104243
chr7	55241613	55241736

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chr7	55242414	55242514
chr7	55248985	55249171
chr7	55259411	55259567
chr7	55260446	55260546
chr7	55266409	55266556
chr7	55268007	55268107
chr7	88962743	88966270
chr7	100385559	100385717
chr7	116411902	116412043
chr7	116418829	116419011
chr7	116423357	116423523
chr7	116435940	116436178
chr7	119914694	119915701
chr7	126173011	126173905
chr7	136699632	136701008
chr7	140453074	140453193
chr7	140481375	140481493
chr7	146829409	146829584
chr19	1206998	1207145
chr19	1220371	1220504
chr19	1221211	1221339
chr19	10597327	10597494
chr19	10599867	10600044
chr19	10600329	10600478
chr19	10602291	10602939
chr19	10610098	10610667
chr19	30934468	30936638
chr19	31025753	31025906
chr19	31038885	31040384
chr19	31767487	31770649
chr19	46627149	46627249
chr19	56538557	56539874
chr19	57325171	57328936
chr8	10464569	10470796
chr8	19809301	19809452
chr8	52320663	52322097
chr8	77616324	77618911
chr8	77690475	77690656
chr8	77763155	77768514
chr8	88885085	88886198
chr8	113256632	113256798
chr8	113301593	113301767
chr8	113304765	113304939
chr8	113569005	113569166
chr8	113694669	113694858
chr8	113697651	113697959
chr8	133905936	133906139
chr8	139151228	139151339
chr8	139163464	139165439
chr8	139606265	139606411
chr5	15928015	15928580
chr5	19473457	19473820
chr5	21751861	21752328
chr5	22078560	22078762
chr5	24487792	24488227
chr5	24509698	24509911
chr5	24537494	24537781
chr5	26881252	26881725
chr5	26906067	26906235
chr5	26915758	26916024
chr5	33576159	33577170
chr5	33683975	33684142
chr5	33947278	33947473
chr5	45262059	45262891
chr5	45267190	45267355
chr5	63256356	63257448
chr5	82937336	82937508
chr5	127648325	127648487
chr5	161128506	161128739
chr5	168134980	168135126
chr6	57512507	57512692
chr6	117609655	117609965
chr6	117622137	117622300
chr6	117629957	117630091
chr6	117631244	117631444

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chr6	117632182	117632282
chr6	117638306	117638435
chr6	117639333	117639433
chr6	117641031	117658503
chr6	165715074	165715671
chr1	37271747	37271863
chr1	37346246	37346445
chr1	46085194	46085368
chr1	74575077	74575237
chr1	75036850	75039127
chr1	92185494	92185654
chr1	99771278	99772551
chr1	158627269	158627431
chr1	158632505	158632685
chr1	167095142	167097827
chr1	175086109	175086340
chr1	175372326	175372736
chr1	176563667	176564723
chr1	176709118	176709331
chr1	176915086	176915253
chr1	177001609	177001965
chr1	190028421	190029374
chr1	190067151	190068180
chr1	190203501	190203607
chr1	196227361	196227562
chr1	237729889	237730050
chr1	237777345	237778139
chr1	237886427	237886562
chr1	237947028	237948233
chr1	247587188	247588870
chr1	248039226	248039722
chr9	21970900	21971204
chr9	119976663	119977014
chr9	120474685	120476922
chr4	46252333	46252605
chr4	48622655	48622795
chr4	96761310	96762445
chr4	114274318	114280137
chr4	134071418	134073888
chr4	134084153	134084390
chr4	153247224	153247368
chr4	164534470	164534652
chr11	30032312	30034192
chr11	40136053	40137815
chr11	59828633	59828789
chr11	92531030	92535026
chr11	113102894	113103059
chr11	132081915	132082041
chrX	32429868	32430030
chrX	54497746	54497920
chrX	111195282	111195616
chrX	112024167	112024328
chrX	125298529	125299762
chrX	125685235	125686525
chrX	135426563	135432565
chrX	144904002	144906476
chr20	1961024	1961489
chr20	9546563	9546971
chr20	57766243	57769791
chr10	25886709	25888201
chr10	87628766	87628937
chr10	89692769	89693008
chr10	89711874	89712016
chr10	89717609	89717776
chr14	42355848	42357213
chr14	99183496	99183609
chr18	22804314	22807568
chr18	31322918	31326558
chr18	42529889	42533273
chr18	43249311	43249421
chr13	58206760	58209200
chr13	70681341	70681820
chr13	84453624	84455615
chr15	23810969	23812451
chr16	49669610	49672754

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chr16	51172603	51176051
chr21	44524418	44524518
track name = 169393__3__NSCLC__cfDNA__P1__tiled__region description = "169393__3__NSCLC__cfDNA__P1__tiled__region"		
chr1	37271737	37271884
chr1	37346222	37346474
chr1	46085160	46085373
chr1	74575056	74575253
chr1	75036823	75039145
chr1	92185462	92185681
chr1	99771257	99772585
chr1	158627234	158627465
chr1	158632484	158632728
chr1	167095120	167095403
chr1	167095405	167095612
chr1	167095615	167097864
chr1	175086076	175086211
chr1	175086221	175086297
chr1	175372301	175372766
chr1	176563636	176564760
chr1	176709091	176709343
chr1	176915056	176915285
chr1	177001586	177002014
chr1	190028399	190029107
chr1	190029289	190029393
chr1	190067124	190068224
chr1	190203479	190203619
chr1	196227340	196227582
chr1	237729868	237730075
chr1	237777319	237778165
chr1	237886394	237886578
chr1	237946994	237948243
chr1	247587162	247588364
chr1	247588367	247588817
chr1	247588817	247588895
chr1	248039192	248039764
chr2	21228028	21231818
chr2	21231828	21235395
chr2	29416068	29416797
chr2	29419608	29419751
chr2	29420373	29420567
chr2	29430003	29430159
chr2	29432613	29432759
chr2	29436818	29436971
chr2	29443538	29443734
chr2	29445168	29445304
chr2	29445348	29445487
chr2	29446178	29446811
chr2	29446818	29448463
chr2	40655614	40657081
chr2	40657084	40657440
chr2	77745452	77746964
chr2	79384672	79384750
chr2	79384762	79384840
chr2	79385437	79385543
chr2	80085113	80085323
chr2	80101198	80101492
chr2	80136708	80136972
chr2	80529410	80530860
chr2	107423111	107423415
chr2	125261856	125262161
chr2	141242897	141243115
chr2	141665412	141665650
chr2	155711255	155711854
chr2	167759985	167760404
chr2	168099060	168101401
chr2	168101415	168104825
chr2	168104835	168105005
chr2	168105020	168108371
chr2	178098777	178098985
chr2	185800493	185801915

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chr2	185801923	185803706
chr2	228881110	228882939
chr2	228882950	228884909
chr2	237172815	237172971
chr3	41266046	41266203
chr3	73432596	73433959
chr3	147108706	147109059
chr3	147113616	147114276
chr3	147127941	147128721
chr3	147128786	147128884
chr3	158982995	158983175
chr3	164905695	164908611
chr3	178935989	178936159
chr3	178951854	178952172
chr4	46252300	46252618
chr4	48622620	48622811
chr4	96761279	96762480
chr4	114274285	114280173
chr4	134071388	134073569
chr4	134073573	134073927
chr4	134084118	134084405
chr4	153247190	153247405
chr4	164534436	164534688
chr5	15927993	15928621
chr5	19473429	19473853
chr5	21751830	21752361
chr5	22078530	22078784
chr5	24487765	24488112
chr5	24488125	24488267
chr5	24509670	24509951
chr5	24537460	24537819
chr5	26881229	26881749
chr5	26906034	26906250
chr5	26915729	26916048
chr5	33576138	33577189
chr5	33683948	33684170
chr5	33947248	33947491
chr5	45262028	45262905
chr5	45267168	45267374
chr5	63256324	63257490
chr5	82937302	82937528
chr5	127648290	127648512
chr5	161128482	161128768
chr5	168134946	168135167
chr6	57512480	57512720
chr6	117609633	117609983
chr6	117622113	117622313
chr6	117629933	117630106
chr6	117631223	117631463
chr6	117632148	117632288
chr6	117638273	117638470
chr6	117639298	117639453
chr6	117641008	117641428
chr6	117641438	117642993
chr6	117643003	117643174
chr6	117643188	117645141
chr6	117645158	117646127
chr6	117646173	117647264
chr6	117647288	117648778
chr6	117648783	117648918
chr6	117648943	117650620
chr6	117650623	117650824
chr6	117650848	117651171
chr6	117651198	117651335
chr6	117651393	117651470
chr6	117651563	117651698
chr6	117651783	117651861
chr6	117652003	117652075
chr6	117652093	117652174
chr6	117652488	117652591
chr6	117654168	117654233
chr6	117657138	117657333
chr6	117657883	117658535
chr6	165715051	165715324
chr6	165715336	165715696

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chr7	18705854	18706045
chr7	53103391	53104267
chr7	55241591	55241759
chr7	55242381	55242526
chr7	55248951	55249200
chr7	55259376	55259601
chr7	55260416	55260574
chr7	55266386	55266601
chr7	55267986	55268123
chr7	88962713	88966297
chr7	100385525	100385744
chr7	116411893	116412071
chr7	116418808	116419051
chr7	116423323	116423536
chr7	116435918	116436202
chr7	119914661	119915728
chr7	126172979	126173946
chr7	136699601	136701045
chr7	140453047	140453121
chr7	140453152	140453225
chr7	140481432	140481507
chr7	146829378	146829605
chr8	10464537	10465023
chr8	10465042	10465142
chr8	10465302	10465600
chr8	10465637	10465932
chr8	10465982	10466059
chr8	10466072	10469014
chr8	10469017	10470834
chr8	19809280	19809487
chr8	52320630	52322120
chr8	77616289	77618931
chr8	77690454	77690692
chr8	77763124	77765281
chr8	77765309	77766540
chr8	77766554	77768548
chr8	88885057	88885301
chr8	88885302	88885455
chr8	88885462	88885546
chr8	88885557	88886235
chr8	113256609	113256816
chr8	113301569	113301781
chr8	113304734	113304955
chr8	113568979	113569192
chr8	113694634	113694887
chr8	113697629	113697972
chr8	133905914	133906161
chr8	139151207	139151354
chr8	139163432	139165467
chr8	139606232	139606449
chr9	21970869	21971023
chr9	21971074	21971146
chr9	119976629	119976988
chr9	120474664	120476938
chr10	25886682	25888217
chr10	87628744	87628948
chr10	89692737	89692810
chr10	89692877	89692951
chr10	89692972	89693037
chr10	89711887	89711966
chr10	89717577	89717711
chr11	30032280	30033827
chr11	30033840	30034213
chr11	40136022	40137848
chr11	59828607	59828816
chr11	92530995	92535065
chr11	113102866	113103090
chr11	132081890	132082058
chr12	22068669	22068839
chr12	25378518	25378628
chr12	25378668	25378736
chr12	25380138	25380301
chr12	25380308	25380385
chr12	25398223	25398302
chr12	78400185	78400668

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chr12	78400730	78401224
chr13	58206731	58209217
chr13	70681316	70681856
chr13	84453597	84455634
chr14	42355817	42357220
chr14	99183471	99183646
chr15	23810934	23812061
chr15	23812104	23812496
chr16	49669576	49672771
chr16	51172578	51173068
chr16	51173088	51174493
chr16	51174583	51174969
chr16	51174978	51175198
chr16	51175198	51175317
chr16	51175353	51175532
chr16	51175583	51175663
chr16	51175678	51175785
chr16	51175823	51175893
chr16	51175943	51176086
chr17	7573894	7574051
chr17	7576984	7577173
chr17	7577469	7577648
chr17	7578154	7578326
chr17	7578339	7578478
chr17	7579289	7579578
chr17	26684291	26684509
chr17	37880143	37880277
chr17	37880943	37881204
chr17	37881538	37881678
chr17	50008315	50008526
chr17	51900371	51902443
chr18	22804281	22807572
chr18	31322885	31325872
chr18	31325880	31326588
chr18	42529861	42533288
chr18	43249282	43249460
chr19	1206975	1207181
chr19	1220350	1220519
chr19	1221180	1221361
chr19	10597293	10597511
chr19	10599833	10600090
chr19	10600303	10600514
chr19	10602263	10602970
chr19	10610073	10610710
chr19	30934446	30936220
chr19	30936236	30936655
chr19	31025721	31025947
chr19	31038856	31040406
chr19	31767466	31769840
chr19	31769851	31770211
chr19	31770246	31770670
chr19	46627115	46627261
chr19	56538529	56539902
chr19	57325149	57325572
chr19	57325594	57325696
chr19	57325734	57328948
chr20	1960990	1961521
chr20	9546528	9546985
chr20	57766211	57769826
chr21	44524394	44524462
chr21	44524464	44524541
chrX	32429836	32430057
chrX	54497725	54497925
chrX	111195250	111195640
chrX	112024145	112024364
chrX	125298579	125298835
chrX	125298859	125299288
chrX	125299329	125299403
chrX	125299449	125299801
chrX	125685269	125685520
chrX	125685544	125685751
chrX	125685759	125685831
chrX	125685854	125685972
chrX	125686009	125686084
chrX	125686129	125686562

TABLE 12-continued

Chromosome	Start (bp)	End (bp)
chrX	135426542	135432592
chrX	144903967	144906492

TABLE 13

Chromosome	Start (bp)	Stop (bp)
chr12	22068691	22068802
chr12	25378548	25378707
chr12	25380167	25380346
chr12	25398207	25398318
chr12	25479184	25479284
chr12	25549002	25549102
chr12	25619069	25619169
chr12	25702320	25702420
chr12	49415559	49415659
chr12	49415825	49415934
chr12	49416049	49416149
chr12	49416372	49416658
chr12	49418360	49418491
chr12	49418592	49418729
chr12	49419964	49421105
chr12	49421585	49421713
chr12	49421791	49421924
chr12	49422610	49422741
chr12	49422843	49423019
chr12	49423171	49423271
chr12	49424062	49424222
chr12	49424383	49424551
chr12	49424675	49424816
chr12	49424957	49427747
chr12	49427849	49428082
chr12	49428176	49428276
chr12	49428357	49428457
chr12	49428594	49428718
chr12	49430907	49432772
chr12	49433004	49433141
chr12	49433217	49433400
chr12	49433506	49435318
chr12	49435413	49435513
chr12	49435686	49435786
chr12	49435871	49436113
chr12	49436336	49436436
chr12	49436523	49436661
chr12	49436858	49436969
chr12	49437128	49437228
chr12	49437417	49437565
chr12	49437650	49437781
chr12	49437982	49438087
chr12	49438185	49438305
chr12	49438526	49438748
chr12	49439676	49439776
chr12	49439847	49439957
chr12	49440042	49440207
chr12	49440391	49440573
chr12	49441747	49441852
chr12	49442441	49442552
chr12	49442887	49443001
chr12	49443464	49444573
chr12	49444668	49446207
chr12	49446346	49446492
chr12	49446697	49446855
chr12	49446989	49447104
chr12	49447258	49447424
chr12	49447760	49447923
chr12	49448089	49448199
chr12	49448310	49448534
chr12	49448682	49448809
chr12	49449033	49449133
chr12	69219731	69219831
chr12	69225913	69226013
chr12	69233291	69233391

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr12	69240320	69240420
chr12	78400210	78401193
chr17	1011270	1011370
chr17	1028560	1028660
chr17	1059293	1059393
chr17	1083413	1083513
chr17	7572917	7573017
chr17	7573926	7574033
chr17	7576510	7576691
chr17	7576839	7576939
chr17	7577018	7577155
chr17	7577498	7577608
chr17	7578176	7578289
chr17	7578361	7578554
chr17	7579311	7579590
chr17	7579660	7579760
chr17	7579825	7579925
chr17	26684313	26684473
chr17	37879790	37879913
chr17	37880164	37880264
chr17	37880978	37881164
chr17	37881301	37881457
chr17	37881567	37881667
chr17	37881959	37882106
chr17	37882813	37882913
chr17	50008349	50008492
chr17	51900393	51902406
chr2	15760376	15760476
chr2	15804642	15804742
chr2	15908988	15909088
chr2	16012401	16012501
chr2	16082531	16082631
chr2	21228063	21235357
chr2	29416090	29416788
chr2	29419631	29419731
chr2	29420408	29420542
chr2	29430037	29430138
chr2	29432648	29432748
chr2	29436849	29436949
chr2	29443572	29443701
chr2	29445192	29445292
chr2	29445378	29445478
chr2	29446207	29448431
chr2	40655642	40657411
chr2	77745476	77746913
chr2	79384697	79384824
chr2	79385451	79385589
chr2	80085138	80085305
chr2	80101224	80101466
chr2	80136739	80136923
chr2	80529444	80530834
chr2	107423137	107423369
chr2	125261887	125262126
chr2	125502734	125502919
chr2	141242924	141243077
chr2	141665445	141665615
chr2	155711287	155711820
chr2	167760006	167760383
chr2	168099081	168108352
chr2	178095513	178096736
chr2	178097120	178097290
chr2	178097973	178098073
chr2	178098733	178098996
chr2	185800516	185803687
chr2	198267280	198267550
chr2	212286730	212286830
chr2	212288879	212289026
chr2	212293120	212293220
chr2	212295669	212295825
chr2	212426627	212426813
chr2	212483901	212484001
chr2	212488646	212488769
chr2	225338962	225339093
chr2	225342917	225343062
chr2	225346609	225346795

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr2	225360549	225360683
chr2	225362468	225362568
chr2	225365080	225365204
chr2	225367682	225367789
chr2	225368369	225368539
chr2	225370673	225370849
chr2	225371575	225371720
chr2	225376071	225376299
chr2	225378241	225378355
chr2	225379329	225379489
chr2	225400245	225400358
chr2	225422376	225422573
chr2	225449644	225449744
chr2	228881145	228884872
chr2	237172849	237172949
chr3	11635131	11635231
chr3	11679363	11679463
chr3	11722418	11722518
chr3	11761268	11761368
chr3	11806354	11806454
chr3	12626012	12626156
chr3	12632296	12632473
chr3	12633199	12633299
chr3	38182623	38182777
chr3	41266080	41266180
chr3	70303419	70303519
chr3	70586835	70586935
chr3	71015074	71015174
chr3	71159348	71159448
chr3	71444358	71444458
chr3	73432629	73433920
chr3	78286439	78286539
chr3	78766444	78766544
chr3	79472272	79472372
chr3	80063205	80063305
chr3	80653452	80653552
chr3	81242598	81242698
chr3	89259009	89259670
chr3	89390065	89390221
chr3	89390904	89391240
chr3	147108740	147109014
chr3	147113642	147114236
chr3	147127972	147128848
chr3	158983022	158983162
chr3	164905717	164908582
chr3	168840391	168840491
chr3	169501256	169501356
chr3	169646255	169646355
chr3	169896593	169896693
chr3	170140983	170141083
chr3	170716033	170716133
chr3	178916538	178916965
chr3	178921331	178921577
chr3	178927973	178928126
chr3	178935997	178936122
chr3	178951881	178952152
chr3	181430148	181431102
chr3	182584093	182584193
chr3	182733240	182733340
chr3	183014809	183014909
chr3	183273245	183273345
chr3	183818306	183818406
chr3	189455528	189455657
chr3	189526060	189526315
chr3	189586368	189586505
chr7	13894226	13894326
chr7	18705889	18706013
chr7	53103417	53104243
chr7	54617645	54617745
chr7	55241613	55241736
chr7	55242414	55242514
chr7	55248985	55249171
chr7	55259411	55259567
chr7	55260446	55260546
chr7	55266409	55266556

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr7	55268007	55268107
chr7	55492985	55493085
chr7	55750380	55750480
chr7	55990868	55990968
chr7	57398678	57398928
chr7	88962743	88966270
chr7	100385559	100385717
chr7	116411902	116412043
chr7	116417433	116417533
chr7	116418829	116419011
chr7	116422041	116422151
chr7	116423357	116423523
chr7	116435708	116435845
chr7	116435940	116436178
chr7	119914694	119915701
chr7	126173011	126173905
chr7	136699632	136701008
chr7	140434396	140434570
chr7	140439611	140439746
chr7	140449086	140449218
chr7	140453074	140453193
chr7	140453960	140454060
chr7	140476711	140476888
chr7	140477783	140477883
chr7	140481375	140481493
chr7	146133409	146133607
chr7	146829409	146829584
chr7	152109145	152110115
chr19	1206912	1207202
chr19	1218407	1218507
chr19	1219317	1219417
chr19	1220371	1220504
chr19	1220579	1220716
chr19	1221211	1221339
chr19	1221926	1222026
chr19	1222983	1223171
chr19	1226452	1226646
chr19	4099198	4099412
chr19	4110506	4110653
chr19	4117416	4117627
chr19	10597327	10597494
chr19	10599867	10600044
chr19	10600329	10600478
chr19	10602291	10602939
chr19	10610098	10610667
chr19	30934468	30936638
chr19	31025753	31025906
chr19	31038885	31040384
chr19	31767487	31770649
chr19	46627149	46627249
chr19	56538557	56539874
chr19	57325171	57328936
chr8	2855569	2855669
chr8	3382862	3382962
chr8	4021464	4021564
chr8	4660066	4660166
chr8	5301234	5301334
chr8	5936810	5936910
chr8	10464569	10470796
chr8	13733843	13733943
chr8	13959896	13959996
chr8	14338894	14338994
chr8	14640268	14640368
chr8	14942769	14942869
chr8	15244538	15244638
chr8	19809301	19809452
chr8	38173445	38173545
chr8	38179041	38179141
chr8	38182800	38182900
chr8	38186559	38186659
chr8	38271435	38271541
chr8	38271669	38271807
chr8	38272062	38272162
chr8	38272296	38272419
chr8	38273387	38273578

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr8	38274823	38274934
chr8	38275387	38275509
chr8	52320663	52322097
chr8	77616324	77618911
chr8	77690475	77690656
chr8	77763155	77768514
chr8	88885085	88886198
chr8	113256632	113256798
chr8	113301593	113301767
chr8	113304765	113304939
chr8	113569005	113569166
chr8	113694669	113694858
chr8	113697651	113697959
chr8	114668600	114668774
chr8	128360232	128360332
chr8	128377618	128377718
chr8	128394799	128394899
chr8	128411949	128412049
chr8	128718569	128718669
chr8	128750829	128750929
chr8	128766379	128766479
chr8	128790280	128790380
chr8	129171307	129171407
chr8	129177137	129177237
chr8	129181775	129181875
chr8	129187690	129187790
chr8	133905936	133906139
chr8	139151228	139151339
chr8	139163464	139165439
chr8	139606265	139606411
chr5	917120	917220
chr5	1034347	1034447
chr5	1083915	1084015
chr5	1216932	1217032
chr5	1295105	1295279
chr5	12091527	12091718
chr5	15928015	15928580
chr5	19473457	19473820
chr5	21751861	21752328
chr5	22078560	22078762
chr5	24487792	24488227
chr5	24509698	24509911
chr5	24537494	24537781
chr5	26881252	26881725
chr5	26906067	26906235
chr5	26915758	26916024
chr5	29809544	29809723
chr5	33576159	33577170
chr5	33683975	33684142
chr5	33947278	33947473
chr5	36037958	36038058
chr5	36183977	36184077
chr5	36679795	36679895
chr5	37370951	37371051
chr5	38352315	38352415
chr5	39306756	39306856
chr5	45262059	45262891
chr5	45267190	45267355
chr5	45292575	45292675
chr5	45321584	45321684
chr5	45353227	45353327
chr5	63256356	63257448
chr5	82937336	82937508
chr5	127648325	127648487
chr5	149498309	149498415
chr5	149499029	149499129
chr5	149499574	149499686
chr5	149500450	149500573
chr5	149500766	149500885
chr5	149501442	149501603
chr5	149502604	149502764
chr5	149503812	149503923
chr5	149504289	149504394
chr5	149505007	149505140
chr5	161128506	161128739

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr5	168134980	168135126
chr6	57512507	57512692
chr6	117609655	117609965
chr6	117622137	117622300
chr6	117629957	117630091
chr6	117631244	117631444
chr6	117632182	117632282
chr6	117638306	117638435
chr6	117639333	117639433
chr6	117641031	117658503
chr6	161969910	161970010
chr6	162225660	162225760
chr6	162490501	162490601
chr6	162753766	162753866
chr6	163149295	163149395
chr6	165715074	165715671
chr1	37271747	37271863
chr1	37346246	37346445
chr1	39927582	39927682
chr1	40035554	40035654
chr1	40124925	40125025
chr1	40363293	40363393
chr1	40627140	40627240
chr1	46085194	46085368
chr1	74575077	74575237
chr1	75036850	75039127
chr1	92185494	92185654
chr1	99771278	99772551
chr1	115256420	115256599
chr1	115258670	115258781
chr1	150477108	150477208
chr1	150550793	150550893
chr1	150727501	150727601
chr1	151108103	151108203
chr1	151316207	151316307
chr1	153177282	153177382
chr1	153430314	153430414
chr1	153907288	153907388
chr1	154246293	154246393
chr1	154401746	154401846
chr1	155264358	155264458
chr1	158627269	158627431
chr1	158632505	158632685
chr1	162743258	162743386
chr1	162745441	162745633
chr1	162745925	162746160
chr1	162748369	162748519
chr1	162749901	162750036
chr1	167095142	167097827
chr1	175086109	175086340
chr1	175372326	175372736
chr1	176563667	176564723
chr1	176709118	176709331
chr1	176915086	176915253
chr1	177001609	177001965
chr1	190028421	190029374
chr1	190067151	190068180
chr1	190203501	190203607
chr1	195246938	195247988
chr1	195899530	195899738
chr1	196227361	196227562
chr1	237729889	237730050
chr1	237777345	237778139
chr1	237886427	237886562
chr1	237947028	237948233
chr1	247587188	247588870
chr1	248039226	248039722
chr9	8528635	8528735
chr9	9659339	9659439
chr9	10332505	10332605
chr9	11005703	11005803
chr9	11677898	11677998
chr9	12352199	12352299
chr9	21901383	21901483
chr9	21925971	21926071

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr9	21954943	21955043
chr9	21968184	21968284
chr9	21968697	21968797
chr9	21970900	21971207
chr9	21974475	21974826
chr9	21994137	21994330
chr9	24503905	24504079
chr9	119976663	119977014
chr9	120474685	120476922
chr9	133738149	133738422
chr9	133747508	133747608
chr9	133748246	133748424
chr9	133750254	133750439
chr9	133753801	133753954
chr9	133755449	133755549
chr9	139390522	139392010
chr9	139396723	139396940
chr9	139397633	139397782
chr9	139399124	139399556
chr4	1803561	1803752
chr4	1805418	1805563
chr4	1806056	1806247
chr4	1806550	1806696
chr4	1807081	1807203
chr4	1807285	1807396
chr4	1807476	1807667
chr4	1807777	1807900
chr4	1807969	1808069
chr4	1808272	1808410
chr4	1808555	1809018
chr4	46252333	46252605
chr4	46329605	46329705
chr4	48622655	48622795
chr4	55139703	55139897
chr4	55140695	55140795
chr4	55141007	55141140
chr4	55143554	55143659
chr4	55144062	55144173
chr4	55144528	55144682
chr4	55146482	55146649
chr4	55151537	55151653
chr4	55152007	55152130
chr4	55153596	55153708
chr4	55154965	55155065
chr4	55155175	55155281
chr4	55592022	55592216
chr4	55593383	55593490
chr4	55593581	55593708
chr4	55593988	55594093
chr4	55594176	55594287
chr4	55595500	55595651
chr4	55597489	55597589
chr4	55598036	55598164
chr4	55599235	55599358
chr4	55602663	55602775
chr4	55602886	55602986
chr4	55603340	55603446
chr4	55955035	55955140
chr4	55962396	55962509
chr4	55964304	55964439
chr4	96761310	96762445
chr4	114274318	114280137
chr4	133331354	133332060
chr4	134071418	134073888
chr4	134084153	134084390
chr4	153247224	153247368
chr4	164534470	164534652
chr4	180440924	180441134
chr4	190551538	190551712
chr4	190596829	190597498
chr4	190626448	190626746
chr11	533765	533944
chr11	534211	534322

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr11	30032312	30034192
chr11	40136053	40137815
chr11	59828633	59828789
chr11	68747922	68748022
chr11	68822681	68822781
chr11	69063409	69063509
chr11	69458629	69458729
chr11	69631089	69631189
chr11	69880510	69880610
chr11	69887113	69887213
chr11	69893509	69893609
chr11	69894990	69895090
chr11	92531030	92535026
chr11	113102894	113103059
chr11	132081915	132082041
chrX	32429868	32430030
chrX	54497746	54497920
chrX	111195282	111195616
chrX	112024167	112024328
chrX	125298529	125299762
chrX	125685235	125686525
chrX	135426563	135432565
chrX	144904002	144906476
chr20	1961024	1961489
chr20	9546563	9546971
chr20	57766243	57769791
chr10	17193296	17193396
chr10	25886709	25888201
chr10	43606655	43612179
chr10	43613820	43613928
chr10	43614978	43615193
chr10	43615528	43615651
chr10	43617379	43617479
chr10	43619118	43619256
chr10	43620330	43620430
chr10	87628766	87628937
chr10	89624216	89624316
chr10	89653774	89653874
chr10	89685242	89685342
chr10	89690774	89690874
chr10	89692769	89693008
chr10	89711874	89712016
chr10	89717609	89717776
chr10	89720650	89720875
chr10	89725043	89725229
chr10	123243211	123243317
chr10	123244908	123245046
chr10	123246853	123246953
chr10	123247504	123247627
chr10	123256045	123256236
chr10	123258008	123258119
chr10	123260339	123260461
chr14	36934833	36934933
chr14	36944809	36944909
chr14	36954643	36954743
chr14	36964548	36964648
chr14	42355848	42357213
chr14	99183496	99183609
chr14	99712930	99713169
chr14	105246424	105246553
chr18	22804314	22807568
chr18	29310984	29311084
chr18	31322918	31326558
chr18	40503582	40503682
chr18	40850409	40850509
chr18	42529889	42533273
chr18	43204654	43204754
chr18	43249311	43249421
chr18	48604681	48604781
chr18	50683754	50683854
chr18	54398671	54398771
chr18	60985557	60985657

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr18	61328318	61328418
chr18	63477037	63477137
chr18	67563059	67563159
chr18	70526135	70526235
chr18	74620352	74620452
chr13	19748101	19748201
chr13	20039376	20039476
chr13	20240592	20240692
chr13	20346482	20346582
chr13	20412932	20413032
chr13	48877646	48877746
chr13	48878048	48878185
chr13	48881415	48881542
chr13	48916734	48916850
chr13	48916903	48917003
chr13	48919215	48919335
chr13	48921930	48922030
chr13	48923075	48923175
chr13	48934152	48934263
chr13	48936950	48937093
chr13	48939018	48939118
chr13	48941629	48941739
chr13	48942651	48942751
chr13	48947534	48947634
chr13	48951053	48951170
chr13	48953708	48953808
chr13	48954154	48954254
chr13	48954289	48954389
chr13	48955382	48955579
chr13	48985992	48986092
chr13	49027128	49027247
chr13	49030339	49030485
chr13	49033823	49033969
chr13	49037866	49037971
chr13	49039133	49039247
chr13	49039340	49039504
chr13	49047461	49047561
chr13	49050836	49050979
chr13	49051465	49051565
chr13	49054120	49054220
chr13	58206760	58209200
chr13	70681341	70681820
chr13	84453624	84455615
chr15	23810969	23812451
chr15	66727364	66727575
chr15	66729083	66729230
chr15	66735606	66735706
chr15	66736969	66737069
chr15	66774092	66774217
chr15	66777327	66777529
chr15	66779548	66779648
chr15	66781533	66781633
chr15	66782028	66782128
chr15	66782839	66782953
chr16	34982573	34982747
chr16	49669610	49672754
chr16	51172603	51176051
chr21	11044261	11044435
chr21	11180809	11182067
chr21	21044289	21044463
chr21	44524418	44524518
chr22	33559458	33559558
chr22	47892673	47892773
chr22	48212160	48212260
chr22	48532012	48532112
chr22	48851913	48852013
chr22	49168045	49168145
chr22	49820007	49820181
chrY	2712158	2712258
chrY	2722676	2722776
chrY	2733157	2733257
chrY	2843160	2843260
chrY	2844737	2844837

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
track name = 169403__4__NSCLC__CLIN__P1__tiled__region description = "169403__4__NSCLC__CLIN__P1__tiled__region"		
chr1	37271737	37271884
chr1	37346222	37346474
chr1	39927559	39927709
chr1	40035524	40035667
chr1	40124904	40125044
chr1	40363259	40363431
chr1	40627119	40627271
chr1	46085160	46085373
chr1	74575056	74575253
chr1	75036823	75039145
chr1	92185462	92185681
chr1	99771257	99772585
chr1	115256398	115256631
chr1	115258638	115258818
chr1	150477075	150477229
chr1	150550760	150550919
chr1	150727470	150727631
chr1	151108075	151108232
chr1	151316185	151316324
chr1	153177253	153177394
chr1	153430293	153430426
chr1	153907318	153907429
chr1	154246270	154246346
chr1	154401715	154401878
chr1	155264325	155264472
chr1	158627234	158627465
chr1	158632484	158632728
chr1	162743224	162743417
chr1	162745419	162745656
chr1	162745904	162746192
chr1	162748334	162748548
chr1	162749879	162750051
chr1	167095120	167095403
chr1	167095405	167095612
chr1	167095615	167097864
chr1	175086076	175086211
chr1	175086221	175086297
chr1	175372301	175372766
chr1	176563636	176564760
chr1	176709091	176709343
chr1	176915056	176915285
chr1	177001586	177002014
chr1	190028399	190029107
chr1	190029289	190029393
chr1	190067124	190068224
chr1	190203479	190203619
chr1	195246910	195248015
chr1	195899500	195899671
chr1	195899700	195899775
chr1	196227340	196227582
chr1	237729868	237730075
chr1	237777319	237778165
chr1	237886394	237886578
chr1	237946994	237948243
chr1	247587162	247588364
chr1	247588367	247588817
chr1	247588817	247588895
chr1	248039192	248039764
chr2	15760345	15760498
chr2	15804615	15804765
chr2	15908965	15909095
chr2	16012370	16012525
chr2	16082505	16082642
chr2	21228028	21231818
chr2	21231828	21235395
chr2	29416068	29416797
chr2	29419608	29419751
chr2	29420373	29420567
chr2	29430003	29430159

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr2	29432613	29432759
chr2	29436818	29436971
chr2	29443538	29443734
chr2	29445168	29445304
chr2	29445348	29445487
chr2	29446178	29446811
chr2	29446818	29448463
chr2	40655614	40657081
chr2	40657084	40657440
chr2	77745452	77746964
chr2	79384672	79384750
chr2	79384762	79384840
chr2	79385437	79385543
chr2	80085113	80085323
chr2	80101198	80101492
chr2	80136708	80136972
chr2	80529410	80530860
chr2	107423111	107423415
chr2	125261856	125262161
chr2	125502712	125502953
chr2	141242897	141243115
chr2	141665412	141665650
chr2	155711255	155711854
chr2	167759985	167760404
chr2	168099060	168101401
chr2	168101415	168104825
chr2	168104835	168105005
chr2	168105020	168108371
chr2	178095482	178096587
chr2	178096587	178096759
chr2	178097127	178097309
chr2	178097952	178098091
chr2	178098702	178099022
chr2	185800493	185801915
chr2	185801923	185803706
chr2	198267256	198267583
chr2	212286709	212286845
chr2	212288844	212289051
chr2	212293094	212293234
chr2	212295644	212295847
chr2	212426604	212426850
chr2	212483874	212484009
chr2	212488614	212488802
chr2	225338940	225339104
chr2	225342895	225343098
chr2	225346585	225346828
chr2	225360525	225360695
chr2	225362445	225362591
chr2	225365045	225365224
chr2	225367655	225367800
chr2	225368335	225368547
chr2	225370650	225370855
chr2	225371545	225371754
chr2	225376050	225376300
chr2	225378220	225378390
chr2	225379295	225379508
chr2	225400220	225400397
chr2	225422350	225422599
chr2	225449620	225449754
chr2	228881110	228882939
chr2	228882950	228884909
chr2	237172815	237172971
chr3	11635122	11635268
chr3	11679337	11679483
chr3	11722397	11722541
chr3	11761247	11761391
chr3	11806327	11806485
chr3	12625987	12626199
chr3	12632272	12632487
chr3	12633172	12633310
chr3	38182589	38182802
chr3	41266046	41266203
chr3	70303387	70303539
chr3	70586847	70586927
chr3	71015043	71015201

