



(43) International Publication Date  
2 October 2014 (02.10.2014)

(10) International Publication Number  
**WO 2014/160046 A1**

- (51) **International Patent Classification:**  
*C12Q 1/68* (2006.01) *C07H 21/02* (2006.01)
- (21) **International Application Number:**  
PCT/US2014/025703
- (22) **International Filing Date:**  
13 March 2014 (13.03.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
61/785,498 14 March 2013 (14.03.2013) US
- (71) **Applicant:** THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA [US/US]; 3160 Chestnut Street, Suite 200, Philadelphia, Pennsylvania 19104 (US).
- (72) **Inventors:** RAJ, Arjun; 1138 Morris Road, Wynnewood, Pennsylvania 19096 (US). LEVESQUE, Marshall; 80 E. Milton Avenue, Unit 907, Rahway, New Jersey 07065 (US).
- (74) **Agents:** DOYLE, Kathryn et al.; Saul Ewing, LLP, Centre Square West, 1500 Market Street, 38th Floor, Philadelphia, Pennsylvania 19102 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**

- with international search report (Art. 21(3))  
— with sequence listing part of description (Rule 5.2(a))



WO 2014/160046 A1

(54) **Title:** A METHOD FOR DETECTING MUTATIONS IN SINGLE CELLS OR SINGLE MOLECULES

(57) **Abstract:** A high efficiency fluorescence *in situ* hybridization (FISH) method for detecting mutations on individual RNA transcripts, including both exonic and intronic RNA transcripts. The method may be used to quantify allelic expression at the population and single cell level, and also to distinguish maternal chromosomes from paternal chromosomes in single cells.

## A METHOD FOR DETECTING MUTATIONS IN SINGLE CELLS OR SINGLE MOLECULES

### 5 CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application 61/785,498, filed March 14, 2013, which application is hereby incorporated by reference in its entirety herein.

### 10 STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was made with government support under Grant No. 1DP2OD008514 awarded by the National Institute of Health (NIH). The government has certain rights in the invention.

15

### BACKGROUND OF THE INVENTION

Genome-wide association studies have highlighted the need for tools to quantify the expression of genes in an allele-specific manner to show how disease-associated single nucleotide variants (SNVs) affect transcription. Advances in single cell imaging have enabled researchers to detect individual RNAs with single molecule resolution (Femino et al., 20 1998, *Science* 280:585–590; Raj et al., 2008, *Nat Methods* 5:877–879), or in conjunction with single chromosomes (Levesque and Raj, 2013, *Nat Methods*, doi:10.1038/nmeth.2372). However, such methods are typically unable to distinguish SNVs in these molecules, and the few methods available for *in situ* SNV detection tend to be complex and suffer from low efficiency (Larsson et al., 2004, *Nat Methods* 1:227–232; Larsson et al., 2010, *Nat Methods* 25 7:395–397). In one method, a complex enzymatic scheme was used to amplify signals from SNVs on RNA molecules to the point where SNVs could be detected *in situ* (Larsson et al., 2010, *Nat Methods* 7:395–397). However, the detection efficiency of this method was very low, likely on the order of 1 % or lower in many cases. Accordingly, development of a method for SNV detection that is able to measure allele-specific gene expression at the 30 single-cell and single-molecule level would be of great utility in a variety of fields, including genetics and transcription (Ferguson-Smith, 2011, *Nat Rev Genet* 12:565–575; Gimelbrant et al., 2007, *Science* 318:1136–1140; Gregg et al., 2010, *Science* 329: 643–648).

One of the primary difficulties associated with detecting a difference of a single base via RNA fluorescence *in situ* hybridization (FISH) is that an oligonucleotide probe on the order of 20 bases or more often hybridizes to the RNA despite the presence of a single mismatch. On the other hand, very short oligonucleotide probes, while able to  
5 discriminate between single-base differences, often fail to remain bound to the target due to reduced binding energy. In either case, distinguishing legitimate signals from false positives can be problematic when using just a single probe.

Thus, there is a need in the art for a method for SNV detection that can efficiently measure gene expression at the population, single cell or single molecule level.

10 The present invention addresses this unmet need in the art.

#### BRIEF SUMMARY OF THE INVENTION

The invention includes a method for detecting a mutation in a target nucleic acid. The method comprises hybridizing at least one labeled, masked detection probe to a  
15 target nucleic acid having at least one mutation, wherein the hybridization region of the detection probe and target nucleic acid includes the mutation. In one embodiment, the mutation is a small-scale mutation selected from the group consisting of a single nucleotide variant (SNV), insertion, and deletion. In another embodiment, the method further comprises hybridizing the target nucleic acid with at least one labeled guide probe. In yet another  
20 embodiment, the hybridization region of the at least one guide probe and target nucleic acid does not include the mutation. In a further embodiment, the at least one guide probe and the detection probe are each labeled with a fluorophore. In another embodiment, the fluorophore for the at least one guide probe and the fluorophore for the detection probe are distinguishable when visualized.

25 In another embodiment, the method of the invention further comprises detecting the at least one guide probe fluorophore in a first image; detecting the detection probe fluorophore in a second image; and detecting the mutation in the nucleotide sequence of the target nucleic acid by comparing said first image with said second image to determine a colocation of the at least one guide probe fluorophore and the detection probe fluorophore.

30 In yet another embodiment, the detection probe is masked by an oligonucleotide that hybridizes to a portion of a nucleotide sequence of the detection probe. In a further embodiment, the detection probe hybridizes to the target nucleic acid on the unmasked region of the detection probe. In yet another embodiment, the oligonucleotide masking a portion of the nucleotide sequence of the detection probe denatures from the

detection probe when the detection probe hybridizes to the target nucleic acid. In another embodiment, the previously masked portion of the detection probe hybridizes to the target nucleic acid subsequent to the oligonucleotide denaturing from the detection probe.

In an additional embodiment, the target nucleic acid is a ribonucleic acid (RNA). In another embodiment, the RNA is selected from the group consisting of messenger RNA, intronic RNA, exonic DNA, and non-coding RNA. In yet another embodiment, the detection probe comprises a nucleotide sequence of at least 10 bases. In a further embodiment, the at least one guide probe comprises a nucleotide sequence of at least 10 bases. In yet a further embodiment, the oligonucleotide masking the detection probe has a nucleotide sequence of at least 5 bases. In another embodiment, the detection probe comprises an unmasked nucleotide sequence of at least 2 bases when the oligonucleotide mask is hybridized to the detection probe.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The following detailed description of preferred embodiments of the invention will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there are shown in the drawings embodiments which are presently preferred. It should be understood, however, that the invention is not limited to the precise arrangements and instrumentalities of the embodiments shown in the drawings.

Figure 1, comprising Figs. 1A through Fig. 1D, is a set of diagrams, images, and graphs showing that toehold probes enable SNV detection on individual RNA molecules *in situ*. Fig. 1A: Schematic of the principle behind *in situ* SNV detection, using the T1799A mutation of BRAF as an example. Fig. 1B: Visualization of the guide probe detecting BRAF mRNA (ATTO488, left panel) and the wild-type and mutant detection probes (Cy5, Cy3, middle and right panels, respectively). Fig. 1C: Classification of RNA as being either wild-type or mutant using the detection probes. Fig. 1D: Quantification and classification of RNA as wild-type or mutant in a group of single cells. Left: cells with only wild-type BRAF; middle: cells that are heterozygous for BRAF; right: cells that are mutant for BRAF. Scale bars are 5µm long.

Figure 2, comprising Figs. 2A through 2C, is a set of diagrams and graphs, showing allele-specific expression at the population and single-cell level in GM12878 cells. Fig. 2A: Allelic imbalance was quantified in the population of the indicated genes by measuring the probability that a transcript comes from either the maternal or paternal allele.

Error bars reflect 99% confidence intervals. Fig. 2B: Using a statistical model, the number of RNAs required to determine whether an allelic imbalance exists for a given actual imbalance was examined. Fig. 2C: Statistical examination of the likelihood of finding a population with the measured number of maternal vs. paternal transcripts given a null hypothesis in which each mRNA was randomly assigned as being either maternal or paternal. As depicted, DNMT1 shows single cell imbalance, whereas SUZ12 and SKA3 do not.

Figure 3, comprising Figs. 3A and 3B, is a set of images and diagrams showing the detection of maternal and paternal chromosomes *in situ* using SNV detection. Fig. 3A: Illustration of the chromosome detection method. iceFISH probes that target chromosome 19 and SNV detection probes targeting 19 SNPs within 15 of these genes on the paternal chromosome were designed. Fig. 3B: Example images showing the two copies of chromosome 19 (dashed regions) with the computationally identified colocalized detection probes labeled with circles.

Figure 4, comprising Figs. 4A and 4B, is a set of graphs showing that the addition of a mask is required for proper discrimination of single nucleotide variant targets. Fig. 4A: The WM3918 cell line was probed, which is homozygous for the wild-type allele, with probes targeting both the mutant and wild type allele, either with or without the mask (left, right, respectively). Each bar represents the mRNA counts from a single cell. It was found that in the presence of mask, the vast majority of transcripts are wild-type, whereas without mask, a large fraction of the mutant probe spuriously bound to the target. Fig. 4B: Other targets also showed single base mismatch discrimination. Sequences were targeted as shown with both perfect match and mismatch detection probes, and found that the perfect match probe was far more likely to bind.

Figure 5 is a graph showing that changing the toehold length can change the detection efficiency without dramatically increasing off-target binding. Toehold length is in nucleotides, with the total probe length remaining constant (toehold length changed by changing the mask probe length). The free energy change of the toehold binding (given in kcal/mol) was computed using parameters from Sugimoto et al. Biochemistry 34: 11211–11216 (1995).

#### DETAILED DESCRIPTION

The present invention relates to a fluorescence *in situ* hybridization method for detecting a mutation on an individual RNA transcript. The method provides high-efficiency detection of mutations in the sequence of individual RNA molecules in single cells.

### Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although any methods and materials similar or equivalent to those described herein can be used in the practice for testing of the present invention, the preferred materials and methods are described herein. In describing and claiming the present invention, the following terminology will be used.

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 be limiting.

The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

“About” as used herein when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass non-limiting variations of  $\pm 20\%$  or  $\pm 10\%$ ,  $\pm 5\%$ ,  $\pm 1\%$ , or  $\pm 0.1\%$  from the specified value, as such variations are appropriate.

The term “mutation” as used herein refers to any change of one or more nucleotides in a nucleotide sequence.

The term “single nucleotide variant” (SNV) refers to a single nucleotide in a DNA or RNA sequence that is different in comparison to a control or reference sequence. The term “single nucleotide variant” includes and encompasses single nucleotide polymorphisms (SNPs), which are DNA sequence variations characterized by a single nucleotide in the genome sequence that is altered in at least 1% of the human population.

“Homologous” as used herein, refers to the subunit sequence similarity between two polymeric molecules, *e.g.*, between two nucleic acid molecules, *e.g.*, two DNA molecules or two RNA molecules, or between two polypeptide molecules. When a subunit position in both of the two molecules is occupied by the same monomeric subunit, *e.g.*, if a position in each of two DNA molecules is occupied by adenine, then they are homologous at that position. The homology between two sequences is a direct function of the number of matching or homologous positions, *e.g.*, if half (*e.g.*, five positions in a polymer ten subunits in length) of the positions in two compound sequences are homologous then the two sequences are 50% homologous, if 90% of the positions, *e.g.*, 9 of 10, are matched or homologous, the two sequences share 90% homology. By way of example, the DNA

sequences 3'-ATTGCC-5' and 3'-TATGGC-5' share 75% homology.

As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a polypeptide of the invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the invention.

A "coding region" of a gene consists of the nucleotide residues of the coding strand of the gene and the nucleotides of the non-coding strand of the gene which are homologous with or complementary to, respectively, the coding region of an mRNA molecule which is produced by transcription of the gene. A "coding region" of an mRNA molecule also consists of the nucleotide residues of the mRNA molecule which are matched with an anti-codon region of a transfer RNA molecule during translation of the mRNA molecule or which encode a stop codon. The coding region may thus include nucleotide residues corresponding to amino acid residues which are not present in the mature protein encoded by the mRNA molecule (*e.g.*, amino acid residues in a protein export signal sequence).

"Encoding" refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (*i.e.*, rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

Unless otherwise specified, a "nucleotide sequence encoding an amino acid sequence" includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

An "isolated nucleic acid" refers to a nucleic acid segment or fragment which has been separated from sequences which flank it in a naturally occurring state, e.g., a DNA fragment which has been removed from the sequences which are normally adjacent to the fragment, e.g., the sequences adjacent to the fragment in a genome in which it naturally  
5 occurs. The term also applies to nucleic acids, which have been substantially purified from other components, which naturally accompany the nucleic acid, e.g., RNA or DNA or proteins, which naturally accompany it in the cell. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a  
10 separate molecule (e.g., as a cDNA or a genomic or cDNA fragment produced by PCR or restriction enzyme digestion) independent of other sequences. It also includes a recombinant DNA, which is part of a hybrid gene encoding additional polypeptide sequence.

As used herein, the term "fragment," as applied to a nucleic acid, refers to a subsequence of a larger nucleic acid. A "fragment" of a nucleic acid can be at least about 15  
15 nucleotides in length; for example, at least about 50 nucleotides to about 100 nucleotides; at least about 100 to about 500 nucleotides, at least about 500 to about 1000 nucleotides, at least about 1000 nucleotides to about 1500 nucleotides; or about 1500 nucleotides to about 2500 nucleotides; or about 2500 nucleotides (and any integer value in between).

A "portion" of a polynucleotide means at least at least about five to about fifty  
20 sequential nucleotide residues of the polynucleotide. It is understood that a portion of a polynucleotide may include every nucleotide residue of the polynucleotide.

"Isolated" means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not "isolated," but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its  
25 natural state is "isolated." An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

"Naturally occurring" as used herein describes a composition that can be found in nature as distinct from being artificially produced. For example, a nucleotide sequence present in an organism, which can be isolated from a source in nature and which  
30 has not been intentionally modified by a person in the laboratory, is naturally occurring.

Unless otherwise specified, a "nucleotide sequence encoding an amino acid sequence" includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. The phrase nucleotide sequence that encodes a protein or an RNA may also include introns to the extent that the nucleotide sequence

encoding the protein may in some version contain an intron(s).