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr3	71159326	71159453
chr3	71444326	71444466
chr3	73432596	73433959
chr3	78286416	78286555
chr3	78766421	78766566
chr3	79472296	79472404
chr3	80063171	80063319
chr3	80653421	80653569
chr3	81242566	81242721
chr3	89258987	89259688
chr3	89390042	89390244
chr3	89390882	89391281
chr3	147108706	147109059
chr3	147113616	147114276
chr3	147127941	147128721
chr3	147128786	147128884
chr3	158982995	158983175
chr3	164905695	164908611
chr3	168840358	168840506
chr3	169501222	169501373
chr3	169646232	169646375
chr3	169896647	169896721
chr3	170140951	170141100
chr3	170716001	170716143
chr3	178916514	178916999
chr3	178921304	178921594
chr3	178927964	178928132
chr3	178935989	178936159
chr3	178951854	178952172
chr3	181430124	181430303
chr3	181430354	181430630
chr3	181430674	181430922
chr3	181430929	181431073
chr3	182584062	182584219
chr3	182733212	182733363
chr3	183014776	183014924
chr3	183273223	183273359
chr3	183818283	183818415
chr3	189455494	189455678
chr3	189526039	189526349
chr3	189586344	189586525
chr4	1803536	1803773
chr4	1805396	1805574
chr4	1806031	1806284
chr4	1806526	1806731
chr4	1807046	1807223
chr4	1807251	1807431
chr4	1807451	1807697
chr4	1807751	1807922
chr4	1807941	1808085
chr4	1808241	1808437
chr4	1808531	1809058
chr4	46252300	46252618
chr4	46329575	46329722
chr4	48622620	48622811
chr4	55139671	55139926
chr4	55140661	55140810
chr4	55140976	55141165
chr4	55143521	55143671
chr4	55144041	55144214
chr4	55144496	55144724
chr4	55146461	55146680
chr4	55151506	55151679
chr4	55151986	55152174
chr4	55153571	55153751
chr4	55154941	55155077
chr4	55155151	55155322
chr4	55592001	55592232
chr4	55593361	55593499
chr4	55593551	55593724
chr4	55593961	55594111
chr4	55594151	55594327
chr4	55595471	55595693
chr4	55597456	55597604
chr4	55598001	55598189

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr4	55599211	55599391
chr4	55602631	55602806
chr4	55602856	55602999
chr4	55603311	55603491
chr4	55955013	55955154
chr4	55962373	55962551
chr4	55964278	55964468
chr4	96761279	96762480
chr4	114274285	114280173
chr4	133331323	133331919
chr4	133331923	133332097
chr4	134071388	134073569
chr4	134073573	134073927
chr4	134084118	134084405
chr4	153247190	153247405
chr4	164534436	164534688
chr4	180440900	180441164
chr4	190551511	190551715
chr4	190596801	190597538
chr4	190626426	190626777
chr5	917097	917247
chr5	1034312	1034466
chr5	1083892	1084043
chr5	1216902	1217054
chr5	1295072	1295271
chr5	12091505	12091636
chr5	12091640	12091739
chr5	15927993	15928621
chr5	19473429	19473853
chr5	21751830	21752361
chr5	22078530	22078784
chr5	24487765	24488112
chr5	24488125	24488267
chr5	24509670	24509951
chr5	24537460	24537819
chr5	26881229	26881749
chr5	26906034	26906250
chr5	26915729	26916048
chr5	29809511	29809740
chr5	33576138	33577189
chr5	33683948	33684170
chr5	33947248	33947491
chr5	36037928	36038001
chr5	36038013	36038091
chr5	36183948	36184102
chr5	36679773	36679913
chr5	37370928	37371073
chr5	38352293	38352437
chr5	39306733	39306870
chr5	45262028	45262905
chr5	45267168	45267374
chr5	45292543	45292684
chr5	45321563	45321717
chr5	45353198	45353353
chr5	63256324	63257490
chr5	82937302	82937528
chr5	127648290	127648512
chr5	149498288	149498434
chr5	149499008	149499153
chr5	149499553	149499723
chr5	149500428	149500605
chr5	149500738	149500927
chr5	149501418	149501626
chr5	149502573	149502798
chr5	149503788	149503961
chr5	149504258	149504418
chr5	149504978	149505164
chr5	161128482	161128768
chr5	168134946	168135167
chr6	57512480	57512720
chr6	117609633	117609983
chr6	117622113	117622313
chr6	117629933	117630106
chr6	117631223	117631463
chr6	117632148	117632288

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr6	117638273	117638470
chr6	117639298	117639453
chr6	117641008	117641428
chr6	117641438	117642993
chr6	117643003	117643174
chr6	117643188	117645141
chr6	117645158	117646127
chr6	117646173	117647264
chr6	117647288	117648778
chr6	117648783	117648918
chr6	117648943	117650620
chr6	117650623	117650824
chr6	117650848	117651171
chr6	117651198	117651335
chr6	117651393	117651470
chr6	117651563	117651698
chr6	117651783	117651861
chr6	117652003	117652075
chr6	117652093	117652174
chr6	117652488	117652591
chr6	117654168	117654233
chr6	117657138	117657333
chr6	117657883	117658535
chr6	161969889	161970031
chr6	162225634	162225776
chr6	162490477	162490612
chr6	162753732	162753882
chr6	163149267	163149420
chr6	165715051	165715324
chr6	165715336	165715696
chr7	13894204	13894353
chr7	18705854	18706045
chr7	53103391	53104267
chr7	54617611	54617769
chr7	55241591	55241759
chr7	55242381	55242526
chr7	55248951	55249200
chr7	55259376	55259601
chr7	55260416	55260574
chr7	55266386	55266601
chr7	55267986	55268123
chr7	55492961	55493101
chr7	55750356	55750509
chr7	55990836	55990988
chr7	57398716	57398826
chr7	57398831	57398960
chr7	88962713	88966297
chr7	100385525	100385744
chr7	116411893	116412071
chr7	116417408	116417550
chr7	116418808	116419051
chr7	116422018	116422192
chr7	116423323	116423536
chr7	116435673	116435857
chr7	116435918	116436202
chr7	119914661	119915728
chr7	126172979	126173946
chr7	136699601	136701045
chr7	140434372	140434446
chr7	140434482	140434561
chr7	140439682	140439757
chr7	140449052	140449119
chr7	140449177	140449255
chr7	140453047	140453121
chr7	140453152	140453225
chr7	140453937	140454069
chr7	140476747	140476828
chr7	140476842	140476919
chr7	140477762	140477913
chr7	140481432	140481507
chr7	146133383	146133624
chr7	146829378	146829605
chr7	152109111	152109218
chr7	152109491	152110155
chr8	2855544	2855681

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr8	3382869	3382970
chr8	4021439	4021575
chr8	4660034	4660187
chr8	5301200	5301353
chr8	5936775	5936913
chr8	10464537	10465023
chr8	10465042	10465142
chr8	10465302	10465600
chr8	10465637	10465932
chr8	10465982	10466059
chr8	10466072	10469014
chr8	10469017	10470834
chr8	13733819	13733959
chr8	13959864	13960019
chr8	14338869	14339001
chr8	14640234	14640402
chr8	14942739	14942888
chr8	15244509	15244647
chr8	19809280	19809487
chr8	38173423	38173569
chr8	38179018	38179158
chr8	38182778	38182902
chr8	38186538	38186679
chr8	38271413	38271567
chr8	38271643	38271843
chr8	38272033	38272174
chr8	38272268	38272446
chr8	38273358	38273606
chr8	38274798	38274969
chr8	38275353	38275526
chr8	52320630	52322120
chr8	77616289	77618931
chr8	77690454	77690692
chr8	77763124	77765281
chr8	77765309	77766540
chr8	77766554	77768548
chr8	88885057	88885301
chr8	88885302	88885455
chr8	88885462	88885546
chr8	88885557	88886235
chr8	113256609	113256816
chr8	113301569	113301781
chr8	113304734	113304955
chr8	113568979	113569192
chr8	113694634	113694887
chr8	113697629	113697972
chr8	114668574	114668781
chr8	128360202	128360350
chr8	128377597	128377732
chr8	128394777	128394933
chr8	128411927	128412069
chr8	128718537	128718694
chr8	128750807	128750953
chr8	128766352	128766511
chr8	128790247	128790364
chr8	129171274	129171417
chr8	129177109	129177255
chr8	129181749	129181892
chr8	129187659	129187804
chr8	133905914	133906161
chr8	139151207	139151354
chr8	139163432	139165467
chr8	139606232	139606449
chr9	8528610	8528754
chr9	9659305	9659445
chr9	10332480	10332631
chr9	11005680	11005810
chr9	11677875	11678026
chr9	12352175	12352317
chr9	21901349	21901500
chr9	21926024	21926096
chr9	21954914	21955054
chr9	21968154	21968293
chr9	21968669	21968817
chr9	21970869	21971023

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr9	21971074	21971146
chr9	21974444	21974836
chr9	21994114	21994361
chr9	24503879	24504095
chr9	119976629	119976988
chr9	120474664	120476938
chr9	133738125	133738431
chr9	133747485	133747622
chr9	133748215	133748439
chr9	133750230	133750477
chr9	133753770	133753998
chr9	133755415	133755561
chr9	139390492	139392023
chr9	139396697	139396976
chr9	139397602	139397822
chr9	139399092	139399590
chr10	17193286	17193418
chr10	25886682	25888217
chr10	43606627	43609247
chr10	43609247	43609666
chr10	43609672	43612193
chr10	43613787	43613935
chr10	43614952	43615241
chr10	43615497	43615675
chr10	43617347	43617495
chr10	43619092	43619274
chr10	43620297	43620443
chr10	87628744	87628948
chr10	89624272	89624350
chr10	89653752	89653825
chr10	89653832	89653909
chr10	89685272	89685376
chr10	89690752	89690894
chr10	89692737	89692810
chr10	89692877	89692951
chr10	89692972	89693037
chr10	89711887	89711966
chr10	89717577	89717711
chr10	89720642	89720880
chr10	89725022	89725169
chr10	123243190	123243329
chr10	123244880	123245060
chr10	123246830	123246966
chr10	123247475	123247652
chr10	123256010	123256272
chr10	123257975	123258163
chr10	123260315	123260496
chr11	533740	533979
chr11	534190	534303
chr11	30032280	30033827
chr11	30033840	30034213
chr11	40136022	40137848
chr11	59828607	59828816
chr11	68747893	68748040
chr11	68822653	68822798
chr11	69063388	69063542
chr11	69458601	69458742
chr11	69631066	69631206
chr11	69880486	69880621
chr11	69887081	69887231
chr11	69893486	69893621
chr11	69894961	69895103
chr11	92530995	92535065
chr11	113102866	113103090
chr11	132081890	132082058
chr12	22068669	22068839
chr12	25378518	25378628
chr12	25378668	25378736
chr12	25380138	25380301
chr12	25380308	25380385
chr12	25398223	25398302
chr12	25479158	25479260
chr12	25548968	25549122
chr12	25619048	25619187
chr12	25702298	25702436

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr12	49415531	49415687
chr12	49415801	49415972
chr12	49416026	49416141
chr12	49416346	49416704
chr12	49418331	49418518
chr12	49418571	49418758
chr12	49419931	49421135
chr12	49421561	49421733
chr12	49421756	49421953
chr12	49422586	49422756
chr12	49422821	49423041
chr12	49423136	49423286
chr12	49424031	49424237
chr12	49424351	49424575
chr12	49424641	49424859
chr12	49424926	49425825
chr12	49425836	49426605
chr12	49426781	49426896
chr12	49426911	49427265
chr12	49427271	49427660
chr12	49427681	49427786
chr12	49427821	49428116
chr12	49428141	49428301
chr12	49428326	49428477
chr12	49428571	49428752
chr12	49430876	49432815
chr12	49432981	49433163
chr12	49433191	49433443
chr12	49433481	49435340
chr12	49435381	49435537
chr12	49435661	49435805
chr12	49435841	49436114
chr12	49436301	49436457
chr12	49436491	49436676
chr12	49436826	49437005
chr12	49437096	49437263
chr12	49437386	49437599
chr12	49437616	49437809
chr12	49437956	49438106
chr12	49438161	49438346
chr12	49438491	49438785
chr12	49439641	49439795
chr12	49439826	49440010
chr12	49440021	49440238
chr12	49440361	49440615
chr12	49441721	49441865
chr12	49442411	49442597
chr12	49442866	49443039
chr12	49443431	49444590
chr12	49444641	49445584
chr12	49445586	49446222
chr12	49446316	49446530
chr12	49446671	49446878
chr12	49446961	49447139
chr12	49447231	49447445
chr12	49447726	49447945
chr12	49448056	49448235
chr12	49448276	49448568
chr12	49448651	49448852
chr12	49449011	49449147
chr12	69219699	69219844
chr12	69225919	69226030
chr12	69233264	69233407
chr12	69240289	69240435
chr12	78400185	78400668
chr12	78400730	78401224
chr13	19748070	19748146
chr13	20039425	20039497
chr13	20240560	20240714
chr13	20346455	20346607
chr13	20412927	20413034
chr13	48877615	48877767
chr13	48878015	48878209
chr13	48881385	48881558
chr13	48916700	48917022

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr13	48919190	48919357
chr13	48921900	48922036
chr13	48923050	48923182
chr13	48934130	48934295
chr13	48936925	48937127
chr13	48938995	48939127
chr13	48941595	48941782
chr13	48942630	48942761
chr13	48947500	48947645
chr13	48951030	48951209
chr13	48953715	48953810
chr13	48954170	48954235
chr13	48954295	48954404
chr13	48955375	48955626
chr13	48985970	48986111
chr13	49027105	49027282
chr13	49030315	49030517
chr13	49033800	49034007
chr13	49037835	49037984
chr13	49039110	49039277
chr13	49039315	49039518
chr13	49047430	49047576
chr13	49050815	49050983
chr13	49051440	49051572
chr13	49054085	49054245
chr13	58206731	58209217
chr13	70681316	70681856
chr13	84453597	84455634
chr14	36934811	36934957
chr14	36944781	36944922
chr14	36954611	36954757
chr14	36964516	36964668
chr14	42355817	42357220
chr14	99183471	99183646
chr14	99712896	99713180
chr14	105246390	105246570
chr15	23810934	23812061
chr15	23812104	23812496
chr15	66727329	66727406
chr15	66727494	66727602
chr15	66729049	66729123
chr15	66729139	66729253
chr15	66735574	66735650
chr15	66735664	66735738
chr15	66737009	66737088
chr15	66774059	66774241
chr15	66777294	66777550
chr15	66779539	66779676
chr15	66781504	66781654
chr15	66781994	66782146
chr15	66782804	66782984
chr16	34982539	34982782
chr16	49669576	49672771
chr16	51172578	51173068
chr16	51173088	51174493
chr16	51174583	51174969
chr16	51174978	51175198
chr16	51175198	51175317
chr16	51175353	51175532
chr16	51175583	51175663
chr16	51175678	51175785
chr16	51175823	51175893
chr16	51175943	51176086
chr17	1011246	1011383
chr17	1028526	1028680
chr17	1059266	1059405
chr17	1083386	1083525
chr17	7572889	7573037
chr17	7573894	7574051
chr17	7576519	7576721
chr17	7576809	7576957
chr17	7576984	7577173
chr17	7577469	7577648
chr17	7578154	7578326
chr17	7578339	7578589

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr17	7579289	7579605
chr17	7579639	7579772
chr17	7579804	7579954
chr17	26684291	26684509
chr17	37879768	37879938
chr17	37880143	37880277
chr17	37880943	37881204
chr17	37881268	37881488
chr17	37881538	37881678
chr17	37881938	37882156
chr17	37882788	37882929
chr17	50008315	50008526
chr17	51900371	51902443
chr18	22804281	22807572
chr18	29310950	29311096
chr18	31322885	31325872
chr18	31325880	31326588
chr18	40503560	40503699
chr18	40850388	40850535
chr18	42529861	42533288
chr18	43204627	43204777
chr18	43249282	43249460
chr18	48604649	48604802
chr18	50683732	50683864
chr18	54398650	54398796
chr18	60985524	60985667
chr18	61328289	61328432
chr18	63477009	63477155
chr18	67563037	67563171
chr18	70526104	70526248
chr18	74620319	74620472
chr19	1206890	1207229
chr19	1218385	1218525
chr19	1219295	1219431
chr19	1220350	1220519
chr19	1220550	1220724
chr19	1221180	1221361
chr19	1221905	1222046
chr19	1222960	1223204
chr19	1226420	1226679
chr19	4099176	4099272
chr19	4099376	4099456
chr19	4110481	4110621
chr19	4117451	4117524
chr19	4117591	4117663
chr19	10597293	10597511
chr19	10599833	10600090
chr19	10600303	10600514
chr19	10602263	10602970
chr19	10610073	10610710
chr19	30934446	30936220
chr19	30936236	30936655
chr19	31025721	31025947
chr19	31038856	31040406
chr19	31767466	31769840
chr19	31769851	31770211
chr19	31770246	31770670
chr19	46627115	46627261
chr19	56538529	56539902
chr19	57325149	57325572
chr19	57325594	57325696
chr19	57325734	57328948
chr20	1960990	1961521
chr20	9546528	9546985
chr20	57766211	57769826
chr21	11044276	11044355
chr21	11180816	11180887
chr21	11181036	11181190
chr21	11181246	11181330
chr21	11181441	11181525
chr21	11181681	11181756
chr21	11181766	11182005
chr21	21044257	21044478
chr21	44524394	44524462
chr21	44524464	44524541

TABLE 13-continued

Chromosome	Start (bp)	Stop (bp)
chr22	33559434	33559565
chr22	47892639	47892789
chr22	48212134	48212283
chr22	48531984	48532124
chr22	48851889	48852027
chr22	49168014	49168162
chr22	49819974	49820191
chrX	32429836	32430057
chrX	54497725	54497925
chrX	111195250	111195640
chrX	112024145	112024364
chrX	125298579	125298835
chrX	125298859	125299288
chrX	125299329	125299403
chrX	125299449	125299801
chrX	125685269	125685520
chrX	125685544	125685751
chrX	125685759	125685831
chrX	125685854	125685972
chrX	125686009	125686084
chrX	125686129	125686562
chrX	135426542	135432592
chrX	144903967	144906492
chrY	2712193	2712272
chrY	2722643	2722721
chrY	2733178	2733258
chrY	2843138	2843272
chrY	2844743	2844868

TABLE 14

Chromosome	Start (bp)	End (bp)	Gene
chr13	32929222	32929396	BRCA2
chr17	7576999	7577173	TP53
chr17	7578384	7578558	TP53
chr17	7579311	7579561	TP53
chr17	7577466	7577640	TP53
chr17	7578145	7578319	TP53
chr17	41234419	41234593	BRCA1
chr17	7576802	7576976	TP53
chr12	25398175	25398349	KRAS
chr17	41275986	41276160	BRCA1
chr17	7573892	7574066	TP53
chr2	233990481	233990655	INPP5D
chrX	153130773	153130955	L1CAM
chr13	32910624	32915006	BRCA2
chr14	96730550	96730769	BDKRB1
chr4	55604571	55604745	KIT
chr12	57910672	57910846	DDIT3
chr12	52981452	52981626	KRT72
chr6	43306871	43307045	ZNF318
chr6	136597137	136597311	BCLAF1
chr12	38714202	38714376	ALG10B
chr19	17435740	17435914	ANO8
chr13	32972466	32972640	BRCA2
chrX	70823742	70823916	ACRC
chr7	107824864	107825038	NRCAM
chr6	26156789	26157000	HIST1H1E
chr7	100275114	100275288	GNB2
chr11	72945430	72945604	P2RY2
chrX	149938722	149938896	CD99L2
chr19	22271072	22271246	ZNF257
chr18	28714531	28714705	DSC1
chr6	169008790	169008964	SMOC2
chr20	2778780	2778954	CPXM1
chr18	32398123	32398297	DTNA
chr22	19241534	19241708	CLTCL1
chr17	74878219	74878393	MGAT5B
chr8	104709387	104709561	RIMS2
chr16	15844000	15844174	MYH11
chr17	41223007	41223181	BRCA1

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr13	32906479	32907422	BRCA2
chr17	41243479	41246667	BRCA1
chr17	41209023	41209197	BRCA1
chr14	65260180	65260354	SPTB
chr15	23811019	23811193	MKRN3
chr10	37430801	37430975	ANKRD30A
chr8	139144840	139145014	FAM135B
chr6	170870945	170871119	TBP
chr1	148753247	148753421	NBPF16
chr17	38569103	38569277	TOP2A
chr7	146829298	146829472	CNTNAP2
chr14	102028203	102028378	DIO3
chr8	122640982	122641166	HAS2
chr3	46307409	46307597	CCR3
chr6	26251882	26252105	HIST1H2BH
chr2	240982125	240982350	PRR21
chr4	81967024	81967254	BMP3
chr1	247420018	247420261	VN1R5
chrX	12938669	12939057	TLR8
chr19	31038886	31039138	ZNF536
chr12	40076580	40076833	C12orf40
chrX	123517632	123518244	ODZ1
chr17	41256121	41256295	BRCA1
chr9	94486985	94487159	ROR2
chr9	136507373	136507547	DBH
chr14	88654294	88654468	KCNK10
chr6	167271630	167271804	RPS6KA2
chr8	53554964	53555138	RB1CC1
chr11	64428266	64428440	NRXN2
chr6	27798986	27799160	HIST1H4K
chr8	2000266	2000440	MYOM2
chr10	26385271	26385445	MYO3A
chr11	47869743	47869917	NUP160
chr2	29287839	29288013	C2orf71
chr7	74005154	74005328	GTF2IRD1
chr19	40398326	40398500	FCGBP
chr13	38211363	38211537	TRPC4
chr20	36030827	36031001	SRC
chr4	189068354	189068528	TRIML1
chr1	147126287	147126461	ACP6
chr17	3427487	3427661	TRPV3
chr21	44836620	44836794	SIK1
chr2	170493291	170493465	PPIG
chr6	133004283	133004457	VNN1
chr13	48986112	48986286	LPAR6
chr22	40825603	40825777	MKL1
chr10	76781752	76781926	KAT6B
chr4	4199402	4199576	OTOP1
chr6	55119985	55120159	HCRTR2
chrX	29938051	29938225	IL1RAPL1
chr12	20890036	20890210	SLCO1C1
chr13	32953452	32953626	BRCA2
chr22	29083842	29084016	CHEK2
chr17	10350307	10350481	MYH4
chr1	176837967	176838141	ASTN1
chr13	37015224	37015398	CCNA1
chr8	27293729	27293903	PTK2B
chr12	114793625	114793799	TBX5
chr3	9512454	9512628	SETD5
chr5	139743618	139743792	SLC4A9
chr1	231401748	231401922	GNPAT
chr9	37442028	37442202	ZBTB5
chr1	156815447	156815621	INSRR
chr12	132502750	132502924	EP400
chr5	5318233	5318407	ADAMTS16
chr9	133936421	133936595	LAMC3
chr22	17684479	17684653	CECR1
chr9	111624625	111624799	ACTL7A
chr6	130475999	130476173	SAMD3
chr10	60549035	60549209	BICC1
chr1	203821227	203821401	ZC3H11A
chr5	13770786	13770960	DNAH5
chr3	13896085	13896260	WNT7A
chrX	34961819	34962307	FAM47B
chr1	159558178	159558367	APCS

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	40997342	40997691	AOC2
chr18	43534627	43534825	EPG5
chr5	24487959	24488158	CDH10
chr4	69816889	69817093	UGT2A3
chr19	52448710	52448914	ZNF613
chr5	9629405	9629772	TAS2R1
chr2	210517906	210518116	MAP2
chr9	990519	990733	DMRT3
chrX	53577910	53578128	HUWE1
chr15	91424680	91424899	FURIN
chr14	77272844	77273066	ANGEL1
chr11	134252641	134252868	B3GAT1
chr19	57640873	57641100	USP29
chr1	206224927	206225336	AVPR1B
chr15	83932594	83932825	BNC1
chr6	146350691	146351339	GRM1
chr2	202900305	202900539	FZD7
chr6	26273207	26273442	HIST1H2BI
chr7	127222557	127222793	GCC1
chr5	31323026	31323264	CDH6
chrX	127185082	127185520	ACTRT1
chr1	114483214	114483663	HIPK1
chr17	29654577	29654837	NF1
chr22	40283477	40283743	ENTHD1
chr6	169648554	169648826	THBS2
chr12	129558548	129559050	TMEM132D
chr3	134670329	134670616	EPHB1
chr2	223917650	223917940	KCNE4
chr6	127797196	127797486	C6orf174
chr17	15234414	15234711	TEKT3
chr22	40814593	40814891	MKL1
chr7	150417156	150417457	GIMAP1
chr19	40433543	40433852	FCGBP
chr5	140718608	140718920	PCDHGA2
chr8	21766910	21767224	DOK2
chrX	129518508	129518825	GPR119
chrX	12736567	12736891	FRMPD4
chr10	98714824	98715150	LCOR
chr6	26056314	26056647	HIST1H1C
chr11	128680372	128680706	FLI1
chr10	26463052	26463392	MYO3A
chr1	18691777	18692119	IGSF21
chr1	248039432	248039775	TRIM58
chrX	30873210	30873558	TAB3
chr6	26158403	26158752	HIST1H2BD
chr7	27135133	27135486	HOXA1
chr21	31538435	31538793	CLDN17
chr2	46985990	46986360	SOCS5
chr19	58967128	58967499	ZNF324B
chr21	38309122	38309498	HLCS
chr3	142840773	142841153	CHST2
chr3	88040207	88040589	HTR1F
chr1	237777437	237777821	RYR2
chrX	23411675	23412063	PTCHD1
chr1	117122025	117122414	IGSF3
chr16	55532193	55532367	MMP2
chr18	30260367	30260541	KLHL14
chr2	157186233	157186407	NR4A2
chr11	121391435	121391609	SORL1
chr21	14982944	14983118	POTED
chr17	41226303	41226477	BRCA1
chr4	113570679	113570853	LARP7
chr22	39112731	39112905	GTPBP1
chr1	180023506	180023680	CEP350
chrX	62898319	62898493	ARHGEF9
chr8	89128777	89128951	MMP16
chr21	15872896	15873070	SAMSN1
chr1	32381454	32381628	PTP4A2
chr3	138383874	138384048	PIK3CB
chr1	46193299	46193473	IPP
chr4	68619760	68619934	GNRHR
chr4	74853659	74853833	PPBP
chrX	44937610	44937784	KDM6A
chr12	109536165	109536339	UNG
chr1	151755349	151755523	TDRKH

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr19	54744228	54744402	LILRA6
chr6	31937640	31937814	DOM3Z
chr3	50005006	50005180	RBM6
chrX	100617525	100617699	BTK
chr7	48017994	48018168	HUS1
chr19	19371633	19371807	HAPLN4
chr7	99000987	99001161	PDAP1
chr17	7358601	7358775	CHRNB1
chr7	29980244	29980418	SCRN1
chr2	62067255	62067429	FAM161A
chr13	95121123	95121297	DCT
chr10	134008343	134008517	DPYSL4
chr4	524344	524518	PIGG
chr20	30753085	30753259	TM9SF4
chr1	72076681	72076855	NEGR1
chr19	52520314	52520488	ZNF614
chr6	137322921	137323095	IL20RA
chr9	103212884	103213058	C9orf30
chr14	76905712	76905886	ESRRB
chr15	41687069	41687243	NDUFAF1
chr22	50869686	50869860	PPP6R2
chr2	207804270	207804444	CPO
chr8	37690507	37690681	GPR124
chr6	5771505	5771679	FARS2
chr7	31124313	31124487	ADCYAP1R1
chr1	207785038	207785212	CR1
chr14	51087293	51087467	ATL1
chr8	124195401	124195575	FAM83A
chr11	30255107	30255281	FSHB
chr12	2788578	2788752	CACNA1C
chr1	179013079	179013253	FAM20B
chrX	5821584	5821758	NLGN4X
chr2	114500190	114500364	SLC35F5
chr12	101490282	101490456	ANO4
chr5	148392134	148392308	SH3TC2
chr12	10962009	10962183	TAS2R9
chr2	32640289	32640463	BIRC6
chr18	70417719	70417893	NETO1
chr18	70450979	70451153	NETO1
chr9	95400367	95400541	IPPK
chr13	35615170	35615344	NBEA
chr7	55259402	55259576	EGFR
chr7	55273137	55273311	EGFR
chr4	186066207	186066381	SLC25A4
chr19	47197125	47197299	PRKD2
chr6	127608323	127608497	RNF146
chr17	37868153	37868327	ERBB2
chr17	37881530	37881704	ERBB2
chr17	74289654	74289828	QRICH2
chr9	4117788	4117962	GLIS3
chr2	131785500	131785674	ARHGEF4
chrX	153760785	153760959	G6PD
chr13	113803617	113803791	F10
chr18	33848484	33848658	MOCOS
chr19	55106319	55106493	LILRA1
chr6	152658007	152658181	SYNE1
chr3	5024988	5025162	BHLHE40
chr6	53518901	53519075	KLHL31
chr1	11078772	11078946	TARDBP
chr5	54581095	54581269	DHX29
chr21	45987677	45987851	TSPEAR
chrX	107403742	107403916	COL4A6
chr2	125530315	125530489	CNTNAP5
chr10	135076611	135076785	ADAM8
chr12	85517858	85518032	LRRIQ1
chr10	105330702	105330876	NEURL
chr9	35107570	35107744	FAM214B
chr7	16505155	16505329	SOSTDC1
chrX	31165427	31165601	DMD
chrX	32583852	32584026	DMD
chr5	7626267	7626441	ADCY2
chr5	7695833	7696007	ADCY2
chr5	7802316	7802490	ADCY2
chr3	8609117	8609291	LMCD1
chr10	117026251	117026425	ATRNL1

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr1	55172071	55172245	HEATR8
chr20	35862373	35862547	RPN2
chr17	56557443	56557617	HSF5
chr10	120920366	120920540	SFXN4
chr2	65559065	65559239	SPRED2
chr11	108277762	108277936	C11orf65
chr1	89730488	89730662	GBP5
chr16	46781757	46781931	MYLK3
chr20	44507103	44507277	ZSWIM3
chr6	27861270	27861444	HIST1H2BO
chr12	103984654	103984828	STAB2
chr22	46327046	46327220	WNT7B
chr1	36645452	36645626	MAP7D1
chr13	36049712	36049886	MAB21L1
chr1	149857811	149857985	HIST2H2BE
chr17	18003840	18004014	DRG2
chr12	53663669	53663843	ESPL1
chr12	53676046	53676220	ESPL1
chr3	48040190	48040364	MAP4
chr6	12122606	12122780	HIVEP1
chr19	54849375	54849549	LILRA4
chr11	94731658	94731832	KDM4D
chr2	109964146	109964320	SH3RF3
chr2	96040018	96040192	KCNIP3
chr7	23296510	23296684	GPNMB
chr14	24845534	24845708	NFATC4
chr22	22324675	22324849	TOP3B
chr4	71114706	71114880	CSN3
chr11	117302257	117302431	DSCAML1
chr17	37676187	37676361	CDK12
chr4	88766974	88767148	MEPE
chr1	181745213	181745387	CACNA1E
chr9	463514	463688	DOCK8
chr20	40081366	40081540	CHD6
chr20	40111927	40112101	CHD6
chr1	186113302	186113476	HMCN1
chr15	64791917	64792091	ZNF609
chr3	184001547	184001721	ECE2
chrX	53423377	53423551	SMC1A
chrX	53432438	53432612	SMC1A
chr5	135692360	135692534	TRPC7
chr1	225706994	225707168	ENAH
chr1	216850514	216850688	ESRRG
chr2	68882394	68882568	PROKR1
chr7	87144562	87144736	ABCB1
chr10	75276690	75276864	USP54
chr8	95172209	95172383	CDH17
chr8	72233954	72234128	EYA1
chr2	200137240	200137414	SATB2
chrX	134706738	134706912	DDX26B
chr17	10535809	10535983	MYH3
chr15	75188492	75188666	MPI
chr12	5708635	5708809	ANO2
chr18	644905	645079	CLUL1
chr2	85628906	85629080	CAPG
chr3	78987947	78988121	ROBO1
chr7	2257550	2257724	MAD1L1
chr1	11561350	11561524	PTCHD2
chr12	104171567	104171741	NT5DC3
chr14	21024722	21024896	RNASE9
chr7	107342250	107342424	SLC26A4
chr14	72205719	72205893	SIPA1L1
chr5	3599655	3599829	IRX1
chr1	24077339	24077513	TCEB3
chr11	47333250	47333424	MADD
chr4	46305465	46305639	GABRA2
chr9	136405705	136405879	ADAMTSL2
chr6	30572351	30572525	PPP1R10
chr5	40976809	40976983	C7
chr6	117010462	117010636	KPNA5
chr1	145440001	145440175	TXNIP
chr1	236918334	236918508	ACTN2
chr20	30915324	30915498	KIF3B
chr4	175598247	175598421	GLRA3
chr4	70512921	70513095	UGT2A1

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	7636373	7636547	DNAH2
chr2	183960180	183960354	DUSP19
chrX	105011258	105011432	IL1RAPL2
chr2	220115767	220115941	TUBA4A
chr8	144942360	144942534	EPPK1
chr3	89259371	89259545	EPHA3
chr3	89456382	89456556	EPHA3
chr20	34241968	34242142	RBM12
chr6	33245175	33245349	B3GALT4
chr17	59560332	59560506	TBX4
chr12	56355086	56355260	PMEL
chr10	51224954	51225128	AGAP8
chr11	56949740	56949914	LRRC55
chrX	47497448	47497622	ELK1
chrX	92927506	92927680	NAP1L3
chr10	26315296	26315470	MYO3A
chr19	36002300	36002474	DMKN
chr19	12986845	12987019	DNASE2
chr6	31727842	31728016	MSH5
chr17	42745332	42745506	C17orf104
chrX	18234688	18234862	BEND2
chr21	41414479	41414653	DSCAM
chr21	41457523	41457697	DSCAM
chr12	51092093	51092267	DIP2B
chr6	161027513	161027687	LPA
chr17	63554329	63554503	AXIN2
chr4	155505472	155505646	FGA
chr4	155506836	155507010	FGA
chr12	7456950	7457124	ACSM4
chr19	58016002	58016176	ZNF773
chr6	150001356	150001530	LATS1
chr3	96706157	96706331	EPHA6
chr7	107204269	107204443	COG5
chr14	65262131	65262305	SPTB
chr1	70502178	70502352	LRRC7
chr6	145956389	145956563	EPM2A
chr3	5249775	5249949	EDEM1
chr8	143961043	143961217	CYP11B1
chr20	44678257	44678431	SLC12A5
chr6	27222972	27223146	PRSS16
chr9	125014101	125014275	RBM18
chr3	193042640	193042814	ATP13A5
chr3	193052715	193052889	ATP13A5
chr11	76900364	76900538	MYO7A
chr3	30729852	30730026	TGFBR2
chr5	112769819	112769993	TSSK1B
chr8	145153751	145153925	SHARPIN
chr12	29648216	29648390	OVCH1
chr6	33236263	33236437	VPS52
chr22	22277480	22277654	PPM1F
chr7	101844836	101845010	CUX1
chr7	101882677	101882851	CUX1
chr10	61967835	61968009	ANK3
chr17	34328406	34328580	CCL15
chr7	73944013	73944187	GTF2IRD1
chr5	167928970	167929144	RARS
chr2	170393706	170393880	FASTKD1
chr3	136708257	136708431	IL20RB
chr3	51399271	51399445	DOCK3
chr3	56667149	56667323	FAM208A
chr19	51649056	51649230	SIGLEC7
chr6	47649589	47649763	GPR111
chr20	60419714	60419888	CDH4
chr1	32671739	32671913	IQCC
chr1	32673139	32673313	IQCC
chr4	87730927	87731101	PTPN13
chr20	1960986	1961160	PDYN
chr4	8465677	8465851	METTL19
chr2	167094603	167094777	SCN9A
chr3	42739739	42739913	HHATL
chr14	92343897	92344071	FBLN5
chr5	36035812	36035986	UGT3A2
chr6	33165518	33165692	RXRB
chr3	183860533	183860707	EIF2B5
chr11	121000695	121000869	TECTA

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	26205002	26205176	HIST1H4E
chr22	24583132	24583306	SUSD2
chr13	32731388	32731562	FRY
chr15	28474329	28474503	HERC2
chr1	46080972	46081146	NASP
chr13	92408524	92408698	GPC5
chr16	57787029	57787203	KATNB1
chr8	145763063	145763237	ARHGAP39
chr5	179228536	179228710	MGAT4B
chr19	54867898	54868072	LAIR1
chr4	146058683	146058857	OTUD4
chr11	7984761	7984935	NLRP10
chr19	7550769	7550943	PEX11G
chr20	35127947	35128121	DLGAP4
chr9	34564547	34564721	CNTFR
chr19	40368393	40368567	FCGBP
chr14	77237473	77237647	VASH1
chrX	128724097	128724271	OCRL
chr4	70346323	70346497	UGT2B4
chr13	52604229	52604403	UTP14C
chr8	27327346	27327520	CHRNA2
chr6	53989304	53989478	MLIP
chr6	54095543	54095717	MLIP
chr2	166797528	166797702	TTC21B
chr17	78168973	78169147	CARD14
chr10	61574374	61574548	CCDC6
chr20	46277713	46277887	NCOA3
chr6	108214685	108214859	SEC63
chr8	145689615	145689789	CYHR1
chr17	47590051	47590225	NGFR
chr7	37907379	37907553	TXNDC3
chr6	87725173	87725347	HTR1E
chr3	123695682	123695856	ROPN1
chr7	29546825	29546999	CHN2
chrX	119077195	119077369	NKAP
chr1	201182610	201182784	IGFN1
chrX	23723842	23724016	ACOT9
chr8	98289303	98289477	TSPYL5
chr2	26696267	26696441	OTOF
chr6	97051481	97051655	FHL5
chr20	17434466	17434640	PCSK2
chr1	192128313	192128487	RGS18
chr15	43438674	43438848	TMEM62
chr20	43942089	43942263	RBPJL
chrX	24226323	24226497	ZFX
chr1	152195573	152195747	HRNR
chrX	15305982	15306156	ASB11
chr19	15079126	15079300	SLC1A6
chr9	88937766	88937940	ZCCHC6
chr19	12384451	12384625	ZNF44
chr7	75959322	75959496	YWHAG
chr6	33154432	33154606	COL11A2
chr10	75557566	75557740	KIAA0913
chr14	60903582	60903756	C14orf39
chr22	22160094	22160268	MAPK1
chr12	11338742	11338916	TAS2R42
chr15	90145016	90145190	C15orf42
chr12	21693338	21693512	GYS2
chr2	197737131	197737305	PGAP1
chr17	8215460	8215634	ARHGEF15
chr6	49427010	49427184	MUT
chr3	52525895	52526069	NISCH
chr12	49087834	49088008	CCNT1
chr3	195295787	195295961	APOD
chr19	52001316	52001490	SIGLEC12
chr10	18940007	18940181	NSUN6
chr7	134135508	134135682	AKR1B1
chrX	135579769	135579943	HTATSF1
chr4	5843007	5843181	CRMP1
chrX	21674122	21674296	KLHL34
chrX	13727247	13727421	RAB9A
chr5	147820656	147820830	FBXO38
chr16	16208588	16208762	ABCC1
chr17	17962145	17962319	C17orf39
chr20	43384832	43385006	RIMS4