The term “polynucleotide” as used herein is defined as a chain of nucleotides. Furthermore, nucleic acids are polymers of nucleotides. Thus, nucleic acids and polynucleotides as used herein are interchangeable. One skilled in the art has the general knowledge that nucleic acids are polynucleotides, which can be hydrolyzed into the monomeric “nucleotides.” The monomeric nucleotides can be hydrolyzed into nucleosides. As used herein polynucleotides include, but are not limited to, all nucleic acid sequences which are obtained by any means available in the art, including, without limitation, recombinant means, i.e., the cloning of nucleic acid sequences from a recombinant library or a cell genome, using ordinary cloning technology and PCR™, and the like, and by synthetic means.

The terms “patient,” “subject,” “individual,” and the like are used interchangeably herein, and refer to any animal, or cells thereof whether *in vitro* or *in situ*, amenable to the methods described herein. Preferably, the patient, subject or individual is a mammal, and more preferable, a human.

“Variant” as the term is used herein, is a nucleic acid sequence or a peptide sequence that differs in sequence from a reference nucleic acid sequence or peptide sequence respectively, but retains essential properties of the reference molecule. Changes in the sequence of a nucleic acid variant may not alter the amino acid sequence of a peptide encoded by the reference nucleic acid, or may result in amino acid substitutions, additions, deletions, fusions and truncations. Changes in the sequence of peptide variants are typically limited or conservative, so that the sequences of the reference peptide and the variant are closely similar overall and, in many regions, identical. A variant and reference peptide can differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A variant of a nucleic acid or peptide can be a naturally occurring such as an allelic variant, or can be a variant that is not known to occur naturally. Non-naturally occurring variants of nucleic acids and peptides may be made by mutagenesis techniques or by direct synthesis.

Ranges: throughout this disclosure, various aspects of the invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be

considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. This applies regardless of the breadth of the range.

5

#### Description

The present invention provides compositions and methods for detecting nucleotide mutations. For example, the invention can be used to detect small-scale mutations including, but not limited to single nucleotide variants (SNVs), point mutations small  
10 insertions, and small deletions.

Single nucleotide polymorphisms (SNPs) are changes to a single DNA base in the genetic sequence that can lead to differences between individuals. Often, these SNPs appear in the transcribed portion of a gene and as such also appear as single nucleotide variants (SNVs) in the sequence of the transcribed RNA. The present invention relates to a  
15 method for detecting these SNVs on individual RNA molecules in single cells via fluorescence microscopy using *in situ* hybridization.

*In situ* hybridization (ISH) is a method that utilizes nucleic acid probes to detect DNA or RNA targets in cells via Watson-Crick base pairing of the probe to the target. A version of fluorescence-based ISH targeting RNA (RNA FISH) in which tens of  
20 fluorescently-labeled DNA oligonucleotide probes are used, each of which bind to different segments of the same RNA target, has been previously described (Raj et al., 2008, Nat Methods 5:877–879) and is incorporated by reference herein in its entirety. The method in Raj leads to a signal concentration at the target RNA, appearing as a bright spot in a fluorescence microscope. However, the RNA FISH method described in Raj does not provide  
25 the necessary discriminating power to distinguish small-scale mutations at the RNA level. This is because each DNA oligonucleotide probe will typically bind to a target despite the presence of a single or small number of mismatched, inserted or deleted nucleotides.

One method relating to *in situ* detection of SNVs on RNA molecules has been described demonstrating the use of signal amplification via a complex enzymatic scheme  
30 (Larsson et al., 2010, Nat Methods 7:395–397). However, the detection efficiency of this method is reported to be highly variable, leading to unreliable quantification.

One issue associated with detecting small-scale mutations, including a difference of a single base, via RNA FISH is that hybridization of relatively long oligonucleotide probes to the RNA may occur, despite not being fully complementary to the

target. Conversely, very short oligonucleotide probes, while able to discriminate between single base differences, will often fail to remain bound to the target due to reduced binding energy. Additionally, in either case, distinguishing signals indicating the presence of a mutation on a target RNA from a false positive result can be difficult or impossible when  
5 using only a single oligonucleotide probe. The method described herein employs a strategy to overcome these issues.

In order to distinguish between single-base mismatches, the method of the present invention includes a “toehold probe” strategy in which a single-stranded DNA oligonucleotide mutation-detection probe, designed to bind to the target RNA region  
10 including the mutation, is hybridized to a shorter DNA “mask” oligonucleotide. In one embodiment, the mask oligonucleotide blocks a portion of the detection probe from binding to the targeted RNA (Fig. 1A) (Zhang & Winfree, 2009, J Am Chem Soc 131:17303–17314; Zhang et al., 2012, Nat Chem 4:208–214; Li et al., 2002, Nucleic Acids Res 30:E5). The remaining single-stranded portion of the oligonucleotide mutation-detection probe may be of  
15 a short enough length to confer selectivity based on single base mismatches in the targeted RNA to which it selectively binds. Once the oligonucleotide mutation-detection probe is bound to the target RNA, the mask oligonucleotide dissociates from the mutation-detection probe, enabling the remainder of the mutation-detection probe length that was previously masked to bind to the target RNA. Thus, the method of the present invention comprises  
20 probes that have specificity for the desired mutation while still retaining a sufficient binding energy to prevent the detection probe from rapidly dissociating from the target after hybridization.

#### Targeted Nucleic Acid Sample

25 As contemplated herein, the present invention may be used in the analysis of any nucleic acid sample for which mutational analysis may be applied, as would be understood by those having ordinary skill in the art. For example, in certain embodiments, the nucleic acid can be mRNA. However, it should be appreciated that there is no limitation to the type of nucleic acid sample, which may include without limitation, any type of RNA, cDNA,  
30 genomic DNA, fragmented RNA or DNA and the like. In certain embodiments, the nucleic acid sample comprises at least one of messenger RNA, intronic RNA, exonic DNA, and non-coding RNA. The nucleic acid may be prepared for hybridization according to any manner as would be understood by those having ordinary skill in the art. It should also be appreciated that the sample may be an isolated nucleic acid sample, or it may form part of a lysed cell, or

it may be an intact living cell. Samples may further be individual cells, or a population of cells, such as a population of cells corresponding to a particular tissue. It should be appreciated that there is no limitation to the size or type of sample, provided the sample includes at least one nucleic acid therein. For example, the sample may be derived or  
5 obtained from one or more eukaryotic cells, prokaryotic cells, bacteria, virus, exosome, liposome, and the like.

#### RNA FISH Method

The present invention includes use of a highly sensitive and specific RNA  
10 FISH method to identify mutations in a targeted nucleic acid sequence. Additional description and explanation of RNA FISH methodologies can be found in copending patent application publication numbers WO/2010/030818 and WO 2012/106711, the entire contents of each are incorporated by reference herein in their entirety.

Probes useful in this invention may be DNA, RNA or mixtures of DNA and  
15 RNA. They may include non-natural nucleotides, and they may include non-natural internucleotide linkages. Non-natural nucleotides that increase the binding affinity of probes include 2' -O-methyl ribonucleotides, for example. The lengths of probes useful in this invention can be about 15-40 nucleotides for typical DNA or RNA probes of average binding affinity. In certain embodiments, the guide probes are about 10-50 nucleotides long, detection  
20 probes are about 10-50 nucleotides long, and masks (short oligonucleotides) are about 5-48 nucleotides long. If means are included to increase a probe's binding affinity, the probe can be shorter, as short as seven nucleotides, as persons in the art will appreciate. A fluorophore can be attached to a probe at any position, including, without limitation, attaching a fluorophore to one end of a probe, preferably to the 3' end. The probes may be included in a  
25 hybridization solution that contains the probes in excess.

A single cell can be probed simultaneously for multiple RNA target sequences, either more than one target sequence of one RNA molecule, or one or more sequences of different RNA molecules. Additionally, one target sequence of an RNA molecule can be probed with more than one set of probes, wherein each set is labeled with a  
30 distinguishable fluorophore, and the fluorophores are distinguishable. In one embodiment, the guide probe and the detection probe will have distinguishable fluorophores. Using more than one color for each of multiple targets permits the use of color-coding schemes in highly multiplexed probing methods, according to the present invention.

Methods of the present invention may also include determining if one or more

spots representing a target mutation sequence is present. Methods according to the present invention also include counting spots of a given color corresponding to a given RNA mutation species. When it is desired to detect more than one mutation, different sets of probes labeled with distinct fluorophores can be used in the same hybridization mixture.

5                   Spots can be detected utilizing microscopic methods. A confocal microscope, or a wide-field fluorescence microscope is sufficient. There is no limitation to the type of microscope used.

                  In another aspect, the present invention includes sets of probes for *in situ* hybridization that enable detection of SNVs on individual RNA molecules in cells. The probes render each molecule so intensely fluorescent that it can be seen as a fine fluorescent spot in fluorescence microscopy. A computer program can be used to identify and count all the RNA molecules in the cell from the microscopic image.

                  The invention also provides a kit, generally comprising a set of probes, an instruction manual for performing any of the methods contemplated herein, and optionally the computer -readable media as described herein.

#### Method

                  Accordingly, the present invention relates to a method for reliably detecting mutations using RNA FISH. The method can be generally described as including the following steps.

                  The method comprises detecting the target RNA of interest in a sample via a conventional RNA FISH method, without having any oligonucleotide probes binding to the region of the RNA containing the mutation. These initial oligonucleotide probes serve as control probes, and are referred to herein as guide probes, since the use of these probes determines the location of the RNA of interest, regardless of whether or not it contains a mutation. The guide probes of the present invention are labeled with a fluorophore that is spectrally distinct from any other types of fluorophores associated with other probes described herein.

                  In one embodiment, the method comprises construction of a second oligonucleotide probe that serves as a mutation detection oligonucleotide probe (the detection probe). In one embodiment, the detection probe is a 28-base probe. The detection probe of the present invention targets a mutation, and is labeled with a fluorophore that is spectrally-distinct from the guide probe. In order to make sure that the mutation detection oligonucleotide probe only binds to the target sequence containing the mutation, and not to

RNA molecules that do not contain the mutation, a DNA oligonucleotide mask is introduced that binds to a portion of the mutation detection oligonucleotide probe, as shown in Fig. 1. The unmasked region of the mutation detection oligonucleotide probe left exposed is relatively short, and will not bind to a sequence if there is at least one base mismatch in the target. The mutation detection oligonucleotide probe can be used to detect mutations in the RNA of interest.

In certain embodiments, the method comprises hybridization of the detection probe and mask. In certain embodiments, the hybridized detection probe/mask complex is added to the sample under reaction conditions suitable for the unmasked region of the detection probe to hybridize to the targeted RNA of interest.

In certain embodiments, the guide probe, detection probe, and oligonucleotide mask are simultaneously applied to the RNA of interest. For example in one embodiment, the method comprises applying a hybridization solution comprising the guide probe, detection probe, and oligonucleotide mask, to a sample. The method further comprise providing suitable conditions to allow for hybridization of the guide probe to the RNA, hybridization of the mask to the detection mask, and hybridization of the unmasked portion of the detection probe to a mutation-containing portion of the RNA.

After the detection probe hybridizes to the targeted RNA of interest (containing the mutation), the mask oligonucleotide denatures from the detection probe allowing the previously masked portion of the detection probe sequence to hybridize to the targeted RNA of interest.

In certain embodiments, the method of the invention comprises detecting more than one mutation in a sample. For example, in one embodiment, the method comprises contacting the sample with more than one detection probe, where each detection probe targets a different mutation, and where each detection probe is labeled with a fluorophore that is spectrally-distinct from the guide probe and spectrally-distinct from all other detection probes used.

The method of the present invention has a much higher efficiency than any methods presently described in the art for detecting small-scale mutations. Further, the method provides significantly more quantitative and reliable results than existing methods. In addition, as compared with biochemical methods such as sequencing, the method of the present invention may provide considerable cost advantages for detecting a particular mutation cheaply and effectively in a particular population of cells. Additionally, the method can work well in tissue sections.

For example, in existing methods, the use of a single probe can lead to a large number of false-positive signals, because every off-target binding event is indistinguishable from on-target binding. Such false-positives can be avoided by relying on the colocalization of multiple probes (Raj et al., 2008, Nat Methods 5:877–879), but colocalization is typically not possible when only a single probe can be used, as is the case in detection of a mutation in a single nucleotide or small number of nucleotides. Therefore, the method of the present invention uses multiple oligonucleotide control probes, collectively referred to as “guide” probes, that bind to the target RNA, thereby robustly identifying the target RNA. Detection probe signals are only considered as legitimate if they colocalize with a guide probe signal, thereby clearly distinguishing false positive signals from true positives (Fig. 1A).

As with any *in situ* hybridization probe, the mutation detection oligonucleotide probes, i.e. detection probes, of the present invention may have a considerable amount of off-target binding, which would be impossible to distinguish from the signals indicating the presence of a mutation on the RNA of interest. However, the method of the present invention utilizes the fact that the RNA of interest, or target RNA, has been separately labeled using at least one guide probe, i.e. an initial oligonucleotide control probe. Therefore, any detection probes that are bound to the correct target will colocalize with a guide probe, whereas detection probes that do not colocalize with a guide probe are binding off-target. These colocalized spots, comprising a fluorophore from at least one guide probe and a fluorophore from a detection probe, can be determined using computational methods described herein for detecting spots in microscope images, thereby allowing identification of target RNA that have the mutation of interest.

## EXPERIMENTAL EXAMPLES

The invention is further described in detail by reference to the following experimental examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified. Thus, the invention should in no way be construed as being limited to the following examples, but rather, should be construed to encompass any and all variations which become evident as a result of the teaching provided herein.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the present invention and practice the claimed methods. The following working examples

therefore, specifically point out the preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

Example 1: Single cell allele-specific expression via single nucleotide variant detection *in*

5 *situ*

To demonstrate the efficacy of the method of the present invention, a series of melanoma cell lines harboring a well-known mutation in the BRAF oncogene were utilized. Cell lines used were homozygous mutant, heterozygous mutant/wild-type and homozygous wild-type in a mutation of the 1799 position from T to A. Two detection probes were  
10 designed for this particular SNV, one targeting the mutant and one targeting wild-type transcripts. A mask oligonucleotide was utilized that was common to both. In the homozygous mutant cell line, it was found that roughly 56% of the RNA identified by the guide probe colocalized with signals from the mutant detection probe, whereas only 7% of the guide probe signals colocalized with the wild-type detection probe. Conversely, in the  
15 homozygous wild-type cell line, it was found that 58% of guide probe signals colocalized with the wild-type detection probe whereas only 7% of the guide probe signals colocalized with the mutant detection probe. In the heterozygous mutant/wild-type cell line, it was found that 36% of BRAF transcripts colocalized with the wild-type detection probe while 29% colocalized with the mutant detection probe, indicating that both copies of the genes  
20 transcribe equivalently in these cells. Overall, it was found that the colocalization efficiency was around 60%, roughly in line with other estimates of efficiency of hybridization of DNA oligonucleotides to RNA (Lubeck & Cai, 2012, Nat Methods 9:743–748). It was also found that the presence of the wild-type probe improves specificity of the mutant detection probe and vice-versa. The mask oligonucleotide is preferred for maintaining this specificity, as the  
25 number of false positive detections observed increased when detection was performed without the mask present (Fig. 4A). This approach works for a variety of different target sequence mismatches (Fig. 4B).