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr2	200820452	200820626	C2orf47
chr10	104679436	104679610	CNNM2
chr14	64954556	64954730	ZBTB25
chr4	80246377	80246551	NAA11
chr6	90642071	90642245	BACH2
chr17	79478311	79478485	ACTG1
chr3	111672739	111672913	PHLDB2
chr19	50939845	50940019	MYBPC2
chr9	91616992	91617166	S1PR3
chr2	165550757	165550931	COBLL1
chr17	45299047	45299221	MYL4
chr1	46489385	46489559	MAST2
chr1	46501611	46501785	MAST2
chr15	65499237	65499411	CILP
chr4	57220203	57220377	AASDH
chr2	10186272	10186446	KLF11
chr5	169483642	169483816	DOCK2
chr15	85383879	85384053	ALPK3
chr1	27720852	27721026	GPR3
chr1	173961961	173962135	RC3H1
chr7	126746533	126746707	GRM8
chr8	119391834	119392008	SAMD12
chr7	12691408	12691582	SCIN
chr12	8083882	8084056	SLC2A3
chr12	57032917	57033091	ATP5B
chr8	139180127	139180301	FAM135B
chr1	211486062	211486236	RCOR3
chr2	206641089	206641263	NRP2
chr1	209964082	209964256	IRF6
chr10	75107866	75108040	TTC18
chr1	150483510	150483684	ECM1
chr11	28134957	28135131	METTL15
chr1	45243348	45243522	RPS8
chr16	28913111	28913285	ATP2A1
chr7	154760639	154760813	PAXIP1
chr3	113955346	113955520	ZNF80
chr10	98133355	98133529	TLL2
chr8	8998346	8998520	PPP1R3B
chr19	16314260	16314434	AP1M1
chr9	75435793	75435967	TMC1
chr19	19790906	19791080	ZNF101
chr6	40399994	40400168	LRFN2
chr1	176668469	176668643	PAPPA2
chr19	34900072	34900246	PDCD2L
chr15	66850053	66850227	LCTL
chr20	40727039	40727213	PTPRT
chr8	2819987	2820161	CSMD1
chr8	2875998	2876172	CSMD1
chr8	3266941	3267115	CSMD1
chr2	43969878	43970052	PLEKHH2
chr14	105212560	105212734	ADSSL1
chr9	98209437	98209611	PTCH1
chr9	98239819	98239993	PTCH1
chr2	165349531	165349705	GRB14
chr11	77937720	77937894	GAB2
chr1	12409267	12409441	VPS13D
chr6	31931738	31931912	SKIV2L
chr12	123276530	123276704	CCDC62
chr11	76174928	76175102	C11orf30
chr1	36367522	36367696	EIF2C1
chr7	149517954	149518128	SSPO
chr6	28472037	28472211	GPX6
chr9	128083689	128083863	GAPVD1
chr2	108478035	108478209	RGPD4
chr13	75868982	75869156	TBC1D4
chr1	110950222	110950396	HBXIP
chr19	8491478	8491652	MARCH2
chr7	99711228	99711402	TAF6
chr5	39383033	39383207	DAB2
chr11	75282953	75283127	SERPINH1
chr12	53012017	53012191	KRT73
chr11	67225828	67226002	CABP4
chr15	101595261	101595435	LRRK1
chr2	175618241	175618415	CHRNA1
chr10	111624918	111625092	XPNPEP1

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	26107915	26108089	HIST1H1T
chr2	96781270	96781444	ADRA2B
chr19	55263807	55263981	KIR2DL3
chr18	24496257	24496431	CHST9
chr15	42041402	42041576	MGA
chr7	104783598	104783772	SRPK2
chr19	48922466	48922640	GRIN2D
chr4	54256654	54256828	FIP1L1
chr16	24358009	24358183	CACNG3
chr19	52715925	52716099	PPP2R1A
chr8	133763964	133764138	TMEM71
chr17	73490958	73491132	KIAA0195
chr3	119219537	119219711	TIMMDC1
chrX	54472681	54472855	FGD1
chr20	52644908	52645082	BCAS1
chr6	30309744	30309918	TRIM39
chr1	237659923	237660097	RYR2
chr1	237863502	237863676	RYR2
chr7	98553760	98553934	TRRAP
chr7	50611556	50611730	DDC
chr11	92495033	92495207	FAT3
chr6	56497682	56497856	DST
chr4	46994816	46994990	GABRA4
chr14	57858158	57858332	NAA30
chr2	178936438	178936612	PDE11A
chr11	60889086	60889260	CD5
chr9	4663050	4663224	PPAPDC2
chr20	58448885	58449059	SYCP2
chr15	81585182	81585356	IL16
chr1	32202161	32202335	BAI2
chr1	32221862	32222036	BAI2
chr1	12939612	12939786	PRAMEF4
chr2	225266098	225266272	FAM124B
chr17	10317200	10317374	MYH8
chr2	178082415	178082589	HNRNPA3
chr6	132171100	132171274	ENPP1
chr6	132211478	132211652	ENPP1
chr10	48370965	48371139	ZNF488
chr12	52093330	52093504	SCN8A
chr12	52115441	52115615	SCN8A
chr11	116730074	116730248	SIK3
chr6	31541057	31541231	LTA
chr1	12837214	12837388	PRAMEF12
chr15	41099812	41099986	ZFYVE19
chr17	33312991	33313165	LIG3
chr16	58711191	58711365	SLC38A7
chr3	137717742	137717916	CLDN18
chr5	160047616	160047790	ATP10B
chr3	130290015	130290189	COL6A6
chrX	142596647	142596821	SPANXN3
chr2	88387334	88387508	SMYD1
chr12	4479674	4479848	FGF23
chr1	153004877	153005051	SPRR1B
chrX	48678500	48678674	HDAC6
chr12	7842774	7842948	GDF3
chr7	121943781	121943955	FEZF1
chr1	156264564	156264738	C1orf85
chr16	57957143	57957317	CNGB1
chr5	16478940	16479114	FAM134B
chrX	107930747	107930921	COL4A5
chr9	74840559	74840733	GDA
chr7	116339781	116339955	MET
chr4	4285321	4285495	LYAR
chr12	6838414	6838588	COPS7A
chr5	72469046	72469220	TMEM174
chr12	116406736	116406910	MED13L
chr19	39321984	39322158	ECH1
chr15	43571307	43571481	TGM7
chr17	76522932	76523106	DNAH17
chr5	454021	454195	EXOC3
chr1	53540222	53540396	PODN
chr2	198363398	198363572	HSPD1
chr10	70502159	70502333	CCAR1
chr1	70715584	70715758	SRSF11
chr2	234652205	234652379	DNAJB3

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr15	52571701	52571875	MYO5C
chrX	153540967	153541141	TKTL1
chr16	74499555	74499729	GLG1
chr1	85397107	85397281	MCOLN2
chr6	3015733	3015907	NQO2
chr6	73787041	73787215	KCNQ5
chrX	40539998	40540172	MED14
chr11	93754532	93754706	HEPHL1
chr9	112899635	112899809	PALM2-AKAP2
chr20	30414583	30414757	MYLK2
chr11	58604481	58604655	GLYATL2
chr10	105362405	105362579	SH3PXD2A
chr4	154702644	154702818	SFRP2
chr4	72994361	72994535	NPFFR2
chr17	48595964	48596138	MYCBPAP
chr16	84035388	84035562	NECAB2
chr9	23692564	23692738	ELAVL2
chr8	113562968	113563142	CSMD3
chr9	12694106	12694280	TYRP1
chr15	102226094	102226268	TARSL2
chr2	86255002	86255176	POLR1A
chr9	112899635	112899809	AKAP2
chr8	57228704	57228878	SDR16C5
chrX	123040810	123040984	XIAP
chr19	14862190	14862364	EMR2
chr1	215963468	215963642	USH2A
chr1	216595236	216595410	USH2A
chr17	29485982	29486156	NF1
chr17	29550436	29550610	NF1
chr20	46365379	46365553	SULF2
chr9	108424841	108425015	TAL2
chr3	142511650	142511824	TRPC1
chr19	15794318	15794492	CYP4F12
chr12	72893259	72893433	TRHDE
chr16	65005813	65005987	CDH11
chr16	22278014	22278188	EEF2K
chrX	100276916	100277090	TRMT2B
chr12	44913848	44914022	NELL2
chr19	6750496	6750670	TRIP10
chr10	98824510	98824684	SLIT1
chr6	74517801	74517975	CD109
chr7	45697317	45697491	ADCY1
chr12	20885899	20886073	SLCO1C1
chr2	220315842	220316016	SPEG
chr4	13617017	13617191	BOD1L
chr11	68703650	68703824	IGHMBP2
chr14	70245109	70245283	SLC10A1
chr13	32950780	32950954	BRCA2
chr11	102587004	102587178	MMP8
chr15	72170402	72170576	MYO9A
chr2	187558916	187559090	FAM171B
chr2	187615858	187616032	FAM171B
chr19	45992597	45992771	RTN2
chr7	18067160	18067334	PRPS1L1
chr4	124323250	124323424	SPRY1
chr20	3127359	3127533	FASTKD5
chrX	19380824	19380998	MAP3K15
chr19	35719285	35719459	FAM187B
chr2	1906858	1907032	MYT1L
chr12	41407977	41408151	CNTN1
chr5	1335143	1335317	CLPTM1L
chr2	131904190	131904364	PLEKHB2
chr20	5294571	5294745	PROKR2
chr1	211749217	211749391	SLC30A1
chr4	52938151	52938325	SPATA18
chr12	108011953	108012127	BTBD11
chr22	29091692	29091866	CHEK2
chr1	37346324	37346498	GRIK3
chr1	37356489	37356663	GRIK3
chr7	54612310	54612484	VSTM2A
chr1	1684333	1684507	NADK
chrX	69646940	69647114	GDPD2
chr3	151105648	151105822	MED12L
chr11	64627490	64627664	EHD1
chr16	22926315	22926489	HS3ST2

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chrX	130220498	130220672	ARHGAP36
chr8	133144381	133144555	KCNQ3
chr14	26917857	26918031	NOVA1
chr10	79576296	79576470	DLG5
chr10	79595459	79595633	DLG5
chr2	27375540	27375714	TCF23
chr16	10721376	10721550	TEKT5
chr16	67974118	67974292	LCAT
chr1	154987548	154987722	ZBTB7B
chr10	21414826	21415000	C10orf113
chr13	42407499	42407673	KIAA0564
chr11	68552295	68552469	CPT1A
chr1	110168907	110169081	AMPD2
chrX	51639534	51639708	MAGED1
chr5	40716351	40716525	TTC33
chr17	21207714	21207888	MAP2K3
chr17	29622546	29622720	OMG
chr1	19442025	19442199	UBR4
chr1	19443804	19443978	UBR4
chr1	15793869	15794043	CELA2A
chr2	120194564	120194738	TMEM37
chr16	30507349	30507523	ITGAL
chr20	61907410	61907584	ARFGAP1
chr2	226273609	226273783	NYAP2
chr20	55033408	55033582	CASS4
chr1	245704099	245704273	KIF26B
chr16	67576751	67576925	FAM65A
chr11	62365743	62365917	MTA2
chr16	343549	343723	AXIN1
chr19	50548092	50548266	ZNF473
chr9	132631949	132632123	USP20
chr1	33236139	33236313	KIAA1522
chr10	390907	391081	DIP2C
chr6	137019657	137019831	MAP3K5
chr14	78161074	78161248	ALKBH1
chr3	35778726	35778900	ARPP21
chr3	45942952	45943126	CCR9
chr13	42772585	42772759	DGKH
chr2	27600860	27601034	ZNF513
chr11	89424102	89424276	FOLH1B
chr12	60164984	60165158	SLC16A7
chr2	179718181	179718355	CCDC141
chr1	156214551	156214725	PAQR6
chr6	32006162	32006336	CYP21A2
chr7	107423636	107423810	SLC26A3
chr19	55748033	55748207	PPP6R1
chr15	92663689	92663863	SLCO3A1
chr22	42049526	42049700	XRCC6
chr20	45891014	45891188	ZMYND8
chr8	143994696	143994870	CYP11B2
chr4	155491574	155491748	FGB
chr13	79175686	79175860	POU4F1
chr11	118772305	118772479	BCL9L
chr6	33272099	33272273	TAPBP
chr19	53912197	53912371	ZNF765
chr13	109318331	109318505	MYO16
chr19	56423803	56423977	NLRP13
chr7	129929427	129929601	CPA2
chr1	66036244	66036418	LEPR
chr1	145281532	145281706	NOTCH2NL
chr6	100838676	100838850	SIM1
chr8	77618010	77618184	ZFHX4
chr5	19483416	19483590	CDH18
chr6	117130543	117130717	GPRC6A
chr14	94120241	94120415	UNC79
chr4	114253089	114253263	ANK2
chr2	152293722	152293896	RIF1
chr19	36214546	36214720	MLL4
chr19	36218419	36218593	MLL4
chr12	39733988	39734162	KIF21A
chr12	25672851	25673025	IFLTD1
chr12	25679005	25679179	IFLTD1
chr1	161163732	161163906	ADAMTS4
chr12	13768037	13768211	GRIN2B
chrX	153662568	153662742	ATP6AP1

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr12	81205280	81205454	LIN7A
chr19	49116355	49116529	FAM83E
chr2	20205962	20206136	MATN3
chr2	159519407	159519581	PKP4
chr6	146256160	146256334	SHPRH
chr9	101900165	101900339	TGFBR1
chr8	130760723	130760897	GSDMC
chr2	158399208	158399382	ACVR1C
chr1	43786872	43787046	TIE1
chr6	26271310	26271484	HIST1H3G
chr6	54214591	54214765	TINAG
chr13	80094940	80095114	NDFIP2
chr1	222717091	222717265	HHIPL2
chr2	219498318	219498492	PLCD4
chr1	43825356	43825530	CDC20
chr1	60505676	60505850	C1orf87
chr12	57501916	57502090	STAT6
chr17	78063535	78063709	CCDC40
chr6	30513916	30514090	GNL1
chr6	30521123	30521297	GNL1
chr21	34882071	34882245	GART
chr19	2799650	2799824	THOP1
chr19	43375863	43376037	PSG1
chr2	135711761	135711935	CCNT2
chr22	38877292	38877466	KDELR3
chr5	161117195	161117369	GABRA6
chr5	161118994	161119168	GABRA6
chr1	236751208	236751382	HEATR1
chr4	165118140	165118314	ANP32C
chr8	17400875	17401049	SLC7A2
chr17	61560782	61560956	ACE
chr7	128415710	128415884	OPN1SW
chrX	153129767	153129941	L1CAM
chr1	173839479	173839653	ZBTB37
chr17	40854849	40855023	EZH1
chr15	75042230	75042404	CYP1A2
chr2	27427641	27427815	SLC5A6
chr20	2636009	2636183	NOP56
chr3	19384055	19384229	KCNH8
chr14	25043532	25043706	CTSG
chr3	178935972	178936146	PIK3CA
chr8	120118077	120118251	COLEC10
chr12	56487159	56487333	ERBB3
chr12	56495315	56495489	ERBB3
chr2	20403715	20403889	SDC1
chr1	79093604	79093778	IFI44L
chr17	49824999	49825173	CA10
chr17	11672436	11672610	DNAH9
chr11	60687162	60687336	TMEM109
chr8	41615517	41615691	ANK1
chr7	87068980	87069154	ABCB4
chr18	9887055	9887229	TXNDC2
chr20	43034707	43034881	HNF4A
chrX	153418409	153418583	OPN1LW
chr6	82924148	82924322	IBTK
chr20	54578958	54579132	CBLN4
chr7	95442491	95442665	DYNC1I1
chr22	44011666	44011840	EFCAB6
chr21	43327743	43327917	C2CD2
chr17	43213839	43214013	ACBD4
chr5	35876286	35876460	IL7R
chr3	38520603	38520777	ACVR2B
chr17	72368438	72368612	GPR142
chr8	25261069	25261243	DOCK5
chr19	51519278	51519452	KLK10
chr2	238249485	238249659	COL6A3
chr2	238274538	238274712	COL6A3
chr8	10480565	10480739	RP1L1
chr18	13438225	13438399	C18orf1
chr12	126004061	126004235	TMEM132B
chr12	126135275	126135449	TMEM132B
chr1	55472663	55472837	BSND
chr4	90816143	90816317	MMRN1
chr9	137716397	137716571	COL5A1
chr7	5540135	5540309	FBXL18

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr1	173873047	173873221	SERPINC1
chr18	8819056	8819230	CCDC165
chr5	149435557	149435731	CSF1R
chr14	39650083	39650257	PNN
chr8	1812483	1812657	ARHGEF10
chr9	139390924	139391098	NOTCH1
chr16	66420661	66420835	CDH5
chr7	72420355	72420529	NSUN5P2
chr5	41153950	41154124	C6
chr7	72727151	72727325	TRIM50
chr12	101723050	101723224	UTP20
chr17	42330511	42330685	SLC4A1
chr11	122726396	122726570	CRTAM
chr3	113329835	113330009	SIDT1
chr18	77063572	77063746	ATP9B
chr3	70014056	70014230	MITE
chr19	49485484	49485658	GYS1
chr3	155560227	155560401	SLC33A1
chr12	51868091	51868265	SLC4A8
chr4	155338	155512	ZNF718
chr6	144783714	144783888	UTRN
chr10	92509125	92509299	HTR7
chr12	132561969	132562143	EP400
chr6	26406085	26406259	BTN3A1
chr1	151400732	151400906	POGZ
chr22	32253397	32253571	DEPDC5
chr19	38980763	38980937	RYR1
chr12	40692897	40693071	LRRK2
chr9	117139290	117139464	AKNA
chr6	51882228	51882402	PKHD1
chr11	85685715	85685889	PICALM
chr12	110878067	110878241	ARPC3
chr19	36333279	36333453	NPHS1
chr13	78335104	78335278	SLAIN1
chr15	44951341	44951515	SPG11
chr15	44952639	44952813	SPG11
chr16	84203438	84203612	DNAAF1
chr19	17132828	17133002	CPAMD8
chrX	53279451	53279625	IQSEC2
chr6	29589474	29589648	GABBR1
chr11	132016119	132016293	NTM
chr10	5247671	5247845	AKR1C4
chr6	99796959	99797133	C6orf168
chr8	143570660	143570834	BAI1
chr6	94120655	94120829	EPHA7
chr2	49381393	49381567	FSHR
chr19	53344324	53344498	ZNF468
chr3	53851996	53852170	CHDH
chr10	55568373	55568547	PCDH15
chr10	55955408	55955582	PCDH15
chr2	112566559	112566733	ANAPC1
chr3	66431018	66431192	LRIG1
chr2	207636592	207636766	FASTKD2
chr3	161221163	161221337	OTOL1
chr6	72984006	72984180	RIMS1
chrX	24329584	24329758	FAM48B2
chr2	24345259	24345433	PFN4
chr16	12145820	12145994	SNX29
chr16	27788923	27789097	KIAA0556
chr8	97251691	97251865	MTERFD1
chr6	99857047	99857221	PNISR
chr15	63964637	63964811	HERC1
chr10	121586905	121587079	INPP5F
chr1	156496247	156496421	IQGAP3
chr17	45216093	45216267	CDC27
chr6	71011653	71011827	COL9A1
chr8	119122814	119122988	EXT1
chr16	86601243	86601417	FOXC2
chr19	56895623	56895797	ZNF582
chr5	37815979	37816153	GDNF
chr14	20851339	20851513	TEP1
chr16	67000610	67000784	CES3
chr4	1742545	1742719	TACC3
chr1	44069358	44069532	PTPRF
chr19	16000247	16000421	CYP4F2

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr1	34076610	34076784	CSMD2
chr5	131606566	131606740	PDLIM4
chr12	7585995	7586169	CD163L1
chr12	54893106	54893280	NCKAP1L
chr8	22064343	22064517	BMP1
chr13	48955419	48955593	RB1
chrX	108652224	108652398	GUCY2F
chr1	32163448	32163622	COL16A1
chr4	153273751	153273925	FBXW7
chr12	75816629	75816803	GLIPR1L2
chr22	35481460	35481634	ISX
chr21	10944640	10944814	TPTE
chrX	7268137	7268311	STS
chr2	121708832	121709006	GLI2
chr4	160264407	160264581	RAPGEF2
chr10	100017735	100017909	LOXL4
chr16	30982824	30982998	SETD1A
chr17	36070503	36070677	HNF1B
chr17	36093557	36093731	HNF1B
chr1	200843017	200843191	GPR25
chr17	54912192	54912366	DGKE
chr16	20974568	20974742	DNAH3
chr16	20981170	20981344	DNAH3
chr2	217142438	217142612	MARCH4
chr11	126306719	126306893	KIRREL3
chr14	79746658	79746832	NRXN3
chr12	109660573	109660747	ACACB
chr1	149761698	149761872	FCGR1A
chr2	163144642	163144816	IFIH1
chr1	156131135	156131309	SEMA4A
chr12	93881267	93881441	MRPL42
chrX	135572470	135572644	BRS3
chr11	62519888	62520062	ZBTB3
chr13	47466518	47466692	HTR2A
chr11	68478266	68478440	MTL5
chr6	47976543	47976717	C6orf138
chr16	4700337	4700511	MGRN1
chr17	47246905	47247079	B4GALNT2
chr12	11139352	11139526	TAS2R50
chr7	150068715	150068889	REPIN1
chr6	117710580	117710754	ROS1
chr19	54377174	54377348	MYADM
chrX	77369231	77369405	PGK1
chr2	15358944	15359118	NBAS
chr2	15468292	15468466	NBAS
chr16	24902178	24902352	SLC5A11
chr3	47098628	47098802	SETD2
chr14	64457120	64457294	SYNE2
chr6	69348908	69349082	BAI3
chr6	69703678	69703852	BAI3
chr9	123199555	123199729	CDK5RAP2
chr14	47389201	47389375	MDGA2
chr1	17275270	17275444	CROCC
chr11	47290086	47290260	NR1H3
chr20	58330260	58330434	PHACTR3
chr1	57476794	57476968	DAB1
chr2	162881294	162881468	DPP4
chr1	11918335	11918509	NPPB
chr1	177901799	177901973	SEC16B
chr9	101829150	101829324	COL15A1
chr19	47151866	47152040	DACT3
chr6	30457597	30457771	HLA-E
chr4	5624280	5624454	EVC2
chr19	55179309	55179483	LILRB4
chrX	130408531	130408705	IGSF1
chr1	110019372	110019546	SYPL2
chr21	16337294	16337468	NRIP1
chr2	160086537	160086711	TANC1
chr17	65026809	65026983	CACNG4
chr3	160120511	160120685	SMC4
chr1	115256403	115256577	NRAS
chr19	36369960	36370134	APLP1
chr18	74563725	74563899	ZNF236
chr5	132270198	132270372	AFF4
chr6	31557561	31557735	NCR3

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chrX	154019995	154020169	MPP1
chr6	50696878	50697052	TFAP2D
chr6	50740350	50740524	TFAP2D
chr19	36237591	36237765	PSENEN
chr4	6374233	6374407	PPP2R2C
chr12	6077225	6077399	VWF
chr8	43152132	43152306	POTEA
chr1	145681971	145682145	RNF115
chr1	173545762	173545936	SLC9A11
chr1	47401169	47401343	CYP4A11
chrX	142795485	142795659	SPANXN2
chr7	19184656	19184830	FERD3L
chr1	234367158	234367332	SLC35F3
chr16	70954630	70954804	HYDIN
chr16	70977696	70977870	HYDIN
chr10	70101685	70101859	HNRNPH3
chr10	71562298	71562472	COL13A1
chr10	30629135	30629309	MTPAP
chr19	49685959	49686133	TRPM4
chr1	226784510	226784684	C1orf95
chr18	28993305	28993479	DSG4
chrX	85950015	85950189	DACH2
chr14	60591751	60591925	C14orf135
chr3	119133904	119134078	ARHGAP31
chr13	48664433	48664607	MED4
chr4	54231540	54231714	SCFD2
chr3	108205234	108205408	MYH15
chr8	133911011	133911185	TG
chr8	133935580	133935754	TG
chr8	134034228	134034402	TG
chr1	246704331	246704505	TFB2M
chr10	134261298	134261472	C10orf91
chr9	90343521	90343695	CTSL1
chr5	26988336	26988510	CDH9
chr7	77756573	77756747	MAGI2
chr17	79614896	79615070	TSPAN10
chr9	123877336	123877510	CNTRL
chr9	127998961	127999135	HSPA5
chr2	166170517	166170691	SCN2A
chr2	166172007	166172181	SCN2A
chr2	225639691	225639865	DOCK10
chr6	34512096	34512270	SPDEF
chr7	128587279	128587453	IRF5
chr6	7373630	7373804	CAGE1
chr12	53553899	53554073	CSAD
chr8	42287587	42287761	SLC20A2
chr22	46772901	46773075	CELSR1
chr2	1507698	1507872	TPO
chr1	175063177	175063351	TNN
chr2	107460198	107460372	ST6GAL2
chr18	59894496	59894670	KIAA1468
chr6	32188161	32188335	NOTCH4
chr17	63221250	63221424	RGS9
chr3	52822217	52822391	ITIH1
chr17	40832523	40832697	CCR10
chr11	7660944	7661118	PPFIBP2
chr3	77599967	77600141	ROBO2
chr2	10904416	10904590	ATP6V1C2
chr5	53606198	53606372	ARL15
chr2	201399752	201399926	SGOL2
chr10	102059324	102059498	PKD2L1
chr22	50659438	50659612	TUBGCP6
chr19	50461883	50462057	SIGLEC11
chr17	74152259	74152433	RNF157
chr10	5948282	5948456	FBXO18
chr18	76753167	76753341	SALL3
chr18	76757139	76757313	SALL3
chr21	39672105	39672279	KCNJ15
chr15	59752176	59752350	FAM81A
chr1	35480306	35480480	ZMYM6
chr1	197886953	197887127	LHX9
chr17	42152661	42152835	G6PC3
chr20	31656606	31656780	BPIFB3
chr2	207621940	207622114	MDH1B
chr19	6495231	6495405	TUBB4A

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr19	6772785	6772959	VAV1
chr22	37334110	37334284	CSF2RB
chr12	103248999	103249173	PAH
chr8	53071507	53071681	ST18
chr9	33386941	33387115	AQP7
chr3	125271291	125271466	OSBPL11
chr7	48318316	48318491	ABCA13
chr4	73956405	73956581	ANKRD17
chr2	235951005	235951182	SH3BP4
chr5	167689431	167689608	ODZ2
chr5	80409389	80409566	RASGRF2
chr20	76820	76998	DEFB125
chr1	33354678	33354858	HPCA
chr20	42788454	42788634	JPH2
chr10	27687472	27687652	PTCHD3
chr17	3920844	3921024	ZZEF1
chr7	119915387	119915569	KCND2
chr3	172165394	172165576	GHSR
chr1	197396583	197396765	CRB1
chr3	134851573	134851755	EPHB1
chrX	123787437	123787620	ODZ1
chr2	48873703	48873886	STON1-GTF2A1L
chr4	187510153	187510336	FAT1
chr3	100413608	100413791	GPR128
chr11	61511045	61511229	DAGLA
chr10	127424274	127424458	C10orf137
chr2	136567005	136567189	LCT
chr6	56417763	56417947	DST
chr7	151845894	151846078	MLL3
chr5	110819727	110819912	CAMK4
chr11	49597896	49598085	LOC440040
chr12	53039024	53039213	KRT2
chr14	61113057	61113248	SIX1
chr2	56144829	56145020	EFEMP1
chr19	52569724	52569915	ZNF841
chr8	145059235	145059426	PARP10
chr7	77885425	77885617	MAGI2
chr1	152538457	152538650	LCE3E
chr8	12947763	12947956	DLC1
chrX	134494199	134494392	ZNF449
chr2	220099648	220099842	ANKZF1
chr12	53509189	53509383	SOAT2
chr2	43520109	43520304	THADA
chr16	62055071	62055266	CDH8
chr9	35674150	35674345	CA9
chr1	149858596	149858791	HIST2H2AC
chr15	73635778	73635974	HCN4
chr19	38055893	38056089	ZNF571
chr2	176982002	176982199	HOXD10
chr6	36297863	36298060	C6orf222
chr1	151773708	151773907	LINGO4
chr4	158224736	158224935	GRIA2
chr6	26031938	26032139	HIST1H3B
chr16	71004448	71004649	HYDIN
chr16	50733557	50733759	NOD2
chr12	12871004	12871206	CDKN1B
chr2	166535513	166535716	CSRNP3
chr20	2397926	2398129	TGM6
chr18	11851694	11851897	CHMP1B
chr12	81111035	81111239	MYF5
chr12	125397960	125398165	UBC
chr12	103696126	103696331	C12orf42
chr2	95537486	95537692	TEKT4
chrX	105855747	105855953	CXorf57
chrX	136113602	136113809	GPR101
chr8	32505636	32505844	NRG1
chr13	29599076	29599285	MTUS2
chr12	52845450	52845661	KRT6B
chrX	64749506	64749717	LAS1L
chr10	93294	93506	TUBB8
chrX	57020661	57020873	SPIN3
chr1	12941995	12942208	PRAMEF4
chr19	21991219	21991433	ZNF43
chr19	4174679	4174894	SIRT6
chr2	72359489	72359704	CYP26B1

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr21	43221549	43221764	PRDM15
chr1	161161088	161161303	ADAMTS4
chr5	140249646	140249862	PCDHA11
chr5	140798422	140798638	PCDHGB7
chr21	35821632	35821848	KCNE1
chr6	121768857	121769074	GJA1
chr19	51361311	51361529	KLK3
chr8	65528283	65528501	CYP7B1
chr6	108882612	108882831	FOXO3
chr1	160460935	160461155	SLAMF6
chr11	62996900	62997120	SLC22A25
chr10	103988730	103988951	ELOVL3
chr12	14976319	14976541	C12orf60
chr10	124390548	124390772	DMBT1
chr22	38273748	38273973	EIF3L
chr14	21052116	21052341	RNASE11
chrX	152612476	152612702	ZNF275
chr4	38829757	38829984	TLR6
chr17	39919262	39919489	JUP
chr11	77884984	77885211	KCTD21
chr1	109394764	109394992	AKNAD1
chr6	90383966	90384195	MDN1
chr5	175110248	175110477	HRH2
chr3	49050041	49050271	WDR6
chr14	71275546	71275777	MAP3K9
chr20	31023268	31023501	ASXL1
chr11	104877809	104878042	CASP5
chr16	90001761	90001994	TUBB3
chr6	97561811	97562045	KLHL32
chr7	6370336	6370571	C7orf70
chr16	19883570	19883805	GPRC5B
chr4	151504973	151505208	MAB21L2
chr8	125989526	125989761	ZNF572
chr7	123152132	123152368	IQUB
chr14	102900802	102901038	TECPR2
chr2	105472800	105473037	POU3F3
chr12	41966488	41966725	PDZRN4
chr11	6231583	6231820	C11orf42
chr1	119964948	119965186	HSD3B2
chr11	6239043	6239281	FAM160A2
chr16	19194848	19195087	SYT17
chr6	56483631	56483871	DST
chr6	27860666	27860908	HIST1H2AM
chr19	12244148	12244391	ZNF20
chr11	22646799	22647042	FANCF
chr11	19970310	19970555	NAV2
chr6	27419768	27420014	ZNF184
chr17	10303801	10304047	MYH8
chr16	24582324	24582570	RBBP6
chr1	193038385	193038631	TROVE2
chr4	169432916	169433163	PALLD
chr6	27777875	27778122	HIST1H3H
chr1	216062066	216062313	USH2A
chr7	31617735	31617985	CCDC129
chr22	38823295	38823545	KCNJ4
chr1	70504432	70504683	LRRC7
chr16	53190437	53190690	CHD9
chr16	30594010	30594264	ZNF785
chr15	51696681	51696935	GLDN
chr1	11082198	11082453	TARDBP
chr6	32188797	32189052	NOTCH4
chr7	92844714	92844970	HEPACAM2
chr10	53458635	53458891	CSTF2T
chr10	102763426	102763684	LZTS2
chr18	22057169	22057427	HRH4
chr4	118005644	118005904	TRAM1L1
chr5	140567090	140568461	PCDHB9
chr2	187626638	187627429	FAM171B
chrX	34148258	34150182	FAM47A
chr3	64084735	64085393	PRICKLE2
chr18	13884638	13885323	MC2R
chr14	96707139	96707829	BDKRB2
chr17	21318947	21319722	KCNJ12
chr5	7867011	7867786	FASTKD3
chrX	78010490	78011286	LPAR4

TABLE 14-continued

Chromosome	Start (bp)	End (bp)	Gene
chr15	84651191	84652001	ADAMTSL3
chrX	100911500	100912314	ARMCX2
chr6	116599936	116600467	TSPYL1

TABLE 15

Chromosome	Start (bp)	End (bp)
chrX	8138109	8138209
chrX	48206376	48206476
chrX	91090521	91090621
chrX	140984443	140984543
chrX	151908881	151908981
chrX	153040953	153041053
chrX	153631424	153631524
chr1	7724882	7724982
chr1	24083476	24083576
chr1	32841896	32841996
chr1	34006830	34006930
chr1	155174859	155174959
chr1	230841801	230841901
chr1	248737649	248737749
chr2	102083166	102083266
chr2	179398718	179398818
chr2	216973855	216973956
chr3	4818945	4819045
chr3	36896651	36896751
chr3	38627116	38627216
chr3	49699564	49699664
chr3	52535094	52535194
chr3	56627536	56627636
chr3	62355799	62355899
chr3	196529877	196530028
chr4	9698387	9698487
chr4	42403164	42403264
chr4	46329605	46329705
chr4	114135195	114135295
chr5	94204043	94204143
chr5	167625953	167626053
chr5	178540908	178541008
chr6	150385805	150385905
chr7	13894226	13894326
chr7	44118318	44118418
chr7	64349834	64349934
chr7	128294440	128294540
chr8	7435213	7435313
chr8	42294507	42294607
chr9	15108	15208
chr9	69847617	69847717
chr10	17193296	17193396
chr10	68979399	68979499
chr11	2356849	2356949
chr11	7982062	7982162
chr11	11987362	11987462
chr11	49032122	49032222
chr11	64055365	64055465
chr11	64678071	64678171
chr11	85456674	85456774
chr11	117395696	117395796
chr11	117789268	117789417
chr11	134605814	134605914
chr12	10659455	10659555
chr12	13719926	13720026
chr12	91501855	91501955
chr13	19041955	19042055
chr15	23258090	23258190
chr15	28954627	28954727
chr15	42978135	42978235
chr15	85054943	85055043
chr15	85788527	85788627
chr16	31272956	31273056
chr16	33783262	33783362

TABLE 15-continued

Chromosome	Start (bp)	End (bp)
chr17	7572917	7573017
chr17	7573926	7574033
chr17	7576510	7576691
chr17	7576839	7576939
chr17	7577018	7577155
chr17	7577490	7577608
chr17	7578176	7578289
chr17	7578361	7578554
chr17	7579311	7579590
chr17	7579660	7579760
chr17	7579825	7579925
chr17	18332963	18333063
chr17	41243965	41244164
chr18	180251	180351
chr18	29310984	29311084
chr19	3011018	3011118
chr19	14091507	14091607
chr19	39421292	39421392
chr19	47137855	47137955
chr19	51960611	51960711
chr20	32336701	32336801
chr20	33586350	33586450
chr21	10212863	10212968
chr21	42613818	42613918
chr22	20710164	20710264
chr22	30803370	30803470
chr22	33559458	33559558
chr22	35713817	35713917
track name = 169373__1__OVCA__VIP__P1__tiled__region description = "169373__1__OVCA__VIP__P1__tiled__region"		
chr1	7724858	7724997
chr1	24083442	24083603
chr1	32841867	32842012
chr1	34006797	34006946
chr1	155174825	155174978
chr1	230841774	230841925
chr1	248737627	248737765
chr2	102083143	102083272
chr2	179398692	179398841
chr2	216973829	216973909
chr2	216973909	216973998
chr3	4818920	4819066
chr3	36896617	36896783
chr3	38627083	38627237
chr3	49699538	49699686
chr3	52535072	52535215
chr3	56627509	56627660
chr3	62355774	62355903
chr3	196529855	196529933
chr3	196529945	196530052
chr4	42403130	42403288
chr4	46329575	46329722
chr5	94204009	94204158
chr5	167625926	167626079
chr5	178540874	178541028
chr6	150385770	150385926
chr7	13894204	13894353
chr7	44118287	44118446
chr7	64349812	64349948
chr8	42294483	42294582
chr10	17193286	17193418
chr10	68979377	68979529
chr11	2356869	2356972
chr11	7982041	7982190
chr11	11987341	11987480
chr11	64055338	64055490
chr11	64678038	64678188
chr11	85456646	85456787
chr11	117395668	117395810
chr11	117789243	117789350
chr11	134605849	134605962
chr12	10659427	10659564