The method for detecting SNVs on RNA molecules provided the ability to measure differences in transcription from the maternal vs. paternal copy of a gene, both at the  
30 population and single cell level. This was explored using the GM12878 cell line, for which complete genetic phase information is available (1000 Genomes Project Consortium et al., 2010, Nature 467:1061–1073), making it ideal for studies involving allele-specific expression (Gertz et al., 2011, PLoS Genet 7:e1002228; Rozowsky et al., 2011, Mol Syst Biol 7:522). First, population-level imbalances were examined in maternal vs. paternal transcript

abundance. It was found that two of the genes (SKA3 and DNMT1) displayed no imbalance, whereas EBF1 and SUZ12 favored transcription from the maternal and paternal chromosome, respectively (Fig. 2A). Through the use of a statistical model, the number of mRNAs that must be classified with this method in order to make the determination that an imbalance exists were computed (Fig. 2B). Strikingly, the ability to make this determination does not depend strongly on the detection efficiency of the SNV detection probes of this method (Fig. 2B).

A search for single cell imbalances that may not be visible at the population level was then conducted (Fig. 2C). A statistical test was used to demonstrate whether the ensemble of individual cells displayed any signs of single cell imbalance.

To examine the effect of the toehold length on detection efficiency, the mask probe length was varied. As shown in Figure 5, changing the toehold length can change the detection efficiency without dramatically increasing off-target binding. Toehold length is in nucleotides, with the total probe length remaining constant (toehold length changed by changing the mask probe length). The free energy change of the toehold binding (given in kcal/mol) was computed using parameters from Sugimoto et al. Biochemistry 34: 11211–11216 (1995).

Another application of the method is to distinguish transcription from the maternal vs. paternal chromosomes *in situ*. A set of probes were developed targeting introns of a set of 31 genes along chromosome 19, yielding an RNA-based chromosome “paint” (Levesque & Raj, 2013, Nat Methods, doi:10.1038/nmeth.2372). A database of SNVs in GM12878 cells was used to find SNVs in the introns of these genes and created a set of detection probes designed to label 15 of the introns from the paternal chromosomes in a distinct color. In this manner, chromosomes were visualized and classified as being either maternal or paternal *in situ*.

Accordingly, the ability to distinguish SNVs with high efficiency and specificity at the level of individual RNA molecules has been demonstrated. The method is simple to implement and uses readily available reagents. Without wishing to be bound by any particular theory, using different nucleic acid chemistries for the detection probe increases the detection efficiency. Among other applications, the method may be used to study the effects of SNVs on gene expression.

#### *Cell culture and fixation*

Melanoma cell lines were grown with the BRAF V600E mutation, SK-MEL-28 (Mut/Mut, ATCC #HTB-72), WM3918 (WT/WT) and WM398b & WM9 (both WT/Mut) (gifts from the lab of Meenhard Herlyn, Wistar Institute), using the recommended cell culture guidelines for each line. The cells were grown on Lab-Tek chambered coverglass (Lab-Tek) and fixed the cells following the protocol in Raj (Raj et al., 2008, Nat Methods 5:877–879). GM12878 cells were obtained from the Coriell Cell Repositories and were grown according to guidelines. Fixed cells were stored in 70% ethanol at 4°C for up to 4 weeks before hybridization; the duration of storage did not affect hybridization efficiency.

#### 10 *Probe design and synthesis*

Detection probes were designed with the single nucleotide difference located at the 5th base position from their 5' end. The total length of the detection oligonucleotide was adjusted to ensure the hybridization energy with target RNA was similar or greater than that of the guide probe oligonucleotides. Mask oligonucleotides complementary to the detection probes were designed that, upon binding to the detection probe, left a 5 to 11 base toehold regions available to target RNAs regions with SNVs. Guide probe oligonucleotides were conjugated to ATTO 488 dye (ATTO-TEC) and Cy3 and Cy5 (GE Healthcare) dyes were used interchangeably for the SNV detection probes. Changes to detection efficiency were not observed when swapping the Cy3/Cy5 dyes. The choice of dyes was influenced by dye stability after a post-fixation step described below and affinities of some dyes that cause excessive binding to the incorrect target. The detection, mask, and guide probe sequences are listed in Example 2 and Example 3.

#### *RNA FISH*

RNA fluorescence in situ hybridization (FISH) was performed as outlined in Raj (Raj et al., 2008, Nat Methods 5:877–879) with some modifications as outlined presently. Firstly, the hybridization buffer consisted of 10% dextran sulfate, 2x saline-sodium citrate (SSC) and 10% formamide (Lubeck & Cai, 2012, Nat Methods 9:743–748). The hybridization was performed as in Raj, using final concentrations of 5nM for the guide probe, wild-type and mutant detection probe, and 10nM for the mask, thereby leading to 1:1 mask:detection oligonucleotide ratios. The hybridization was allowed to proceed overnight at 37°C. For Lab-Tek chamber samples, 50µL hybridization solution was used with a coverslip and included a moistened paper towel to prevent excessive evaporation in a para-filmed culture dish. For suspension cells, 50uL hybridization solution was used in a 1.5mL

Eppendorf tube. In the morning, the samples were washed twice with a 2X SSC and 10% formamide wash buffer. Suspension cells included 0.1% Triton-X in the wash and fixation buffers. A postfixation step was then performed using 4% formaldehyde in 2X SSC for 30 minutes at 25°C to crosslink the detection probes and thereby prevent dissociation during  
5 imaging, followed by 2 washes in 2X SSC. The cells were then put into anti-fade buffer with catalase and glucose oxidase as described (Raj et al., 2008, Nat Methods 5:877–879) to prevent photobleaching of Cy5 during imaging. For the chromosome 19 paints, probes against introns of 31 genes with 12-16 oligonucleotides per gene were used, each at 0.1nM, for the guide probe in Cy3 (Levesque & Raj, 2013, Nat Methods, doi:10.1038/nmeth.2372).  
10 Maternal and paternal probes were added, in Cy3 and Cy5 respectively, for 19 SNV sites within 15 of the chromosome 19 paint genes, added masks, and performed hybridization as described elsewhere herein.

### *Imaging*

15 All images were taken on a Leica DMI600B automated widefield fluorescence microscope equipped with a 100x Plan Apo objective, a Pixis 1024BR cooled CCD camera and a Prior Lumen 220 light source. Image stacks were taken in each fluorescence channel consisting of sets of images separated by 0.35µm. The exposure times were 1500ms and 3500ms for guide and detection probes respectively. Longer exposure times were used for  
20 the wild-type and mutant detection probes owing to the low signal afforded by single dye molecules relative to the dozens of fluorophores typically used in the guide probes.

### *Image analysis*

The image analysis involved first manually segmenting the cells using custom  
25 software written in MATLAB (Mathworks), after which spots were identified using algorithms. Relatively permissive thresholds were chosen for spots in the channels for the SNV detection probe channels, thereby trying to avoid false negatives due to overly stringent criteria for spot detection. Once the spots were located, spots were denoted as colocalized if two spots from different fluorescence channels were within 4 pixels of each other in order to  
30 account for a ~2 pixel chromatic aberration in portions of the images from the different channels. In the event of a colocalization event in which spots appeared in more than 2 channels or in which more than 2 spots were in the neighborhood of the guide probe, colocalized pairs in the rest of the image were used as a measure of how to correct for shifts between channels, thereby allowing the colocalization window to be tightened.

Example 2: Guide Probe Sequences

| Probe         | Sequence              | Identifier    |
|---------------|-----------------------|---------------|
| SUZ12_mRNA_2  | gaggaaaagctcgtggtcag  | SEQ ID NO: 1  |
| SUZ12_mRNA_3  | gatctgtgttggttctcaa   | SEQ ID NO: 2  |
| SUZ12_mRNA_4  | gagattccgagttcgaagaa  | SEQ ID NO: 3  |
| SUZ12_mRNA_5  | tgtgcaaaaatattggtgct  | SEQ ID NO: 4  |
| SUZ12_mRNA_6  | tctggagtttcgatgagaca  | SEQ ID NO: 5  |
| SUZ12_mRNA_7  | tgagattcttgcctccttt   | SEQ ID NO: 6  |
| SUZ12_mRNA_8  | ctgcaaatgagctgacaagc  | SEQ ID NO: 7  |
| SUZ12_mRNA_9  | tggagaaccagtaaacgt    | SEQ ID NO: 8  |
| SUZ12_mRNA_10 | aagagtgaactgcaacgtag  | SEQ ID NO: 9  |
| SUZ12_mRNA_11 | gcaataggagccgtagattt  | SEQ ID NO: 10 |
| SUZ12_mRNA_12 | atttctagtggcaagaggtt  | SEQ ID NO: 11 |
| SUZ12_mRNA_13 | taactgaaccaggcttggtt  | SEQ ID NO: 12 |
| SUZ12_mRNA_14 | acagcaatagtttgagtagg  | SEQ ID NO: 13 |
| SUZ12_mRNA_15 | tgttgccttgattgtgtt    | SEQ ID NO: 14 |
| SUZ12_mRNA_16 | caggtcatctcttgcctcag  | SEQ ID NO: 15 |
| SUZ12_mRNA_17 | tcagagtacaccaagggcaa  | SEQ ID NO: 16 |
| SUZ12_mRNA_18 | aaactataaagtttgcggca  | SEQ ID NO: 17 |
| SUZ12_mRNA_19 | tggcagagttaagatgctt   | SEQ ID NO: 18 |
| SUZ12_mRNA_20 | cctagcaccttttgatgat   | SEQ ID NO: 19 |
| SUZ12_mRNA_21 | ggagccatcataacactcat  | SEQ ID NO: 20 |
| SUZ12_mRNA_22 | tatcctgaggatttctgca   | SEQ ID NO: 21 |
| SUZ12_mRNA_23 | tgcgactaaaagcaaatcca  | SEQ ID NO: 22 |
| SUZ12_mRNA_24 | ggtgttctcttaactgtgcc  | SEQ ID NO: 23 |
| SUZ12_mRNA_25 | gcctgcacacaagaatatgt  | SEQ ID NO: 24 |
| SUZ12_mRNA_26 | catgcttgcttttgctggt   | SEQ ID NO: 25 |
| SUZ12_mRNA_27 | ctttgctgttctactccccc  | SEQ ID NO: 26 |
| SUZ12_mRNA_28 | aaatacagacgattgtggcc  | SEQ ID NO: 27 |
| SUZ12_mRNA_29 | agaggtgaagcaggtatcact | SEQ ID NO: 28 |
| SUZ12_mRNA_30 | gacatggagattccagagtt  | SEQ ID NO: 29 |
| SUZ12_mRNA_31 | cagcaataaacccatgcttc  | SEQ ID NO: 30 |
| SUZ12_mRNA_32 | caggcatgattcatttgatt  | SEQ ID NO: 31 |
| SUZ12_mRNA_33 | tgaagcatgaagtttcgaca  | SEQ ID NO: 32 |
| SUZ12_mRNA_34 | aaagtcatgcatgctgacta  | SEQ ID NO: 33 |
| SUZ12_mRNA_35 | catttcacggagcttggttaa | SEQ ID NO: 34 |
| SUZ12_mRNA_36 | tatttctcgtttgcagggg   | SEQ ID NO: 35 |
| SUZ12_mRNA_37 | ccatttgctgtcccattttg  | SEQ ID NO: 36 |
| SUZ12_mRNA_38 | ctgtttgaaacccctgaga   | SEQ ID NO: 37 |
| SUZ12_mRNA_39 | acatgggggttagagcttttc | SEQ ID NO: 38 |
| SUZ12_mRNA_40 | agaggatgaattccctaaaa  | SEQ ID NO: 39 |
| SUZ12_mRNA_41 | tgaagtagaacctgataca   | SEQ ID NO: 40 |
| SUZ12_mRNA_42 | cctcccaagaaaatgtctc   | SEQ ID NO: 41 |