TABLE 15-continued

Chromosome	Start (bp)	End (bp)
chr12	13719897	13720051
chr12	91501825	91501965
chr15	42978103	42978261
chr16	31272942	31273080
chr17	7572889	7573037
chr17	7573894	7574051
chr17	7576519	7576721
chr17	7576809	7576957
chr17	7576984	7577173
chr17	7577469	7577644
chr17	7578154	7578326
chr17	7578339	7578589
chr17	7579289	7579605
chr17	7579639	7579772
chr17	7579804	7579954
chr17	41243941	41244199
chr18	180225	180362
chr18	29310950	29311096
chr19	3010990	3011131
chr19	14091474	14091628
chr19	39421257	39421413
chr19	47137832	47137963
chr19	51960576	51960733
chr20	32336677	32336822
chr20	33586318	33586464
chr21	42613785	42613925
chr22	30803344	30803489
chr22	33559434	33559565
chr22	35713793	35713944
chrX	8138074	8138152
chrX	91090488	91090640
chrX	140984422	140984566
chrX	151908888	151908988
chrX	153040929	153041077
chrX	153631403	153631545

TABLE 16

Chromosome	Start (bp)	End (bp)
chr12	347084	347184
chr12	416903	417003
chr12	2064597	2064726
chr12	2760857	2760957
chr12	5603686	5603962
chr12	6649648	6649754
chr12	6711158	6711258
chr12	6711541	6711663
chr12	6858008	6858108
chr12	7061155	7061308
chr12	11139382	11139482
chr12	11338798	11338899
chr12	18841022	18841152
chr12	20832994	20833143
chr12	21644458	21644558
chr12	25362729	25362845
chr12	25368375	25368494
chr12	25378548	25378707
chr12	25380226	25380326
chr12	25398207	25398318
chr12	29450060	29450160
chr12	40044040	40044156
chr12	40704272	40704372
chr12	46318554	46318654
chr12	46320706	46321116
chr12	49087385	49087485
chr12	50452516	50452616
chr12	52715016	52715116
chr12	53097052	53097152
chr12	54367341	54367441
chr12	56628940	56629110
chr12	57422538	57422665

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr12	57605705	57605805
chr12	57883039	57883139
chr12	57919252	57919352
chr12	58024982	58025147
chr12	65856934	65857102
chr12	66531887	66531987
chr12	68707423	68707533
chr12	70070695	70070848
chr12	72057233	72057333
chr12	72070631	72070776
chr12	72094611	72094775
chr12	88524044	88524197
chr12	88566373	88566522
chr12	98921663	98921790
chr12	101680107	101680207
chr12	102056179	102056308
chr12	109278860	109278960
chr12	110765384	110765516
chr12	111311645	111311765
chr12	111758234	111758479
chr12	120595688	120595788
chr12	121017118	121017218
chr12	122812642	122812742
chr12	123794256	123794403
chr12	130647686	130648006
chr12	132281685	132281785
chr12	133219992	133220146
chr14	20852549	20852667
chr14	21861648	21861748
chr14	21961011	21961111
chr14	23312918	23313079
chr14	23341479	23341579
chr14	23845008	23845108
chr14	23869951	23870051
chr14	24646889	24646989
chr14	24785073	24785421
chr14	31355161	31355271
chr14	35331373	35331473
chr14	35592699	35593374
chr14	39871604	39871715
chr14	45642257	45642413
chr14	45693598	45693723
chr14	51094835	51094995
chr14	53558502	53558650
chr14	60903515	60903615
chr14	74824321	74824464
chr14	75514336	75514605
chr14	75590716	75590816
chr14	76112729	76112829
chr14	86087922	86089483
chr14	89629100	89629200
chr14	94517547	94517647
chr14	94545646	94545824
chr14	100367265	100367376
chr14	101005222	101005322
chr14	103599698	103599854
chr14	103996521	103996621
chr14	105174198	105174298
chr14	105241988	105242136
chr19	1037596	1037696
chr19	1486948	1487060
chr19	4329957	4330058
chr19	6222271	6222535
chr19	6477211	6477311
chr19	7734212	7734330
chr19	10262073	10262221
chr19	11541718	11541840
chr19	11618784	11618884
chr19	12384398	12384498
chr19	12430167	12430267
chr19	12461691	12461791
chr19	14208172	14208295
chr19	14262079	14262179
chr19	16024567	16024667
chr19	16633944	16634044

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr19	17160657	17160757
chr19	17943412	17943512
chr19	18376903	18377003
chr19	18420571	18420671
chr19	18887987	18888087
chr19	33353366	33353492
chr19	33666370	33666470
chr19	34710284	34710384
chr19	35773475	35773575
chr19	36050008	36050108
chr19	36050723	36050823
chr19	36053402	36053541
chr19	36054285	36054429
chr19	36255934	36256077
chr19	36583590	36583713
chr19	38948136	38948270
chr19	38976413	38976523
chr19	39898899	39898999
chr19	39972522	39972673
chr19	40711865	40711994
chr19	41063118	41063286
chr19	42752815	42753348
chr19	42795774	42795874
chr19	42797189	42797366
chr19	43766151	43766251
chr19	43969634	43969734
chr19	44612213	44612313
chr19	45655720	45655820
chr19	47572352	47572452
chr19	47883109	47883209
chr19	47935503	47935684
chr19	49218062	49218165
chr19	49850446	49850620
chr19	49971705	49971805
chr19	49978947	49979058
chr19	50310434	50310534
chr19	51133049	51133284
chr19	51189493	51189612
chr19	54327355	54327455
chr19	54544214	54544334
chr19	54649645	54649779
chr19	54675698	54675798
chr19	55607418	55607546
chr19	55711564	55711664
chr19	55815034	55815194
chr19	56171881	56171985
chr19	58083493	58084580
chr19	58596589	58596689
chr22	19373087	19373187
chr22	24583181	24583281
chr22	24717388	24717488
chr22	26906134	26906234
chr22	29913251	29913355
chr22	30742279	30742379
chr22	31011287	31011460
chr22	31535980	31536136
chr22	36689374	36689527
chr22	36696899	36696999
chr22	40816886	40817022
chr22	41257113	41257836
chr22	41650387	41650487
chr22	41753358	41753458
chr22	42271587	42271687
chr22	43213734	43213863
chr22	43218275	43218415
chr22	44083350	44083461
chr22	44559724	44559824
chr22	46327193	46327293
chr22	49042408	49042558
chr22	50506861	50506984
chr17	4619761	4619861
chr17	4937214	4937374
chr17	6364673	6364773
chr17	7106511	7106648
chr17	7193549	7193649

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr17	7495819	7495919
chr17	7572917	7573017
chr17	7573926	7574033
chr17	7576511	7576691
chr17	7576840	7576940
chr17	7577018	7577155
chr17	7577498	7577608
chr17	7578176	7578289
chr17	7578361	7578554
chr17	7579310	7579590
chr17	7579661	7579761
chr17	7579826	7579926
chr17	7606668	7606768
chr17	7796757	7796857
chr17	7798715	7798815
chr17	7801813	7801913
chr17	7843413	7843560
chr17	8397050	8397203
chr17	8415771	8415871
chr17	11650871	11650971
chr17	11924204	11924318
chr17	11958206	11958308
chr17	11984673	11984847
chr17	11998892	11999011
chr17	12011107	12011226
chr17	12013668	12013768
chr17	12016550	12016677
chr17	12028600	12028700
chr17	12032456	12032604
chr17	12043129	12043229
chr17	12044464	12044577
chr17	17394656	17394756
chr17	18167754	18167854
chr17	19232867	19232967
chr17	21094282	21094382
chr17	26653715	26653815
chr17	27001300	27001459
chr17	27027179	27027279
chr17	27889785	27889885
chr17	32483179	32483325
chr17	33520308	33520408
chr17	33749443	33749543
chr17	34077107	34077207
chr17	37879790	37879913
chr17	37880164	37880264
chr17	37880978	37881164
chr17	37881301	37881457
chr17	37881567	37881667
chr17	37881959	37882106
chr17	37882813	37882913
chr17	38421173	38421340
chr17	39022873	39022973
chr17	39122863	39122963
chr17	40837254	40837354
chr17	42992591	42992762
chr17	45219562	45219662
chr17	46622081	46622181
chr17	48433917	48434017
chr17	49156944	49157065
chr17	55028068	55028168
chr17	56389919	56390037
chr17	56434857	56434957
chr17	56435073	56435352
chr17	56435382	56435482
chr17	56435847	56435947
chr17	57247113	57247241
chr17	59668385	59668537
chr17	61899082	61899203
chr17	64092676	64092776
chr17	65905707	65905807
chr17	67522710	67522850
chr17	71354217	71354343
chr17	72469662	72469762
chr17	72943166	72943313
chr17	73239143	73239247

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr17	73481951	73482079
chr17	73732131	73732235
chr17	74077962	74078130
chr17	76458992	76459133
chr17	78201616	78201759
chr4	661644	661795
chr4	3430284	3430438
chr4	3443728	3443845
chr4	4204174	4204305
chr4	10080514	10080625
chr4	15995602	15995702
chr4	39462409	39462582
chr4	41648459	41648559
chr4	46060233	46060386
chr4	56336878	56336978
chr4	57179453	57179553
chr4	70599128	70599228
chr4	71522103	71522203
chr4	76539530	76539630
chr4	83785564	83785678
chr4	85611658	85611817
chr4	87622492	87622608
chr4	88344057	88344164
chr4	88986525	88986647
chr4	89381234	89381334
chr4	90844318	90844423
chr4	105412045	105412145
chr4	106863633	106863733
chr4	109784459	109784559
chr4	110756521	110756621
chr4	123302188	123302288
chr4	128564867	128564967
chr4	134072527	134073572
chr4	146077068	146077168
chr4	168155242	168155342
chr4	169182015	169182140
chr4	170926870	170926970
chr4	177100610	177100730
chr4	190873316	190873442
chr10	7212889	7213018
chr10	17363164	17363264
chr10	22498435	22498535
chr10	24821998	24822166
chr10	26575273	26575423
chr10	27040575	27040675
chr10	27964175	27964310
chr10	33018259	33018385
chr10	46969352	46969452
chr10	50732090	50732190
chr10	55826516	55826645
chr10	61847978	61848078
chr10	63958099	63958199
chr10	64952649	64952749
chr10	70156536	70156638
chr10	70182461	70182561
chr10	70509280	70509442
chr10	75673297	75673488
chr10	81070738	81070838
chr10	81072398	81072506
chr10	82036208	82036308
chr10	93247437	93247537
chr10	93711159	93711323
chr10	96331115	96331215
chr10	98336425	98336525
chr10	101558963	101559127
chr10	102107787	102107887
chr10	102265117	102265252
chr10	103916945	103917085
chr10	104836779	104836930
chr10	105048222	105048322
chr10	105727503	105727653
chr10	116062097	116062243
chr10	116444030	116444130
chr10	118424288	118424388
chr10	125528116	125528216

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr10	129913954	129914054
chr9	732426	732526
chr9	2837246	2837346
chr9	5743663	5743763
chr9	20414277	20414380
chr9	21968185	21968285
chr9	21968698	21968798
chr9	21970900	21971207
chr9	21974475	21974826
chr9	21994137	21994330
chr9	32420853	32421025
chr9	35095211	35095311
chr9	35236465	35236583
chr9	72131028	72131128
chr9	77416863	77416963
chr9	77422949	77423090
chr9	94486025	94486742
chr9	95077956	95078056
chr9	113166731	113166831
chr9	115969484	115969584
chr9	119976940	119977040
chr9	123253584	123253755
chr9	124522388	124522509
chr9	129455461	129455561
chr9	130575652	130575823
chr9	130580994	130581111
chr9	131022868	131022968
chr9	131591013	131591139
chr9	132687193	132687293
chr9	135941916	135942047
chr9	136226831	136226931
chr9	137653747	137653847
chr9	139008443	139008679
chr9	139397633	139397782
chr9	140056855	140056968
chr9	140218176	140218308
chr9	140952471	140952571
chr1	1222151	1222263
chr1	6535990	6536096
chr1	6727768	6727870
chr1	7811247	7811347
chr1	7838153	7838253
chr1	7980889	7980989
chr1	7998252	7998390
chr1	9790591	9790691
chr1	12052611	12052747
chr1	12785443	12785543
chr1	16264312	16264412
chr1	19477028	19477128
chr1	20005564	20005664
chr1	22838512	22838612
chr1	25889552	25889652
chr1	26349532	26349756
chr1	27057828	27057928
chr1	27087322	27087422
chr1	27088638	27088738
chr1	27088744	27088844
chr1	27089460	27089560
chr1	27092949	27093049
chr1	27099898	27099998
chr1	27105786	27106229
chr1	27106416	27106516
chr1	27106566	27106717
chr1	28800223	28800323
chr1	29475123	29475223
chr1	32196435	32196612
chr1	34666397	34666547
chr1	35370323	35370423
chr1	38003354	38003454
chr1	38078426	38078593
chr1	38227550	38227650
chr1	39788581	39788681
chr1	39878456	39878556
chr1	40713659	40713759
chr1	40756525	40756669

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr1	43395320	43395420
chr1	44071897	44071997
chr1	45111071	45111171
chr1	46184874	46185012
chr1	46494439	46494605
chr1	52940825	52941047
chr1	57378115	57378215
chr1	65306925	65307025
chr1	67390344	67390514
chr1	70819823	70819923
chr1	74575077	74575237
chr1	74957780	74957950
chr1	85331663	85331765
chr1	85736461	85736561
chr1	87029343	87029452
chr1	90470691	90470807
chr1	91967273	91967388
chr1	97771732	97771853
chr1	100377960	100378073
chr1	100534028	100534142
chr1	103355010	103355118
chr1	109276088	109276188
chr1	109477357	109477457
chr1	110031507	110031607
chr1	110300565	110300680
chr1	110906387	110906535
chr1	112020604	112020745
chr1	114968115	114968228
chr1	115537502	115537640
chr1	120056676	120056801
chr1	145415368	145415659
chr1	150444622	150444722
chr1	150530456	150530556
chr1	150551491	150551722
chr1	151773421	151774034
chr1	152732617	152732717
chr1	154245815	154245915
chr1	154680537	154680637
chr1	154917496	154917596
chr1	154962035	154962183
chr1	155886415	155886528
chr1	156693973	156694073
chr1	156761486	156761586
chr1	156849790	156849949
chr1	157805857	157805957
chr1	159163212	159163350
chr1	159273798	159273898
chr1	161044006	161044163
chr1	161058997	161059097
chr1	161202958	161203128
chr1	161254128	161254271
chr1	165177284	165177384
chr1	167384967	167385067
chr1	169833509	169833649
chr1	176762695	176762805
chr1	179337934	179338107
chr1	181732541	181732667
chr1	183616777	183616877
chr1	196227348	196227480
chr1	200842727	200842827
chr1	202568345	202568479
chr1	203054859	203055000
chr1	203786175	203786275
chr1	204135322	204135422
chr1	204416563	204416706
chr1	212115142	212115242
chr1	214556938	214557228
chr1	215960116	215960216
chr1	218609321	218609421
chr1	228467831	228467931
chr1	230839005	230839105
chr1	231131517	231131617
chr1	231829954	231830346
chr1	237024424	237024577
chr3	433357	433480

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr3	3888107	3888207
chr3	9027235	9027335
chr3	10291028	10291128
chr3	14219965	14220068
chr3	18462306	18462406
chr3	20216000	20216100
chr3	33602299	33602463
chr3	33866678	33866832
chr3	41265517	41265617
chr3	41266017	41266244
chr3	42739621	42739721
chr3	46306947	46307344
chr3	48200833	48200945
chr3	48587550	48587667
chr3	48677592	48677692
chr3	49321898	49322017
chr3	53220615	53220715
chr3	57509273	57509373
chr3	57542719	57542819
chr3	66436616	66436716
chr3	71247308	71247408
chr3	100148573	100148729
chr3	109049556	109049656
chr3	113737603	113737703
chr3	122433108	122433276
chr3	122634318	122634418
chr3	123411581	123411698
chr3	125876185	125876351
chr3	129370341	129370593
chr3	130095308	130095408
chr3	137880694	137880794
chr3	138121011	138121111
chr3	148563300	148563400
chr3	149260145	149260245
chr3	151148077	151148177
chr3	154032928	154033028
chr3	157081149	157081249
chr3	180666134	180666283
chr3	183822574	183822744
chr3	184104303	184104403
chr3	185364812	185364926
chr3	186362539	186362709
chr3	188933076	188933176
chr3	194181386	194181560
chr3	194790696	194790813
chr3	195595179	195595279
chr3	196214337	196214438
chr5	6754964	6755064
chr5	13919333	13919433
chr5	16694498	16694614
chr5	24492925	24493034
chr5	32419902	32420002
chr5	33577063	33577163
chr5	39382968	39383091
chr5	40947724	40947835
chr5	41160228	41160328
chr5	44388666	44388766
chr5	68692223	68692373
chr5	79372725	79372825
chr5	92923761	92923925
chr5	112175114	112176032
chr5	113698630	113698896
chr5	115202362	115202462
chr5	122881445	122881545
chr5	124079764	124079864
chr5	127420206	127420310
chr5	130782282	130782382
chr5	131676257	131676393
chr5	135692811	135692926
chr5	137627658	137627805
chr5	137801497	137801597
chr5	139905626	139905726
chr5	140052235	140052335
chr5	140431915	140432015
chr5	140502437	140502537

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr5	140559621	140559744
chr5	141694312	141694412
chr5	145886673	145886773
chr5	156378717	156378817
chr5	156535915	156536015
chr5	168244304	168244468
chr5	172578568	172578697
chr5	175306916	175307016
chr5	176301253	176301353
chr5	179149870	179149970
chr5	180039505	180039611
chr7	720192	720363
chr7	897464	897594
chr7	1586613	1586713
chr7	2963866	2963999
chr7	5410117	5410274
chr7	5427420	5427520
chr7	6621799	6621899
chr7	7679950	7680050
chr7	8198156	8198282
chr7	11075266	11075413
chr7	16655368	16655529
chr7	21659573	21659696
chr7	23775253	23775353
chr7	27135296	27135396
chr7	27222460	27222563
chr7	30661932	30662078
chr7	33312673	33312773
chr7	44684935	44685097
chr7	44805030	44805130
chr7	53103800	53104078
chr7	55241613	55241736
chr7	55242414	55242514
chr7	55248985	55249171
chr7	55259411	55259567
chr7	55260446	55260546
chr7	55266409	55266556
chr7	55268007	55268107
chr7	57528603	57528763
chr7	73123363	73123463
chr7	76112199	76112299
chr7	77326171	77326271
chr7	77423410	77423510
chr7	77569466	77569582
chr7	87074177	87074291
chr7	87195385	87195557
chr7	91671359	91671500
chr7	91752444	91752544
chr7	91870306	91870466
chr7	92146671	92146771
chr7	94879372	94879472
chr7	98995484	98995636
chr7	99032556	99032656
chr7	100028451	100029194
chr7	100281642	100281780
chr7	100283596	100283702
chr7	100479279	100479426
chr7	100481991	100482091
chr7	102964919	102965019
chr7	106508895	106508995
chr7	123508655	123508828
chr7	127670197	127670326
chr7	135418734	135418834
chr7	139094304	139094404
chr7	142460718	142460871
chr7	144097276	144097376
chr7	149479933	149480085
chr7	150775932	150776055
chr7	152346170	152346270
chr7	155531024	155531124
chr7	158704240	158704370
chr16	677434	677581
chr16	709056	709156
chr16	1824257	1824357
chr16	3350422	3350522

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr16	4796927	4797027
chr16	4910642	4910742
chr16	15131881	15131981
chr16	17211715	17211832
chr16	19580745	19580913
chr16	20370651	20370751
chr16	24807194	24807294
chr16	27373786	27373989
chr16	28931106	28931264
chr16	29998236	29999167
chr16	30365504	30365643
chr16	30531177	30531288
chr16	30736293	30736393
chr16	30793022	30793122
chr16	57486712	57486848
chr16	61851469	61851569
chr16	66603838	66603997
chr16	67267798	67267898
chr16	67299990	67300128
chr16	67693601	67693701
chr16	67963836	67963998
chr16	68718454	68718554
chr16	69782929	69783029
chr16	70814692	70814842
chr16	77334159	77334301
chr16	77356232	77356363
chr16	84797797	84797897
chr16	89950977	89951077
chr6	5771523	5771662
chr6	6002585	6002730
chr6	10755370	10755470
chr6	15497081	15497191
chr6	17688669	17688769
chr6	20490463	20490618
chr6	24145856	24146061
chr6	25670215	25670315
chr6	27100315	27100415
chr6	29454914	29455158
chr6	29573345	29573473
chr6	29694753	29694853
chr6	29797146	29797246
chr6	30157194	30157310
chr6	30166696	30166796
chr6	30572770	30572871
chr6	31083900	31084623
chr6	31597407	31597507
chr6	31939778	31939878
chr6	32063512	32063628
chr6	36368230	36368361
chr6	36824346	36824446
chr6	36867321	36867421
chr6	41168669	41168769
chr6	41621120	41621220
chr6	42611951	42612074
chr6	43323452	43323552
chr6	44269109	44269209
chr6	46129279	46129389
chr6	52883127	52883299
chr6	55039361	55039461
chr6	66063350	66063510
chr6	80751827	80751927
chr6	84896183	84896283
chr6	88144650	88144750
chr6	89891623	89891776
chr6	89975331	89975484
chr6	90660238	90660536
chr6	91296482	91296602
chr6	100382273	100382393
chr6	111982992	111983092
chr6	129959553	129959653
chr6	131191024	131191124
chr6	131919776	131919876
chr6	144086412	144086913
chr6	146350670	146350782
chr6	155743827	155743990

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr6	161560496	161560596
chr6	165809805	165809948
chr6	168461473	168461614
chr6	170852688	170852818
chr21	28296424	28296739
chr21	33043862	33043973
chr21	34799190	34799339
chr21	37510122	37510230
chr21	37617677	37617852
chr21	43221366	43221466
chr21	45402179	45402279
chr21	45472225	45472325
chr21	46276107	46276279
chr21	47532247	47532347
chr2	3691370	3691470
chr2	10917710	10917848
chr2	17898007	17898172
chr2	20482928	20483028
chr2	23977516	23977644
chr2	25466997	25467097
chr2	26022253	26022404
chr2	26534363	26534463
chr2	37454708	37454909
chr2	39074138	39074238
chr2	39440527	39440627
chr2	44051454	44051561
chr2	54093225	54093360
chr2	54133924	54134024
chr2	70903931	70904031
chr2	73315288	73315388
chr2	74687496	74687596
chr2	86075243	86075343
chr2	96919732	96919832
chr2	96992434	96992796
chr2	97526796	97526896
chr2	98427590	98427695
chr2	99226307	99226448
chr2	99778731	99778831
chr2	100055052	100055152
chr2	103149087	103149187
chr2	103324696	103324796
chr2	108863650	108863822
chr2	109087882	109088537
chr2	109380401	109380689
chr2	109524315	109524475
chr2	113416558	113416658
chr2	128046912	128047077
chr2	128471199	128471490
chr2	128712606	128712706
chr2	141641447	141641589
chr2	143959687	143959787
chr2	157186341	157186486
chr2	160801392	160801492
chr2	165551295	165551409
chr2	167760151	167760306
chr2	170493716	170493871
chr2	176044816	176044916
chr2	178988853	178989017
chr2	179192982	179193105
chr2	191161538	191161679
chr2	196788321	196788421
chr2	197649567	197649667
chr2	201436954	201437054
chr2	202352302	202352402
chr2	204000757	204000857
chr2	204150334	204150455
chr2	207827229	207827329
chr2	210887630	210887730
chr2	211532909	211533009
chr2	216263977	216264077
chr2	219252271	219252371
chr2	233546294	233546429
chr2	233675953	233676061
chr2	238253015	238253143
chr2	242738474	242738574

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chrX	16850745	16850865
chrX	32361250	32361403
chrX	35989816	35989916
chrX	37312561	37312661
chrX	47058878	47059013
chrX	54011356	54011456
chrX	54578264	54578407
chrX	69478724	69478845
chrX	70367813	70367913
chrX	100880103	100880203
chrX	102004353	102004453
chrX	111090413	111090513
chrX	112048174	112048320
chrX	119072683	119072783
chrX	119694068	119694168
chrX	130409150	130409250
chrX	132458510	132458610
chrX	135313710	135314195
chrX	135584936	135585100
chrX	142605130	142605230
chrX	142803673	142803773
chrX	149937490	149937590
chrX	149984465	149984565
chrX	153627827	153627935
chr11	281504	281604
chr11	592558	592674
chr11	864424	864524
chr11	970168	970311
chr11	3720339	3720439
chr11	6412812	6412912
chr11	6432281	6432381
chr11	6622388	6622563
chr11	6650980	6651111
chr11	6661246	6662162
chr11	16822520	16822622
chr11	18127453	18127588
chr11	19251411	19251511
chr11	26619903	26620003
chr11	34161946	34162119
chr11	35218292	35218421
chr11	35513592	35513721
chr11	46702189	46702297
chr11	46829580	46829694
chr11	47359052	47359152
chr11	57564170	57564343
chr11	61183716	61183816
chr11	61607836	61607936
chr11	62303416	62303570
chr11	62649364	62649538
chr11	63149630	63149749
chr11	63306987	63307096
chr11	64004626	64004726
chr11	64888152	64888278
chr11	66454969	66455069
chr11	66468053	66468445
chr11	68190977	68191128
chr11	72528800	72528900
chr11	73555851	73556021
chr11	74336452	74336618
chr11	75694430	75694557
chr11	76506624	76506724
chr11	85967428	85967554
chr11	102668103	102668203
chr11	102738638	102738799
chr11	104879533	104879707
chr11	107535750	107535923
chr11	108559662	108559789
chr11	113281505	113281641
chr11	118220533	118220633
chr11	118770650	118770899
chr11	118967851	118968007
chr11	119149307	119149407
chr11	124740060	124740199
chr11	125505323	125505428
chr11	126174112	126174212

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr8	1626395	1626542
chr8	6289013	6289113
chr8	6378748	6378848
chr8	12957610	12957920
chr8	21985047	21985147
chr8	22020086	22020245
chr8	24813395	24813506
chr8	37728883	37728983
chr8	41166589	41166689
chr8	41798371	41798471
chr8	48701466	48701610
chr8	48746749	48746849
chr8	48790284	48790412
chr8	52321507	52321943
chr8	52359563	52359722
chr8	59404128	59404295
chr8	59750747	59750847
chr8	68062017	68062170
chr8	70513976	70514076
chr8	70617374	70617474
chr8	73480174	73480441
chr8	74507400	74507532
chr8	81733696	81733796
chr8	86129614	86129728
chr8	92406163	92406267
chr8	95952360	95952460
chr8	100990144	100990303
chr8	101724879	101725017
chr8	103274142	103274298
chr8	105456610	105456710
chr8	113240984	113241120
chr8	124219641	124219741
chr8	124368623	124368723
chr8	124664069	124665023
chr8	128750556	128750656
chr8	133150131	133150263
chr8	139809031	139809131
chr8	143310842	143310942
chr8	145006567	145006729
chr13	26975602	26975761
chr13	27255211	27255388
chr13	31715258	31715396
chr13	32376379	32376479
chr13	43934078	43934178
chr13	47243164	47243291
chr13	50235132	50235232
chr13	51948358	51948458
chr13	73337588	73337745
chr13	79918803	79918929
chr13	95695935	95696041
chr13	100617846	100617948
chr13	103524562	103524662
chr18	909503	909603
chr18	5398019	5398142
chr18	5415791	5415891
chr18	5891943	5892043
chr18	8792995	8793118
chr18	10677743	10677872
chr18	13681701	13681801
chr18	21745024	21745124
chr18	21750290	21750417
chr18	43685124	43685224
chr18	48573417	48573665
chr18	48575056	48575230
chr18	48575630	48575730
chr18	48581151	48581363
chr18	48584495	48584614
chr18	48584710	48584826
chr18	48586212	48586312
chr18	48591868	48591968
chr18	48593389	48593557
chr18	48603007	48603146
chr18	48604651	48604751
chr18	55992227	55992394
chr18	74962884	74962984

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr15	26825465	26825603
chr15	28200282	28200382
chr15	28515840	28516014
chr15	29346290	29346390
chr15	31196844	31196944
chr15	33962619	33962757
chr15	34546661	34546761
chr15	37385781	37385931
chr15	40282473	40282578
chr15	40631712	40631820
chr15	42439823	42439960
chr15	43038022	43038314
chr15	43641104	43641275
chr15	45710754	45710880
chr15	59182495	59182595
chr15	64972946	64973046
chr15	65678280	65678380
chr15	68937474	68937574
chr15	72208710	72208833
chr15	74636145	74636262
chr15	75122509	75122609
chr15	91019894	91020050
chr15	91550178	91550278
chr20	4850551	4850699
chr20	5903282	5903662
chr20	7962931	7963031
chr20	17462207	17462307
chr20	17581630	17581730
chr20	21492785	21492920
chr20	23065522	23066663
chr20	25422325	25422425
chr20	30064293	30064393
chr20	30232570	30232707
chr20	30309541	30309641
chr20	35060132	35060246
chr20	39990779	39990879
chr20	40033354	40033454
chr20	43933081	43933181
chr20	45874914	45875073
chr20	56099137	56099237
chr20	57478787	57478887
chr20	57480445	57480545
chr20	57484391	57484491
chr20	57484547	57484647
chr20	58452439	58452610
chr20	58466997	58467097
chr20	58490512	58490612
track		
name = 169413__5__PANC__		
v2__P1__tiled__region		
description = "169413__5__		
PANC__v2__P1__tiled__region"		
chr1	1222128	1222293
chr1	6536023	6536127
chr1	6727733	6727908
chr1	7811223	7811356
chr1	7838123	7838205
chr1	7838218	7838293
chr1	7980913	7980990
chr1	7998228	7998398
chr1	9790562	9790713
chr1	12052590	12052760
chr1	12785415	12785556
chr1	16264279	16264427
chr1	19476998	19477152
chr1	20005533	20005690
chr1	22838478	22838558
chr1	22838568	22838647
chr1	25889521	25889665
chr1	26349541	26349790
chr1	27057796	27057948
chr1	27087296	27087444
chr1	27088606	27088863
chr1	27089431	27089580
chr1	27092926	27093065

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr1	27099876	27100016
chr1	27105756	27106252
chr1	27106381	27106527
chr1	27106541	27106748
chr1	28800191	28800344
chr1	29475096	29475215
chr1	32196407	32196648
chr1	34666362	34666583
chr1	35370302	35370443
chr1	38003377	38003480
chr1	38078392	38078607
chr1	38227515	38227674
chr1	39788559	39788697
chr1	39878424	39878579
chr1	40713629	40713784
chr1	40756504	40756711
chr1	43395299	43395448
chr1	44071865	44072028
chr1	45111050	45111198
chr1	46184850	46185021
chr1	46494415	46494633
chr1	52940800	52941043
chr1	57378090	57378229
chr1	65306902	65307035
chr1	67390312	67390529
chr1	70819794	70819945
chr1	74575056	74575253
chr1	74957746	74957961
chr1	85331672	85331786
chr1	85736427	85736577
chr1	87029322	87029488
chr1	90470668	90470839
chr1	91967242	91967417
chr1	97771697	97771879
chr1	100377937	100378107
chr1	100534007	100534171
chr1	103354987	103355153
chr1	109276053	109276203
chr1	109477333	109477480
chr1	110031476	110031617
chr1	110300536	110300710
chr1	110906356	110906567
chr1	112020582	112020752
chr1	114968123	114968267
chr1	115537479	115537671
chr1	120056642	120056769
chr1	145415374	145415689
chr1	150444590	150444745
chr1	150530425	150530568
chr1	150551460	150551747
chr1	151773398	151774056
chr1	152732593	152732726
chr1	154245880	154245950
chr1	154680510	154680590
chr1	154680595	154680681
chr1	154917515	154917627
chr1	154962010	154962225
chr1	155886424	155886565
chr1	156693944	156694076
chr1	156761464	156761602
chr1	156849764	156849973
chr1	157805829	157805970
chr1	159163177	159163361
chr1	159273772	159273919
chr1	161043973	161044048
chr1	161044063	161044202
chr1	161059033	161059134
chr1	161202923	161203159
chr1	161254163	161254294
chr1	165177249	165177331
chr1	165177339	165177425
chr1	167384939	167385096
chr1	169833485	169833662
chr1	176762661	176762845
chr1	179337909	179338106

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr1	181732509	181732688
chr1	183616744	183616816
chr1	183616829	183616907
chr1	196227315	196227503
chr1	200842702	200842781
chr1	200842782	200842866
chr1	202568322	202568500
chr1	203054827	203055012
chr1	203786227	203786298
chr1	204135297	204135455
chr1	204416537	204416719
chr1	212115109	212115270
chr1	214556914	214557243
chr1	215960094	215960240
chr1	218609296	218609439
chr1	228467801	228467947
chr1	230838979	230839127
chr1	231131492	231131625
chr1	231829932	231830387
chr1	237024408	237024593
chr2	3691337	3691499
chr2	10917688	10917823
chr2	17897984	17898181
chr2	20482893	20482976
chr2	20482983	20483053
chr2	23977483	23977658
chr2	25466973	25467111
chr2	26022228	26022298
chr2	26022308	26022445
chr2	26534328	26534470
chr2	37454684	37454918
chr2	39074109	39074245
chr2	39440499	39440637
chr2	44051432	44051571
chr2	54093201	54093368
chr2	54133901	54134038
chr2	70903898	70904041
chr2	73315263	73315329
chr2	73315343	73315416
chr2	86075222	86075361
chr2	96919709	96919857
chr2	96992399	96992798
chr2	97526764	97526909
chr2	98427563	98427713
chr2	99226283	99226486
chr2	99778696	99778803
chr2	100055021	100055096
chr2	100055111	100055181
chr2	103149053	103149193
chr2	103324663	103324807
chr2	108863616	108863839
chr2	109087851	109088569
chr2	109380371	109380727
chr2	109524281	109524509
chr2	113416533	113416682
chr2	128046877	128046951
chr2	128046952	128047089
chr2	128471171	128471522
chr2	128712581	128712732
chr2	141641422	141641629
chr2	143959666	143959809
chr2	157186320	157186525
chr2	160801361	160801509
chr2	165551270	165551446
chr2	167760120	167760339
chr2	170493690	170493851
chr2	176044792	176044930
chr2	178988832	178989032
chr2	179192952	179193128
chr2	191161514	191161720
chr2	196788296	196788437
chr2	197649536	197649677
chr2	201436928	201437069
chr2	202352268	202352424
chr2	204000734	204000887

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr2	204150309	204150475
chr2	207827204	207827345
chr2	210887605	210887743
chr2	211532899	211533036
chr2	216263944	216264106
chr2	219252245	219252385
chr2	233546262	233546444
chr2	233675937	233676092
chr2	238252990	238253170
chr2	242738497	242738598
chr3	433335	433511
chr3	3888085	3888234
chr3	9027204	9027279
chr3	9027289	9027366
chr3	10290999	10291116
chr3	14219968	14220108
chr3	18462284	18462426
chr3	20215974	20216109
chr3	33602277	33602490
chr3	33866657	33866859
chr3	41265496	41265642
chr3	41265996	41266268
chr3	42739586	42739733
chr3	46306916	46307374
chr3	48200806	48200981
chr3	48587523	48587711
chr3	48677558	48677709
chr3	49321873	49322059
chr3	53220583	53220731
chr3	57509244	57509390
chr3	57542684	57542840
chr3	66436585	66436736
chr3	71247321	71247438
chr3	100148595	100148766
chr3	109049529	109049685
chr3	113737569	113737727
chr3	122433085	122433298
chr3	122634290	122634440
chr3	123411560	123411737
chr3	125876160	125876363
chr3	129370311	129370559
chr3	130095276	130095412
chr3	137880712	137880831
chr3	138120978	138121142
chr3	148563271	148563414
chr3	149260114	149260200
chr3	149260204	149260287
chr3	151148044	151148201
chr3	154032896	154032975
chr3	154032986	154033061
chr3	157081114	157081264
chr3	180666159	180666232
chr3	180666234	180666309
chr3	183822548	183822773
chr3	184104268	184104424
chr3	185364888	185364960
chr3	186362518	186362729
chr3	188933049	188933194
chr3	194181361	194181588
chr3	194790662	194790847
chr3	195595152	195595286
chr3	196214302	196214468
chr4	661620	661760
chr4	3430251	3430358
chr4	3430376	3430466
chr4	3443696	3443769
chr4	3443801	3443882
chr4	4204256	4204329
chr4	10080479	10080669
chr4	15995581	15995722
chr4	39462385	39462587
chr4	41648435	41648585
chr4	46060200	46060408
chr4	56336848	56336994
chr4	57179428	57179569