|               |                       |               |
|---------------|-----------------------|---------------|
| SUZ12_mRNA_43 | aggatcaaagtttgactgca  | SEQ ID NO: 42 |
| SUZ12_mRNA_44 | gggtgagcaatgcactaaaa  | SEQ ID NO: 43 |
| SUZ12_mRNA_46 | caaatgcgttcttctctgg   | SEQ ID NO: 44 |
| SUZ12_mRNA_47 | ttctcccttataagtgaca   | SEQ ID NO: 45 |
| SUZ12_mRNA_49 | acacatataacacagggcaa  | SEQ ID NO: 46 |
| SUZ12_mRNA_50 | caactgcaaatatgtgcgtg  | SEQ ID NO: 47 |
| SUZ12_mRNA_51 | tgcttgtaatgtgccagta   | SEQ ID NO: 48 |
| SUZ12_mRNA_52 | cggagttggaataaaaacct  | SEQ ID NO: 49 |
| SUZ12_mRNA_53 | gatgttactcaaccacagtg  | SEQ ID NO: 50 |
| SUZ12_mRNA_54 | acacatcttaaagaccagtc  | SEQ ID NO: 51 |
| SUZ12_mRNA_55 | tcgttaaatagcctcacagt  | SEQ ID NO: 52 |
| SUZ12_mRNA_56 | tgacaaatcacatccacact  | SEQ ID NO: 53 |
| SUZ12_mRNA_57 | aatgaaagctgcagttccc   | SEQ ID NO: 54 |
| SUZ12_mRNA_58 | gcttaccaatcaaggaatct  | SEQ ID NO: 55 |
| SUZ12_mRNA_59 | ccagaggcaaaaatcagagt  | SEQ ID NO: 56 |
| SUZ12_mRNA_60 | cgagataaacgctcgagatc  | SEQ ID NO: 57 |
| SUZ12_mRNA_61 | tatgtgcacagcttagcaa   | SEQ ID NO: 58 |
| SUZ12_mRNA_62 | ttctacactacatctcccc   | SEQ ID NO: 59 |
| SUZ12_mRNA_63 | agcattaagagcataactgc  | SEQ ID NO: 60 |
| SUZ12_mRNA_64 | gcaaacaatgctagccttct  | SEQ ID NO: 61 |
| SUZ12_mRNA_65 | gggtgggaatcaccaactttt | SEQ ID NO: 62 |
| DNMT1_mRNA_1  | gataccctgtgcagaaggat  | SEQ ID NO: 63 |
| DNMT1_mRNA_2  | gatgtaccaaacggagagag  | SEQ ID NO: 64 |
| DNMT1_mRNA_3  | cattcacttcccgttgtaa   | SEQ ID NO: 65 |
| DNMT1_mRNA_4  | ttggttcccgtttctagac   | SEQ ID NO: 66 |
| DNMT1_mRNA_5  | ctctacgggcttcacttctt  | SEQ ID NO: 67 |
| DNMT1_mRNA_6  | gaggtttgaaaggggttg    | SEQ ID NO: 68 |
| DNMT1_mRNA_7  | ctggtctttgtcttctct    | SEQ ID NO: 69 |
| DNMT1_mRNA_8  | cgttctctggatgtaactct  | SEQ ID NO: 70 |
| DNMT1_mRNA_9  | ttctcgtctccatcttctg   | SEQ ID NO: 71 |
| DNMT1_mRNA_10 | gttttgcgtctcttctctc   | SEQ ID NO: 72 |
| DNMT1_mRNA_11 | agttcatgactgtttggcg   | SEQ ID NO: 73 |
| DNMT1_mRNA_12 | agcttctcattgtcagcat   | SEQ ID NO: 74 |
| DNMT1_mRNA_13 | gactcgttggcatcaaagat  | SEQ ID NO: 75 |
| DNMT1_mRNA_14 | ctgaagcaggctcagttgtg  | SEQ ID NO: 76 |
| DNMT1_mRNA_15 | gtgaccgtgcttacagtaca  | SEQ ID NO: 77 |
| DNMT1_mRNA_16 | tattcttctcgatgaggccg  | SEQ ID NO: 78 |
| DNMT1_mRNA_18 | ctccatcaaagccagtgatc  | SEQ ID NO: 79 |
| DNMT1_mRNA_19 | ccatcagaatgtattcgga   | SEQ ID NO: 80 |
| DNMT1_mRNA_20 | caaatatggcgcatactcg   | SEQ ID NO: 81 |
| DNMT1_mRNA_21 | tcataggtcgagtcggaatt  | SEQ ID NO: 82 |
| DNMT1_mRNA_22 | ctcgatcttgtgatcaggt   | SEQ ID NO: 83 |
| DNMT1_mRNA_23 | gttcaagttgaggccagaag  | SEQ ID NO: 84 |
| DNMT1_mRNA_24 | ctcgtcataactctccacct  | SEQ ID NO: 85 |
| DNMT1_mRNA_25 | gtagaatgcctgatggtctg  | SEQ ID NO: 86 |

|               |                       |                |
|---------------|-----------------------|----------------|
| DNMT1_mRNA_26 | cctttcaatttgcctcgcg   | SEQ ID NO: 87  |
| DNMT1_mRNA_27 | cttaaaggcgttctcctgt   | SEQ ID NO: 88  |
| DNMT1_mRNA_28 | tccactgccaccaaatttaa  | SEQ ID NO: 89  |
| DNMT1_mRNA_29 | cttcattggccatattgggac | SEQ ID NO: 90  |
| DNMT1_mRNA_30 | gacttctcatcgtcatctg   | SEQ ID NO: 91  |
| DNMT1_mRNA_31 | ggcatctctgggatgttacc  | SEQ ID NO: 92  |
| DNMT1_mRNA_32 | cttctccgacccaagagatg  | SEQ ID NO: 93  |
| BRAF_mRNA_2   | ttgataatccacacctctc   | SEQ ID NO: 94  |
| BRAF_mRNA_3   | tgtccaatagggcctctata  | SEQ ID NO: 95  |
| BRAF_mRNA_5   | gctgggtgtattctcatagg  | SEQ ID NO: 96  |
| BRAF_mRNA_6   | tctttgttgagtgcatcta   | SEQ ID NO: 97  |
| BRAF_mRNA_7   | gttccccagagattccaata  | SEQ ID NO: 98  |
| BRAF_mRNA_8   | gtatccattgatgcagagct  | SEQ ID NO: 99  |
| BRAF_mRNA_9   | ttgttgggcaggaagactct  | SEQ ID NO: 100 |
| BRAF_mRNA_10  | actgtaactccacaccttgc  | SEQ ID NO: 101 |
| BRAF_mRNA_11  | agtgccttctttagactgtc  | SEQ ID NO: 102 |
| BRAF_mRNA_12  | ttctctccatcctgaattct  | SEQ ID NO: 103 |
| BRAF_mRNA_13  | agtgtcccaaccaattgggt  | SEQ ID NO: 104 |
| BRAF_mRNA_14  | tccacatgcaattcttctcc  | SEQ ID NO: 105 |
| BRAF_mRNA_15  | tggattgggtggtgttcaa   | SEQ ID NO: 106 |
| BRAF_mRNA_16  | cttcatctgctggctggaag  | SEQ ID NO: 107 |
| BRAF_mRNA_18  | ggggttagcagacaaacctgt | SEQ ID NO: 108 |
| BRAF_mRNA_19  | tcacgttagttagtgagcca  | SEQ ID NO: 109 |
| BRAF_mRNA_20  | tgaggctctggagattctg   | SEQ ID NO: 110 |
| BRAF_mRNA_21  | attctgtctctgaggatg    | SEQ ID NO: 111 |
| BRAF_mRNA_22  | ctcccaatcatcactcgagt  | SEQ ID NO: 112 |
| BRAF_mRNA_23  | tctttgtcccaactgtaatt  | SEQ ID NO: 113 |
| BRAF_mRNA_24  | ttcactgccacatcaccatg  | SEQ ID NO: 114 |
| BRAF_mRNA_25  | gtaggtgctgtcacattcaa  | SEQ ID NO: 115 |
| BRAF_mRNA_26  | tcacatgctgtgtttctctg  | SEQ ID NO: 116 |
| BRAF_mRNA_27  | caatagccagttgtggcttt  | SEQ ID NO: 117 |
| BRAF_mRNA_28  | tgggtatacaagctggagcc  | SEQ ID NO: 118 |
| BRAF_mRNA_32  | gcattctgatgactctggt   | SEQ ID NO: 119 |
| BRAF_mRNA_33  | tccctgttgatgtttga     | SEQ ID NO: 120 |
| BRAF_mRNA_34  | gcttttggacagttactccg  | SEQ ID NO: 121 |
| BRAF_mRNA_35  | tcaaggagggttctgatgca  | SEQ ID NO: 122 |
| BRAF_mRNA_36  | tctctgtttggaaaccagc   | SEQ ID NO: 123 |
| BRAF_mRNA_37  | tctgaactctcactcat     | SEQ ID NO: 124 |
| BRAF_mRNA_38  | aagcctctagaagagctct   | SEQ ID NO: 125 |
| BRAF_mRNA_40  | atctgttcagttgccttat   | SEQ ID NO: 126 |
| SKA3_mRNA_1   | atcttcaaagtcgcttcct   | SEQ ID NO: 127 |
| SKA3_mRNA_2   | catcatcctttagagtctga  | SEQ ID NO: 128 |
| SKA3_mRNA_3   | ccttcttgatttccaatct   | SEQ ID NO: 129 |
| SKA3_mRNA_4   | gtacttttggcctttatg    | SEQ ID NO: 130 |
| SKA3_mRNA_5   | gacacgtggactatatccat  | SEQ ID NO: 131 |

|              |                       |                |
|--------------|-----------------------|----------------|
| SKA3_mRNA_6  | tcttgctcgtgtactgaatt  | SEQ ID NO: 132 |
| SKA3_mRNA_7  | ctctgggtcagagttaatgg  | SEQ ID NO: 133 |
| SKA3_mRNA_8  | atcttcacatcagtcttct   | SEQ ID NO: 134 |
| SKA3_mRNA_9  | cttgcaacaggaggatcaga  | SEQ ID NO: 135 |
| SKA3_mRNA_10 | actacgtggagacttctcag  | SEQ ID NO: 136 |
| SKA3_mRNA_11 | tgtggagggttggtagaac   | SEQ ID NO: 137 |
| SKA3_mRNA_12 | ctcttccttatagtgttca   | SEQ ID NO: 138 |
| SKA3_mRNA_13 | agtgattgttggtagggtgg  | SEQ ID NO: 139 |
| SKA3_mRNA_14 | gtgcacattttggagttttt  | SEQ ID NO: 140 |
| SKA3_mRNA_15 | tccattgtgtaattctcat   | SEQ ID NO: 141 |
| SKA3_mRNA_16 | cctcacttttattatctc    | SEQ ID NO: 142 |
| SKA3_mRNA_17 | ctggattctgtatctatggc  | SEQ ID NO: 143 |
| SKA3_mRNA_18 | gctgggagtggtcaaaacat  | SEQ ID NO: 144 |
| SKA3_mRNA_19 | ccaggagtacagaatgtagg  | SEQ ID NO: 145 |
| SKA3_mRNA_20 | gtggataccaaagctatgct  | SEQ ID NO: 146 |
| SKA3_mRNA_21 | agtacgatcttcaactcca   | SEQ ID NO: 147 |
| SKA3_mRNA_22 | cgtagggtgaagagggtctg  | SEQ ID NO: 148 |
| SKA3_mRNA_23 | gcagggcaatgtgaatgta   | SEQ ID NO: 149 |
| SKA3_mRNA_24 | gggtcaacgtttaaggggga  | SEQ ID NO: 150 |
| SKA3_mRNA_25 | atgactgggtacatgtcaa   | SEQ ID NO: 151 |
| SKA3_mRNA_26 | ctactctctgttactactta  | SEQ ID NO: 152 |
| SKA3_mRNA_27 | cagacagcactgagcaagta  | SEQ ID NO: 153 |
| SKA3_mRNA_28 | ctccaatgggtcacagtatat | SEQ ID NO: 154 |
| SKA3_mRNA_29 | aactctcaacctgacagcag  | SEQ ID NO: 155 |
| SKA3_mRNA_30 | aagagaaaagacccccacac  | SEQ ID NO: 156 |
| SKA3_mRNA_31 | cagctgccatgtgatagaaa  | SEQ ID NO: 157 |
| SKA3_mRNA_32 | aaattcagggtctactgggt  | SEQ ID NO: 158 |
| SKA3_mRNA_33 | catagctaacagtgcacagg  | SEQ ID NO: 159 |
| SKA3_mRNA_34 | taagcagggtcaacctgaagc | SEQ ID NO: 160 |
| SKA3_mRNA_35 | aatcagaaggatcgtaggcc  | SEQ ID NO: 161 |
| SKA3_mRNA_36 | cgggtacacagtgaaggctg  | SEQ ID NO: 162 |
| EBF1_mRNA_1  | ctctgctcaaaactgagcga  | SEQ ID NO: 163 |
| EBF1_mRNA_2  | aagttggattccgcagatt   | SEQ ID NO: 164 |
| EBF1_mRNA_3  | tattggcttttcgctgttg   | SEQ ID NO: 165 |
| EBF1_mRNA_4  | tatccattgctgtagagaa   | SEQ ID NO: 166 |
| EBF1_mRNA_5  | tcaatgaggcgcacgtagaa  | SEQ ID NO: 167 |
| EBF1_mRNA_6  | acactatggcttgtttgtc   | SEQ ID NO: 168 |
| EBF1_mRNA_7  | caagactcggcacatttctg  | SEQ ID NO: 169 |
| EBF1_mRNA_8  | ttttctgtcacaacagcgg   | SEQ ID NO: 170 |
| EBF1_mRNA_9  | atctgagggtgtctcatttc  | SEQ ID NO: 171 |
| EBF1_mRNA_10 | tgatggcttgatacaggga   | SEQ ID NO: 172 |
| EBF1_mRNA_12 | aatgaccttctgtaacctct  | SEQ ID NO: 173 |
| EBF1_mRNA_13 | tgagtattacttctttggc   | SEQ ID NO: 174 |
| EBF1_mRNA_14 | atacagtgttctaccagat   | SEQ ID NO: 175 |
| EBF1_mRNA_15 | tccactgaacgaattcacgc  | SEQ ID NO: 176 |