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr4	70599102	70599244
chr4	71522072	71522232
chr4	76539504	76539639
chr4	83785539	83785683
chr4	85611637	85611856
chr4	87622457	87622637
chr4	88344030	88344183
chr4	88986490	88986675
chr4	89381200	89381357
chr4	90844295	90844372
chr4	105412024	105412179
chr4	106863612	106863684
chr4	106863692	106863763
chr4	109784431	109784576
chr4	110756491	110756635
chr4	123302156	123302317
chr4	128564846	128564994
chr4	134072493	134073557
chr4	146077040	146077105
chr4	146077130	146077203
chr4	168155211	168155356
chr4	169181982	169182163
chr4	170926847	170926987
chr4	177100589	177100759
chr5	6754930	6755002
chr5	6755015	6755098
chr5	13919310	13919459
chr5	16694463	16694644
chr5	24492890	24493075
chr5	32419878	32419946
chr5	33577038	33577182
chr5	39382943	39383122
chr5	40947690	40947873
chr5	41160205	41160281
chr5	44388640	44388715
chr5	44388720	44388800
chr5	68692198	68692376
chr5	79372704	79372858
chr5	92923737	92923967
chr5	112175091	112176075
chr5	113698607	113698935
chr5	115202425	115202496
chr5	122881416	122881559
chr5	124079732	124079805
chr5	124079822	124079892
chr5	127420214	127420358
chr5	130782253	130782404
chr5	131676227	131676405
chr5	135692786	135692962
chr5	137627627	137627838
chr5	137801472	137801608
chr5	139905598	139905735
chr5	140052273	140052356
chr5	140431893	140432028
chr5	140502408	140502552
chr5	140559613	140559686
chr5	141694278	141694358
chr5	141694368	141694448
chr5	145886642	145886717
chr5	145886732	145886809
chr5	156378682	156378840
chr5	156535952	156536053
chr5	168244281	168244510
chr5	172578559	172578730
chr5	175306884	175306959
chr5	175306974	175307051
chr5	176301219	176301300
chr5	176301309	176301383
chr5	179149844	179149987
chr5	180039476	180039659
chr6	5771502	5771686
chr6	6002562	6002771
chr6	10755337	10755407
chr6	15497047	15497229
chr6	17688638	17688792

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr6	20490498	20490569
chr6	20490578	20490656
chr6	24145831	24146079
chr6	25670194	25670335
chr6	27100294	27100396
chr6	29454891	29455182
chr6	29573311	29573434
chr6	29694761	29694834
chr6	29797196	29797268
chr6	30157172	30157341
chr6	30166667	30166822
chr6	30572738	30572912
chr6	31083868	31084599
chr6	31597376	31597513
chr6	31939756	31939884
chr6	32063481	32063664
chr6	36368209	36368386
chr6	36824324	36824473
chr6	36867289	36867427
chr6	41168721	41168801
chr6	41621086	41621268
chr6	42611941	42612105
chr6	43323426	43323566
chr6	44269086	44269226
chr6	46129252	46129403
chr6	52883094	52883312
chr6	55039334	55039475
chr6	66063315	66063527
chr6	80751806	80751945
chr6	84896156	84896302
chr6	88144618	88144762
chr6	89891593	89891733
chr6	89975298	89975525
chr6	90660213	90660562
chr6	91296458	91296558
chr6	100382249	100382412
chr6	111982970	111983114
chr6	129959531	129959675
chr6	131190996	131191067
chr6	131191081	131191157
chr6	131919751	131919882
chr6	144086380	144086944
chr6	146350645	146350827
chr6	155743803	155744017
chr6	161560524	161560629
chr6	165809781	165809982
chr6	168461438	168461661
chr6	170852667	170852850
chr7	720164	720381
chr7	897429	897618
chr7	1586589	1586724
chr7	2963843	2964018
chr7	5410085	5410269
chr7	5427395	5427533
chr7	6621775	6621921
chr7	7679925	7680071
chr7	8198135	8198300
chr7	11075244	11075450
chr7	16655339	16655550
chr7	21659550	21659719
chr7	23775226	23775376
chr7	27135262	27135330
chr7	27135337	27135411
chr7	27222467	27222604
chr7	30661907	30662124
chr7	33312644	33312789
chr7	44684912	44685124
chr7	44805007	44805147
chr7	53103766	53104085
chr7	55241591	55241759
chr7	55242381	55242526
chr7	55248951	55249200
chr7	55259376	55259601
chr7	55260416	55260574
chr7	55266386	55266601

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr7	55267986	55268123
chr7	57528571	57528714
chr7	73123341	73123418
chr7	73123426	73123509
chr7	76112171	76112316
chr7	77326136	77326219
chr7	77326226	77326300
chr7	77423386	77423529
chr7	77569436	77569612
chr7	87074152	87074236
chr7	87074247	87074317
chr7	87195357	87195570
chr7	91671337	91671518
chr7	91752497	91752574
chr7	91870282	91870413
chr7	91870427	91870499
chr7	92146637	92146783
chr7	94879337	94879498
chr7	98995450	98995521
chr7	98995600	98995678
chr7	99032530	99032601
chr7	99032615	99032689
chr7	100028420	100029207
chr7	100281610	100281817
chr7	100283575	100283740
chr7	100479250	100479463
chr7	100481960	100482040
chr7	100482045	100482118
chr7	102964895	102965033
chr7	106508865	106509013
chr7	123508620	123508832
chr7	127670169	127670364
chr7	135418701	135418854
chr7	139094280	139094351
chr7	142460687	142460761
chr7	142460792	142460893
chr7	144097242	144097401
chr7	149479911	149480123
chr7	150775906	150776087
chr7	152346136	152346282
chr7	155530999	155531143
chr7	158704214	158704398
chr8	1626365	1626581
chr8	6288981	6289102
chr8	6378716	6378860
chr8	12957619	12957930
chr8	21985022	21985156
chr8	22020057	22020161
chr8	22020162	22020272
chr8	24813363	24813549
chr8	37728858	37729010
chr8	41166558	41166712
chr8	41798343	41798416
chr8	41798428	41798507
chr8	48701435	48701617
chr8	48746720	48746865
chr8	48790250	48790435
chr8	52321485	52321980
chr8	52359530	52359755
chr8	59404096	59404313
chr8	59750725	59750871
chr8	68061992	68062192
chr8	70513952	70514089
chr8	70617352	70617492
chr8	73480152	73480472
chr8	74507415	74507561
chr8	81733668	81733808
chr8	86129583	86129762
chr8	92406142	92406285
chr8	95952337	95952413
chr8	95952417	95952500
chr8	100990111	100990320
chr8	101724961	101725031
chr8	103274111	103274326
chr8	105456586	105456724

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr8	113240959	113241128
chr8	124219617	124219743
chr8	124368592	124368737
chr8	124664037	124665069
chr8	128750532	128750678
chr8	133150109	133150244
chr8	139809002	139809155
chr8	145006545	145006623
chr8	145006640	145006743
chr9	732478	732554
chr9	2837221	2837357
chr9	5743634	5743776
chr9	21968154	21968310
chr9	21968669	21968817
chr9	21970869	21971023
chr9	21971074	21971146
chr9	21974444	21974836
chr9	21994114	21994361
chr9	32420823	32421033
chr9	35095204	35095357
chr9	35236444	35236619
chr9	72130998	72131143
chr9	77416842	77416983
chr9	77422922	77423098
chr9	94486031	94486773
chr9	95077931	95078069
chr9	113166707	113166850
chr9	115969454	115969600
chr9	119976909	119976988
chr9	119976999	119977080
chr9	123253559	123253768
chr9	124522395	124522523
chr9	129455432	129455569
chr9	130575617	130575854
chr9	130580972	130581146
chr9	131022836	131022917
chr9	131590978	131591165
chr9	132687168	132687306
chr9	135941895	135942068
chr9	136226810	136226952
chr9	137653713	137653852
chr9	139008448	139008690
chr9	139397602	139397822
chr9	140056823	140056992
chr9	140218148	140218334
chr9	140952438	140952595
chr10	7212857	7213037
chr10	17363131	17363274
chr10	22498401	22498568
chr10	24821968	24822114
chr10	24822128	24822205
chr10	26575246	26575457
chr10	27040626	27040702
chr10	27964141	27964333
chr10	33018233	33018400
chr10	46969324	46969408
chr10	46969409	46969489
chr10	50732066	50732140
chr10	50732146	50732226
chr10	55826495	55826669
chr10	61847946	61848088
chr10	63958066	63958141
chr10	63958151	63958220
chr10	64952621	64952766
chr10	70156512	70156654
chr10	70182427	70182576
chr10	70509297	70509464
chr10	75673263	75673530
chr10	81070709	81070784
chr10	81070794	81070870
chr10	81072369	81072545
chr10	82036174	82036342
chr10	93247411	93247568
chr10	93711126	93711339
chr10	96331166	96331233

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr10	98336401	98336540
chr10	101558983	101559162
chr10	102107758	102107836
chr10	102107838	102107913
chr10	102265083	102265269
chr10	103916914	103916988
chr10	103917049	103917123
chr10	104836748	104836970
chr10	105048193	105048269
chr10	105048278	105048346
chr10	105727468	105727673
chr10	116062068	116062278
chr10	116444005	116444147
chr10	118424265	118424405
chr10	125528085	125528163
chr10	125528175	125528249
chr10	129913929	129914061
chr11	281480	281552
chr11	281560	281634
chr11	592525	592705
chr11	864445	864545
chr11	970145	970357
chr11	3720311	3720459
chr11	6412781	6412937
chr11	6432251	6432388
chr11	6622366	6622585
chr11	6650951	6651138
chr11	6661216	6662201
chr11	16822494	16822646
chr11	18127429	18127534
chr11	18127544	18127616
chr11	19251379	19251525
chr11	26619878	26620010
chr11	34161913	34162140
chr11	35218258	35218448
chr11	35513558	35513748
chr11	46702161	46702338
chr11	46829551	46829728
chr11	47359026	47359168
chr11	57564135	57564363
chr11	61183692	61183768
chr11	61183772	61183842
chr11	61607807	61607880
chr11	61607892	61607971
chr11	62303395	62303607
chr11	62649335	62649550
chr11	63149678	63149779
chr11	63306958	63307120
chr11	64004593	64004751
chr11	64888123	64888261
chr11	66454947	66455082
chr11	66468022	66468487
chr11	68190943	68191166
chr11	72528770	72528852
chr11	72528865	72528933
chr11	73555825	73556051
chr11	74336422	74336636
chr11	75694437	75694571
chr11	76506592	76506671
chr11	76506682	76506756
chr11	85967394	85967572
chr11	102668076	102668224
chr11	102738617	102738819
chr11	104879508	104879708
chr11	107535728	107535942
chr11	108559738	108559811
chr11	113281481	113281664
chr11	118220498	118220650
chr11	118770623	118770935
chr11	118967828	118968044
chr11	119149273	119149349
chr11	119149363	119149450
chr11	124740025	124740205
chr11	125505300	125505375
chr11	125505380	125505450

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr11	126174088	126174223
chr12	347049	347194
chr12	416869	417012
chr12	2064569	2064742
chr12	2760834	2760969
chr12	5603663	5603978
chr12	6649623	6649789
chr12	6711518	6711694
chr12	6857978	6858125
chr12	7061134	7061344
chr12	11139347	11139449
chr12	11338777	11338922
chr12	18840997	18841186
chr12	20832966	20833150
chr12	21644445	21644588
chr12	25362768	25362865
chr12	25368353	25368527
chr12	25378518	25378628
chr12	25378668	25378736
chr12	25380198	25380313
chr12	25398223	25398302
chr12	29450098	29450189
chr12	40044008	40044190
chr12	40704243	40704386
chr12	46318589	46318702
chr12	46320714	46321133
chr12	49087351	49087425
chr12	49087441	49087522
chr12	50452481	50452628
chr12	52715022	52715133
chr12	53097017	53097165
chr12	54367312	54367453
chr12	56628907	56629135
chr12	57422517	57422685
chr12	57605672	57605825
chr12	57883012	57883155
chr12	57919227	57919364
chr12	58024947	58025169
chr12	65856902	65857128
chr12	66531861	66532002
chr12	68707389	68707570
chr12	70070674	70070884
chr12	72057264	72057366
chr12	72070609	72070811
chr12	72094580	72094790
chr12	88524011	88524224
chr12	88566366	88566533
chr12	98921641	98921743
chr12	101680085	101680228
chr12	102056154	102056335
chr12	109278836	109278980
chr12	110765355	110765526
chr12	111311635	111311768
chr12	111758240	111758518
chr12	120595667	120595742
chr12	120595747	120595823
chr12	121017083	121017236
chr12	122812608	122812761
chr12	123794223	123794440
chr12	130647654	130648042
chr12	132281651	132281727
chr12	132281736	132281809
chr12	133219967	133220177
chr13	26975697	26975800
chr13	27255177	27255424
chr13	31715236	31715403
chr13	32376346	32376420
chr13	32376436	32376506
chr13	43934053	43934183
chr13	47243133	47243319
chr13	50235103	50235242
chr13	51948333	51948473
chr13	73337566	73337767
chr13	79918781	79918949
chr13	95695906	95696015

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr13	100617825	100617972
chr13	103524540	103524678
chr14	20852522	20852698
chr14	21861622	21861784
chr14	21960977	21961059
chr14	21961067	21961151
chr14	23312897	23313105
chr14	23341457	23341530
chr14	23341537	23341607
chr14	23844982	23845130
chr14	23870012	23870093
chr14	24646867	24646971
chr14	24785047	24785428
chr14	31355130	31355318
chr14	35331341	35331496
chr14	35592666	35593413
chr14	39871577	39871748
chr14	45642233	45642446
chr14	45693573	45693743
chr14	51094809	51095027
chr14	53558479	53558682
chr14	60903490	60903622
chr14	74824296	74824501
chr14	75514311	75514613
chr14	75590681	75590844
chr14	76112766	76112869
chr14	86087899	86089516
chr14	89629085	89629160
chr14	94517522	94517597
chr14	94517602	94517684
chr14	94545622	94545831
chr14	100367231	100367417
chr14	101005191	101005268
chr14	101005281	101005356
chr14	103599674	103599890
chr14	103996500	103996652
chr14	105174175	105174328
chr14	105241955	105242184
chr15	26825438	26825614
chr15	28200261	28200406
chr15	28515831	28515913
chr15	29346298	29346373
chr15	31196823	31196955
chr15	33962654	33962797
chr15	34546639	34546790
chr15	37385752	37385967
chr15	40282438	40282588
chr15	40631678	40631858
chr15	42439798	42439975
chr15	43037988	43038350
chr15	43641083	43641289
chr15	45710731	45710900
chr15	59182460	59182534
chr15	64972923	64972996
chr15	64973003	64973077
chr15	65678245	65678396
chr15	68937449	68937585
chr15	72208682	72208855
chr15	74636115	74636294
chr15	75122475	75122560
chr15	75122565	75122641
chr15	91019861	91020076
chr15	91550156	91550229
chr16	91550236	91550314
chr16	677400	677619
chr16	709025	709186
chr16	1824226	1824384
chr16	3350396	3350543
chr16	4796966	4797064
chr16	4910616	4910696
chr16	4910701	4910772
chr16	15131846	15131917
chr16	15131936	15132007
chr16	17211687	17211868
chr16	19580722	19580792

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr16	19580802	19580945
chr16	20370627	20370695
chr16	20370707	20370781
chr16	24807159	24807296
chr16	27373792	27374008
chr16	28931082	28931293
chr16	29998212	29998288
chr16	29998307	29998488
chr16	29998507	29999198
chr16	30365477	30365649
chr16	30531147	30531223
chr16	30531252	30531336
chr16	30736272	30736399
chr16	30792987	30793144
chr16	57486680	57486865
chr16	61851439	61851577
chr16	66603816	66604018
chr16	67267764	67267913
chr16	67300029	67300163
chr16	67693569	67693726
chr16	67963814	67964016
chr16	68718427	68718563
chr16	69782986	69783060
chr16	70814658	70814883
chr16	77334134	77334346
chr16	77356204	77356390
chr16	84797828	84797902
chr16	89950943	89951021
chr16	89951033	89951110
chr17	4619726	4619873
chr17	4937191	4937382
chr17	6364651	6364792
chr17	7106484	7106663
chr17	7193514	7193591
chr17	7193609	7193677
chr17	7495794	7495934
chr17	7572889	7573037
chr17	7573894	7574051
chr17	7576519	7576721
chr17	7576819	7576963
chr17	7576984	7577173
chr17	7577469	7577648
chr17	7578154	7578326
chr17	7578339	7578589
chr17	7579289	7579605
chr17	7579639	7579772
chr17	7579804	7579954
chr17	7606634	7606779
chr17	7796749	7796891
chr17	7798684	7798832
chr17	7801789	7801860
chr17	7801864	7801941
chr17	7843379	7843604
chr17	8397029	8397104
chr17	8397109	8397213
chr17	8415824	8415899
chr17	11650845	11650990
chr17	11924178	11924359
chr17	11958173	11958241
chr17	11958273	11958347
chr17	11984748	11984821
chr17	11998858	11999038
chr17	12011073	12011147
chr17	12011188	12011258
chr17	12013633	12013769
chr17	12016558	12016637
chr17	12016638	12016723
chr17	12028578	12028721
chr17	12032423	12032495
chr17	12032528	12032635
chr17	12043108	12043240
chr17	12044473	12044617
chr17	17394625	17394706
chr17	17394715	17394798
chr17	18167789	18167891

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr17	19232843	19232984
chr17	21094248	21094404
chr17	26653686	26653791
chr17	27001306	27001491
chr17	27027146	27027294
chr17	27889751	27889904
chr17	32483150	32483364
chr17	33520305	33520387
chr17	33749415	33749560
chr17	34077072	34077226
chr17	37879768	37879938
chr17	37880143	37880277
chr17	37880943	37881204
chr17	37881268	37881488
chr17	37881538	37881678
chr17	37881938	37882156
chr17	37882788	37882929
chr17	38421138	38421328
chr17	39022843	39022987
chr17	39122838	39122981
chr17	40837221	40837371
chr17	42992558	42992767
chr17	45219538	45219608
chr17	46622058	46622130
chr17	46622138	46622213
chr17	48433885	48434029
chr17	49156920	49157096
chr17	55028036	55028155
chr17	56389889	56390072
chr17	56434824	56434979
chr17	56435049	56435504
chr17	56435824	56435979
chr17	57247089	57247259
chr17	59668355	59668485
chr17	61899060	61899235
chr17	64092655	64092757
chr17	65905685	65905757
chr17	65905765	65905840
chr17	67522678	67522861
chr17	71354193	71354380
chr17	72469628	72469709
chr17	72943133	72943370
chr17	73239108	73239255
chr17	73481928	73482105
chr17	73732098	73732255
chr17	74077933	74078144
chr17	76458968	76459179
chr17	78201588	78201804
chr18	909475	909606
chr18	5397994	5398169
chr18	5415759	5415908
chr18	5891919	5892065
chr18	8792971	8793145
chr18	10677709	10677884
chr18	13681675	13681816
chr18	21745000	21745137
chr18	21750255	21750437
chr18	43685102	43685207
chr18	48573394	48573701
chr18	48575029	48575253
chr18	48575604	48575742
chr18	48581129	48581385
chr18	48584464	48584644
chr18	48584679	48584858
chr18	48586189	48586326
chr18	48591844	48591994
chr18	48593364	48593571
chr18	48602979	48603166
chr18	48604619	48604762
chr18	55992200	55992275
chr18	55992300	55992401
chr18	74962849	74963010
chr19	1037565	1037640
chr19	1486925	1487090
chr19	4329936	4330073

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr19	6222276	6222408
chr19	6222426	6222560
chr19	6477176	6477334
chr19	7734186	7734359
chr19	10262043	10262260
chr19	11541689	11541867
chr19	11618749	11618898
chr19	12384459	12384532
chr19	12430219	12430289
chr19	12461699	12461804
chr19	14208179	14208311
chr19	14262044	14262183
chr19	16024574	16024676
chr19	16633919	16634058
chr19	17160629	17160770
chr19	17943378	17943537
chr19	18376879	18377011
chr19	18420549	18420684
chr19	18887964	18888074
chr19	33353341	33353512
chr19	33666341	33666487
chr19	34710251	34710329
chr19	35773442	35773597
chr19	36049982	36050123
chr19	36050702	36050840
chr19	36053372	36053556
chr19	36054257	36054469
chr19	36255957	36256095
chr19	36583567	36583752
chr19	38948102	38948293
chr19	38976382	38976566
chr19	39898864	39898940
chr19	39898954	39899025
chr19	39972489	39972708
chr19	40711839	40712026
chr19	41063094	41063312
chr19	42752790	42753141
chr19	42753150	42753358
chr19	42795740	42795897
chr19	42797165	42797402
chr19	43969655	43969759
chr19	44612180	44612341
chr19	45655690	45655763
chr19	45655775	45655861
chr19	47572330	47572409
chr19	47572410	47572489
chr19	47883075	47883154
chr19	47883165	47883241
chr19	47935470	47935583
chr19	47935610	47935683
chr19	49218040	49218183
chr19	49850420	49850644
chr19	49971680	49971830
chr19	49978970	49979074
chr19	50310400	50310477
chr19	50310490	50310569
chr19	51133019	51133307
chr19	51189464	51189643
chr19	54327333	54327466
chr19	54544183	54544376
chr19	54649623	54649801
chr19	54675668	54675742
chr19	54675753	54675826
chr19	55607388	55607564
chr19	55711543	55711691
chr19	55815008	55815221
chr19	56171859	56171994
chr19	58083470	58083773
chr19	58083780	58084301
chr19	58084310	58084414
chr19	58084455	58084590
chr19	58596555	58596694
chr20	4850517	4850733
chr20	5903287	5903676
chr20	7962897	7963043

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chr20	17462175	17462321
chr20	17581650	17581761
chr20	21492750	21492931
chr20	23065487	23066683
chr20	25422301	25422399
chr20	30064258	30064411
chr20	30232538	30232723
chr20	30309518	30309657
chr20	35060105	35060280
chr20	39990751	39990902
chr20	40033321	40033396
chr20	40033411	40033494
chr20	43933056	43933200
chr20	45874882	45875104
chr20	56099104	56099263
chr20	57478766	57478903
chr20	57480416	57480561
chr20	57484356	57484497
chr20	57484516	57484673
chr20	58452416	58452518
chr20	58452526	58452624
chr20	58466971	58467118
chr20	58490491	58490596
chr21	28296393	28296754
chr21	33043828	33043980
chr21	34799155	34799365
chr21	37510098	37510269
chr21	37617643	37617868
chr21	43221339	43221415
chr21	43221424	43221505
chr21	45402170	45402306
chr21	45472195	45472276
chr21	45472280	45472346
chr21	46276075	46276186
chr21	46276200	46276305
chr21	47532215	47532368
chr22	19373054	19373133
chr22	19373144	19373230
chr22	24583167	24583309
chr22	24717367	24717510
chr22	26906111	26906264
chr22	29913229	29913371
chr22	30742244	30742319
chr22	30742334	30742412
chr22	31011259	31011466
chr22	31535959	31536185
chr22	36689343	36689415
chr22	36689433	36689545
chr22	36696873	36696954
chr22	36696958	36697034
chr22	40816890	40817031
chr22	41257085	41257858
chr22	41753325	41753473
chr22	42271561	42271704
chr22	43213706	43213880
chr22	43218251	43218423
chr22	44083341	44083474
chr22	44559691	44559835
chr22	46327164	46327304
chr22	49042379	49042586
chr22	50506837	50507013
chrX	16850722	16850828
chrX	32361221	32361427
chrX	35989786	35989933
chrX	37312526	37312680
chrX	47058852	47059024
chrX	54011325	54011408
chrX	54011415	54011492
chrX	54578240	54578439
chrX	69478691	69478889
chrX	70367786	70367931
chrX	100880079	100880240
chrX	102004344	102004417
chrX	111090390	111090535
chrX	112048150	112048367

TABLE 16-continued

Chromosome	Start (bp)	End (bp)
chrX	119072675	119072814
chrX	119694035	119694119
chrX	119694125	119694200
chrX	130409129	130409260
chrX	132458481	132458583
chrX	135313677	135314212
chrX	135584907	135585121
chrX	142605167	142605243
chrX	142803661	142803788
chrX	149937455	149937613
chrX	149984430	149984502
chrX	149984530	149984601
chrX	153627898	153627975

TABLE 17

Chromosome	Start (bp)	End (bp)	Gene
chr17	47696342	47696470	SPOP
chr20	29628226	29628331	FRG1B
chr9	20414279	20414380	MLLT3
chr1	145367713	145367822	NBPF10
chr20	29625872	29625984	FRG1B
chr1	145302664	145302763	NBPF10
chr10	47207779	47207878	AGAP9
chr3	41266067	41266166	CTNNB1
chr17	7577498	7577608	TP53
chr17	47696595	47696747	SPOP
chr19	12187274	12187476	ZNF844
chr12	132547064	132547163	EP400
chr20	33345694	33345793	NCOA6
chr2	207025311	207025434	EEF1B2
chr7	57187717	57187816	ZNF479
chr6	45390414	45390513	RUNX2
chr1	74575077	74575237	LRRIQ3
chr1	145296361	145296460	NBPF10
chr17	7578403	7578527	TP53
chr14	71275725	71275824	MAP3K9
chr7	131241018	131241118	PODXL
chr9	127790664	127790763	SCAI
chr22	43213734	43213863	ARFGAP3
chr10	89692848	89692969	PTEN
chr6	29760313	29760412	LOC554223
chr2	242794863	242794962	PDCD1
chr2	209113063	209113162	IDH1
chr6	170871016	170871115	TBP
chr11	47788617	47788716	FNBP4
chr22	29091697	29091861	CHEK2
chr11	61161352	61161451	TMEM216
chr2	107049581	107049680	RGPD3
chr20	46279761	46279860	NCOA3
chr4	147560412	147560511	POU4F2
chr22	42538728	42538881	CYP2D7P1
chr3	137880694	137880793	DBR1
chr5	156378525	156378748	TIMD4
chrX	128599495	128599699	SMARCA1
chr19	12501392	12501650	ZNF799
chr10	52595832	52596044	A1CF
chr7	127235458	127235735	FSCN3
chr8	23538908	23539136	NKX3-1
chr7	154667616	154667768	DPP6
chr11	117863955	117864125	IL10RA
chr1	1850612	1850711	TMEM52
chrX	70349165	70349279	MED12
chr19	13423482	13423595	CACNA1A
chr7	114269946	114270045	FOXP2
chr19	53958800	53958899	ZNF761
chr19	53116795	53116894	ZNF83
chr18	51795916	51796031	POLI
chr17	40847582	40847681	CNTNAP1
chr7	75130882	75130988	SPDYE5
chr11	102738638	102738799	MMP12

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	118790284	118790444	CEP85L
chr17	42293021	42293177	UBTF
chr11	106558306	106558448	GUCY1A2
chr12	122812650	122812749	CLIP1
chr17	8138395	8138519	CTC1
chr3	12046075	12046174	SYN2
chr12	124171409	124171508	TCTN2
chr5	153144021	153144160	GRIA1
chr2	233620932	233621044	GIGYF2
chr4	62813829	62813928	LPHN3
chr9	136340557	136340656	SLC2A6
chr6	100390835	100390959	MCHR2
chr10	89720772	89720871	PTEN
chr1	12887549	12887687	PRAMEF11
chr1	152975658	152975782	SPRR3
chrX	151869696	151869995	MAGEA6
chr6	26216685	26216865	HIST1H2BG
chr1	186275975	186276982	PRG4
chr6	54735044	54735358	FAM83B
chr10	128973672	128973909	FAM196A
chr19	7810516	7810836	CD209
chr5	179264556	179264808	C5orf45
chr1	75038566	75038907	C1orf173
chr1	237753954	237754211	RYR2
chr12	12870802	12871146	CDKN1B
chr14	38060571	38061706	FOXA1
chrX	144337190	144337334	SPANXN1
chr5	121356087	121356220	SRFBP1
chr12	8290735	8290883	CLEC4A
chr19	8398906	8399005	KANK3
chr8	135612714	135612813	ZFAT
chr7	31862739	31862845	PDE1C
chr20	25436308	25436422	NINL
chr8	89128848	89128947	MMP16
chr5	52779941	52780041	FST
chr8	69002813	69002950	PREX2
chr14	24530710	24530809	LRRC16B
chr3	124132292	124132391	KALRN
chr16	4910801	4910929	UBN1
chr10	124273706	124273875	HTRA1
chr1	158057795	158057944	KIRREL
chr4	79792109	79792208	BMP2K
chr15	79051762	79051920	ADAMTS7
chr15	58306055	58306196	ALDH1A2
chr12	10233905	10234004	CLEC1A
chr16	15045627	15045732	NPIP
chr4	68719785	68719904	TMPRSS11D
chr1	26608782	26608881	UBXN11
chr12	49434173	49434308	MLL2
chr5	149776098	149776197	TCOF1
chr1	44450608	44450707	B4GALT2
chr15	65041611	65041710	RBPMS2
chr6	28227391	28227528	NKAPL
chr1	240255520	240255619	FMN2
chr15	51507266	51507429	CYP19A1
chr7	82784421	82784520	PCLO
chr11	108205695	108205836	ATM
chr6	27114416	27114574	HIST1H2BK
chr3	44672553	44672713	ZNF197
chr9	79634891	79635215	FOXB2
chr7	82763828	82764232	PCLO
chr19	2917732	2917874	ZNF57
chr19	49926468	49926596	PTH2
chr1	147380241	147380400	GJA8
chr19	58384330	58386285	ZNF814
chr11	120008137	120008335	TRIM29
chr7	100349723	100350362	ZAN
chr1	214556763	214557052	PTPN14
chr8	100866041	100866334	VPS13B
chr5	427983	428082	AHRR
chr6	26027263	26027362	HIST1H4B
chr1	204409319	204409449	PIK3C2B
chr20	1433675	1433785	NSFL1C
chr9	73152075	73152174	TRPM3
chr11	117374641	117374740	DSCAML1

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	7788130	7788229	CHD3
chr19	51217054	51217206	SHANK1
chr6	397107	397252	IRF4
chr14	105996001	105996100	TMEM121
chr1	149783601	149783719	HIST2H2BF
chr16	55362954	55363123	IRX6
chr17	16843653	16843775	TNFRSF13B
chr13	19748101	19748254	TUBA3C
chr3	10976714	10976885	SLC6A11
chr9	12775812	12775911	LURAP1L
chr19	54646838	54646937	CNOT3
chr1	91967273	91967388	CDC7
chr3	178935997	178936122	PIK3CA
chr1	160654784	160654893	CD48
chr18	8609822	8609921	RAB12
chr6	28554156	28554275	SCAND3
chr3	123419028	123419195	MYLK
chrX	50350643	50350742	SHROOM4
chr17	7578176	7578289	TP53
chr10	89717609	89717776	PTEN
chr7	101988883	101989029	SPDYE6
chr12	52680984	52681083	KRT81
chr19	17943403	17943502	JAK3
chr20	57415296	57415471	GNAS
chr12	49432197	49432375	MLL2
chr1	156565213	156565531	GPATCH4
chr3	49395481	49395680	GPX1
chr15	90320120	90320477	MESP2
chr12	49391304	49391666	DDN
chr1	149857820	149858039	HIST2H2BE
chr1	152636614	152636836	LCE2D
chr1	111216033	111217283	KCNA3
chr17	18022688	18023608	MYO15A
chr14	69256454	69256864	ZFP36L1
chr12	11546149	11546791	PRB2
chr2	238274413	238274669	COL6A3
chr11	128781512	128781969	KCNJ5
chr17	16593776	16594037	CCDC144A
chr5	114466298	114466560	TRIM36
chr19	45911703	45911968	CD3EAP
chr7	5352527	5352794	TNRC18
chr7	92146658	92147137	PEX1
chrX	37026829	37028758	FAM47C
chr4	1389053	1389324	CRIPAK
chr4	81123290	81123568	PRDM8
chr8	139890022	139890302	COL22A1
chr11	124857494	124857795	CCDC15
chr12	66725027	66725339	HELB
chr1	12921089	12921406	PRAMEF2
chr22	50658758	50659326	TUBGCP6
chr20	39990959	39991281	EMILIN3
chr17	18874709	18875038	FAM83G
chr1	152127347	152129152	RPTN
chr3	128181413	128182005	DNAJB8
chr6	87725480	87726075	HTR1E
chr6	51612954	51613290	PKHD1
chr19	57065051	57065990	ZFP28
chr7	26224956	26225295	NFE2L3
chr19	21239843	21240183	ZNF430
chr19	44103049	44103390	ZNF576
chr5	45262466	45262808	HCN1
chr11	57077241	57077613	TNKS1BP1
chr9	111617061	111618108	ACTL7B
chr4	94750558	94750938	ATOH1
chr7	94293245	94293632	PEG10
chr7	94293245	94293632	PEG10
chr9	135073410	135073797	NTNG2
chr19	53855195	53856762	ZNF845
chr20	49620846	49620945	KCNG1
chr4	126242026	126242125	FAT4
chr21	15011835	15011959	POTED
chr17	48264041	48264185	COL1A1
chr7	43659195	43659322	STK17A
chr22	29885743	29885869	NEFH
chr12	94763686	94763812	CCDC41

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr3	126730814	126730913	PLXNA1
chr8	89058896	89059012	MMP16
chr2	207425844	207425943	ADAM23
chr15	75013006	75013105	CYP1A1
chr13	115047495	115047602	UPF3A
chr12	6494392	6494545	LTBR
chr6	135787137	135787236	AHI1
chr17	39673059	39673216	KRT15
chr1	150297400	150297545	PRPF3
chr3	38139234	38139393	DLEC1
chr4	162680562	162680683	FSTL5
chr16	2903127	2903248	PRSS22
chr3	186362539	186362709	FETUB
chr12	57674140	57674313	R3HDM2
chr9	43818040	43818184	CNTNAP3B
chr9	43915811	43915910	CNTNAP3B
chr2	141473541	141473671	LRP1B
chr17	12032455	12032604	MAP2K4
chr18	5397338	5397437	EPB41L3
chr3	93646093	93646251	PROS1
chr22	26884093	26884192	SRRD
chr19	367064	367224	THEG
chr1	224621717	224621816	WDR26
chr1	109823577	109823676	PSRC1
chr1	109824257	109824424	PSRC1
chr17	80401692	80401838	C17orf62
chr7	31126537	31126636	ADCYAP1R1
chr3	125824596	125824695	ALDH1L1
chrX	5821234	5821333	NLGN4X
chr14	55818553	55818655	FBXO34
chr12	62777620	62777779	USP15
chr1	207818577	207818676	CR1L
chr1	207870839	207870938	CR1L
chr14	24619759	24619858	RNF31
chr1	157103871	157104019	ETV3
chr12	57566909	57567008	LRP1
chr5	14711305	14711419	ANKH
chr19	47935632	47935731	SLC8A2
chr16	31427825	31427967	ITGAD
chr4	89618409	89618508	NAP1L5
chr1	55280586	55280712	C1orf177
chr1	11090220	11090319	MASP2
chrX	69261657	69261812	AWAT2
chr12	104487249	104487348	HCFC2
chr12	64491020	64491155	SRGAP1
chr6	152647463	152647581	SYNE1
chr17	34079527	34079626	GAS2L2
chr3	122002719	122002818	CASR
chr21	45953556	45953655	TSPEAR
chr2	220396479	220396624	ASIC4
chr9	73235184	73235283	TRPM3
chr2	125521553	125521724	CNTNAP5
chrX	148798035	148798187	MAGEA11
chr19	4538279	4538378	LRG1
chr1	158063128	158063236	KIRREL
chr12	75601571	75601722	KCNC2
chr10	133949433	133949580	JAKMIP3
chr5	80548514	80548613	CKMT2
chr3	38755448	38755571	SCN10A
chr4	20255544	20255643	SLIT2
chr17	65688748	65688847	PITPNC1
chr12	104107420	104107546	STAB2
chr17	9739687	9739792	GLP2R
chr3	127399112	127399211	ABTB1
chr22	24621212	24621382	GGT5
chr1	12979714	12979813	PRAMEF8
chr1	27023450	27023622	ARID1A
chr1	27057926	27058086	ARID1A
chr10	15145322	15145421	RPP38
chr2	96795562	96795717	ASTL
chr5	88027590	88027689	MEF2C
chr4	26622222	26622321	TBC1D19
chr10	44104061	44104188	ZNF485
chr15	89760369	89760468	RLBP1
chr9	113547878	113547977	MUSK

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr12	117768454	117768562	NOS1
chr8	103564110	103564209	ODF1
chr3	180359821	180359981	CCDC39
chr20	47845251	47845425	DDX27
chr16	67876758	67876857	THAP11
chr1	181701977	181702076	CACNA1E
chr1	181708282	181708389	CACNA1E
chr1	181767489	181767588	CACNA1E
chr19	36046577	36046714	ATP4A
chr6	89974139	89974238	GABRR2
chr9	33135275	33135376	B4GALT1
chr18	47808962	47809109	CXXC1
chr1	63999765	63999887	EFCAB7
chr3	52412605	52412704	DNAH1
chr4	177058666	177058765	WDR17
chr22	29628244	29628391	EMID1
chr11	58391812	58391922	CNTF
chr17	7106228	7106327	DLG4
chr1	186915768	186915906	PLA2G4A
chr6	29910643	29910742	HLA-A
chr17	40474417	40474516	STAT3
chr11	68822681	68822780	TPCN2
chr20	33875187	33875286	FAM83C
chr1	24080589	24080745	TCEB3
chr19	10602357	10602518	KEAP1
chr12	121471395	121471535	OASL
chr2	210993754	210993896	KANSL1L
chr17	7680785	7680923	DNAH2
chrX	70472885	70472984	ZMYM3
chr1	156937778	156937916	ARHGEF11
chr12	121434064	121434216	HNF1A
chr5	131544933	131545032	P4HA2
chr3	112546244	112546343	CD200R1L
chr10	26454995	26455107	MYO3A
chr2	224849584	224849683	SERPINE2
chr12	113748038	113748160	SLC24A6
chr14	21960902	21961063	TOX4
chr2	39053657	39053831	DHX57
chr1	20005665	20005831	HTR6
chr10	60994116	60994215	PHYHIPL
chr12	10223942	10224041	CLEC1A
chr1	10713917	10714016	CASZ1
chr2	120438864	120438963	TMEM177
chr1	10207020	10207148	UBE4B
chr1	157558964	157559063	FCRL4
chr1	9662257	9662356	TMEM201
chr2	67631942	67632041	ETAA1
chr10	105798163	105798286	COL17A1
chr19	19136340	19136439	SUGP2
chr19	20807998	20808097	ZNF626
chr22	24982242	24982341	FAM211B
chr9	78686641	78686814	PCSK5
chr8	104340495	104340644	FZD6
chr1	159033262	159033361	AIM2
chr14	61115589	61115688	SIX1
chr1	76397666	76397765	ASB17
chr5	55471978	55472109	ANKRD55
chr21	28327057	28327190	ADAMTS5
chr20	60448783	60448956	CDH4
chr3	154042007	154042106	DHX36
chr1	148594406	148594508	NBPF15
chr20	44004087	44004186	TP53TG5
chr12	31237902	31238060	DDX11
chrX	96139971	96140070	RPA4
chr22	47095209	47095362	CERK
chr11	134253690	134253789	B3GAT1
chr17	80121075	80121233	CCDC57
chr5	180651193	180651292	TRIM41
chr4	69870619	69870718	UGT2B10
chr1	12726521	12726654	AADACL4
chr4	184567615	184567714	RWDD4
chr5	37038703	37038840	NIPBL
chr19	40367791	40367890	FCGBP
chr1	159683404	159683568	CRP
chr20	61525185	61525284	DIDO1