|              |                        |                |
|--------------|------------------------|----------------|
| EBF1_mRNA_16 | gagttatagtggctctgctg   | SEQ ID NO: 177 |
| EBF1_mRNA_17 | aattggacattgcggcagag   | SEQ ID NO: 178 |
| EBF1_mRNA_18 | ggtgagaaggagaagatgcc   | SEQ ID NO: 179 |
| EBF1_mRNA_19 | tctgacgactggtgcgaaag   | SEQ ID NO: 180 |
| EBF1_mRNA_20 | acaatcatgccagatatcgc   | SEQ ID NO: 181 |
| EBF1_mRNA_21 | ctctgggacttgatcagat    | SEQ ID NO: 182 |
| EBF1_mRNA_22 | cctcttaaaaaggcctgagt   | SEQ ID NO: 183 |
| EBF1_mRNA_23 | ccttgatagagctttacgg    | SEQ ID NO: 184 |
| EBF1_mRNA_24 | tcttcctttacacagcttta   | SEQ ID NO: 185 |
| EBF1_mRNA_25 | tttaaggcgcaaaagccgac   | SEQ ID NO: 186 |
| EBF1_mRNA_26 | taaggggtggcatgttaagt   | SEQ ID NO: 187 |
| EBF1_mRNA_27 | acatcgcgttttaactttcc   | SEQ ID NO: 188 |
| EBF1_mRNA_28 | tagaggcactggattttcga   | SEQ ID NO: 189 |
| EBF1_mRNA_29 | cagagtgtggaattctgtgc   | SEQ ID NO: 190 |
| EBF1_mRNA_30 | gtgaggttttgcttgtaa     | SEQ ID NO: 191 |
| EBF1_mRNA_31 | gtctgcatttaggacgagta   | SEQ ID NO: 192 |
| EBF1_mRNA_32 | cgaaaatacctgccacgttg   | SEQ ID NO: 193 |
| EBF1_mRNA_33 | ctgatagaggcagtatctgg   | SEQ ID NO: 194 |
| EBF1_mRNA_34 | ccctttaactcttaattcca   | SEQ ID NO: 195 |
| EBF1_mRNA_35 | aatgcacctcagagcttt     | SEQ ID NO: 196 |
| EBF1_mRNA_36 | cgcacttttcacgtagcaaa   | SEQ ID NO: 197 |
| EBF1_mRNA_37 | tccaaaggagagattccat    | SEQ ID NO: 198 |
| EBF1_mRNA_38 | ctatgaagtttctccctaga   | SEQ ID NO: 199 |
| EBF1_mRNA_39 | tgcattgttaaggcatccaa   | SEQ ID NO: 200 |
| EBF1_mRNA_40 | ctggtgcacagttacaatgt   | SEQ ID NO: 201 |
| EBF1_mRNA_41 | agcctagtgtgaaaaccattgc | SEQ ID NO: 202 |
| EBF1_mRNA_42 | tttctgagtaccgagaagca   | SEQ ID NO: 203 |
| EBF1_mRNA_43 | tgacagatgggtagtgtctg   | SEQ ID NO: 204 |
| EBF1_mRNA_44 | agttctaaccactgcacatg   | SEQ ID NO: 205 |
| EBF1_mRNA_45 | aggctctttggactttcaag   | SEQ ID NO: 206 |
| EBF1_mRNA_46 | agtcatccagaagctgtatt   | SEQ ID NO: 207 |
| EBF1_mRNA_47 | ttaagctgcacctagtaca    | SEQ ID NO: 208 |
| EBF1_mRNA_48 | gaactgtacagtgtgtgtct   | SEQ ID NO: 209 |
| DHPS_int_1   | ctagtaccgcgttggttcta   | SEQ ID NO: 210 |
| DHPS_int_2   | cagagtgcaaaatccccttc   | SEQ ID NO: 211 |
| DHPS_int_3   | taccgtaccagcatgtaact   | SEQ ID NO: 212 |
| DHPS_int_4   | cacaccaagaatcagtcctc   | SEQ ID NO: 213 |
| DHPS_int_5   | gcccattccagaaagcttta   | SEQ ID NO: 214 |
| DHPS_int_6   | cctccatatectcccttaa    | SEQ ID NO: 215 |
| DHPS_int_7   | tctaaattcaagaaccgccc   | SEQ ID NO: 216 |
| DHPS_int_8   | atctcctttcagatccggtc   | SEQ ID NO: 217 |
| DHPS_int_9   | ggctccagaaacagatttca   | SEQ ID NO: 218 |
| DHPS_int_10  | taaatcccagactcaggact   | SEQ ID NO: 219 |
| DHPS_int_11  | aagggccaaagtaagttaag   | SEQ ID NO: 220 |
| DHPS_int_12  | ataactgcattgccattga    | SEQ ID NO: 221 |

|              |                       |                |
|--------------|-----------------------|----------------|
| DHPS_int_13  | tgactcttatgagggagctc  | SEQ ID NO: 222 |
| DHPS_int_14  | agtgaatttgcccaagaag   | SEQ ID NO: 223 |
| DHPS_int_15  | cagagatagtctgggaggag  | SEQ ID NO: 224 |
| DHPS_int_16  | gctaagtgttcctctactg   | SEQ ID NO: 225 |
| DNMT1_int_1  | gaatccacggccattttgg   | SEQ ID NO: 226 |
| DNMT1_int_2  | cttgctgtattggggatca   | SEQ ID NO: 227 |
| DNMT1_int_3  | catcgagatgcacagctttg  | SEQ ID NO: 228 |
| DNMT1_int_4  | gtgacatccgtctctggagg  | SEQ ID NO: 229 |
| DNMT1_int_5  | aaggagcaagaaccacacag  | SEQ ID NO: 230 |
| DNMT1_int_6  | aatgcacggttaaagttcct  | SEQ ID NO: 231 |
| DNMT1_int_7  | caggcacagatttacaggaa  | SEQ ID NO: 232 |
| DNMT1_int_8  | agccagttctcattagcaag  | SEQ ID NO: 233 |
| DNMT1_int_9  | acacactaaagaacacaccc  | SEQ ID NO: 234 |
| DNMT1_int_10 | gatccttgtgcacggaagtt  | SEQ ID NO: 235 |
| DNMT1_int_11 | aatgaactgatggcgttcat  | SEQ ID NO: 236 |
| DNMT1_int_12 | cacacctcacttgaacaagt  | SEQ ID NO: 237 |
| DNMT1_int_13 | gtgaggggttcctctgactca | SEQ ID NO: 238 |
| DNMT1_int_14 | tttcacaaatccagctggaa  | SEQ ID NO: 239 |
| DNMT1_int_15 | cccaaagacccaaatcagaa  | SEQ ID NO: 240 |
| DNMT1_int_16 | ggggttgaaccaaatacca   | SEQ ID NO: 241 |
| EGLN2_int_1  | aactgcctaaaccttctgtg  | SEQ ID NO: 242 |
| EGLN2_int_2  | gtccccacaagtaagcatac  | SEQ ID NO: 243 |
| EGLN2_int_3  | atcaggtgcacacattaagg  | SEQ ID NO: 244 |
| EGLN2_int_5  | gaatgtcagcagctctcatg  | SEQ ID NO: 245 |
| EGLN2_int_7  | gatggactagaacatgggc   | SEQ ID NO: 246 |
| EGLN2_int_8  | ccacccatgaagacaatgat  | SEQ ID NO: 247 |
| EGLN2_int_9  | cagaagcagaaccaagatg   | SEQ ID NO: 248 |
| EGLN2_int_10 | gctcagctatcaagtaacgg  | SEQ ID NO: 249 |
| EGLN2_int_11 | gtattccgtggatcagcaaa  | SEQ ID NO: 250 |
| EGLN2_int_12 | gtcccaaccacatagaaag   | SEQ ID NO: 251 |
| EGLN2_int_13 | tccttacttcctaggacaa   | SEQ ID NO: 252 |
| EGLN2_int_14 | cccatctaaaagcgggaaag  | SEQ ID NO: 253 |
| EGLN2_int_15 | gagtacaggagagagtcag   | SEQ ID NO: 254 |
| EGLN2_int_16 | cagaacgactaagaagcacg  | SEQ ID NO: 255 |
| EIF3K_int_1  | gattctctcgcttctaggcc  | SEQ ID NO: 256 |
| EIF3K_int_2  | cgaagagactgagtggtacc  | SEQ ID NO: 257 |
| EIF3K_int_3  | aaggaaaccttaaggcaatt  | SEQ ID NO: 258 |
| EIF3K_int_4  | atgtccacctgaacactctg  | SEQ ID NO: 259 |
| EIF3K_int_5  | tcctgaatgtctctgctact  | SEQ ID NO: 260 |
| EIF3K_int_6  | tcagtcactgcagcttgtag  | SEQ ID NO: 261 |
| EIF3K_int_7  | taggatgcctcctcaacctc  | SEQ ID NO: 262 |
| EIF3K_int_8  | aagctctaaactccactgga  | SEQ ID NO: 263 |
| EIF3K_int_9  | ttcttatcccagacctctcg  | SEQ ID NO: 264 |
| EIF3K_int_10 | gctctatccaggtagtgaat  | SEQ ID NO: 265 |
| EIF3K_int_11 | cagccaccttatggagcaag  | SEQ ID NO: 266 |

|               |                       |                |
|---------------|-----------------------|----------------|
| EIF3K_int_12  | agacagagagctagacactt  | SEQ ID NO: 267 |
| EIF3K_int_13  | ctagtgtctgcaatgggagt  | SEQ ID NO: 268 |
| EIF3K_int_14  | gctgcattgttcaggatact  | SEQ ID NO: 269 |
| EIF3K_int_15  | ctagtcttgacaccaagag   | SEQ ID NO: 270 |
| FBL_int_1     | gacctgctggaatcagaatc  | SEQ ID NO: 271 |
| FBL_int_2     | cctattagacggcctcaatg  | SEQ ID NO: 272 |
| FBL_int_3     | ctcctgccaatatccaaaa   | SEQ ID NO: 273 |
| FBL_int_4     | cagatgcctgaatccaaact  | SEQ ID NO: 274 |
| FBL_int_5     | caagcctgattccaaaaca   | SEQ ID NO: 275 |
| FBL_int_6     | ggtggaaatcttaatcccca  | SEQ ID NO: 276 |
| FBL_int_7     | cgagctgttaagtctctgc   | SEQ ID NO: 277 |
| FBL_int_8     | gagtggtttcagcagaatct  | SEQ ID NO: 278 |
| FBL_int_9     | accaccgagaaggattctaa  | SEQ ID NO: 279 |
| FBL_int_10    | tttcacacagatgagtgcg   | SEQ ID NO: 280 |
| FBL_int_11    | taggaaaacagacccttgg   | SEQ ID NO: 281 |
| FBL_int_12    | tcaagagatcccaaacacg   | SEQ ID NO: 282 |
| FBL_int_13    | atcacagaccagaatgcctg  | SEQ ID NO: 283 |
| FBL_int_14    | cattetaccacacatggagg  | SEQ ID NO: 284 |
| FBL_int_15    | gagctaacacctgacaactt  | SEQ ID NO: 285 |
| FBL_int_16    | ctcactcaggctaaaatcct  | SEQ ID NO: 286 |
| MARK4_int_1   | gaccttgaagaagccagaaa  | SEQ ID NO: 287 |
| MARK4_int_2   | cctgaagctgagaagttgat  | SEQ ID NO: 288 |
| MARK4_int_3   | caagggaaaagggttgaaa   | SEQ ID NO: 289 |
| MARK4_int_4   | gagaaagcttcagcagatt   | SEQ ID NO: 290 |
| MARK4_int_5   | aggtaaggggtctagaaat   | SEQ ID NO: 291 |
| MARK4_int_6   | agatgaataaaggctgagcc  | SEQ ID NO: 292 |
| MARK4_int_7   | ctggaagtatgggtaggaa   | SEQ ID NO: 293 |
| MARK4_int_8   | tcctaggaatcagagaaggg  | SEQ ID NO: 294 |
| MARK4_int_9   | ggaatggtggaaagtgacaa  | SEQ ID NO: 295 |
| MARK4_int_10  | tgatcagagacacaggagat  | SEQ ID NO: 296 |
| MARK4_int_11  | gcaggctcttggaaagtgatc | SEQ ID NO: 297 |
| MARK4_int_12  | tggggagaagtctaggattg  | SEQ ID NO: 298 |
| MARK4_int_13  | gatctgcaagatgaggaagg  | SEQ ID NO: 299 |
| MARK4_int_14  | actccaaattggagtcttg   | SEQ ID NO: 300 |
| MARK4_int_15  | attgtagtaccaaggaaca   | SEQ ID NO: 301 |
| MARK4_int_16  | ctgaatcgagtaagccttgg  | SEQ ID NO: 302 |
| PPP2R1A_int_1 | caaccggggagataagagac  | SEQ ID NO: 303 |
| PPP2R1A_int_2 | cctactggagcaagtcatg   | SEQ ID NO: 304 |
| PPP2R1A_int_3 | aattagatggcaggccttc   | SEQ ID NO: 305 |
| PPP2R1A_int_4 | aaaatgagaggcggaggaag  | SEQ ID NO: 306 |
| PPP2R1A_int_5 | cgctctcttaggacacctaa  | SEQ ID NO: 307 |
| PPP2R1A_int_6 | gctcctaaacttgctagtc   | SEQ ID NO: 308 |
| PPP2R1A_int_7 | tatcctggtcaatgggagga  | SEQ ID NO: 309 |
| PPP2R1A_int_8 | gcttagcaaatccctcaacc  | SEQ ID NO: 310 |
| PPP2R1A_int_9 | catcccataaccaggaatgt  | SEQ ID NO: 311 |

|                |                       |                |
|----------------|-----------------------|----------------|
| PPP2R1A_int_10 | cctctttaatcaccactccc  | SEQ ID NO: 312 |
| PPP2R1A_int_11 | aacagacctaaagggaggat  | SEQ ID NO: 313 |
| PPP2R1A_int_12 | gtttggcaggttaccagtg   | SEQ ID NO: 314 |
| PPP2R1A_int_13 | tataccaggaacctaggagg  | SEQ ID NO: 315 |
| PPP2R1A_int_14 | tccccagcatcatatctcat  | SEQ ID NO: 316 |
| PPP2R1A_int_15 | tatagcaactggtgtctcca  | SEQ ID NO: 317 |
| PPP2R1A_int_16 | cctgtttcacatctggatcc  | SEQ ID NO: 318 |
| PTBP1_int_1    | gaatgcgaacatctccagc   | SEQ ID NO: 319 |
| PTBP1_int_2    | aaacttctcaggaaaacgga  | SEQ ID NO: 320 |
| PTBP1_int_3    | ctcttctgacaccacagact  | SEQ ID NO: 321 |
| PTBP1_int_5    | gaacacagcctcagttactg  | SEQ ID NO: 322 |
| PTBP1_int_6    | ctgaaactggcaaactcaca  | SEQ ID NO: 323 |
| PTBP1_int_7    | gtgctttccagtaagttgga  | SEQ ID NO: 324 |
| PTBP1_int_8    | cacgttccaagacaaaagaca | SEQ ID NO: 325 |
| PTBP1_int_9    | ccacttgactgcaacttgaa  | SEQ ID NO: 326 |
| PTBP1_int_10   | cgctagagaaaagctcagaag | SEQ ID NO: 327 |
| PTBP1_int_11   | cggagaagcaaagtgagaag  | SEQ ID NO: 328 |
| PTBP1_int_12   | aactccagattccagaccaa  | SEQ ID NO: 329 |
| PTBP1_int_13   | tgacagcaagaaccgaagag  | SEQ ID NO: 330 |
| PTBP1_int_14   | taggctggattctatccagg  | SEQ ID NO: 331 |
| PTBP1_int_15   | tcaaccagtaaatgcccatc  | SEQ ID NO: 332 |
| PTBP1_int_16   | cccttctcacaatgctgag   | SEQ ID NO: 333 |
| RPL18A_int_1   | gttgggtgcaacaagagaag  | SEQ ID NO: 334 |
| RPL18A_int_2   | ctatgctgcgcgacttatc   | SEQ ID NO: 335 |
| RPL18A_int_3   | tttcatctgcttctcacagc  | SEQ ID NO: 336 |
| RPL18A_int_4   | tatgtaccacagcgttaagc  | SEQ ID NO: 337 |
| RPL18A_int_5   | ccatagagccgttgattct   | SEQ ID NO: 338 |
| RPL18A_int_6   | agtccaggttctctatctc   | SEQ ID NO: 339 |
| RPL18A_int_7   | agctggagatctggacataa  | SEQ ID NO: 340 |
| RPL18A_int_8   | cccttgacagcaaggaaacc  | SEQ ID NO: 341 |
| RPL18A_int_9   | gttcagggaagggaacaatgg | SEQ ID NO: 342 |
| RPL18A_int_10  | ttttactgtgaacctgaccc  | SEQ ID NO: 343 |
| RPL18A_int_11  | aaaccacctctgaaactgac  | SEQ ID NO: 344 |
| RPL18A_int_12  | aatctttggtaagtccagg   | SEQ ID NO: 345 |
| RPL18A_int_13  | ggtttacagatgcagaggtg  | SEQ ID NO: 346 |
| RPL18A_int_14  | aactgcaatccaaacgtttg  | SEQ ID NO: 347 |
| RPL18A_int_15  | agaactaggacaagacctca  | SEQ ID NO: 348 |
| RPL18A_int_16  | catcttctttcacctgagg   | SEQ ID NO: 349 |
| RPS19_int_1    | tctggatcgactaacagag   | SEQ ID NO: 350 |
| RPS19_int_2    | catcctaaaccgtgtgtaccc | SEQ ID NO: 351 |
| RPS19_int_3    | ggagaaagtcaagcatgtga  | SEQ ID NO: 352 |
| RPS19_int_4    | tttgaacctcagtccecaaa  | SEQ ID NO: 353 |
| RPS19_int_5    | gtacaaaagagaggtggaac  | SEQ ID NO: 354 |
| RPS19_int_6    | cctcaacacaaactatgctgt | SEQ ID NO: 355 |
| RPS19_int_7    | ctaccccatatcccaaatgc  | SEQ ID NO: 356 |

|               |                       |                |
|---------------|-----------------------|----------------|
| RPS19_int_8   | cacgattcagtcattccac   | SEQ ID NO: 357 |
| RPS19_int_10  | gaagtatggttgccagg     | SEQ ID NO: 358 |
| RPS19_int_12  | caagtgtgacacaaccaag   | SEQ ID NO: 359 |
| RPS19_int_13  | tcgaatgtcacatcacaaa   | SEQ ID NO: 360 |
| RPS19_int_14  | aaaaacttgagtagcaagt   | SEQ ID NO: 361 |
| RPS19_int_15  | ttcatctgtcttggttcc    | SEQ ID NO: 362 |
| RPS19_int_16  | aaaccacctgtaagcaaat   | SEQ ID NO: 363 |
| RPS9_int_1    | gcttactcatggaaactcgg  | SEQ ID NO: 364 |
| RPS9_int_2    | tcatagtcagtattgcccc   | SEQ ID NO: 365 |
| RPS9_int_3    | atccgatctcgcgagaataa  | SEQ ID NO: 366 |
| RPS9_int_4    | gagagaagtgtgagcgtaag  | SEQ ID NO: 367 |
| RPS9_int_5    | atagagtacatgggcacctt  | SEQ ID NO: 368 |
| RPS9_int_6    | gtaccaaatttagggacgg   | SEQ ID NO: 369 |
| RPS9_int_7    | tggaatcacaaaaccttct   | SEQ ID NO: 370 |
| RPS9_int_8    | cgtactggcacaacaactag  | SEQ ID NO: 371 |
| RPS9_int_9    | gggagaatgaacctcacaa   | SEQ ID NO: 372 |
| RPS9_int_10   | gacacaactctcactcgg    | SEQ ID NO: 373 |
| RPS9_int_11   | aaggctggtccatttacc    | SEQ ID NO: 374 |
| RPS9_int_12   | tttctacttcacaagtccc   | SEQ ID NO: 375 |
| RPS9_int_13   | acctaagaaacaggcacaag  | SEQ ID NO: 376 |
| RPS9_int_14   | aggcctatcttagctctgg   | SEQ ID NO: 377 |
| RPS9_int_15   | gcccagaaattccactcat   | SEQ ID NO: 378 |
| RPS9_int_16   | aacctgttatcaccatcac   | SEQ ID NO: 379 |
| SLC1A5_int_1  | gaggactcactgagcgaaag  | SEQ ID NO: 380 |
| SLC1A5_int_2  | tgcattttccaggaaactaa  | SEQ ID NO: 381 |
| SLC1A5_int_3  | tgagcccgattctcattga   | SEQ ID NO: 382 |
| SLC1A5_int_4  | aattaaaactcacaggagcc  | SEQ ID NO: 383 |
| SLC1A5_int_5  | atgccaagctaacaatgctc  | SEQ ID NO: 384 |
| SLC1A5_int_6  | gtgtccatcgttaccaggcc  | SEQ ID NO: 385 |
| SLC1A5_int_7  | taggcaaagaggtagagccc  | SEQ ID NO: 386 |
| SLC1A5_int_8  | aaggactgcagagtgtaaat  | SEQ ID NO: 387 |
| SLC1A5_int_9  | acaaagtagagacctatcca  | SEQ ID NO: 388 |
| SLC1A5_int_10 | cacctgggggtgggaaaagag | SEQ ID NO: 389 |
| SLC1A5_int_11 | gagagggcagcatggaatgg  | SEQ ID NO: 390 |
| SLC1A5_int_12 | cagtttgagcaggttgaggg  | SEQ ID NO: 391 |
| SLC1A5_int_13 | aaaggacactcagtctacct  | SEQ ID NO: 392 |
| SLC1A5_int_14 | ctgtgggcaaggaacagatc  | SEQ ID NO: 393 |
| SLC1A5_int_15 | caaacagaatgccccgcacc  | SEQ ID NO: 394 |
| SLC1A5_int_16 | gtgaatagagggtgccccat  | SEQ ID NO: 395 |
| SUPT5H_int_1  | tctcttgagacaatctggga  | SEQ ID NO: 396 |
| SUPT5H_int_2  | ccttgcttgactcgaactt   | SEQ ID NO: 397 |
| SUPT5H_int_3  | tgttatctcactggacctga  | SEQ ID NO: 398 |
| SUPT5H_int_4  | cttttttagggggtggtgg   | SEQ ID NO: 399 |
| SUPT5H_int_5  | ctctgatecaaccaaagtgg  | SEQ ID NO: 400 |
| SUPT5H_int_6  | ggaacacagtagtagatgca  | SEQ ID NO: 401 |

|               |                       |                |
|---------------|-----------------------|----------------|
| SUPT5H_int_7  | accagacacctgagaagtaa  | SEQ ID NO: 402 |
| SUPT5H_int_8  | ctggttgtgattgctacct   | SEQ ID NO: 403 |
| SUPT5H_int_9  | gtgcatcaaacaagggatct  | SEQ ID NO: 404 |
| SUPT5H_int_10 | ggcaactaacatatacctggg | SEQ ID NO: 405 |
| SUPT5H_int_11 | acacagccacatgaaatctt  | SEQ ID NO: 406 |
| SUPT5H_int_12 | gcactctccatctaccaaac  | SEQ ID NO: 407 |
| SUPT5H_int_13 | taagtgactgggacaagtca  | SEQ ID NO: 408 |
| SUPT5H_int_14 | ggcaatcaaatgtccacaga  | SEQ ID NO: 409 |
| SUPT5H_int_15 | cactctccaaagggtcacaat | SEQ ID NO: 410 |
| SUPT5H_int_16 | atccagtgtacagcttaga   | SEQ ID NO: 411 |
| TOMM40_int_1  | atcacctcctagtgtctgta  | SEQ ID NO: 412 |
| TOMM40_int_2  | ccatcattctaagcccaag   | SEQ ID NO: 413 |
| TOMM40_int_3  | ctgctatgacattccatccc  | SEQ ID NO: 414 |
| TOMM40_int_4  | gatcacaagagaagtctggc  | SEQ ID NO: 415 |
| TOMM40_int_5  | gaacaccagaaacagaactcc | SEQ ID NO: 416 |
| TOMM40_int_6  | cacagaattgtcccaaggat  | SEQ ID NO: 417 |
| TOMM40_int_7  | aagctgagcaactttagggtc | SEQ ID NO: 418 |
| TOMM40_int_8  | ttgcttgctcaaatctcact  | SEQ ID NO: 419 |
| TOMM40_int_9  | tacagcctagttagacaggg  | SEQ ID NO: 420 |
| TOMM40_int_10 | gattccacctgtaagaggc   | SEQ ID NO: 421 |
| TOMM40_int_11 | cacagcagaacatctcacia  | SEQ ID NO: 422 |
| TOMM40_int_12 | gagaagagaaaacgctgtcac | SEQ ID NO: 423 |
| TOMM40_int_13 | tctggcactatatactccgag | SEQ ID NO: 424 |
| TOMM40_int_14 | taggtcccaactgagtag    | SEQ ID NO: 425 |
| TOMM40_int_15 | agagactcagagaaaggagg  | SEQ ID NO: 426 |
| TOMM40_int_16 | tgtgacgatactggacagat  | SEQ ID NO: 427 |
| UBA52_int_1   | ctaaagtcagcacaaccac   | SEQ ID NO: 428 |
| UBA52_int_2   | gagaagcaagggaacacag   | SEQ ID NO: 429 |
| UBA52_int_3   | caaacgttttcagatcaca   | SEQ ID NO: 430 |
| UBA52_int_4   | ggccaactgaggtagaagat  | SEQ ID NO: 431 |
| UBA52_int_5   | cactacccccagtttctcaa  | SEQ ID NO: 432 |
| UBA52_int_6   | tattactggcagtgctctct  | SEQ ID NO: 433 |
| UBA52_int_7   | gaggctcagttagaggctct  | SEQ ID NO: 434 |
| UBA52_int_8   | caatgctcctttcctaggac  | SEQ ID NO: 435 |
| UBA52_int_9   | acactgaattcttgctgctc  | SEQ ID NO: 436 |
| UBA52_int_10  | aaggtcagacactgaagtct  | SEQ ID NO: 437 |
| UBA52_int_11  | ccgacctctaagtgggtcag  | SEQ ID NO: 438 |
| UBA52_int_12  | gcatecatctgggtttctaa  | SEQ ID NO: 439 |
| UBA52_int_13  | ggagtctgagactgacacat  | SEQ ID NO: 440 |
| UBA52_int_14  | cgtgtggaagatacactgtc  | SEQ ID NO: 441 |
| UBA52_int_15  | gcctatagtctgctgcttc   | SEQ ID NO: 442 |
| UBA52_int_16  | caagcatcggagcacacata  | SEQ ID NO: 443 |
| ZNF444_int_1  | gagtcacactgggtttcagg  | SEQ ID NO: 444 |
| ZNF444_int_2  | cttctctgataagccgtgac  | SEQ ID NO: 445 |
| ZNF444_int_3  | gagaggacagctggtaactg  | SEQ ID NO: 446 |

|               |                       |                |
|---------------|-----------------------|----------------|
| ZNF444_int_4  | tttgaacacattgggggtcc  | SEQ ID NO: 447 |
| ZNF444_int_5  | gtgccactactgaaaggatg  | SEQ ID NO: 448 |
| ZNF444_int_6  | gactgctctgactcttcacc  | SEQ ID NO: 449 |
| ZNF444_int_7  | ggtagcacttatgagggaac  | SEQ ID NO: 450 |
| ZNF444_int_8  | tctctgctgctacatctcag  | SEQ ID NO: 451 |
| ZNF444_int_9  | catgagaaggagacggatg   | SEQ ID NO: 452 |
| ZNF444_int_10 | atcccagacaataagagggg  | SEQ ID NO: 453 |
| ZNF444_int_11 | cgaggatagagaagccagag  | SEQ ID NO: 454 |
| ZNF444_int_12 | cccacttttgggaacaatga  | SEQ ID NO: 455 |
| ZNF444_int_13 | gctgcgtttgtgatttgta   | SEQ ID NO: 456 |
| ZNF444_int_14 | ggatgaaagcagaggtaag   | SEQ ID NO: 457 |
| ZNF444_int_15 | ggatgaaagcagaggtaag   | SEQ ID NO: 458 |
| ZNF444_int_16 | taagtgggtcaaggtcagag  | SEQ ID NO: 459 |
| ZNF91_int_1   | gattgtggagctgactgaag  | SEQ ID NO: 460 |
| ZNF91_int_2   | catcttatcgctgaagggga  | SEQ ID NO: 461 |
| ZNF91_int_3   | ctgcacaatctgggagagac  | SEQ ID NO: 462 |
| ZNF91_int_4   | gagttaggctggagggaacag | SEQ ID NO: 463 |
| ZNF91_int_5   | tggttaagatagctgcgtcta | SEQ ID NO: 464 |
| ZNF91_int_6   | actgaagacacatcaccta   | SEQ ID NO: 465 |
| ZNF91_int_7   | tccaagaaaaactgaaggg   | SEQ ID NO: 466 |
| ZNF91_int_8   | agagaatatgaccagaagc   | SEQ ID NO: 467 |
| ZNF91_int_9   | tcaatacctcaggtgtcct   | SEQ ID NO: 468 |
| ZNF91_int_10  | gtccacacttgagaagctaa  | SEQ ID NO: 469 |
| ZNF91_int_11  | cactattttctgcccccta   | SEQ ID NO: 470 |
| ZNF91_int_12  | gcaagtcttacgccateta   | SEQ ID NO: 471 |
| ZNF91_int_13  | gtgcctcaggcacattatac  | SEQ ID NO: 472 |
| ZNF91_int_14  | aggagactctgaactatgcc  | SEQ ID NO: 473 |
| ZNF91_int_15  | ttaagtgtcaataaccccc   | SEQ ID NO: 474 |
| ZNF91_int_16  | tcaagtcaggccattcaatt  | SEQ ID NO: 475 |
| AMH_int_1     | cacctgaaggaagacggcgg  | SEQ ID NO: 476 |
| AMH_int_2     | ctgatctccctaggaggac   | SEQ ID NO: 477 |
| AMH_int_3     | tcctaggaacaagaccacag  | SEQ ID NO: 478 |
| AMH_int_5     | ttcggggacaggacagagtg  | SEQ ID NO: 479 |
| AMH_int_6     | aggggcaagtctaagagctg  | SEQ ID NO: 480 |
| AMH_int_7     | cagggagtatcttaggccag  | SEQ ID NO: 481 |
| AMH_int_9     | ccgggaatcttgggatctg   | SEQ ID NO: 482 |
| AMH_int_10    | tgagccattgaaggccacac  | SEQ ID NO: 483 |
| AMH_int_12    | gacctgcaacaaatccatg   | SEQ ID NO: 484 |
| AMH_int_13    | ctacgtgttctcagtactg   | SEQ ID NO: 485 |
| AMH_int_16    | tggaacccgcaattggagga  | SEQ ID NO: 486 |
| AXL_int_1     | ggagatgagacagagaggag  | SEQ ID NO: 487 |
| AXL_int_2     | aggggagaagccagagaaagg | SEQ ID NO: 488 |
| AXL_int_3     | tctagagacttaacaggtcg  | SEQ ID NO: 489 |
| AXL_int_4     | cgacatagttgcagagtgc   | SEQ ID NO: 490 |
| AXL_int_5     | ggaagagttagcttaggaac  | SEQ ID NO: 491 |