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	10216499	10216635	MYH13
chrX	70348447	70348568	MED12
chr17	10404708	10404807	MYH1
chr2	238737977	238738076	RBM44
chr19	52537921	52538072	ZNF432
chr8	109001323	109001422	RSPO2
chrX	83411082	83411199	RPS6KA6
chr22	19197945	19198044	CLTCL1
chr1	153084995	153085094	SPRR2F
chr18	55992227	55992394	NEDD4L
chr19	24310250	24310349	ZNF254
chr17	40328126	40328225	KCNH4
chr19	8188347	8188473	FBN3
chr4	73951018	73951117	ANKRD17
chr6	146480622	146480721	GRM1
chr6	146755621	146755764	GRM1
chr17	56688554	56688675	TEX14
chr1	32050486	32050637	TINAGL1
chr16	20335227	20335326	GP2
chr12	46246099	46246242	ARID2
chr17	12647643	12647742	MYOCD
chr8	41456658	41456823	AGPAT6
chr1	47904197	47904296	FOXD2
chr19	10088269	10088377	COL5A3
chr6	33137159	33137267	COL11A2
chr1	220253068	220253188	BPNT1
chr1	75072284	75072383	C1orf173
chr2	130832499	130832651	POTEF
chr2	119915150	119915249	C1QL2
chr9	116858330	116858487	KIF12
chr4	190873316	190873442	FRG1
chr4	190878552	190878657	FRG1
chr12	667673	667827	B4GALNT3
chr2	176981838	176981937	HOXD10
chr2	96592948	96593047	ANKRD36C
chr12	52863561	52863660	KRT6C
chr20	1902225	1902324	SIRPA
chr6	132910415	132910514	TAAR5
chr16	314894	314993	ITFG3
chr8	25745357	25745488	EBF2
chr1	158585064	158585164	SPTA1
chr1	158615284	158615384	SPTA1
chr1	158637744	158637843	SPTA1
chr1	158639487	158639586	SPTA1
chr6	85446535	85446704	TBX18
chr19	48620894	48621038	LIG1
chr1	108023224	108023323	NTNG1
chr16	89703633	89703763	DPEP1
chr19	3589440	3589553	GIPC3
chr4	73186430	73186587	ADAMTS3
chr11	803338	803451	PIDD
chr1	183515217	183515316	SMG7
chr19	46020920	46021092	VASP
chr4	189060920	189061024	TRIML1
chr18	72228090	72228253	CNDP1
chr10	37430714	37430813	ANKRD30A
chr10	37508446	37508620	ANKRD30A
chr12	49416383	49416482	MLL2
chr4	103528306	103528476	NFKB1
chr7	20824878	20824977	SP8
chr7	130357573	130357716	TSGA13
chr22	19119463	19119562	TSSK2
chr11	864424	864523	TSPAN4
chr20	3802816	3802915	AP5S1
chr11	58892327	58892426	FAM111B
chr4	3443728	3443845	HGFAC
chr1	156877401	156877522	PEAR1
chr16	19041545	19041691	TMC7
chr16	19058398	19058497	TMC7
chr19	40580502	40580601	ZNF780A
chr6	34985544	34985689	ANKS1A
chr15	65502013	65502112	CILP
chr13	113825968	113826067	PROZ
chrX	150156249	150156387	HMGB3
chr17	1198784	1198917	TUSC5

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr8	134488065	134488182	ST3GAL1
chr17	72736906	72737017	RAB37
chr9	139847343	139847480	LCN12
chr2	55461966	55462098	RPS27A
chr8	139277945	139278085	FAM135B
chr1	207112652	207112751	PIGR
chr2	206562233	206562332	NRP2
chr18	25565571	25565670	CDH2
chr18	25572607	25572706	CDH2
chr12	76740913	76741012	BBS10
chr18	29099765	29099900	DSG2
chr10	75104806	75104905	TTC18
chr5	179564960	179565059	RASGEF1C
chr5	72419528	72419627	TMEM171
chr7	138764409	138764547	ZC3HAV1
chr17	37369247	37369385	STAC2
chr9	139277946	139278045	SNAPC4
chr3	20219751	20219850	SGOL1
chr19	58352925	58353024	ZNF587B
chr10	50835685	50835784	CHAT
chr18	21662876	21663045	TTC39C
chr6	123319045	123319144	CLVS2
chr6	123369766	123369877	CLVS2
chr16	67199621	67199749	HSF4
chr2	100055091	100055190	REV1
chr15	66853341	66853440	LCTL
chr14	24040358	24040526	JPH4
chr11	47198066	47198185	ARFGAP2
chr16	72162964	72163132	PMFBP1
chr8	2830669	2830768	CSMD1
chr8	3000013	3000112	CSMD1
chr15	43714099	43714238	TP53BP1
chr2	128046320	128046419	ERCC3
chr2	128050179	128050278	ERCC3
chr2	165365251	165365362	GRB14
chr1	12331050	12331181	VPS13D
chr6	31930215	31930362	SKIV2L
chr14	23651972	23652123	SLC7A8
chr16	29820927	29821026	MAZ
chr12	121176620	121176719	ACADS
chr1	210796869	210797014	HHAT
chr2	137814398	137814556	THSD7B
chr7	149518485	149518638	SSPO
chr20	61288117	61288287	SLCO4A1
chr7	4824538	4824679	AP5Z1
chr1	165370499	165370647	RXRG
chr1	5964710	5964829	NPHP4
chr16	89799877	89799976	ZNF276
chrX	148037315	148037414	AFF2
chr11	26702601	26702768	SLC5A12
chr15	101528887	101528986	LRRK1
chr2	88425694	88425867	FABP1
chr10	97192204	97192333	SORBS1
chr11	47744540	47744639	FNBP4
chr13	22246171	22246273	FGF9
chr8	52359563	52359722	PXDNL
chr4	110972750	110972849	ELOVL6
chr12	63543656	63543829	AVPR1A
chr10	7780583	7780721	ITIH2
chr16	67913752	67913874	EDC4
chr8	142222364	142222493	SLC45A4
chr5	156592677	156592776	FAM71B
chr9	71628893	71628992	PRKACG
chr19	4682830	4682929	LOC100131094
chr11	92568127	92568226	FAT3
chr11	92577096	92577195	FAT3
chr2	145156348	145156447	ZEB2
chr12	108603924	108604023	WSCD2
chr11	4130853	4130952	RRM1
chr1	94487401	94487507	ABCA4
chr17	74732289	74732457	SRSF2
chr15	52433336	52433469	GNB5
chr19	44099348	44099447	IRGQ
chr19	57089368	57089491	ZNF470
chr11	122774655	122774754	C11orf63

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr22	16287547	16287674	POTEH
chr10	73461839	73461938	CDH23
chr10	73574933	73575032	CDH23
chr14	72976861	72976987	RGS6
chr14	93708963	93709062	BTBD7
chr6	132195392	132195491	ENPP1
chr14	103442219	103442386	CDC42BPB
chr12	52200654	52200822	SCN8A
chr11	116718191	116718324	SIK3
chr16	85936621	85936795	IRF8
chr15	49036436	49036540	CEP152
chr22	40816858	40816957	MKL1
chr2	80874757	80874856	CTNNA2
chr4	9828029	9828128	SLC2A9
chr4	9982215	9982361	SLC2A9
chr4	141543403	141543502	TBC1D9
chrX	125685873	125685972	DCAF12L1
chr20	3214796	3214895	SLC4A11
chr3	26751542	26751672	LRRC3B
chr10	88277653	88277752	WAPAL
chr4	158253970	158254138	GRIA2
chr12	4479821	4479920	FGF23
chr10	78647069	78647237	KCNMA1
chr19	19729319	19729438	PBX4
chr1	46751084	46751183	LRRC41
chr22	37263419	37263518	NCF4
chr17	19235212	19235311	EPN2
chr15	75941774	75941873	SNX33
chr9	99521361	99521460	ZNF510
chr12	58014047	58014190	SLC26A10
chr9	77448913	77449038	TRPM6
chr5	130840318	130840471	RAPGEF6
chr17	5009529	5009628	ZNF232
chr15	48431290	48431389	SLC24A5
chr1	70687321	70687420	SRSF11
chr3	44700524	44700679	ZNF35
chr11	60785253	60785404	CD6
chr1	97564044	97564188	DPYD
chr9	94172204	94172303	NFIL3
chr16	74490546	74490653	GLG1
chr3	111426771	111426941	PLCXD2
chr11	30034012	30034111	KCNA4
chr1	156779027	156779126	SH2D2A
chr5	175110214	175110313	HRH2
chr6	168430261	168430360	KIF25
chr1	43296636	43296735	ERMAP
chr6	134212845	134212944	TCF21
chr20	42747144	42747263	JPH2
chr20	42788362	42788461	JPH2
chr2	219735781	219735880	WNT6
chr2	171862656	171862792	TLK1
chr2	45233390	45233550	SIX2
chr8	113308061	113308235	CSMD3
chr16	28883169	28883268	SH2B1
chr13	21417908	21418025	XPO4
chr19	14877046	14877193	EMR2
chr1	35915958	35916110	KIAA0319L
chr1	215848612	215848711	USH2A
chr16	29996643	29996742	TAOK2
chrX	151138728	151138827	GABRE
chr20	3686553	3686652	SIGLEC1
chr20	46301026	46301125	SULF2
chr19	15794434	15794533	CYP4F12
chr19	23927261	23927360	ZNF681
chrX	29972638	29972809	IL1RAPL1
chr1	55521742	55521841	PCSK9
chr19	18976420	18976575	UPF1
chr1	160466058	160466157	SLAMF6
chr6	6224961	6225060	F13A1
chr1	152800111	152800256	LCE1A
chr7	33312673	33312772	BBS9
chr6	42897308	42897459	CNPY3
chr5	139939910	139940032	APBB3
chr11	43345020	43345119	API5
chr18	14513660	14513784	POTEC

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr22	23523994	23524103	BCR
chr2	129026304	129026420	HS6ST1
chr17	1559941	1560055	PRPF8
chr6	155743827	155743990	NOX3
chr1	44595137	44595236	KLF17
chr3	111795744	111795843	TMPRSS7
chr3	137483866	137483993	SOX14
chr4	42895344	42895493	GRXCR1
chr4	170613374	170613473	CLCN3
chr10	113920449	113920548	GPAM
chr7	54617680	54617802	VSTM2A
chr3	168819847	168820014	MECOM
chr17	78187970	78188127	SGSH
chr19	56459465	56459564	NLRP8
chr19	56466870	56466969	NLRP8
chr3	151148077	151148176	MED12L
chr6	36651924	36652023	CDKN1A
chr1	54359977	54360076	DIO1
chr1	109803674	109803773	CELSR2
chr16	2506573	2506672	CCNF
chr8	133196487	133196614	KCNQ3
chr17	10350361	10350462	MYH4
chr1	177030339	177030438	ASTN1
chr7	75053808	75053911	POM121C
chr12	11461577	11461676	PRB4
chr1	16534594	16534703	ARHGEF19
chr9	5763493	5763655	KIAA1432
chr1	64608117	64608216	ROR1
chr9	18504826	18504954	ADAMTSL1
chr7	140453074	140453193	BRAF
chr7	140481375	140481493	BRAF
chr11	65325188	65325329	LTBP3
chr22	23959753	23959852	C22orf43
chr7	48266858	48267022	ABCA13
chr7	48314998	48315097	ABCA13
chr7	48506547	48506646	ABCA13
chr1	152748923	152749022	LCE1F
chr3	49136770	49136869	QARS
chr17	4805226	4805382	CHRNE
chr5	14504502	14504650	TRIO
chr1	17599834	17599942	PADI3
chr10	73050799	73050898	UNC5B
chr18	3215086	3215185	MYOM1
chr6	136913310	136913479	MAP3K5
chr4	9783978	9784077	DRD5
chr19	7670147	7670246	CAMSAP3
chr3	17413566	17413739	TBC1D5
chr3	122459485	122459584	HSPBAP1
chr10	25160909	25161008	PRTFDC1
chr19	36278554	36278653	ARHGAP33
chr12	7080183	7080282	EMG1
chr22	37480766	37480879	TMPRSS6
chr4	46067416	46067561	GABRG1
chr8	101612573	101612672	SNX31
chr8	102570741	102570840	GRHL2
chr13	95862940	95863039	ABCC4
chr11	47360070	47360230	MYBPC3
chr1	114442764	114442863	AP4B1
chr15	78290585	78290684	TBC1D2B
chr15	78290571	78290670	TBC1D2B
chr14	94004483	94004582	UNC79
chr8	20107495	20107642	LZTS1
chr1	74649255	74649354	LRRIQ3
chr17	41957219	41957318	MPP2
chr1	43783549	43783648	TIE1
chr6	26271419	26271518	HIST1H3G
chr21	30339198	30339297	LTN1
chr19	58132377	58132476	ZNF134
chr2	218683367	218683466	TNS1
chr5	125919614	125919713	ALDH7A1
chr5	159680529	159680628	CCNJL
chr18	48591845	48591944	SMAD4
chr11	3380509	3380679	ZNF195
chr8	131861854	131861953	ADCY8
chr19	58565016	58565115	ZSCAN1

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr1	29475170	29475269	SRSF4
chr2	219353026	219353144	USP37
chr19	803548	803647	PTBP1
chr1	35578983	35579082	ZMYM1
chr1	89206759	89206858	PKN2
chr17	61557709	61557836	ACE
chr16	85743769	85743913	C16orf74
chr19	39995871	39995970	DLL3
chr3	112710021	112710120	GTPBP8
chr1	57173292	57173391	PRKAA2
chr6	34824015	34824186	UHRF1BP1
chr8	145109714	145109816	OPLAH
chr1	231299623	231299722	TRIM67
chr1	155887289	155887463	KIAA0907
chr16	3658438	3658537	SLX4
chr9	139091592	139091726	LHX3
chr3	178916856	178916955	PIK3CA
chr3	178921503	178921602	PIK3CA
chr3	178952018	178952117	PIK3CA
chr1	145327492	145327665	NBPF10
chr1	145359014	145359187	NBPF10
chr1	145368502	145368605	NBPF10
chr4	3234944	3235043	HTT
chr11	20622882	20622995	SLC6A5
chr7	3990554	3990666	SDK1
chr22	21354935	21355076	THAP7
chr15	53992045	53992144	WDR72
chr9	125895123	125895247	STRBP
chr11	6261369	6261468	CNGA4
chr8	41529856	41529955	ANK1
chr17	72915872	72915971	USH1G
chr 19	40331053	40331155	FBL
chr7	95705368	95705509	DYNC1I1
chr22	44083350	44083461	EFCAB6
chr10	21962565	21962664	MLLT10
chr12	15742388	15742524	PTPRO
chr1	44680395	44680510	DMAP1
chr8	25174522	25174647	DOCK5
chr6	17507399	17507543	CAP2
chr18	65179095	65179194	DSEL
chr1	158325230	158325329	CD1E
chr2	242039084	242039187	MTERFD2
chr1	6228209	6228337	CHD5
chr7	139285208	139285307	HIPK2
chr10	81072398	81072506	ZMIZ1
chr9	137657503	137657602	COL5A1
chr7	141672536	141672670	TAS2R38
chr9	72755060	72755204	MAMDC2
chr7	148701023	148701136	PDIA4
chr5	149460466	149460565	CSF1R
chr1	233190095	233190194	PCNXL2
chr17	42390778	42390877	RUNDC3A
chr16	66436886	66436985	CDH5
chr1	151340674	151340795	SELENBP1
chr16	55844428	55844576	CES1
chr16	70287821	70287941	AARS
chr20	42265784	42265893	IFT52
chr17	42333040	42333214	SLC4A1
chr19	22836719	22836818	ZNF492
chr21	18924171	18924270	CXADR
chr12	58022501	58022637	B4GALNT1
chr12	58024982	58025147	B4GALNT1
chr17	79428858	79428957	BAHCC1
chr16	5038149	5038281	SEC14L5
chr17	17250118	17250261	NT5M
chr6	33177763	33177862	RING1
chr16	7568192	7568291	RBFOX1
chr16	7645547	7645646	RBFOX1
chr4	39230176	39230275	WDR19
chr14	19553528	19553627	POTEG
chr7	141952291	141952431	PRSS58
chr19	49646062	49646161	PPFIA3
chr17	3917383	3917482	ZZEF1
chr4	76813003	76813131	PPEF2
chr4	1805418	1805563	FGFR3

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr7	151845676	151845775	MLL3
chr11	132177582	132177717	NTM
chr1	112318708	112318871	KCND3
chr9	35906509	35906608	HRCT1
chr2	215645754	215645853	BARD1
chr3	55508371	55508481	WNT5A
chr10	26575273	26575423	GAD2
chr11	77924691	77924859	USP35
chr8	48771409	48771547	PRKDC
chr10	55587148	55587308	PCDH15
chr11	111591255	111591354	SIK2
chr3	140406822	140406921	TRIM42
chr10	43288403	43288533	BMS1
chr4	1643035	1643134	FAM53A
chr3	55733405	55733540	ERC2
chr1	94054820	94054919	BCAR3
chr22	42264673	42264772	SREBF2
chr3	27763357	27763456	EOMES
chr19	44564607	44564734	ZNF223
chr12	53238342	53238507	KRT78
chr4	62903448	62903574	LPHN3
chr12	113321097	113321207	RPH3A
chr19	43439763	43439888	PSG7
chr7	8125954	8126053	GLCCI1
chr7	99689042	99689143	COPS6
chr7	22985617	22985716	FAM126A
chr17	7340254	7340353	TMEM102
chr7	39610104	39610226	YAE1D1
chr10	60148429	60148579	TFAM
chr14	20846219	20846389	TEP1
chr20	9364886	9364985	PLCB4
chr7	129664105	129664204	ZC3HC1
chr1	34102076	34102175	CSMD2
chr10	51754122	51754221	AGAP6
chr10	106976741	106976840	SORCS3
chr16	630857	630972	PIGQ
chr4	13610153	13610292	BOD1L1
chr18	54591200	54591299	WDR7
chr12	86373503	86373602	MGAT4C
chr12	54903631	54903769	NCKAP1L
chr2	229890606	229890760	PID1
chr8	66753605	66753743	PDE7A
chr2	158272195	158272363	CYTIP
chr11	118986781	118986880	C2CD2L
chr1	35370016	35370115	DLGAP3
chr7	76029671	76029806	SRCRB4D
chr11	8737282	8737381	ST5
chr19	17394176	17394285	ANKLE1
chr11	108788586	108788685	DDX10
chr10	103906597	103906696	PPRC1
chr3	51864402	51864501	IQCF3
chr15	101425471	101425576	ALDH1A3
chr17	7576839	7576938	TP53
chr17	7577018	7577155	TP53
chr17	2290808	2290907	MNT
chr17	38906741	38906840	KRT25
chr1	24019098	24019249	RPL11
chr20	34312491	34312644	RBM39
chr2	99182102	99182229	INPP4A
chr14	94395228	94395393	FAM181A
chr19	19337518	19337626	NCAN
chr22	36876686	36876785	TXN2
chr2	169417701	169417832	CERS6
chr13	47409097	47409211	HTR2A
chr22	37962525	37962624	CDC42EP1
chr1	38197082	38197254	EPHA10
chr7	796442	796544	HEATR2
chr1	147230938	147231037	GJA5
chr17	27380499	27380598	PIPOX
chr17	58700833	58700932	PPM1D
chr3	11744426	11744525	VGLL4
chr16	17221521	17221620	XYLT1
chr11	14808043	14808142	PDE3B
chr4	155241970	155242139	DCHS2
chr4	155298431	155298530	DCHS2

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr4	88047243	88047342	AFF1
chr4	156643189	156643348	GUCY1A3
chr3	47125462	47125561	SETD2
chrX	140969360	140969485	MAGEC3
chr17	27999050	27999149	SSH2
chr5	141304974	141305092	KIAA0141
chr6	35927487	35927586	SLC26A8
chr9	101748265	101748364	COL15A1
chr1	65147687	65147789	CACHD1
chr19	49982165	49982304	FLT3LG
chr22	36689374	36689527	MYH9
chr1	196658587	196658726	CFH
chr1	196709748	196709922	CFH
chr16	767059	767158	METRN
chr7	18688122	18688221	HDAC9
chr3	38038998	38039125	VILL
chr3	38047325	38047429	VILL
chr7	96635388	96635548	DLX6
chr3	149686218	149686317	PFN2
chr1	1148371	1148473	TNFRSF4
chr4	47427806	47427905	GABRB1
chr1	45974587	45974686	MMACHC
chr6	50740404	50740503	TFAP2D
chr6	63990280	63990379	LGSN
chr2	97464793	97464963	CNNM4
chr12	6125254	6125398	VWF
chr12	6161822	6161937	VWF
chr10	12131091	12131190	DHTKD1
chr6	123127377	123127502	SMPDL3A
chr1	228433168	228433267	OBSCN
chr1	228466364	228466463	OBSCN
chr21	27840819	27840950	CYYR1
chr1	207133023	207133142	FCAMR
chr9	140086589	140086702	TPRN
chr4	8221074	8221173	SH3TC1
chr13	24895747	24895846	C1QTNF9
chr13	39357206	39357332	FREM2
chr3	33552112	33552239	CLASP2
chr3	33602299	33602463	CLASP2
chr21	42843752	42843851	TMPRSS2
chr12	49724259	49724358	TROAP
chr5	137721935	137722058	KDM3B
chr5	153830650	153830773	SAP30L
chr17	56060582	56060681	VEZF1
chr20	24523909	24524046	SYNDIG1
chr22	39909947	39910046	SMCR7L
chr1	44878071	44878185	RNF220
chr16	46638270	46638369	SHCBP1
chr9	117852944	117853043	TNC
chr11	94924621	94924720	SESN3
chr2	133489317	133489416	NCKAP5
chr17	38938514	38938671	KRT27
chr9	8449724	8449845	PTPRD
chr7	100210469	100210602	MOSPD3
chr7	107580611	107580710	LAMB1
chr10	134261380	134261479	C10orf91
chrX	99662363	99662462	PCDH19
chr2	241831081	241831180	C2orf54
chr1	151105832	151105931	SEMA6C
chr17	6941869	6941968	SLC16A13
chr16	68855983	68856089	CDH1
chr20	61981680	61981809	CHRNA4
chr12	11506565	11506670	PRB1
chr17	80398916	80399062	HEXDC
chr17	56083168	56083267	SRSF1
chr20	13763317	13763444	ESF1
chr7	84628810	84628963	SEMA3D
chr3	167747601	167747700	GOLIM4
chr11	62444224	62444335	UBXN1
chr14	63453789	63453905	KCNH5
chr9	138395790	138395889	MRPS2
chr11	115080285	115080384	CADM1
chr9	13121751	13121850	MPDZ
chr17	34945767	34945866	GGNBP2
chr12	121134117	121134216	MLEC

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr6	126080680	126080849	HEY2
chr9	112184997	112185132	PTPN3
chr14	73719356	73719483	PAPLN
chr11	46917760	46917886	LRP4
chr6	32726625	32726775	HLA-DQB2
chr16	31498976	31499075	SLC5A2
chr18	51013166	51013328	DCC
chr17	47302341	47302440	PHOSPHO1
chr18	76755211	76755310	SALL3
chr10	51584790	51584889	NCOA4
chr19	7585057	7585156	ZNF358
chr19	38572756	38572931	SIPA1L3
chr12	118198838	118199237	KSR2
chr6	27806475	27806652	HIST1H2BN
chr1	190129809	190129986	FAM5C
chr12	115109859	115110036	TBX3
chr6	3850336	3850736	FAM50B
chr11	116728636	116729351	SIK3
chr1	240370882	240371602	FMN2
chr19	16860826	16861006	NWD1
chr6	32053653	32053833	TNXB
chr9	27948976	27949701	LINGO2
chr10	48389661	48390798	RBP3
chr16	3778252	3778982	CREBBP
chr5	66458986	66459168	MAST4
chr13	45148515	45148697	TSC22D1
chr13	29599442	29600585	MTUS2
chr9	138714197	138714932	CAMSAP1
chr3	187447231	187447646	BCL6
chr5	143586926	143587110	KCTD16
chr1	152082219	152082960	TCHH
chr8	81897132	81897550	PAG1
chr5	3599605	3600023	IRX1
chr4	71472191	71472378	AMBN
chr1	203316604	203316791	FMOD
chr9	118950295	118950482	PAPPA
chr7	23207466	23207657	KLHL7
chr19	2226422	2226858	DOT1L
chr19	44515337	44515532	ZNF230
chr3	78666862	78667057	ROBO1
chr9	37740676	37740872	FRMPD1
chr19	3600347	3600543	TBXA2R
chr19	53762189	53762386	VN1R2
chr18	56246149	56246942	ALPK2
chr22	38822858	38823056	KCNJ4
chr19	51628274	51628472	SIGLEC9
chr11	128844093	128844291	ARHGAP32
chr2	128262414	128262863	IWS1
chr17	37627428	37627878	CDK12
chr2	233246233	233246433	ALPP
chr15	89386656	89386856	ACAN
chr12	49484962	49485162	DHH
chr7	148979000	148979202	ZNF783
chr19	50461936	50462139	SIGLEC11
chr4	146058803	146059007	OTUD4
chr18	74091046	74091251	ZNF516
chr20	278687	279151	ZCCHC3
chr7	41739652	41739858	INHBA
chr3	147128329	147128794	ZIC1
chr3	49697948	49698155	BSN
chr14	24877089	24877296	NYNRIN
chr4	41648507	41648714	LIMCH1
chr4	146823319	146824153	ZNF827
chr6	54805204	54805412	FAM83B
chr1	237947226	237947436	RYR2
chr1	13183307	13183781	LOC440563
chr2	177036308	177036520	HOXD3
chr10	93999535	93999748	CPEB3
chr1	12907827	12908040	LOC649330
chr16	52484189	52484403	TOX3
chr1	152382358	152382842	CRNN
chr8	59409195	59409410	CYP7A1
chr9	108424811	108425026	TAL2
chr13	115090480	115090967	CHAMP1
chr7	2583248	2583465	BRAT1

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr5	44388498	44388718	FGF10
chr14	37132443	37132663	PAX9
chr2	85554392	85554613	TGOLN2
chr12	57920424	57920646	MBD6
chr8	110980472	110980696	KCNV1
chr3	50332155	50332379	HYAL3
chr1	158064570	158064795	KIRREL
chr17	46805441	46805667	HOXB13
chr7	48318293	48318520	ABCA13
chr5	137680779	137681006	FAM53C
chr5	140604527	140605446	PCDHB14
chr12	86373541	86374059	MGAT4C
chr1	18023591	18023821	ARHGEF10L
chr7	123301994	123302928	LMOD2
chr16	9857804	9858738	GRIN2A
chr1	205238669	205238902	TMCC2
chr8	1497601	1497835	DLGAP2
chr3	150931785	150932019	P2RY14
chr5	148406690	148407224	SH3TC2
chr8	88885072	88886041	DCAF4L2
chr6	66205044	66205286	EYS
chr19	56089907	56090152	ZNF579
chr16	30456027	30456580	SEPHS2
chr2	167262387	167262941	SCN7A
chr4	70146234	70146413	UGT2B28
chr7	149430770	149431016	KRBA1
chr5	43039708	43039954	ANXA2R
chr4	158257611	158257857	GRIA2
chr18	19153403	19154391	ESCO1
chr1	70503848	70504095	LRRC7
chr14	26917260	26917507	NOVA1
chr1	67147795	67148042	SGIP1
chr6	69348839	69349087	BAI3
chr3	148458895	148459145	AGTR1
chr16	75690149	75690400	TERF2IP
chr14	51132077	51132329	SAV1
chr4	57181438	57182008	KIAA1211
chr17	7139748	7140002	PHF23
chr19	52919155	52919411	ZNF528
chr18	9887073	9888100	TXNDC2
chr5	148747556	148747814	PCYOX1L
chr3	38592386	38592969	SCN5A
chr22	41573199	41573978	EP300
chr17	51900687	51901273	KIF2B
chr14	93649655	93649915	MOAP1
chr18	5416006	5416267	EPB41L3
chr5	45645336	45645597	HCN1
chr22	30688487	30688749	TBC1D10A
chr11	19077544	19077807	MRGPRX2
chr16	87451065	87451329	ZCCHC14
chr3	194081182	194081449	LRRC15
chr19	12155152	12155758	ZNF878
chr16	1129641	1129912	SSTR5
chr9	100970982	100971253	TBC1D2
chr5	130766662	130766934	RAPGEF6
chr12	57485186	57485458	NAB2
chr20	62839352	62839625	MYT1
chr2	133540001	133540275	NCKAP5
chr17	30348866	30349141	LRRC37B
chr2	136569966	136570243	LCT
chr1	180885313	180885942	KIAA1614
chr3	73111480	73111760	EBLN2
chr5	140563060	140563341	PCDHB16
chr11	130343147	130343429	ADAMTS15
chr7	30491365	30491789	NOD1
chr7	72891713	72891996	BAZ1B
chr5	63256284	63256568	HTR1A
chr5	139060649	139060934	CXXC5
chr6	100841375	100841664	SIM1
chr1	235345028	235345318	ARID4B
chr8	73480144	73480434	KCNB2
chr9	104238561	104239215	TMEM246
chr16	24372742	24373179	CACNG3
chr5	159992484	159992775	ATP10B
chr1	112524315	112524976	KCND3

TABLE 17-continued

Chromosome	Start (bp)	End (bp)	Gene
chr3	124951821	124952116	ZNF148
chr13	84454722	84455387	SLITRK1
chr19	37677024	37677691	ZNF585B
chr14	23828654	23828952	EFS
chr16	3299468	3299766	MEFV
chr10	88768853	88769151	AGAP11
chr3	151545614	151545912	AADAC
chr6	143074326	143074627	HIVEP2
chr4	77818328	77818630	SOWAHB
chr15	33261112	33261414	FMN1
chr1	203134668	203134970	ADORA1
chr1	226924201	226924885	ITPKB
chr20	16359635	16360553	KIF16B
chr17	7366046	7366352	ZBTB4
chr7	89856338	89856644	STEAP2
chr18	8824762	8825068	SOGA2
chr2	237489180	237489880	CXCR7
chr3	184910073	184910385	EHHADH
chr16	830488	830800	MSLNL
chr12	47471281	47471985	AMIGO2
chrX	129518742	129519056	GPR119
chr5	140589805	140590278	PCDHB12
chr1	227152756	227153071	ADCK3
chr6	7229900	7230612	RREB1
chr1	207195319	207195635	C1orf116
chr19	52272404	52272720	FPR2
chr1	12855600	12855917	PRAMEF1
chr1	149884959	149885277	SV2A
chr14	94088049	94088368	UNC79
chr19	46375371	46375690	FOXA3
chr19	38976305	38976784	RYR1
chr19	50102808	50103129	PRR12
chr10	52103413	52103737	SGMS1
chr1	169510827	169511558	F5
chr17	74392305	74392630	UBE2O
chr7	110763904	110764638	LRRN3
chr17	39967832	39968158	LEPREL4
chr1	74507070	74507398	LRRIQ3
chr13	41705439	41706179	KBTBD6
chr12	13716715	13717458	GRIN2B
chr6	76660410	76660740	IMPG1
chr16	83998850	83999181	OSGIN1
chr11	118498112	118498443	PHLDB1
chr3	39229796	39230129	XIRP1
chr2	219507558	219507894	ZNF142
chr12	54756830	54757588	GPR84
chr2	99012645	99012982	CNGA3
chr13	51825913	51826252	FAM124A
chr3	101383902	101384242	ZBTB11
chr2	80529774	80530544	LRRTM1
chr6	128134392	128134735	THEMIS
chr4	155506851	155507195	FGA
chr3	119133913	119134257	ARHGAP31
chr19	49206415	49207195	FUT2

TABLE 18

Chromosome	Start (bp)	End (bp)	Gene
chr7	140453046	140453220	BRAF
chr1	115256441	115256615	NRAS
chr9	21971015	21971189	CDKN2A
chr7	142459644	142459861	PRSS1
chr17	11666754	11666928	DNAH9
chr5	13766095	13766269	DNAH5
chr14	19553511	19553826	POTEG
chr1	241261926	241262100	RGS7
chr16	67694127	67694301	ACD
chr20	41306559	41306779	PTPRT
chr11	99690275	99690470	CNTN5
chr2	141777477	141777651	LRP1B
chr2	107049548	107049722	RGPD3

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr8	121381559	121381733	COL14A1
chr1	153177243	153177438	LELP1
chr1	176915070	176915251	ASTN1
chr19	43699170	43699391	PSG4
chr3	38949439	38949613	SCN11A
chr2	138169212	138169386	THSD7B
chr10	68940056	68940230	CTNNA3
chr5	13864513	13864742	DNAH5
chr8	131916026	131916289	ADCY8
chr4	47746422	47746596	CORIN
chr1	179620032	179620206	TDRD5
chr19	57666600	57666774	DUXA
chr5	101834205	101834544	SLCO6A1
chr6	57512513	57512695	PRIM2
chr21	41648023	41648197	DSCAM
chr8	3081232	3081406	CSMD1
chr12	122812612	122812786	CLIP1
chr7	140481347	140481521	BRAF
chr10	89692828	89693004	PTEN
chr18	14542652	14543063	POTEC
chr19	4902699	4902873	ARRDC5
chr12	11506164	11506863	PRB1
chr1	12887175	12887687	PRAMEF11
chr3	10452304	10452486	ATP2B2
chr21	41450619	41450888	DSCAM
chr11	102738631	102738805	MMP12
chr6	55147022	55147215	HCRTR2
chr7	53103402	53104244	POM121L12
chr16	9857007	9858801	GRIN2A
chr7	82544018	82546173	PCLO
chrX	140993242	140996574	MAGEC1
chr2	228881125	228884868	SPHKAP
chr1	190067150	190068214	FAM5C
chr1	12907287	12908052	LOC649330
chr22	26422410	26423627	MYO18B
chr8	73848206	73850274	KCNB2
chr21	42080411	42080679	DSCAM
chr5	35876083	35876528	IL7R
chr16	26147026	26147562	HS3ST4
chr2	137813991	137814765	THSD7B
chr1	13183059	13183834	LOC440563
chr12	11545876	11546912	PRB2
chr6	49753679	49754855	PGK2
chr7	150174218	150174831	GIMAP8
chr8	57228599	57228862	SDR16C5
chr2	229890370	229890777	PID1
chr12	18891225	18892057	CAPZA3
chr10	37430660	37431196	ANKRD30A
chr7	141672503	141673475	TAS2R38
chr1	75036822	75039123	C1orf173
chr5	145393330	145393609	SH3RF2
chr5	13753349	13753600	DNAH5
chr2	107459496	107460403	ST6GAL2
chr6	54804555	54806798	FAM83B
chr19	56538529	56539860	NLRP5
chr12	46230544	46230718	ARID2
chr2	103274162	103274336	SLC9A2
chr1	196928037	196928211	CFHR2
chr21	10916335	10916509	TPTE
chr7	81381415	81381589	HGF
chr9	121929379	121930447	DBC1
chr5	13762830	13763006	DNAH5
chr20	5282767	5283371	PROKR2
chr2	226446675	226447685	NYAP2
chr1	247587230	247588834	NLRP3
chr2	196749356	196749534	DNAH7
chr8	57353846	57354407	PENK
chr16	24372710	24373179	CACNG3
chr1	143767490	143767838	PPIAL4G
chr13	19748019	19748261	TUBA3C
chr1	240370098	240371828	FMN2
chr11	40135943	40137832	LRRC4C
chr7	150324821	150325587	GIMAP6
chr11	18955375	18956329	MRGPRX1
chr1	152127304	152129413	RPTN

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr5	13793640	13793833	DNAH5
chr19	22362771	22364371	ZNF676
chr1	197390129	197391069	CRB1
chr1	117311112	117311314	CD2
chr2	192700686	192701441	SDPR
chr3	122002547	122004030	CASR
chr2	30966312	30966486	CAPN13
chr3	139297716	139297890	NMNAT3
chr17	10401044	10401218	MYH1
chr8	3216703	3216877	CSMD1
chr13	103698459	103698633	SLC10A2
chr2	103300624	103300798	SLC9A2
chr5	41181486	41181660	C6
chr17	7578145	7578319	TP53
chr8	55534675	55534849	RP1
chr12	11420521	11421056	PRB3
chr8	73479975	73480514	KCNB2
chr1	233807016	233807256	KCNK1
chr4	188924021	188924868	ZFP42
chr7	143175136	143175886	TAS2R41
chr5	13776578	13776798	DNAH5
chr7	136699706	136700966	CHRM2
chr10	50315670	50315893	VSTM4
chr5	156381434	156381696	TIMD4
chr5	140558037	140559873	PCDHB8
chr8	139163459	139165440	FAM135B
chr2	108487224	108489211	RGPD4
chr1	197396609	197397108	CRB1
chr8	52320675	52322010	PXDNL
chr5	45262032	45262839	HCN1
chr3	96706247	96706814	EPHA6
chr3	121980431	121981249	CASR
chr19	31038957	31040239	ZNF536
chr7	150217094	150217919	GIMAP7
chr14	70633362	70635118	SLC8A3
chr7	86394511	86394878	GRM3
chr5	35065372	35066067	PRLR
chr1	157514068	157514311	FCRL5
chr14	94756231	94756929	SERPINA10
chr21	41719668	41719842	DSCAM
chr2	209113025	209113199	IDH1
chr6	55638856	55639030	BMP5
chr7	6426791	6426965	RAC1
chr12	7635211	7635385	CD163
chr7	117175296	117175470	CFTR
chr4	158057627	158057801	GLRB
chr19	43762386	43762596	PSG9
chr17	10399578	10399790	MYH1
chr20	9546581	9547020	PAK7
chr3	54958663	54959241	LRTM1
chrX	151869542	151870061	MAGEA6
chrX	105449517	105451061	MUM1L1
chr9	104432377	104433303	GRIN3A
chrX	139865857	139866502	CDR1
chr11	129306708	129306904	BARX2
chr19	56423088	56424654	NLRP13
chr2	230910665	230911384	SLC16A14
chrX	141290652	141291767	MAGEC2
chr10	27702222	27703028	PTCHD3
chr3	168833183	168834491	MECOM
chr16	19451377	19452048	TMC5
chr6	128134092	128135061	THEMIS
chr12	125834042	125834786	TMEM132B
chr7	150269243	150270062	GIMAP4
chr7	100349366	100350706	ZAN
chr6	63990056	63991126	LGSN
chr12	11461251	11461805	PRB4
chr10	37507967	37508797	ANKRD30A
chr14	63174333	63175165	KCNH5
chr2	132021042	132021875	POTEE
chr6	28542427	28543881	SCAND3
chr5	135692305	135693068	TRPC7
chr12	117768164	117768857	NOS1
chr7	143140576	143141494	TAS2R60
chr20	1616835	1617043	SIRPG

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr20	20033039	20033213	CRNKL1
chr12	81112658	81112832	MYF5
chr19	59010782	59010956	SLC27A5
chr22	16266924	16267098	POTEH
chr5	38881693	38881867	OSMR
chr5	168233434	168233608	SLIT3
chr1	145296319	145296493	NBPF10
chr7	146997220	146997394	CNTNAP2
chr6	28501717	28501891	GPX5
chr12	132547028	132547202	EP400
chr21	10920036	10920210	TPTE
chr3	7188164	7188338	GRM7
chr1	16892122	16892296	NBPF1
chr5	13727603	13727777	DNAH5
chr2	228886430	228886640	SPHKAP
chr1	34208908	34209161	CSMD2
chr1	196952004	196952178	CFHR5
chr2	185798307	185798481	ZNF804A
chr1	57347134	57347308	C8A
chr16	20476858	20477032	ACSM2A
chr4	107845185	107845359	DKK2
chr18	52556476	52556650	RAB27B
chr8	2813104	2813278	CSMD1
chr7	34851341	34851515	NPSR1
chr22	16279160	16279334	POTEH
chr2	196759765	196759939	DNAH7
chr8	131921945	131922119	ADCY8
chr16	20548544	20548718	ACSM2B
chr12	18691080	18691254	PIK3C2G
chr18	28968320	28968494	DSG4
chr19	55417848	55418022	NCR1
chr18	51025713	51025887	DCC
chr20	41419836	41420049	PTPRT
chr3	121712023	121712803	ILDR1
chr3	38888195	38889215	SCN11A
chr8	105405030	105405207	DPYS
chr3	38770041	38770340	SCN10A
chr20	40980728	40980904	PTPRT
chr16	70954500	70955014	HYDIN
chr12	7639970	7640270	CD163
chr10	124457271	124457788	C10orf120
chr6	136599029	136599819	BCLAF1
chr19	38951023	38951200	RYR1
chr4	71275138	71275789	PROL1
chr4	104510866	104511124	TACR3
chr17	12655754	12656628	MYOCD
chr1	176525475	176526349	PAPPA2
chr4	187454896	187455693	MTNR1A
chr3	39307001	39307959	CX3CR1
chr7	146829337	146829554	CNTNAP2
chr17	10434960	10435140	MYH2
chr10	124402647	124402908	DMBT1
chr15	86807596	86808063	AGBL1
chr19	56369134	56370574	NLRP4
chr3	108072295	108072560	HHLA2
chr19	43680035	43680256	PSG5
chr4	9783735	9785082	DRD5
chr5	36049046	36049521	UGT3A2
chr7	123593678	123594502	SPAM1
chr1	175375366	175375846	TNR
chr12	33559743	33560284	SYT10
chr5	41149354	41149542	C6
chr4	80327830	80329316	GK2
chr12	7531616	7531889	CD163L1
chr1	159799732	159799921	SLAMF8
chr10	124358298	124358572	DMBT1
chr21	41414345	41414577	DSCAM
chr5	42718558	42719405	GHR
chr3	169539795	169540644	LRRIQ4
chr5	121786322	121787255	SNCAIP
chr7	150163828	150164384	GIMAP8
chr8	110456922	110457857	PKHD1L1
chr5	13900317	13900510	DNAH5
chrX	151303133	151304072	MAGEA10
chr5	41382017	41382513	PLCXD3

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr7	154875938	154876130	HTR5A
chr18	28576770	28577003	DSC3
chr19	57646301	57647423	ZIM3
chr12	18234136	18234370	RERGL
chr2	141773268	141773463	LRP1B
chr1	152975540	152975922	SPRR3
chr5	13841809	13842004	DNAH5
chr6	165715218	165715663	C6orf118
chr10	124380646	124380883	DMBT1
chr6	100841375	100841762	SIM1
chr20	19665766	19666005	SLC24A3
chr1	152552163	152552402	LCE3D
chr4	111397572	111398147	ENPEP
chr2	234652180	234652466	DNAJB3
chr1	57480637	57481087	DAB1
chr5	26881297	26881689	CDH9
chr2	125261884	125262127	CNTNAP5
chr18	50278457	50278698	DCC
chr1	147380099	147381357	GJA8
chr12	126138150	126139215	TMEM132B
chr1	159557900	159558414	APCS
chr19	55107115	55107359	LILRA1
chr3	96962801	96963090	EPHA6
chr1	177249565	177250636	FAM5B
chr8	56435861	56436755	XKR4
chr12	81110909	81111199	MYF5
chr6	130761677	130762861	TMEM200A
chr1	248039235	248039760	TRIM58
chr19	55106566	55106813	LILRA1
chr6	40399471	40400563	LRFN2
chr1	216496827	216497031	USH2A
chr3	7620124	7620952	GRM7
chr14	26917299	26918130	NOVA1
chr2	196728871	196729703	DNAH7
chr4	100234991	100235199	ADH1B
chr4	71232442	71232695	SMR3A
chr18	61471515	61471767	SERPINB7
chr3	38627256	38627508	SCN5A
chr7	150439300	150440141	GIMAP1-GIMAP5
chr19	43575869	43576077	PSG2
chr6	96651050	96652078	FUT9
chr5	49699025	49699235	EMB
chr3	38768108	38768523	SCN10A
chr7	126173041	126173892	GRM8
chr3	161214629	161214934	OTOL1
chr18	59483146	59483694	RNF152
chr4	70146237	70146931	UGT2B28
chr21	39086552	39087409	KCNJ6
chr6	139094793	139094967	CCDC28A
chr3	2928737	2928911	CNTN4
chr8	69434032	69434206	C8orf34
chr1	179631229	179631403	TDRD5
chr7	34192714	34192888	BMPER
chr8	110509134	110509308	PKHD1L1
chr1	145281429	145281603	NOTCH2NL
chr1	74492487	74492661	LRRIQ3
chr20	57828025	57828199	ZNF831
chr7	146471330	146471504	CNTNAP2
chr7	147092699	147092873	CNTNAP2
chr1	165218684	165218858	LMX1A
chr8	108334119	108334293	ANGPT1
chr5	13871663	13871837	DNAH5
chr5	13931198	13931372	DNAH5
chr6	117113320	117114370	GPRC6A
chr7	31918614	31918788	PDE1C
chr13	20048047	20048221	TPTE2
chr2	119738946	119739120	MARCO
chr4	70156363	70156537	UGT2B28
chr12	8687232	8687406	CLEC4E
chr15	35083333	35083507	ACTC1
chr17	10408460	10408634	MYH1
chr8	25708106	25708280	EBF2
chr7	142650881	142651055	KEL
chr20	40789992	40790166	PTPRT
chr20	40944382	40944556	PTPRT

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr1	12939546	12939720	PRAMEF4
chr3	108682264	108682438	MORC1
chr2	196651727	196651901	DNAH7
chr1	196918572	196918746	CFHR2
chr5	45645489	45645663	HCN1
chr2	219293987	219294161	VIL1
chr21	10914315	10914489	TPTE
chr8	62289153	62289327	CLVS1
chr5	13769579	13769753	DNAH5
chr3	38938386	38938702	SCN11A
chr8	62212397	62212832	CLVS1
chr8	55533562	55534135	RP1
chr12	73012657	73012831	TRHDE
chr18	28725569	28725743	DSC1
chr7	141722066	141722240	MGAM
chr8	118159207	118159389	SLC30A8
chr16	77398090	77398273	ADAMTS18
chr1	152784961	152785194	LCE1B
chr11	58601913	58602311	GLYATL2
chr5	89924432	89924622	GPR98
chr7	70885917	70886091	WBSCR17
chr17	10351211	10351429	MYH4
chr8	110099743	110100525	TRHR
chr4	70455135	70455330	UGT2A1
chr5	160721114	160721433	GABRB2
chr3	130095149	130095628	COL6A5
chr7	86415599	86416389	GRM3
chr5	121758578	121759419	SNCAIP
chr12	2705017	2705191	CACNA1C
chr3	108475881	108476055	RETNLB
chr2	128341744	128341918	MYO7B
chr16	31539847	31540021	AHSP
chr3	38591831	38593038	SCN5A
chr4	20620396	20620618	SLIT2
chr12	118198892	118199317	KSR2
chr6	41165870	41166047	TREML2
chr19	43579535	43579758	PSG2
chr12	33579074	33579406	SYT10
chr19	43233328	43233512	PSG3
chr3	167023493	167023698	ZBBX
chr6	25726519	25726750	HIST1H2AA
chr4	115997240	115998160	NDST4
chr3	38622517	38622852	SCN5A
chr1	47610224	47610404	CYP4A22
chr3	189526071	189526277	TP63
chr16	77401346	77401602	ADAMTS18
chr12	70946573	70946801	PTPRB
chr1	12835081	12835291	PRAMEF12
chr5	31322960	31323361	CDH6
chr10	28409122	28409305	MPP7
chr18	61390288	61390630	SERPINB11
chr7	30795098	30795311	INMT
chr5	26915794	26916005	CDH9
chr12	126128625	126128810	TMEM132B
chr5	13737372	13737605	DNAH5
chr6	55216050	55216369	GFRAL
chr20	57828961	57829780	ZNF831
chr12	70954498	70954695	PTPRB
chr1	82408711	82409445	LPHN2
chr4	138442193	138442744	PCDH18
chr6	73904254	73905119	KCNQ5
chr12	55420245	55421211	NEUROD4
chr1	171251204	171251420	FMO1
chr7	37780104	37780795	GPR141
chr14	95029830	95030429	SERPINA4
chrX	142795147	142795594	SPANXN2
chr1	152748888	152749156	LCE1F
chr5	13901375	13901592	DNAH5
chr10	28023390	28023716	MKX
chrX	151899863	151900744	MAGEA12
chr5	121739436	121739610	SNCAIP
chr2	227945124	227945298	COL4A4
chr4	70359397	70359571	UGT2B4
chr10	28225644	28225818	ARMC4
chr12	79679586	79679760	SYT1

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr17	10300137	10300311	MYH8
chr17	10362565	10362739	MYH4
chr8	106810955	106811129	ZFPM2
chr9	127911991	127912165	PPP6C
chr5	13824287	13824461	DNAH5
chr5	156589595	156590571	FAM71B
chr12	71029485	71029731	PTPRB
chr1	57257765	57258456	C1orf168
chr1	158261892	158262111	CD1C
chr14	92251507	92251699	TC2N
chr9	113703755	113704413	LPAR1
chr1	157494193	157494367	FCRL5
chr3	38798166	38798340	SCN10A
chr5	40964821	40964995	C7
chr21	41465637	41465811	DSCAM
chr11	63173954	63174128	SLC22A9
chr11	100141811	100141985	CNTN5
chr1	75078336	75078510	C1orf173
chr2	183104838	183105012	PDE1A
chr12	100813657	100813831	SLC17A8
chr8	87738734	87738908	CNGB3
chr5	41153927	41154101	C6
chr17	10432901	10433075	MYH2
chr11	113286121	113286295	DRD2
chr4	166924534	166924708	TLL1
chr5	13830127	13830301	DNAH5
chr7	98254233	98254458	NPTX2
chr22	26688485	26689101	SEZ6L
chr16	1270071	1270919	CACNA1H
chr2	196837004	196837182	DNAH7
chrX	151935229	151935936	MAGEA3
chr21	15561359	15561697	LIPI
chr10	105048126	105048323	INA
chr16	10273881	10274211	GRIN2A
chr4	42964912	42965115	GRXCR1
chr12	7651533	7651782	CD163
chr4	189012596	189013034	TRIML2
chr10	52103295	52103773	SGMS1
chr7	63726289	63727145	ZNF679
chr5	82948396	82948601	HAPLN1
chr7	57187662	57188801	ZNF479
chr12	7867798	7868019	DPPA3
chr10	96612489	96612671	CYP2C19
chr4	55139713	55139895	PDGFRA
chrX	35974118	35974300	CXorf22
chr1	12942929	12943184	PRAMEF4
chr14	62462738	62463261	SYT16
chr10	24762204	24762897	KIAA1217
chr1	157516796	157516970	FCRL5
chr8	25718564	25718738	EBF2
chr4	94137888	94138062	GRID2
chr12	21032367	21032541	SLCO1B3
chrX	130218213	130218387	ARHGAP36
chr12	344235	344409	SLC6A13
chr5	13717450	13717624	DNAH5
chr3	189455505	189455679	TP63
chr2	155711256	155711815	KCNJ3
chrX	35993820	35994003	CXorf22
chr1	12854105	12854554	PRAMEF1
chr6	55223696	55223929	GFRAL
chr2	51254666	51255363	NRXN1
chr21	41384997	41385255	DSCAM
chr12	10783683	10783894	STYK1
chr4	40439811	40440698	RBM47
chr6	70070762	70071333	BAI3
chr3	38936083	38936404	SCN11A
chr16	20043063	20043913	GPR139
chr1	201046066	201046245	CACNA1S
chr19	51729080	51729289	CD33
chr4	42895294	42895591	GRXCR1
chr4	44176893	44177191	KCTD8
chr19	52034552	52034742	SIGLEC6
chr19	56515103	56515436	NLRP5
chr8	53084354	53085076	ST18
chr1	18691757	18692086	IGSF21

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr7	120385851	120386064	KCND2
chrX	105280470	105280898	SERPINA7
chr5	36039596	36039789	UGT3A2
chr1	75055374	75055761	C1orf173
chr14	88729551	88729797	KCNK10
chr4	69433497	69434190	UGT2B17
chr16	71570728	71571674	CHST4
chr4	70504801	70505137	UGT2A1
chr1	22973755	22974269	C1QC
chr20	31607383	31607557	BPIFB2
chr7	141635594	141635768	CLEC5A
chr8	39603983	39604157	ADAM2
chr4	73012772	73013480	NPFFR2
chr7	141731449	141731623	MGAM
chr7	141754543	141754717	MGAM
chr9	78848356	78848530	PCSK5
chr10	28420467	28420641	MPP7
chr12	70988298	70988472	PTPRB
chr8	24324306	24324480	ADAM7
chr19	50169028	50169202	BCL2L12
chr5	35957306	35957480	UGT3A1
chr4	46060483	46060657	GABRG1
chr9	21974618	21974792	CDKN2A
chr20	9417636	9417810	PLCB4
chr6	100395680	100395854	MCHR2
chr1	153122394	153122568	SPRR2G
chr16	70926260	70926434	HYDIN
chr9	39171364	39171538	CNTNAP3
chr14	20019836	20020060	POTEM
chrX	65486280	65486506	HEPH
chr19	55106128	55106349	LILRA1
chr19	51728497	51728841	CD33
chr2	102626046	102626247	IL1R2
chr3	107096456	107097221	CCDC54
chr9	21216783	21217278	IFNA16
chr1	78958514	78959151	PTGFR
chr10	95790860	95791925	PLCE1
chr5	35909981	35910155	CAPSL
chr18	57103207	57103381	CCBE1
chr1	181726066	181726240	CACNA1E
chr19	55377963	55378137	KIR3DL2
chr12	43826101	43826275	ADAMTS20
chr9	78547259	78547433	PCSK5
chr11	100169925	100170099	CNTN5
chr18	31538245	31538419	NOL4
chr1	158585010	158585184	SPTA1
chr2	155566125	155566299	KCNJ3
chr13	72063117	72063291	DACH1
chr10	28378597	28378771	MPP7
chr5	100147666	100147840	ST8SIA4
chr12	81205307	81205481	LIN7A
chr5	41203173	41203347	C6
chr19	17088176	17088350	CPAMD8
chr19	17091323	17091497	CPAMD8
chr6	100390822	100390996	MCHR2
chr6	117609734	117609908	ROS1
chr6	70064085	70064259	BAI3
chr15	88423492	88423666	NTRK3
chr4	55956096	55956270	KDR
chr1	47515653	47515827	CYP4X1
chr18	55027303	55027477	ST8SIA3
chr3	189587066	189587240	TP63
chr1	181767431	181767894	CACNA1E
chr1	192335064	192335245	RGS21
chr11	123753849	123754053	TMEM225
chr4	70360874	70361511	UGT2B4
chr14	96706805	96707830	BDKRB2
chr4	42403051	42403226	SHISA3
chr3	46399236	46399940	CCR2
chr5	153190583	153190778	GRIA1
chr10	30336467	30336728	KIAA1462
chr1	38227124	38227754	EPHA10
chr3	169099048	169099284	MECOM
chr12	81101507	81101934	MYF6
chr8	95680195	95680372	ESRP1

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr9	121976230	121976407	DBC1
chr3	38738846	38739954	SCN10A
chrX	140984914	140985121	MAGEC3
chr1	159273742	159273950	FCER1A
chr14	88477275	88478075	GPR65
chr8	39872788	39873122	IDO2
chr12	40114613	40114948	C12orf40
chr5	156479422	156479659	HAVCR1
chr22	17288657	17288962	XKR3
chr10	27687286	27688145	PTCHD3
chr8	88885041	88886170	DCAF4L2

TABLE 18-continued

Chromosome	Start (bp)	End (bp)	Gene
chr5	156816239	156816423	CYFIP2
chr11	62996843	62997107	SLC22A25
chr5	151784008	151784668	NMUR2
chr5	23522741	23522957	PRDM9
chr1	158224895	158225111	CD1A
chr16	82032726	82033761	SDR42E1
chr10	12940434	12940627	CCDC3
chr1	75072302	75072545	C1orf173
chr1	177001591	177001975	ASTN1
chr17	72469698	72469918	CD300A

TABLE 19

Case	Variant class	Variant type	Chr	Position	Tumor allele	Ref. allele	Gene	RefSeq	Residue change	Protein position	Supporting reads (non-deduped)	Total depth (non-deduped)	Percent mutant allele	Confirmed by clinical assay
P1	Indel	frame shift	chr 17	7578474	+G	C	TP53	NM_000546.5	none	NA	41	332	12%	
P1	Indel	frame shift	chr 17	29552244	−A	G	NF1	NM_000267.3	none	NA	117	1010	12%	
P1	Indel	frame shift	chr 17	29553484	+T	C	NF1	NM_000267.3	none	NA	88	643	14%	
P1	Indel	intron	chr 17	29592185	−T	C	NF1	NM_000267.3	none	NA	127	936	14%	
P1	SNV	utr-5	chr 1	156785560	A	G	NTRK1, SH2D2A	NM_001007792.1	none	NA	40	738	5%	
P1	SNV	intron	chr 1	157806043	T	G	CD5L	NM_005894.2	none	NA	44	319	14%	
P1	SNV	coding-synonymous	chr 1	248525206	G	C	OR2T4	NM_001004696.1	none	108/349	47	552	9%	
P1	SNV	intron	chr 2	33500291	C	T	LTBP1	NM_000627.3	none	NA	48	238	20%	
P1	SNV	missense	chr 4	55946307	A	C	KDR	NM_002253.2	ARG > MET	1291/1357	264	1001	26%	
P1	SNV	intron	chr 4	55963949	G	A	KDR	NM_002253.2	none	NA	202	960	21%	
P1	SNV	missense	chr 4	55968672	A	C	KDR	NM_002253.2	ARG > LEU	664/1357	162	982	17%	
P1	SNV	intron	chr 6	117642146	C	T	ROS1	NM_002944.2	none	NA	305	1397	22%	
P1	SNV	missense	chr 9	8376700	T	G	PTPRD	NM_002839.3	SER > ARG	1471/1913	339	1196	28%	
P1	SNV	intron	chr 9	8733625	T	C	PTPRD	NM_001040712.2	none	NA	85	265	32%	
P1	SNV	intron	chr 10	43611663	T	G	RET	NM_020630.4	none	NA	54	588	9%	
P1	SNV	utr-3	chr 15	88522525	T	G	NTRK3	NM_001007156.2	none	NA	67	724	9%	
P2	Indel	intron	chr 2	79314100	+A	C	REG1B	NM_006507.3	none	NA	981	4086	24%	
P2	SNV	splice-5	chr 2	50463926	A	C	NRXN1	NM_001135659.1	none	NA	2904	8529	34%	
P2	SNV	intron	chr 3	89457148	G	A	EPHA3	NM_005233.5	none	NA	2668	4414	60%	
P2	SNV	intron	chr 3	89468286	T	G	EPHA3	NM_005233.5	none	NA	838	4066	21%	
P2	SNV	intron	chr 3	89480240	T	A	EPHA3	NM_005233.5	none	NA	786	3722	21%	
P2	SNV	utr-3	chr 4	66189669	T	A	EPHA5	NM_004439.5	none	NA	575	1632	35%	
P2	SNV	intron	chr 4	66242868	T	G	EPHA5	NM_004439.5	none	NA	1849	2849	65%	
P2	SNV	intron	chr 5	176522747	A	C	FGFR4	NM_002011.3	none	NA	1938	2637	73%	
P2	SNV	intron	chr 6	117648229	C	T	ROS1	NM_002944.2	none	NA	3047	8531	36%	
P2	SNV	missense	chr 12	78400637	A	C	NAV3	NM_014903.4	PRO > HIS	440/2364	1414	8119	17%	
P2	SNV	missense	chr 12	78400910	T	G	NAV3	NM_014903.4	GLY > VAL	531/2364	3069	8571	36%	
P2	SNV	missense	chr 17	7577551	T	C	TP53	NM_000546.5	GLY > SER	244/394	3294	4966	66%	
P2	SNV	intron	chr 19	1207247	T	G	STK11	NM_000455.4	none	NA	1067	2876	37%	
P3	SNV	missense	chr 17	7578253	A	C	TP53	NM_000546.5	GLY > VAL	199/394	455	4409	10%	
P4	SNV	missense	chr 2	212248555	T	C	ERBB4	NM_005235.2	ASP > ASN	1238/1309	1006	4095	25%	
P4	SNV	missense	chr 12	25398281	T	C	KRAS	NM_033360.2	GLY > ASP	13/190	1196	4536	26%	yes
P5	SNV	missense	chr 7	55249071	T	C	EGFR	NM_005228.3	THR > MET	790/1211	659	7660	9%	yes

TABLE 19-continued

Case	Variant class	Variant type	Chr	Position	Tumor allele	Ref. allele	Gene	RefSeq	Residue change	Protein position	Supporting reads (non-de-duped)	Total depth (non-de-duped)	Percent mutant allele	Confirmed by clinical assay
P5	SNV	missense	chr 7	55259515	G	T	EGFR	NM_005228.3	LEU > ARG	858/1211	4170	11863	35%	yes
P5	SNV	near-gene-5	chr 11	55135338	A	G	none	none	none	NA	716	3285	22%	
P5	SNV	missense	chr 17	7577097	T	C	TP53	NM_000546.5	ASP > ASN	281/394	2539	5928	43%	
P6	SNV	coding-synonymous	chr 12	78400791	A	G	NAV3	NM_014903.4	none	491/2364	1223	2615	47%	
P6	SNV	missense	chr 12	129822187	T	G	TMEM132D	NM_133448.2	LEU > MET	431/1100	1595	2989	53%	
P6	SNV	stop-gained	chr 17	7578275	A	G	TP53	NM_000546.5	GLN > stop	192/394	3795	3825	99%	
P6	SNV	coding-synonymous	chr 9	8500803	A	G	PTPRD	NM_002839.3	none	NA	643	8021	8%	
P11	SNV	intron	chr 2	29448209	T	C	ALK	none	none	NA	2011	8410	24%	
P11	SNV	missense	chr 21	44524456	A	G	U2AF1	NM_006758.2	SER > PHE	34/241	1607	7775	21%	
P12	Indel	frame shift	chr 17	7577057	−C	T	TP53	NM_000546.5	none	NA	597	2735	22%	
P12	SNV	intron	chr 4	55973786	T	C	KDR	NM_002253.2	none	NA	349	1439	24%	
P12	SNV	intron	chr 6	117650296	T	G	ROS1	NM_002944.2	none	NA	889	4857	18%	
P12	SNV	missense	chr 7	41729291	G	T	INHBA	NM_002192.2	LYS > THR	413/427	186	3516	5%	
P12	SNV	intron	chr 9	8471102	T	A	PTPRD	NM_001040712.2	none	NA	747	3019	25%	
P12	SNV	missense	chr 12	25380276	G	T	KRAS	NM_033360.2	GLN > PRO	61/190	321	4104	8%	
P12	SNV	missense	chr 19	10602473	A	C	KEAP1	NM_012289.3	VAL > LEU	369/625	619	2783	22%	
P13	SNV	missense	chr 1	190067540	T	C	FAM5C	NM_199051.1	GLY > SER	637/767	404	2983	14%	
P13	SNV	stop-gained	chr 5	45461969	T	C	HCN1	NM_021072.3	TRP > stop	330/891	341	4749	7%	
P13	SNV	intron	chr 8	38276015	G	C	FGFR1	NM_001174063.1	none	NA	543	4016	14%	
P13	SNV	missense	chr 15	88483904	T	C	NTRK3	NM_001012338.2	GLU > LYS	556/840	839	4713	18%	
P13	SNV	missense	chr 17	7577538	T	C	TP53	NM_000546.5	ARG > GLN	248/394	269	2190	12%	
P14	SNV	missense	chr 1	156841521	C	A	NTRK1	NM_002529.3	GLU > ALA	275/797	710	1583	45%	
P14	SNV	intron	chr 3	89176334	T	G	EPHA3	NM_005233.5	none	NA	969	1873	52%	
P14	SNV	coding-synonymous	chr 7	55249159	A	G	EGFR	NM_005228.3	none	819/1211	796	1509	53%	
P14	SNV	missense	chr 7	55259515	G	T	EGFR	NM_005228.3	LEU > ARG	858/1211	251	2044	12%	yes
P14	SNV	intron	chr 10	43607789	T	C	RET	NM_020630.4	none	NA	688	1544	45%	
P14	SNV	missense	chr 17	7577545	C	T	TP53	NM_000546.5	MET > VAL	246/394	213	1452	15%	
P14	SNV	missense	chr 17	29553484	T	C	NF1	NM_001042492.2	PRO > LEU	678/2840	590	1192	50%	
P14	SNV	missense	chr 19	1223125	G	C	STK11	NM_000455.4	PHE > LEU	354/434	968	1659	58%	
P15	Indel	intron	chr 17	29533514	+T	G	NF1	NM_000267.3	none	NA	161	1109	15%	
P15	SNV	missense	chr 1	70226008	T	G	LRRC7	NM_020794.2	VAL > PHE	41/1538	653	6399	10%	
P15	SNV	missense	chr 1	144882833	A	C	PDE4DIP	NM_001198834.2	GLN > HIS	1062/2363	457	8590	5%	
P15	SNV	missense	chr 1	190203515	A	C	FAM5C	NM_199051.1	LYS > ASN	237/767	210	3488	6%	
P15	SNV	missense	chr 1	248525334	A	C	OR2T4	NM_001004696.1	ALA > ASP	151/349	562	3071	18%	
P15	SNV	intron	chr 2	155157911	A	C	GALNT13	NM_052917.2	none	NA	215	3469	6%	
P15	SNV	intron	chr 2	212495103	A	G	ERBB4	NM_001042599.1	none	NA	512	4067	13%	
P15	SNV	utr-3	chr 3	89528742	T	G	EPHA3	NM_005233.5	none	NA	50	710	7%	

TABLE 19-continued

Case	Variant class	Variant type	Chr	Position	Tumor allele	Ref. allele	Gene	RefSeq	Residue change	Protein position	Supporting reads (non-de-duped)	Total depth (non-de-duped)	Percent mutant allele	Confirmed by clinical assay
P15	SNV	coding-synonymous	chr 4	55979517	T	G	KDR	NM_002253.2	none	310/1357	909	4871	19%	
P15	SNV	utr-3	chr 4	66189751	A	C	EPHA5	NM_004439.5	none	NA	120	2226	5%	
P15	SNV	intron	chr 4	66233002	A	C	EPHA5	NM_004439.5	none	NA	391	1427	27%	
P15	SNV	intron	chr 4	66233003	A	C	EPHA5	NM_004439.5	none	NA	487	1523	32%	
P15	SNV	intron	chr 4	66233146	T	G	EPHA5	NM_004439.5	none	NA	553	3459	16%	
P15	SNV	missense	chr 5	176523126	A	C	FGFR4	NM_002011.3	ASP > GLU	630/803	860	4341	20%	
P15	SNV	coding-synonymous	chr 5	176524647	A	C	FGFR4	NM_002011.3	none	793/803	203	3896	5%	
P15	SNV	missense	chr 7	41729339	A	C	INHBA	NM_002192.2	ARG > ILE	397/427	735	4383	17%	
P15	SNV	intron	chr 8	87738607	A	C	CNGB3	NM_019098.4	none	NA	199	1839	11%	
P15	SNV	intron	chr 8	113563115	A	C	CSMD3	NM_052900.2	none	NA	415	4108	10%	
P15	SNV	missense	chr 9	8528716	A	C	PTPRD	NM_002839.3	ARG > LEU	139/1913	735	3641	20%	
P15	SNV	missense	chr 9	138439735	A	T	OBP2A	NM_014582.2	ILE > LYS	99/171	783	3487	22%	
P15	SNV	intron	chr 10	43608292	A	C	RET	NM_020630.4	none	NA	401	3402	12%	
P15	SNV	intron	chr 10	43608755	T	C	RET	NM_020630.4	none	NA	408	4206	10%	
P15	SNV	missense	chr 11	55135855	A	C	OR4A15	NM_001005275.1	ARG > SER	166/345	1143	4667	24%	
P15	SNV	missense	chr 12	25398284	T	C	KRAS	NM_033360.2	GLY > ASP	12/190	254	4577	6%	yes
P15	SNV	missense	chr 13	48954333	T	C	RB1	NM_000321.2	SER > PHE	485/929	251	4856	5%	
P15	SNV	intron	chr 13	48954451	T	G	RB1	NM_000321.2	none	NA	222	2178	10%	
P16	Indel	intron	chr 2	212295977	+T	A	ERBB4	NM_001042599.1	none	NA	160	1138	14%	
P16	Indel	frame shift	chr 19	1220638	−C	T	STK11	NM_000455.4	none	NA	279	3306	8%	
P16	SNV	coding-synonymous	chr 1	156843429	A	G	NTRK1	NM_002529.3	none	285/797	106	2064	5%	
P16	SNV	coding-synonymous	chr 1	181708291	T	C	CACNA1E	NM_001205293.1	none	1207/2314	252	4341	6%	
P16	SNV	coding-synonymous	chr 1	248525326	A	C	OR2T4	NM_001004696.1	none	148/349	208	4051	5%	
P16	SNV	intron	chr 2	125530343	A	C	CNTNAP5	NM_130773.2	none	NA	312	4546	7%	
P16	SNV	coding-synonymous	chr 2	212530083	A	C	ERBB4	NM_005235.2	none	612/1309	322	5104	6%	
P16	SNV	coding-synonymous-near-splice	chr 2	212587119	C	T	ERBB4	NM_005235.2	none	294/1309	442	4704	9%	
P16	SNV	intron	chr 4	55958900	T	G	KDR	NM_002253.2	none	NA	304	4371	7%	
P16	SNV	intron	chr 4	55962358	C	T	KDR	NM_002253.2	none	NA	530	3346	16%	
P16	SNV	missense	chr 4	55968588	A	C	KDR	NM_002253.2	GLY > VAL	692/1357	300	5352	6%	
P16	SNV	coding-synonymous	chr 4	55970963	G	A	KDR	NM_002253.2	none	612/1357	495	5149	10%	
P16	SNV	intron	chr 4	55971241	A	C	KDR	NM_002253.2	none	NA	231	2622	9%	
P16	SNV	intron	chr 5	19473838	T	G	CDH18	NM_001167667.1	none	NA	225	3964	6%	
P16	SNV	missense	chr 5	112176654	A	G	APC	NM_000038.5	ARG > HIS	1788/2844	395	7777	5%	
P16	SNV	intron	chr 5	176520134	T	G	FGFR4	NM_002011.3	none	NA	167	3238	5%	
P16	SNV	intron	chr 7	11501543	T	G	THSD7A	NM_015204.2	none	NA	97	1694	6%	
P16	SNV	utr-5	chr 7	53103357	A	C	POM121L12	NM_182595.3	none	NA	63	1228	5%	
P16	SNV	missense	chr 7	116411990	T	C	MET	NM_001127500.1	THR > ILE	1010/1409	831	7410	11%	
P16	SNV	intron	chr 10	43606641	A	C	RET	NM_020630.4	none	NA	201	2822	7%	
P16	SNV	intron	chr 11	534195	A	G	HRAS	NM_001130442.1	none	NA	252	2619	10%	

TABLE 19-continued

Case	Variant class	Variant type	Chr	Position	Tumor allele	Ref. allele	Gene	RefSeq	Residue change	Protein position	Supporting reads (non-deduped)	Total depth (non-deduped)	Percent mutant allele	Confirmed by clinical assay
P16	SNV	missense	chr 11	108143456	G	C	ATM	NM_000051.3	PRO > ARG	1054/3057	744	6374	12%	
P16	SNV	missense	chr 12	25398284	A	C	KRAS	NM_033360.2	GLY > VAL	12/190	942	7146	13%	yes
P16	SNV	coding-synonymous	chr 13	48947619	A	C	RB1	NM_000321.2	none	402/929	400	7741	5%	
P16	SNV	intron	chr 13	70314492	T	C	KLHL1	NM_020866.2	none	NA	524	4290	12%	
P16	SNV	intron	chr 13	70314809	A	T	KLHL1	NM_020866.2	none	NA	471	2018	23%	
P16	SNV	intron	chr 15	88472337	C	G	NTRK3	NM_001012338.2	none	NA	149	2747	5%	
P16	SNV	intron	chr 17	7578132	A	C	TP53	NM_000546.5	none	NA	76	1470	5%	
P17	SNV	missense	chr 7	81386606	T	G	HGF	NM_000601.4	ASN > LYS	127/729	276	4991	6%	
P17	SNV	missense	chr 12	25398285	A	C	KRAS	NM_033360.2	GLY > CYS	12/190	437	4384	10%	yes

TABLE 20

Non-deduped								
Sample count	Sample description/ patient (P#)/ healthy control (C#)	No. of reads mapped	% reads mapped	% properly paired reads	No. of reads on-target	Selector on-target rate	Median depth	Median fragment length
1	H3122 0.1% into HCC78	24503042	99.0%	96.8%	17041857	69.5%	8688	173
2	H3122 1% into HCC78	19199810	98.9%	96.7%	13173049	69.8%	8657	171
3	H3122 10% into HCC78	19329153	98.9%	96.5%	13486460	69.8%	6890	170
4	H3122 100%	24470094	99.0%	96.8%	16789007	68.6%	6739	174
5	HCC78 100%	21276865	99.0%	96.9%	14835137	69.7%	7602	172
6	HCC78 10% into C1 plasma DNA 4 cycles	9023859	97.5%	83.3%	5351003	59.3%	2682	170
7	HCC78 10% into C1 plasma DNAS 8 cycles SigmaWGA	7852585	79.5%	72.0%	3958384	50.4%	15	158
8	HCC78 10% into C1 plasma DNA 6 cycles	26605244	97.7%	87.2%	16066902	60.4%	8261	169
9	HCC78 10% into C1 plasma DNA 8 cycles NEBNextOvernightBead	19811700	96.9%	91.8%	12098869	61.1%	6258	166
10	HCC78 10% into C1 plasma DNA 8 cycles OrigNEBNext 15 minLig	30672877	98.0%	93.1%	18671777	60.9%	9862	167
11	HCC78 10% into C1 plasma DNA 4 ng 9 cycles	37509063	97.6%	87.6%	22690732	60.5%	11630	169
12	HCC78 0.025% into C1 plasma DNA	17409235	98.2%	87.0%	8055464	46.3%	3913	169
13	HCC78 0.05% into C1 plasma DNA	30253156	98.1%	86.1%	13529312	44.7%	6549	169
14	HCC78 0.1% into C1 plasma DNA	31335854	98.4%	88.1%	14071945	44.9%	6897	169