|                |                      |                |
|----------------|----------------------|----------------|
| AXL_int_6      | tgtatgacgctaggttcac  | SEQ ID NO: 492 |
| AXL_int_7      | acccaaaactctgacattct | SEQ ID NO: 493 |
| AXL_int_8      | taacttatatggactcccag | SEQ ID NO: 494 |
| AXL_int_9      | ctttcaccacagaacatata | SEQ ID NO: 495 |
| AXL_int_10     | tgatgcctaggacatccaac | SEQ ID NO: 496 |
| AXL_int_11     | tatctgaagaccagaatggc | SEQ ID NO: 497 |
| AXL_int_12     | catctcattctatggatgag | SEQ ID NO: 498 |
| AXL_int_13     | cagaatgtctgacagctcat | SEQ ID NO: 499 |
| AXL_int_14     | atacccaacacacacgtgac | SEQ ID NO: 500 |
| AXL_int_15     | attgacttgccttatctgat | SEQ ID NO: 501 |
| AXL_int_16     | caagatgcctgactcaccac | SEQ ID NO: 502 |
| BCL3_int_1     | atatctatataccttggaa  | SEQ ID NO: 503 |
| BCL3_int_2     | taagagctgtactttcttg  | SEQ ID NO: 504 |
| BCL3_int_3     | tcataacgctacacatctgg | SEQ ID NO: 505 |
| BCL3_int_4     | atatcaggatacagggaa   | SEQ ID NO: 506 |
| BCL3_int_5     | tgagattgtgacatcctgag | SEQ ID NO: 507 |
| BCL3_int_6     | taactgagtaccacagaa   | SEQ ID NO: 508 |
| BCL3_int_7     | cctgatcttgacctcaata  | SEQ ID NO: 509 |
| BCL3_int_8     | aggggtgcagaaactac    | SEQ ID NO: 510 |
| BCL3_int_9     | tctactcttggaataggaa  | SEQ ID NO: 511 |
| BCL3_int_10    | actgtttaaaaggggccac  | SEQ ID NO: 512 |
| BCL3_int_11    | ttaccagaagcctgttcaa  | SEQ ID NO: 513 |
| BCL3_int_12    | gtgggataaatgtagatgg  | SEQ ID NO: 514 |
| BCL3_int_13    | gattttgtggctgatgaatt | SEQ ID NO: 515 |
| BCL3_int_14    | gacaaagggtcccttaagaa | SEQ ID NO: 516 |
| BCL3_int_15    | acttctaaaactttgctcct | SEQ ID NO: 517 |
| BCL3_int_16    | aatggaggacagagaaggga | SEQ ID NO: 518 |
| ATF5_int_1     | gagaagagactgtgatccga | SEQ ID NO: 519 |
| ATF5_int_2     | gctgacaaacatcgagtcta | SEQ ID NO: 520 |
| ATF5_int_3     | aacaaccaagcacgaagaat | SEQ ID NO: 521 |
| ATF5_int_4     | ggccaattgcctttacacc  | SEQ ID NO: 522 |
| ATF5_int_5     | ctgtcaatctaagccagtc  | SEQ ID NO: 523 |
| ATF5_int_6     | cacttagagggttgcctctg | SEQ ID NO: 524 |
| ATF5_int_7     | gtgacgacagtgggtctatc | SEQ ID NO: 525 |
| ATF5_int_8     | tcacaatgtgcttccgaaat | SEQ ID NO: 526 |
| ATF5_int_9     | cctcaactaacggttggttt | SEQ ID NO: 527 |
| ATF5_int_10    | gcctctttgtcacatcgaaa | SEQ ID NO: 528 |
| ATF5_int_11    | agcccttcagcaatgttata | SEQ ID NO: 529 |
| ATF5_int_12    | cagaatatttcgctcaagg  | SEQ ID NO: 530 |
| ATF5_int_13    | tcccagggaattaccttag  | SEQ ID NO: 531 |
| ATF5_int_14    | gtcctaaagtaggaccgatg | SEQ ID NO: 532 |
| ATF5_int_15    | actcccaagggaatttcttt | SEQ ID NO: 533 |
| ATF5_int_16    | ctacacgcgtgagacaagag | SEQ ID NO: 534 |
| C19orf48_int_1 | acacagccggaaaaagtaat | SEQ ID NO: 535 |
| C19orf48_int_2 | cagctataggagtcgcaaac | SEQ ID NO: 536 |

|                 |                       |                |
|-----------------|-----------------------|----------------|
| C19orf48_int_3  | gaaagaactcaatgtcgcct  | SEQ ID NO: 537 |
| C19orf48_int_4  | agcacgaaaatggaaacact  | SEQ ID NO: 538 |
| C19orf48_int_5  | ttcctaaggagcacctgta   | SEQ ID NO: 539 |
| C19orf48_int_6  | ctagggtgcacaaaatgcc   | SEQ ID NO: 540 |
| C19orf48_int_7  | actcgataccctatctgtgc  | SEQ ID NO: 541 |
| C19orf48_int_8  | cggcaattcacctagaaact  | SEQ ID NO: 542 |
| C19orf48_int_9  | atttctagggtggtgcctg   | SEQ ID NO: 543 |
| C19orf48_int_10 | aatattgccaaaggacgacg  | SEQ ID NO: 544 |
| C19orf48_int_11 | aggcacaagcagattgtta   | SEQ ID NO: 545 |
| C19orf48_int_12 | ataatgaggagagaccctcg  | SEQ ID NO: 546 |
| C19orf48_int_13 | ttgcagggtttcttagcact  | SEQ ID NO: 547 |
| C19orf48_int_14 | gcactacatgaagatgtccc  | SEQ ID NO: 548 |
| C19orf48_int_15 | actcatccagaaggtaggaa  | SEQ ID NO: 549 |
| C19orf48_int_16 | gaggaaacagtctgcaagtc  | SEQ ID NO: 550 |
| FTL_int_1       | cgaccgcacaaagaaggctg  | SEQ ID NO: 551 |
| FTL_int_2       | atggaggcaacaatggtag   | SEQ ID NO: 552 |
| FTL_int_3       | tgggagatgtagtccattac  | SEQ ID NO: 553 |
| FTL_int_4       | tcaaatcaaggctcctcgcg  | SEQ ID NO: 554 |
| FTL_int_5       | caagaccgaactcaatctcc  | SEQ ID NO: 555 |
| FTL_int_6       | ctattccagcggtaagag    | SEQ ID NO: 556 |
| FTL_int_7       | taagaccacgcagcgggtgtg | SEQ ID NO: 557 |
| FTL_int_8       | ctagctcttacagctatacg  | SEQ ID NO: 558 |
| FTL_int_9       | tgacagcttgattatcac    | SEQ ID NO: 559 |
| FTL_int_10      | aaacctacattccaagag    | SEQ ID NO: 560 |
| FTL_int_11      | cacacactcggcacatagaa  | SEQ ID NO: 561 |
| FTL_int_12      | gaaggggagaaatggctcaga | SEQ ID NO: 562 |
| FTL_int_13      | atggatgcagcgggtacgta  | SEQ ID NO: 563 |
| FTL_int_14      | aaatttgcccaaaggagca   | SEQ ID NO: 564 |
| FTL_int_15      | gtgtgaaatgaggctctgaa  | SEQ ID NO: 565 |
| FTL_int_16      | cagccagttgcagattaaaa  | SEQ ID NO: 566 |
| HNRNPL_int_1    | aactgtgatggttcaagg    | SEQ ID NO: 567 |
| HNRNPL_int_2    | ttggtttatcaaatctcgg   | SEQ ID NO: 568 |
| HNRNPL_int_3    | cgaatgtctccaacatggc   | SEQ ID NO: 569 |
| HNRNPL_int_4    | ctccaaaagcaacgactggt  | SEQ ID NO: 570 |
| HNRNPL_int_5    | caactcgtaaccctaaacc   | SEQ ID NO: 571 |
| HNRNPL_int_6    | gggtagaattggattctatt  | SEQ ID NO: 572 |
| HNRNPL_int_7    | tgagaatcactgggcagaag  | SEQ ID NO: 573 |
| HNRNPL_int_8    | cttcctaagagaagtatctt  | SEQ ID NO: 574 |
| HNRNPL_int_9    | cccataaccattaagaaccga | SEQ ID NO: 575 |
| HNRNPL_int_10   | ttagaaagaccagagactcc  | SEQ ID NO: 576 |
| HNRNPL_int_11   | caaattcctgccctttaac   | SEQ ID NO: 577 |
| HNRNPL_int_12   | tcactcgaccaaggaaaca   | SEQ ID NO: 578 |
| HNRNPL_int_13   | caagtcccaaatcttacgtt  | SEQ ID NO: 579 |
| HNRNPL_int_14   | atgcacaggaaagactaggg  | SEQ ID NO: 580 |
| HNRNPL_int_15   | accagtcagataaaacagt   | SEQ ID NO: 581 |

|               |                      |                |
|---------------|----------------------|----------------|
| HNRNPL_int_16 | ccagaaatcgtagcacaca  | SEQ ID NO: 582 |
| MRPL34_int_1  | ccaaaccacaactcgtagt  | SEQ ID NO: 583 |
| MRPL34_int_2  | agactaaggacacgtacagg | SEQ ID NO: 584 |
| MRPL34_int_3  | gtttgcattccaaaagcaa  | SEQ ID NO: 585 |
| MRPL34_int_4  | gcaacttcgttaagggaact | SEQ ID NO: 586 |
| MRPL34_int_5  | tactcagagtctaattttgt | SEQ ID NO: 587 |
| MRPL34_int_6  | cttaacgttctgatgaccag | SEQ ID NO: 588 |
| MRPL34_int_7  | caagaatgtagctttcctg  | SEQ ID NO: 589 |
| MRPL34_int_8  | tcacatacatatcctgtat  | SEQ ID NO: 590 |
| MRPL34_int_9  | aagtgaacagtcagtaggt  | SEQ ID NO: 591 |
| MRPL34_int_10 | aatggtgtactcatttatct | SEQ ID NO: 592 |
| MRPL34_int_11 | ctgatcagcagacagtataa | SEQ ID NO: 593 |
| MRPL34_int_12 | aaatgagttgcttcaccatt | SEQ ID NO: 594 |
| MRPL34_int_13 | cttttaattccacatctag  | SEQ ID NO: 595 |
| MRPL34_int_14 | tggtctacacttttaacac  | SEQ ID NO: 596 |
| MRPL34_int_15 | tacacttttaacacctccg  | SEQ ID NO: 597 |
| MRPL34_int_16 | cctatgctcatgagaataca | SEQ ID NO: 598 |
| MYADM_int_1   | ctagccacaggaaagtaact | SEQ ID NO: 599 |
| MYADM_int_2   | aacgcagccacaggactaaa | SEQ ID NO: 600 |
| MYADM_int_3   | ttgcaaacctacagcttcc  | SEQ ID NO: 601 |
| MYADM_int_4   | aaccaatttgcttcttagt  | SEQ ID NO: 602 |
| MYADM_int_5   | ttccaactgtccagtaggaa | SEQ ID NO: 603 |
| MYADM_int_6   | ccagcttctgggaaactcag | SEQ ID NO: 604 |
| MYADM_int_7   | aaaaagctagttgagcctc  | SEQ ID NO: 605 |
| MYADM_int_8   | tgagatttaggactggagta | SEQ ID NO: 606 |
| MYADM_int_9   | ttggagattagcaaccttca | SEQ ID NO: 607 |
| MYADM_int_10  | cttctgacaaggagcctaaa | SEQ ID NO: 608 |
| MYADM_int_11  | ggaaaaccaagagatctccc | SEQ ID NO: 609 |
| MYADM_int_12  | gatatcttctcatagagatc | SEQ ID NO: 610 |
| MYADM_int_13  | agggacccacatctaataag | SEQ ID NO: 611 |
| MYADM_int_14  | gaacagtgacctagacattt | SEQ ID NO: 612 |
| MYADM_int_15  | tgtgagcaaaacttccttct | SEQ ID NO: 613 |
| MYADM_int_16  | aaaacgcttgcttccacag  | SEQ ID NO: 614 |
| RPL13A_int_1  | tcattatggatgtaagctc  | SEQ ID NO: 615 |
| RPL13A_int_2  | cttaccacaaaagacttcca | SEQ ID NO: 616 |
| RPL13A_int_3  | aaatgccacattgtagccaa | SEQ ID NO: 617 |
| RPL13A_int_4  | gtgaaatcaaagtagcccg  | SEQ ID NO: 618 |
| RPL13A_int_5  | cggcaatcccaaacgacg   | SEQ ID NO: 619 |
| RPL13A_int_6  | atacttaagtagctacacca | SEQ ID NO: 620 |
| RPL13A_int_7  | caaagaaaccttccagggc  | SEQ ID NO: 621 |
| RPL13A_int_8  | gaacacaactgcctagatc  | SEQ ID NO: 622 |
| RPL13A_int_9  | aaaatccagacagacttct  | SEQ ID NO: 623 |
| RPL13A_int_10 | ttacacatgcttgaagcg   | SEQ ID NO: 624 |
| RPL13A_int_11 | gattaggcagagtagtcaaa | SEQ ID NO: 625 |
| RPL13A_int_12 | ttctctctaattgtatagga | SEQ ID NO: 626 |