TABLE 20-continued

Non-deduped								
Sample count	Sample description/ patient (P#)/ healthy control (C#)	No. of reads mapped	% reads mapped	% properly paired reads	No. of reads on-target	Selector on-target rate	Median depth	Median fragment length
15	HCC78 0.5% into C1 plasma DNA	35236429	98.8%	89.8%	16277998	46.2%	8096	169
16	HCC78 1% into C1 plasma DNA	33272947	98.5%	89.8%	15528745	46.5%	7779	171
17	P1	21702598	99.3%	97.1%	12400852	57.1%	7336	220
18	P2	22430498	99.2%	97.5%	12942388	57.7%	7680	235
19	P3	25961431	99.3%	97.8%	14809108	57.0%	8838	235
20	P4	21912624	99.1%	96.5%	12389268	56.5%	7331	227
21	P5	23357455	99.2%	97.2%	13712765	58.7%	8155	219
22	P6	11356360	96.7%	92.6%	7626499	67.2%	3848	152
23	P7	10342837	97.1%	93.5%	6943003	67.1%	3552	155
24	P8	11888370	96.9%	93.0%	7827674	65.8%	4021	154
25	P9	17626969	97.0%	94.4%	10437704	59.2%	5441	172
26	P10	13290607	96.9%	93.6%	8680450	65.3%	4572	161
27	P11	22496393	96.7%	93.8%	13270664	59.0%	6970	169
28	P12	21230200	98.8%	97.7%	8703464	40.5%	4710	258
29	P13	24801066	97.8%	96.6%	9933117	39.2%	5324	252
30	P14	21873764	97.7%	96.4%	9032079	40.3%	4867	248
31	P15	23130748	97.9%	96.8%	9343153	39.6%	5041	253
32	P16	22245944	98.1%	97.2%	8955379	39.5%	4816	263
33	P17	25906115	97.9%	97.2%	10775948	40.7%	5816	239
34	P1	2916102	94.6%	90.1%	1776887	60.9%	976	192
35	P2	21639699	99.0%	97.1%	13491073	62.3%	7247	204
36	P3	23518792	99.3%	98.0%	15524732	66.0%	9562	204
37	P4	11959399	97.5%	94.1%	7178723	60.0%	3968	189
38	P5	20192824	98.8%	97.0%	12832040	63.5%	6930	187
39	P6	7773013	87.0%	81.8%	5027345	64.7%	2445	158
40	P7	14127683	94.1%	89.3%	9045653	64.0%	4793	162
41	P8	16093442	91.7%	85.4%	10242535	63.6%	5331	151
42	P9	24980306	99.2%	97.3%	13824322	55.3%	7312	239
43	P10	15408447	94.0%	89.6%	10038486	65.1%	5335	157
44	P11	23382212	93.4%	88.3%	14342719	61.3%	7700	156
45	P12	17316416	96.7%	95.9%	7304561	40.8%	3836	230
46	P13	15170651	97.7%	97.4%	6292372	40.5%	3308	241
47	P14	7141267	95.1%	96.2%	3096168	41.2%	1650	187
48	P15	19706548	97.6%	97.4%	8720851	43.2%	4538	209
49	P16	19889232	98.0%	98.3%	9011417	44.4%	4734	220
50	P17	18092543	98.5%	97.7%	7781779	42.4%	4280	238
51	C1	26766224	97.5%	86.7%	16147472	60.3%	8280	168
52	C2	20092668	98.2%	90.2%	9916653	48.5%	5089	176
53	C3	16454970	97.4%	89.2%	8206791	48.6%	4199	175
54	C4	22388109	97.3%	88.0%	11165306	48.5%	5562	175
55	C5	21899643	97.6%	86.4%	11005231	49.1%	5525	170
56	P1 time point 1	14656874	99.0%	85.0%	9475015	64.6%	5079	171
57	P1 time point 2	18861849	99.4%	84.7%	12093175	64.1%	6487	172
58	P1 time point 3	23920634	97.5%	84.7%	11695968	47.7%	5768	173
59	P2 time point 1	18474671	99.4%	86.9%	12436916	67.3%	6876	172
60	P2 time point 2	13894587	99.5%	96.4%	8839565	63.6%	5248	185
61	P2 time point 3	20191825	97.5%	96.5%	9874542	47.7%	5370	182
62	P3 time point 1	20880669	99.2%	86.0%	13261172	63.5%	7057	170
63	P3 time point 2	29631697	99.3%	86.5%	18805559	63.5%	10089	171
64	P4 time point 1	19128070	99.0%	87.4%	12679761	66.3%	6971	169
65	P4 time point 2	27673936	99.4%	85.9%	18257927	66.0%	9926	171
66	P5 time point 1	19610825	99.3%	87.8%	13069492	66.6%	7604	169
68	P5 time point 2	23075293	98.0%	93.9%	11383523	48.3%	6105	176
67	P5 time point 3	28075947	99.4%	88.0%	18938907	67.5%	10451	170
69	P6 time point 1	47768468	98.6%	91.3%	22179023	46.4%	11172	166
70	P6 time point 2	35775847	98.5%	92.0%	16677920	46.6%	8455	166
71	P9 time point 1	19595585	99.1%	84.2%	12848481	65.6%	6839	172
72	P9 time point 2	18474032	98.4%	83.9%	12047199	65.2%	6043	169
73	P9 time point 3	21996272	99.4%	88.7%	14859835	67.6%	8141	167
74	P9 time point 4	24577249	98.0%	90.4%	12087359	48.2%	6256	174
75	P9 time point 5	22592773	97.6%	84.1%	11325418	48.9%	5572	170
76	P12 time point 1	11793847	99.1%	89.1%	7612261	64.0%	3946	168
77	P12 time point 2	18761346	98.6%	85.2%	9483960	49.8%	4704	172
78	P13 time point 1	15097466	98.1%	88.4%	9550125	62.1%	4921	167
79	P13 time point 2	20074378	98.3%	86.7%	12405223	60.8%	6283	171

TABLE 20-continued

Non-deduped								
Sample count	Sample description/ patient (P#)/ healthy control (C#)	No. of reads mapped	% reads mapped	% properly paired reads	No. of reads on-target	Selector on-target rate	Median depth	Median fragment length
80	P14 time point 1	20510385	98.2%	87.8%	12803787	61.3%	6483	168
81	P14 time point 2	20676149	97.5%	87.5%	10489917	49.5%	5275	167
82	P15 time point 1	16113392	97.8%	84.3%	9826356	59.7%	4802	171
83	P15 time point 2	17611896	98.5%	96.7%	10299562	57.6%	5638	184
84	P15 time point 3	21463621	98.2%	87.0%	13024286	59.6%	6534	174
85	P15 time point 4	14616334	97.6%	83.4%	8751266	58.4%	4349	173
86	P15 time point 5	15582630	98.1%	86.4%	9505656	59.8%	4840	175
87	P16 time point 1	16329648	97.3%	85.7%	10088350	60.1%	5069	173
88	P16 time point 2	25438935	98.2%	87.4%	12932279	49.9%	6587	169
89	P16 time point 3	20158925	98.2%	86.5%	12591048	61.4%	6399	169
90	P17 time point 1	13920942	98.5%	97.1%	8358972	59.1%	4521	183

TABLE 21

Deduped (by coordinates & sequence)[a]								
Sample count	Sample description/ patient (P#)/ healthy control (C#)	No. of reads mapped	Duplication rate	Selector on-target rate	Median depth	Fraction of possible genome equivalents (%)[b]	Fold increase in library complexity (het SNPs)[c]	Estimated % possible genome equivalents sequenced (het SNPs)[d]
1	H3122 0.1% into HCC78	9447750	61%	60.2%	2922.5	34%	1.06	36%
2	H3122 1% into HCC78	7363376	62%	58.5%	2263	26%	1.07	28%
3	H3122 10% into HCC78	8585796	56%	61.4%	2711	39%	1.06	42%
4	H3122 100%	9405562	62%	60.7%	2922	43%	1.06	46%
5	HCC78 100%	8433702	60%	60.8%	2649	35%	1.05	37%
6	HCC78 10% into C1 plasma DNA 4 cycles	4864712	46%	56.1%	1364	51%	1.27	65%
7	HCC78 10% into C1 plasma DNA 8 cycles Sigma WGA	1506958	81%	15.4%	8	53%	1.07	57%
8	HCC78 10% into C1 plasma DNA 6 cycles	12258172	54%	51.4%	3107	38%	1.44	54%
9	HCC78 10% into C1 plasma DNA 8 cycles NEBNextOvernightBead	9160482	54%	51.6%	2414	39%	1.40	54%
10	HCC78 10% into C1 plasma DNA 8 cycles OrigNEBNext 15 minLig	12128078	60%	46.3%	2830	29%	1.42	41%
11	HCC78 10% into C1 plasma DNA 4 ng 9 cycles	9488082	75%	32.1%	1447	100%	1.19	100%
12	HCC78 0.025% into C1 plasma DNA	9477184	46%	34.8%	1548	40%	1.26	50%
13	HCC78 0.05% into C1 plasma DNA	15575778	49%	33.1%	2424	37%	1.37	51%
14	HCC78 0.1% into C1 plasma DNA	17236094	45%	32.9%	2703	39%	1.40	55%
15	HCC78 0.5% into C1 plasma DNA	18212006	48%	33.3%	2889	36%	1.41	50%

TABLE 21-continued

Sample count	Sample description/ patient (P#)/ healthy control (C#)	No. of reads mapped	Duplication rate	Selector on-target rate	Median depth	Fraction of possible genome equivalents (%)[b]	Fold increase in library complexity (het SNPs)[c]	Estimated % possible genome equivalents sequenced (het SNPs)[d]
		Deduped (by coordinates & sequence)[a]						
16	HCC78 1% into C1 plasma DNA	17692196	47%	33.6%	2845	37%	1.40	51%
17	P1	9849054	55%	52.1%	3018	41%	1.06	44%
18	P2	12321552	45%	55.1%	3999	52%	1.06	55%
19	P3	13958798	46%	54.1%	4489	51%	1.06	54%
20	P4	10554320	52%	51.9%	3215	44%	1.05	46%
21	P5	12655290	46%	55.9%	4205	52%	1.06	55%
22	P6	5985032	47%	63.0%	1940	50%	1.09	55%
23	P7	5330048	48%	62.5%	1729	49%	1.07	52%
24	P8	6048134	49%	61.6%	1946	48%	1.08	52%
25	P9	10297340	42%	54.4%	2924	54%	1.08	58%
26	P10	6621152	50%	59.6%	2114	46%	1.07	49%
27	P11	12588032	44%	53.2%	3529	51%	1.08	55%
28	P12	11268046	47%	37.0%	2274	48%	1.03	50%
29	P13	12409366	50%	35.9%	2433	46%	1.03	47%
30	P14	11153394	49%	37.2%	2278	47%	1.03	48%
31	P15	12056584	48%	36.6%	2415	48%	1.03	50%
32	P16	12219738	45%	36.7%	2451	51%	1.03	52%
33	P17	12958646	50%	37.2%	2636	45%	1.04	47%
34	P1	1409454	52%	57.1%	435	45%	1.03	46%
35	P2	9764204	55%	56.6%	2976	41%	1.05	43%
36	P3	11211374	52%	62.8%	4308	45%	1.07	48%
37	P4	6149264	49%	56.3%	1912	48%	1.04	50%
38	P5	7456332	63%	54.0%	2095	30%	1.05	32%
39	P6	4146734	47%	60.4%	1247	51%	1.06	54%
40	P7	5946980	58%	53.7%	1709	36%	1.04	37%
41	P8	6173080	62%	51.8%	1695	32%	1.05	33%
42	P9	12548696	50%	50.7%	3395	46%	1.05	49%
43	P10	5951104	61%	52.1%	1657	31%	1.04	32%
44	P11	10862910	54%	50.6%	2938	38%	1.07	41%
45	P12	7950700	54%	34.9%	1479	39%	1.03	40%
46	P13	3922778	74%	15.8%	317	21%	1.03	22%
47	P14	3088542	57%	34.7%	566	34%	1.02	35%
48	P15	4519878	77%	12.4%	284	20%	1.06	21%
49	P16	4361750	78%	12.1%	266	7%	1.03	7%
50	P17	8267660	54%	35.4%	1594	37%	1.03	38%
51	C1	11839302	56%	50.7%	2955	36%	1.43	51%
52	C2	5816892	71%	14.9%	424	53%	1.11	59%
53	C3	8282466	50%	38.1%	1575	38%	1.26	47%
54	C4	6079494	73%	11.9%	341	91%	1.13	100%
55	C5	9758232	55%	33.6%	1546	28%	1.28	36%
56	P1 time point 1	3680488	75%	52.4%	948	22%	1.34	30%
57	P1 time point 2	3733984	80%	46.8%	856	37%	1.25	46%
58	P1 time point 3	11150518	53%	35.0%	1818	32%	1.31	41%
59	P2 time point 1	5340414	71%	57.8%	1608	37%	1.29	48%
60	P2 time point 2	4772686	66%	56.2%	1559	30%	1.21	36%
61	P2 time point 3	10102650	50%	37.2%	2045	38%	1.24	47%
62	P3 time point 1	6710612	68%	50.8%	1702	34%	1.33	46%
63	P3 time point 2	9571240	68%	51.5%	2474	47%	1.42	66%
64	P1 time point 1	5119914	73%	54.4%	1424	43%	1.27	55%
65	P4 time point 2	8288640	70%	55.9%	2351	45%	1.40	62%
66	P5 time point 1	5185064	74%	53.4%	1527	51%	1.32	68%
68	P5 time point 2	11429884	50%	37.2%	2235	37%	1.30	48%
67	P5 time point 3	7875654	72%	56.0%	2255	46%	1.38	63%
69	P6 time point 1	21842910	54%	28.2%	3003	54%	1.41	76%
70	P6 time point 2	18629126	48%	32.6%	3023	46%	1.44	66%
71	P9 time point 1	5114308	74%	52.8%	1316	33%	1.29	43%
72	P9 time point 2	3767226	80%	46.5%	791	14%	1.24	17%
73	P9 time point 3	6988880	68%	59.1%	2153	41%	1.40	57%
74	P9 time point 4	12801394	48%	39.2%	2553	41%	1.34	55%
75	P9 time point 5	11359054	50%	39.1%	2136	38%	1.37	53%
76	P12 time point 1	4998908	58%	53.3%	1307	33%	1.25	41%
77	P12 time point 2	9297216	50%	37.8%	1682	36%	1.29	46%
78	P13 time point 1	6320228	58%	52.9%	1661	34%	1.31	44%
79	P13 time point 2	6366844	68%	45.8%	1441	29%	1.28	37%
80	P14 time point 1	7239082	65%	48.2%	1689	30%	1.33	39%
81	P14 time point 2	10120132	51%	38.5%	1898	36%	1.34	48%
82	P15 time point 1	5848926	64%	50.2%	1453	30%	1.33	40%

TABLE 21-continued

Deduped (by coordinates & sequence)[a]								
Sample count	Sample description/ patient (P#)/ healthy control (C#)	No. of reads mapped	Duplication rate	Selector on-target rate	Median depth	Fraction of possible genome equivalents (%)[b]	Fold increase in library complexity (het SNPs)[c]	Estimated % possible genome equivalents sequenced (het SNPs)[d]
83	P15 time point 2	7756082	56%	49.6%	2093	37%	1.26	47%
84	P15 time point 3	4418526	79%	31.2%	667	29%	1.11	32%
85	P15 time point 4	5921542	59%	49.2%	1416	33%	1.28	42%
86	P15 time point 5	4156694	73%	39.8%	813	32%	1.20	39%
87	P16 time point 1	5626572	66%	46.9%	1282	25%	1.25	32%
88	P16 time point 2	9929984	61%	28.3%	1336	64%	1.34	86%
89	P16 time point 3	8175762	59%	50.9%	2019	32%	1.32	42%
90	P17 time point 1	3945290	72%	40.6%	842	21%	1.18	25%

[a]Statistics for post-duplicate reads
[b]Theoretically maximum number of input genomic equivalents sequenced (minimum of input (Table 3-Expected haploid genome copies) and depth sequenced (Table 20 -Median Depth)
[c]A maximum of 100% is possible.
[d]Maximum number of input genomic equivalents sequenced (Fraction of possible genome equivalent) × fold increase in library complexity. Maximum value is 100%

[0957] All patents, patent publications, and other published references mentioned herein are hereby incorporated by reference in their entireties as if each had been individually and specifically incorporated by reference herein.

[0958] While specific examples have been provided, the above description is illustrative and not restrictive. Any one or more of the features of the previously described embodiments can be combined in any manner with one or more features of any other embodiments in the present invention. Furthermore, many variations of the invention will become apparent to those skilled in the art upon review of the specification. The scope of the invention should, therefore, be determined by reference to the appended claims, along with their full scope of equivalents.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 32

<210> SEQ ID NO 1
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

agaaatacta ataaaatgat taaagaaggt gtgtctttaa ttgaagcatg atttaaagta      60

aatgcaaagc taaaaatcag accactgcac tccagcctgg g                         101


<210> SEQ ID NO 2
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

tactaataaa atgattaaag aaggtgtgtc tttaattgaa gcatgattta aagtaaatgc      60

aaagctaaaa atcagaccac tgcactccag cctggggaac a                         101


<210> SEQ ID NO 3
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

aaatgattaa agaaggtgtg tctttaattg aagcatgatt taaagtaaat gcaaagctaa      60

aaatcagacc actgcactcc agcctgggga acaagagtga a                         101
```

-continued

<210> SEQ ID NO 4
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

gtgtgtcttt aattgaagca tgatttaaag taaatgcaaa gctaaaaatc agaccactgc 60

actccagcct ggggaacaag agtgaaaccc catctcaaaa a 101

<210> SEQ ID NO 5
<211> LENGTH: 100
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

gtgtctttaa ttgaagcatg atttaaagta aatgcaaagc taaaaatcag accactgcac 60

tccagcctgg ggaacaagag tgaaacccca tctcaaaaac 100

<210> SEQ ID NO 6
<211> LENGTH: 100
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

gtctttaatt gaagcatgat ttaaagtaaa tgcaaagcta aaaatcagac cactgcactc 60

cagcctgggg aacaagagtg aaaccccatc tcaaaaacaa 100

<210> SEQ ID NO 7
<211> LENGTH: 92
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

tgaagcatga tttaaagtaa atgcaaagct aaaaatcaga ccactgcact ccagcctggg 60

gaacaagagt gaaaccccat ctcaaaaaca aa 92

<210> SEQ ID NO 8
<211> LENGTH: 92
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

atgatttaaa gtaaatgcaa agctaaaaat cagaccactg cactccagcc tggggaacaa 60

gagtgaaacc ccatctcaaa aacaaacaaa ca 92

<210> SEQ ID NO 9
<211> LENGTH: 92
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

agtaaatgca aagctaaaaa tcagaccact gcactccagc ctggggaaca agagtgaaac 60

cccatctcaa aaacaaacaa acaaaacaaa ac 92

<210> SEQ ID NO 10
<211> LENGTH: 100
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 10

atgcaaagct aaaaatcaga ccactgcact ccagcctggg gaacaagagt gaaaccccat 60

ctcaaaaaca aacaaacaaa acaaaacaaa aaaaactaag 100

<210> SEQ ID NO 11
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

atgcaaagct aaaaatcaga ccactgcact ccagcctggg 40

<210> SEQ ID NO 12
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

tgtcagagta gtggtggttt ataagacggg agaaaatagc acctcacttc cagaaagctt 60

taagacaaaa ggtgagtact agagtaagat tcagtctcag a 101

<210> SEQ ID NO 13
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

gagtagtggt ggtttataag acgggagaaa atagcacctc acttccagaa agctttaaga 60

caaaaggtga gtactagagt aagattcagt ctcagatctg g 101

<210> SEQ ID NO 14
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

gtggtggttt ataagacggg agaaaatagc acctcacttc cagaaagctt taagacaaaa 60

ggtgagtact agagtaagat tcagtctcag atctgggtga cac 103

<210> SEQ ID NO 15
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

gtttataaga cgggagaaaa tagcacctca cttccagaaa gctttaagac aaaaggtgag 60

tactagagta agattcagtc tcagatctgg gtgacacaaa g 101

<210> SEQ ID NO 16
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

ataagacggg agaaaatagc acctcacttc cagaaagctt taagacaaaa ggtgagtact 60

agagtaagat tcagtctcag atctgggtga cacaaaggac c 101

-continued

<210> SEQ ID NO 17
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

agaaaatagc acctcacttc cagaaagctt taagacaaaa ggtgagtact agagtaagat 60

tcagtctcag atctgggtga cacaaaggac catggatttc t 101

<210> SEQ ID NO 18
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

aatagcacct cacttccaga aagctttaag acaaaaggtg agtactagag taagattcag 60

tctcagatct gggtgacaca aaggaccatg gatttctgca a 101

<210> SEQ ID NO 19
<211> LENGTH: 104
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

acctcacttc cagaaagctt taagacaaaa ggtgagtact agagtaagat tcagtctcag 60

atctgggtga cacaaaggac catggatttc tgcaaccctt ggtg 104

<210> SEQ ID NO 20
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

cagaaagctt taagacaaaa ggtgagtact agagtaagat tcagtctcag atctgggtga 60

cacaaaggac catggatttc tgcaaccctt ggtgcctttc t 101

<210> SEQ ID NO 21
<211> LENGTH: 101
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

aagacaaaag gtgagtacta gagtaagatt cagtctcaga tctgggtgac acaaaggacc 60

atggatttct gcaacccttg gtgcctttct tgggaaccca t 101

<210> SEQ ID NO 22
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22

aagacaaaag gtgagtacta gagtaagatt cagtctcaga 40

<210> SEQ ID NO 23
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 23

agacgggaga aaatagcacc 20

<210> SEQ ID NO 24
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 24

accaagggtt gcagaaatcc 20

<210> SEQ ID NO 25
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 25

gagatggagt ttcactcttg ttgc 24

<210> SEQ ID NO 26
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 26

gaacctttcc atcatactta gaaatac 27

<210> SEQ ID NO 27
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 27

tccatggaag ccagaac 17

<210> SEQ ID NO 28
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 28

atgctaagat gtgtctgtca 20

<210> SEQ ID NO 29
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 29

-continued

```
ccttaacaca gatggctctt gatgc                                          25


<210> SEQ ID NO 30
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 30

tcctctttcc accttggctt tcc                                            23


<210> SEQ ID NO 31
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 31

ggttcagaac taccaataac aag                                            23


<210> SEQ ID NO 32
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polynucleotide

<400> SEQUENCE: 32

acctgatgtg tgacctgatt gatg                                           24
```

What is claimed is:

1. A method of detecting, diagnosing, prognosing, or therapy selection of a cancer in a subject in need thereof, the method comprising:

(a) obtaining sequence information of a cell-free DNA (cfDNA) sample derived from the subject; and

(b) using the sequence information derived from (a) to detect circulating tumor DNA (ctDNA) in the sample, wherein the method is capable of detecting a percentage of ctDNA that is less than or equal to 2% of total cfDNA.

2. The method of claim **1**, wherein the method is capable of detecting a percentage of ctDNA that is less than or equal to 1.75%, 1.5%, 1.25%, 1%, 0.75%, 0.50%, 0.25%, 0.1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 0.1%, 0.05%, 0.01%, 0.009%, 0.008%, 0.007%, 0.006%, 0.005%, 0.004%, 0.003%, 0.002%, 0.001%, 0.0005%, or 0.00001% of the total cfDNA.

3. The method of claim **1**, wherein the sample is a plasma, serum, sweat, breath, tears, saliva, urine, stool, amniotic fluid, or cerebral spinal fluid sample.

4. The method of claim **1**, wherein the sample is not a pap smear, cyst fluid, or pancreatic fluid sample.

5. The method of claim **1**, wherein the sequence information comprises information related to at least 2, 3, 5, 8, 10, 20, 30, 40, 100, 200, or 300 genomic regions.

6. The method of claim **5**, wherein the genomic regions comprise two or more of exonic regions, intronic regions, and untranslated regions.

7. The method of claim **5**, wherein the genomic regions comprise less than 1.5 megabases (Mb), 1 Mb, 500 kb, 350 kb, 100 kb, 75 kb, 50 kb or 25 kb of the genome.

8. The method of claim **1**, wherein the sequence information comprises information pertaining to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions from a selector set comprising a plurality of genomic regions.

9. The method of claim **8**, wherein the plurality of genomic regions are based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects.

10. The method of claim **8**, wherein at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the plurality of genomic regions are based on a selector set comprising genomic regions comprising one or more mutations present in one or more subjects from a population of cancer subjects.

11. The method of claim **9** or **10**, wherein the selector set comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more genomic regions selected from any one of Tables 2 and 18.

12. The method of claim **1**, wherein the obtaining sequence information of step (a) comprises performing massively parallel sequencing.

13. The method of claim **1**, wherein the obtaining sequence information of step (a) comprises using one or more adaptors.

14. The method of claim **13**, wherein the one or more adaptors comprise a molecular barcode comprising a randomer sequence.

15. The method of claim **1**, wherein using the sequence information of step (b) comprises detecting one or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome.

16. The method of claim **1**, wherein using the sequence information of step (b) comprises detecting two or more of SNVs, indels, copy number variants, and rearrangements in selected regions of the subject's genome.

17. The method of claim **1**, wherein the detecting of step (b) does not involve performing digital PCR (dPCR).

18. The method of claim **1**, wherein the detecting of step (b) comprises applying an algorithm to the sequence information to determine a quantity of one or more genomic regions from a selector set.

19. The method of claim **1**, further comprising detecting, diagnosing, prognosing or selecting a therapy for a cancer in the subject based on the detection of ctDNA.

20. The method of claim **19**, wherein diagnosing or prognosing the cancer has a sensitivity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%.

21. The method of claim **19**, wherein diagnosing or prognosing the cancer has a specificity of at least about 50%, 52%, 55%, 57%, 60%, 62%, 65%, 67%, 70%, 72%, 75%, 77%, 80%, 82%, 85%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, or 99%.

22. A method of producing a selector set for a cancer comprising:

(a) identifying genomic regions comprising mutations in one or more subjects from a population of subjects suffering from the cancer;

(b) ranking the genomic regions based on a Recurrence Index (RI), wherein the RI of the genomic region is determined by dividing the number of subjects or tumors with mutations in the genomic region by the size of the genomic region; and

(c) producing a selector set based on the RI.

23. The method of claim **22**, wherein at least a subset of the genomic regions are exon regions, intron regions, untranslated regions, or a combination thereof.

24. The method of claim **22**, wherein producing the selector set based on the RI comprises selecting genomic regions that have a recurrence index in the top 70th, 75th, 80th, 85th, 90th, or 95th or greater percentile.

25. The method of claim **22**, wherein producing the selector set comprises applying an algorithm to a subset of the ranked genomic regions.

26. The method of claim **22**, wherein producing the selector set comprises selecting genomic regions that maximize a median number of mutations per subject of the selector set.

27. The method of claim **22**, wherein producing the selector set comprises selecting genomic regions that maximize the number of subjects in the selector set.

28. The method of claim **22**, wherein producing the selector set comprises selecting genomic regions that minimize the total size of the genomic regions.

29. A computer readable medium comprising sequence information for two or more genomic regions wherein:

(a) the two or more genomic regions comprise one or more mutations present in greater than or equal to 80% of tumors from a first population of subjects suffering from a first type of cancer;

(b) the two or more genomic regions represent less than 1.5 Mb of the genome; and

(c) one or more of the following:

(i) the condition is not hairy cell leukemia, ovarian cancer, Waldenstrom's macroglobulinemia;

(ii) a genomic region comprises at least one mutation in at least one subject afflicted with the cancer;

(iii) the two or more genomic regions comprise one or more mutations present in a second population of subjects suffering from a second type of cancer;

(iv) the two or more genomic regions are derived from two or more different genes;

(v) the genomic regions comprise two or more mutations; or

(vi) the two or more genomic regions comprise at least 10 kb.

30. The computer readable medium of claim **29**, wherein the genomic regions comprise one or more mutations present in greater than or equal to 60% of tumors from the second population of subjects suffering from the second type of cancer.

31. The computer readable medium of claim **29**, wherein the genomic regions are derived from 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more different genes.

32. The computer readable medium of claim **29**, wherein the genomic regions comprise at least 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 kb.

33. The computer readable medium of claim **29**, wherein the sequence information comprises genomic coordinates pertaining to the two or more genomic regions.

34. The computer readable medium of claim **29**, wherein the sequence information comprises a nucleic acid sequence pertaining to the two or more genomic regions.

35. The computer readable medium of claim **29**, wherein the sequence information comprises a length of the two or more genomic regions.

36. A composition comprising a set of oligonucleotides that selectively hybridize to a plurality of genomic regions, wherein:

(a) greater than or equal to 80% of tumors from a population of cancer subjects include one or more mutations in the genomic regions;

(b) the plurality of genomic regions represent less than 1.5 Mb of the genome; and

(c) the set of oligonucleotides comprise 5 or more different oligonucleotides that selectively hybridize to the plurality of genomic DNA regions.

37. The composition of claim **36**, wherein the genomic DNA regions comprise at least 2 regions from those identified in any one of Tables 2 and 6-18.

38. The composition of claim **36**, wherein the set of oligonucleotides hybridize to between about 5 kb to 1000 kb of the genome.

39. The composition of claim **36**, wherein the set of oligonucleotides are capable of hybridizing to 5 or more different genomic regions.

40. The composition of claim **36**, wherein the oligonucleotides are attached to a solid support.

41. The composition of claim **40**, wherein the solid support is a bead.

42. The composition of claim **40**, wherein the solid support is an array.

43. A method for preparing a library for sequencing comprising:

(a) conducting an amplification reaction on cell-free DNA (cfDNA) derived from a sample to produce a plurality of amplicons, wherein the amplification reaction comprises 20 or fewer amplification cycles; and

(b) producing a library for sequencing, the library comprising the plurality of amplicons.

44. The method of claim **43**, wherein the amplification reaction comprises 15 or fewer amplification cycles.

45. The method of claim **43**, further comprising attaching adaptors to the cell-free DNA.

46. The method of claim **45**, wherein the adaptors comprise a molecular barcode.

47. The method of claim **45**, wherein the adaptors comprise a sample index.

48. The method of claim **45**, wherein the adaptors comprise a primer sequence.

49. The method of claim **45**, wherein the adaptors comprise a Y-shaped adaptor.

50. The method of claim **43**, further comprising fragmenting the cfDNA.

51. The method of claim **43**, further comprising end-repairing the cfDNA.

52. The method of claim **43**, further comprising A-tailing the cfDNA.

53. A method of determining a statistical significance of a selector set, the method comprising:

(a) detecting a presence of one or more mutations in one or more samples from a subject, wherein the one or more mutations are based on a selector set comprising genomic regions comprising the one or more mutations;

(b) determining a mutation type of the one or more mutations present in the sample; and

(c) determining a statistical significance of the selector set by calculating a ctDNA detection index based on a p-value of the mutation type of mutations present in the one or more samples.

54. The method of claim **53**, wherein if a rearrangement is observed in two or more samples from the subject, then the ctDNA detection index is 0.

55. The method of claim **54**, wherein at least one of the two or more samples is a plasma sample.

56. The method of claim **54**, wherein at least one of the two or more samples is a tumor sample.

57. The method of claim **54**, wherein the rearrangement is a fusion or a breakpoint.

58. The method of claim **53**, wherein if one type of mutation is present, then the ctDNA detection index is the p-value of the one type of mutation.

59. The method of claim **53**, wherein if: (i) two or more types of mutations are present in the sample; (ii) the p-values of the two or more types mutations are less than 0.1; and (iii) a rearrangement is not one of the types of mutations, then the ctDNA detection is calculated based on the combined p-values of the two or more mutations.

60. The method of claim **59**, wherein the p-values of the two or more mutations are combined according to Fisher's method.

61. The method of claim **59**, wherein one of the two or more types of mutations is a SNV.

62. The method of claim **61**, wherein the p-value of the SNV is determined by Monte Carlo sampling.

63. The method of claim **59**, wherein one of the two or more types of mutations is an indel.

64. The method of claim **53**, wherein if: (i) two or more types of mutations are present in the sample; (ii) a p-value of at least one of the two or more types of mutations are greater than 0.1; and (iii) a rearrangement is not one of the types of mutations, then the ctDNA detection is calculated based on the p-value of one of the two or more types mutations.

65. The method of claim **64**, wherein one of the two or more types of mutations is a SNV.

66. The method of claim **65**, wherein the ctDNA detection index is calculated based on the p-value of the SNV.

67. The method of claim **64**, wherein one of the two or more types of mutations is an indel.

68. A method of identifying rearrangements in one or more nucleic acids, the method comprising:

(a) obtaining sequencing information pertaining to a plurality of genomic regions;

(b) producing a list of genomic regions, wherein the genomic regions are adjacent to one or more candidate rearrangement sites or the genomic regions comprise one or more candidate rearrangement sites;

(c) applying an algorithm to the list of genomic regions to validate candidate rearrangement sites, thereby identifying rearrangements.

69. The method of claim **68**, wherein the sequencing information comprises an alignment file.

70. The method of claim **69**, wherein the alignment file comprises an alignment file of pair-end reads, exon coordinates, and a reference genome.

71. The method of claim **68**, wherein the sequencing information is obtained from a database.

72. The method of claim **68**, wherein the sequencing information is obtained from one or more samples from one or more subjects.

73. The method of claim **68**, wherein producing the list of genomic regions comprises identifying discordant read pairs based on the sequencing information.

74. The method of claim **73**, wherein producing the list of genomic regions comprises classifying the discordant read pairs based on the sequencing information.

75. The method of claim **73**, wherein producing the list of genomic regions further comprises ranking the genomic regions.

76. The method of claim **75**, wherein the genomic regions are ranked in decreasing order of discordant read depth.

77. The method of claim **68**, wherein producing the list of genomic regions comprises using an algorithm to analyze properly paired reads in which one of the paired reads is truncated to produce a soft-clipped read.

78. The method of claim **68**, wherein the algorithm analyzes the soft-clipped reads based on a pattern.

79. The method of claim **78**, wherein the pattern is based on x number of skipped bases (Sx) and on y number of contiguous mapped bases (My).

80. The method of claim **79**, wherein the pattern is MySx or SxMy.

81. The method of claim **68**, wherein applying the algorithm to validate the candidate rearrangement sites comprises ranking the candidate rearrangements based on their read frequency.

82. The method of claim **68**, wherein applying the algorithm to validate the candidate rearrangement sites comprises comparing two or more reads of the candidate rearrangement.

83. The method of claim **82**, wherein applying the algorithm to validate the candidate rearrangement sites comprises identifying the candidate rearrangement as a rearrangement if the two or more reads have a sequence alignment.

84. A method of identifying tumor-derived single nucleotide variations (SNVs), the method comprising:

(a) obtaining a sample from a subject suffering from a cancer or suspected of suffering from a cancer;

(b) conducting a sequencing reaction on the sample to produce sequencing information;

(c) applying an algorithm to the sequencing information to produce a list of candidate tumor alleles based on the sequencing information from step (b), wherein a candidate tumor allele comprises a non-dominant base that is not a germline SNP; and

(d) identifying tumor-derived SNVs based on the list of candidate tumor alleles.

85. The method of claim **84**, wherein producing the list of candidate tumor alleles comprises ranking the tumor alleles by their fractional abundance.

86. The method of claim **85**, wherein producing the list of candidate tumor alleles comprises ranking the tumor alleles based on a sequencing depth.

87. The method of claim **86**, wherein producing the list of candidate tumor alleles comprises selecting tumor alleles that meet a minimum sequencing depth.

88. The method of claim **87**, wherein the minimum sequencing depth is at least 100×, 200×, 300×, 400×, 500×, 600×, 700×, 800×, 900×, 1000× or more.

89. A method of producing a selector set comprising:

(a) obtaining sequencing information of a tumor sample from a subject suffering from a cancer;

(b) comparing the sequencing information of the tumor sample to sequencing information from a non-tumor sample from the subject to identify one or more mutations specific to the sequencing information of the tumor sample; and

(c) producing a selector set comprising one or more genomic regions comprising the one or more mutations specific to the sequencing information of the tumor sample.

90. The method of claim **89**, wherein the selector set comprises sequencing information pertaining to the one or more genomic regions.

91. The method of claim **90**, wherein the selector set comprises genomic coordinates pertaining to the one or more genomic regions.

92. The method of claim **90**, wherein the selector set comprises a plurality of oligonucleotides that selectively hybridize the one or more genomic regions.

93. The method of claim **92**, wherein the plurality of oligonucleotides are biotinylated.

94. The method of claim **89**, the one or more mutations comprise SNVs, indels, rearrangements, or a combination thereof.

95. The method of claim **94**, wherein producing the selector set comprises identifying tumor-derived SNVs based on the method of any one of claims **84-88**.

96. The method of claim **94**, wherein producing the selector set comprises identifying tumor-derived rearrangements based on the method of any one of claims **68-83**.

* * * * *