|               |                       |                |
|---------------|-----------------------|----------------|
| RPL13A_int_13 | ccagttacaagacctgttaa  | SEQ ID NO: 627 |
| RPL13A_int_14 | tgaacgaagaccagcaaggc  | SEQ ID NO: 628 |
| RPL13A_int_15 | ctatgagtattttatgctca  | SEQ ID NO: 629 |
| RPL13A_int_16 | gggagacaaggcgaaaagga  | SEQ ID NO: 630 |
| RPL18_int_1   | agtgacaggtccagataaag  | SEQ ID NO: 631 |
| RPL18_int_2   | atcgaacccgactttcata   | SEQ ID NO: 632 |
| RPL18_int_3   | tactgtgtaccaacgaaca   | SEQ ID NO: 633 |
| RPL18_int_4   | tcagactgtgtgacaatta   | SEQ ID NO: 634 |
| RPL18_int_5   | ctttctgagcaacttaagtt  | SEQ ID NO: 635 |
| RPL18_int_6   | catcagtcacccaacaacc   | SEQ ID NO: 636 |
| RPL18_int_7   | cgaatccttattcagaacct  | SEQ ID NO: 637 |
| RPL18_int_8   | atcaaaactgggtccatac   | SEQ ID NO: 638 |
| RPL18_int_9   | gcacacagaaccctgaaaca  | SEQ ID NO: 639 |
| RPL18_int_10  | tgatccctctgtaaaagg    | SEQ ID NO: 640 |
| RPL18_int_11  | ctagaatggaggatctgcaa  | SEQ ID NO: 641 |
| RPL18_int_12  | aacagaaccagagacccgag  | SEQ ID NO: 642 |
| RPL18_int_13  | acacatgatgatctggacta  | SEQ ID NO: 643 |
| RPL18_int_14  | acgggaagacagtgagaagc  | SEQ ID NO: 644 |
| RPL18_int_15  | cgcgactggctgagacaaac  | SEQ ID NO: 645 |
| RPL18_int_16  | agcagagcctttgacataaa  | SEQ ID NO: 646 |
| RPL28_int_1   | catcgtgaataaggcgaccg  | SEQ ID NO: 647 |
| RPL28_int_2   | ggcccaaaccagacaaaact  | SEQ ID NO: 648 |
| RPL28_int_3   | ttctgcctgacagtaaaaa   | SEQ ID NO: 649 |
| RPL28_int_4   | aacaaccaccgaggaagttc  | SEQ ID NO: 650 |
| RPL28_int_5   | gacttgcttgagaatttca   | SEQ ID NO: 651 |
| RPL28_int_6   | tgacagcacaatgtatggag  | SEQ ID NO: 652 |
| RPL28_int_7   | aatccccaaaaaggcctgca  | SEQ ID NO: 653 |
| RPL28_int_8   | ccgactcacacgaacattca  | SEQ ID NO: 654 |
| RPL28_int_9   | gaacacatgtgaagctggac  | SEQ ID NO: 655 |
| RPL28_int_10  | ttgctgccactgagtaaaag  | SEQ ID NO: 656 |
| RPL28_int_11  | catgtattcggcaatccaag  | SEQ ID NO: 657 |
| RPL28_int_12  | caatgggtatgaacagcttc  | SEQ ID NO: 658 |
| RPL28_int_13  | ctgtccacatgaaacact    | SEQ ID NO: 659 |
| RPL28_int_14  | gacagtctcacattgagaag  | SEQ ID NO: 660 |
| RPL28_int_15  | catattgcaaccagtagtgg  | SEQ ID NO: 661 |
| RPL28_int_16  | aggctgaaagcaagaatcca  | SEQ ID NO: 662 |
| RPS28_int_2   | aggaaagaagctgggttcag  | SEQ ID NO: 663 |
| RPS28_int_4   | agaggtcacatcagggttta  | SEQ ID NO: 664 |
| RPS28_int_5   | gaaatgagaatcggtcccg   | SEQ ID NO: 665 |
| RPS28_int_6   | aaagggaaccatcagaaccc  | SEQ ID NO: 666 |
| RPS28_int_7   | cggggccagaatacagatcg  | SEQ ID NO: 667 |
| RPS28_int_8   | agcagatatgaattcagaga  | SEQ ID NO: 668 |
| RPS28_int_10  | tggaggaagaagtaaggcgcg | SEQ ID NO: 669 |
| RPS28_int_12  | caccaaccaaaaggctatca  | SEQ ID NO: 670 |
| RPS28_int_13  | tcacctgggcttaacgaag   | SEQ ID NO: 671 |

|                 |                      |                |
|-----------------|----------------------|----------------|
| RPS28_int_14    | cagcaattggcaaactcccg | SEQ ID NO: 672 |
| RPS28_int_15    | aacatgagatgtgacaggc  | SEQ ID NO: 673 |
| TPM4-tv2_int_1  | gaagccgggaagaaaggagg | SEQ ID NO: 674 |
| TPM4-tv2_int_2  | gtcaaaccttagacaatgga | SEQ ID NO: 675 |
| TPM4-tv2_int_3  | gccgcgaagaggcgagaaa  | SEQ ID NO: 676 |
| TPM4-tv2_int_4  | aggaatcagctggagaaagg | SEQ ID NO: 677 |
| TPM4-tv2_int_5  | ctccataactgcacagaaa  | SEQ ID NO: 678 |
| TPM4-tv2_int_6  | tttattctataagaccagc  | SEQ ID NO: 679 |
| TPM4-tv2_int_7  | tcatacataactttcttgc  | SEQ ID NO: 680 |
| TPM4-tv2_int_8  | taagtgtgagaaaggacca  | SEQ ID NO: 681 |
| TPM4-tv2_int_9  | ccattttgaagccgagttaa | SEQ ID NO: 682 |
| TPM4-tv2_int_10 | tccaaaacatcttctctcc  | SEQ ID NO: 683 |
| TPM4-tv2_int_11 | atcctttgcttttcaagaa  | SEQ ID NO: 684 |
| TPM4-tv2_int_12 | gaaaccagctgtcgggaaac | SEQ ID NO: 685 |
| TPM4-tv2_int_13 | agatgtcagttaatcacaca | SEQ ID NO: 686 |
| TPM4-tv2_int_14 | gctagatacatgtggcaaag | SEQ ID NO: 687 |
| TPM4-tv2_int_15 | ggattcccttataaagtaa  | SEQ ID NO: 688 |
| TPM4-tv2_int_16 | caatcaaagccttcacacg  | SEQ ID NO: 689 |

### Example 3: SNV-detection probes and masks

#### SNP FISH Oligonucleotide sequences

| Probe                 | Sequence                        | Identifier     |
|-----------------------|---------------------------------|----------------|
| BRAF-T1799A-WT        | tttactgttagctagacaaaatcacct     | SEQ ID NO: 690 |
| BRAF-T1799A-MUT       | tttctctgttagctagacaaaatcacct    | SEQ ID NO: 691 |
| BRAF-T1799-Mask       | aggtgattttggtctag               | SEQ ID NO: 692 |
| BRAF-Common-Match     | ggcgtgtaagtaatccatgccctgtg      | SEQ ID NO: 693 |
| BRAF-Common-Mismatch  | ggcgagtaagtaatccatgccctgtg      | SEQ ID NO: 694 |
| BRAF-Common-Mask      | cacagggcatggattac               | SEQ ID NO: 695 |
| DNMT1-Common-Match    | cttggttcccggtttctagacgtc        | SEQ ID NO: 696 |
| DNMT1-Common-MisMatch | cttctgtcccggtttctagacgtc        | SEQ ID NO: 697 |
| DNMT1-Common-Mask     | gacgtctagaaaacg                 | SEQ ID NO: 698 |
| SKA3-GM-SNP-Mat       | cattttgttcagtttctgtgttga        | SEQ ID NO: 699 |
| SKA3-GM-SNP-PAT       | cattctgttcagtttctgtgttga        | SEQ ID NO: 700 |
| SKA3-GM-SNP-Mask      | tccaacacagaaactgaac             | SEQ ID NO: 701 |
| DNMT1-GM-SNP-Mat      | gagacgggtcatcatcatagattgg       | SEQ ID NO: 702 |
| DNMT1-GM-SNP-Pat      | gagatgggtcatcatcatagattgg       | SEQ ID NO: 703 |
| DNMT1-GM-SNP-Mask     | ccaatctatgatgatgacc             | SEQ ID NO: 704 |
| SUZ12-GM-SNP-Mat      | atctctattggatattctattcatgaaacac | SEQ ID NO: 705 |
| SUZ12-GM-SNP-Pat      | atctgtattggatattctattcatgaaacac | SEQ ID NO: 706 |
| SUZ12-GM-SNP-Mask     | gtgtttcatgaatagaataccaat        | SEQ ID NO: 707 |
| EBF1-GM-SNP-Mat       | ctgtataaatgaatctgcctggtgt       | SEQ ID NO: 708 |
| EBF1-GM-SNP-Pat       | ctgtgtaaatgaatctgcctggtgt       | SEQ ID NO: 709 |

|                  |                     |                |
|------------------|---------------------|----------------|
| EBF1-GM-SNP-Mask | acaccaggcagattcattt | SEQ ID NO: 710 |
|------------------|---------------------|----------------|

## SNP Chr19 GM Paint Probes

| Probe          | Sequence                  | Identifier     |
|----------------|---------------------------|----------------|
| PTBP1-1-m      | ggagggcaggaggcgccg        | SEQ ID NO: 711 |
| PTBP1-2-m      | tcccggggccaggagagc        | SEQ ID NO: 712 |
| PTBP1-3-m      | aggaaacaacgttagcctggtggc  | SEQ ID NO: 713 |
| PTBP1-4-m      | ccaccgaacggcacgattccagc   | SEQ ID NO: 714 |
| EEF2-5-m       | caccaccccaggaaataacggg    | SEQ ID NO: 715 |
| P2RY11-6-m     | cccatgttgccaagaggtct      | SEQ ID NO: 716 |
| IER2-7-m       | gacactccccgtgaacacgtgg    | SEQ ID NO: 717 |
| TOMM40-14-m    | acctcgtgacaccttgctgaca    | SEQ ID NO: 718 |
| TOMM40-15-m    | ccactactactactaattgctc    | SEQ ID NO: 719 |
| SLC1A5-16-m    | cagccgaccgacctccaacc      | SEQ ID NO: 720 |
| RPL18-17-m     | aaaagaatctttatccctttcccc  | SEQ ID NO: 721 |
| PPP1R15A-18-m  | ctcctaaggctgcttctgaatca   | SEQ ID NO: 722 |
| LINC00085-19-m | tgaccacccagaacttcatagcca  | SEQ ID NO: 723 |
| RPS9-20-m      | gaaggcgaaccacttcccgaag    | SEQ ID NO: 724 |
| ICAM1-21-m     | tgggacctcagcatacccaatag   | SEQ ID NO: 725 |
| SUPT5H-22-m    | ccccctcaatggcctgcttc      | SEQ ID NO: 726 |
| PVR-23-m       | accagcctgggcagcatagcgaaa  | SEQ ID NO: 727 |
| PPP1R15A-24-m  | tcctctctaagactcaagaatctgg | SEQ ID NO: 728 |
| EGLN2-25-m     | tcgatccgagcccagcggg       | SEQ ID NO: 729 |
| PTBP1-1-p      | ggagagcaggaggcgccg        | SEQ ID NO: 730 |
| PTBP1-2-p      | tcccggggccaggagagc        | SEQ ID NO: 731 |
| PTBP1-3-p      | aggagacaacgttagcctggtggc  | SEQ ID NO: 732 |
| PTBP1-4-p      | ccactgaacggcacgattccagc   | SEQ ID NO: 733 |
| EEF2-5-p       | caccgaccccaggaaataacggg   | SEQ ID NO: 734 |
| P2RY11-6-p     | cccacgttgccaagaggtct      | SEQ ID NO: 735 |
| IER2-7-p       | gacagtcctccgtgaacacgtgg   | SEQ ID NO: 736 |
| TOMM40-14-p    | accttctgacaccttgctgaca    | SEQ ID NO: 737 |
| TOMM40-15-p    | ccaccactactactaattgctc    | SEQ ID NO: 738 |
| SLC1A5-16-p    | cagcagaccgacctccaacc      | SEQ ID NO: 739 |
| RPL18-17-p     | aaaataatctttatccctttcccc  | SEQ ID NO: 740 |
| PPP1R15A-18-p  | ctcccaaggctgcttctgaatca   | SEQ ID NO: 741 |
| LINC00085-19-p | tgacacaccagaacttcatagcca  | SEQ ID NO: 742 |
| RPS9-20-p      | gaagccgaaccacttcccgaag    | SEQ ID NO: 743 |
| ICAM1-21-p     | tggggcctcagcatacccaatag   | SEQ ID NO: 744 |
| SUPT5H-22-p    | ccccctcaatggcctgcttc      | SEQ ID NO: 745 |
| PVR-23-p       | accaacctgggcagcatagcgaaa  | SEQ ID NO: 746 |
| PPP1R15A-24-p  | ttcttctaagactcaagaatctgg  | SEQ ID NO: 747 |
| EGLN2-25-p     | tcgacccgagcccagcggg       | SEQ ID NO: 748 |
| PTBP1-1-t6     | cggccgcctctg              | SEQ ID NO: 749 |

|                 |                     |                |
|-----------------|---------------------|----------------|
| PTBP1-2-t6      | gctctcttgccc        | SEQ ID NO: 750 |
| PTBP1-3-t6      | gccaccaggctaacgttg  | SEQ ID NO: 751 |
| PTBP1-4-t6      | gctggaatcgtgccgtt   | SEQ ID NO: 752 |
| EEF2-5-t6       | cccgttatttctgggg    | SEQ ID NO: 753 |
| P2RY11-6-t6     | agacctcttgccaa      | SEQ ID NO: 754 |
| IER2-7-t6       | ccacgtgttcacgggg    | SEQ ID NO: 755 |
| TOMM40-14-t6    | tgtcagcaaggtgtcag   | SEQ ID NO: 756 |
| TOMM40-15-t6    | gagcaattagagtagtag  | SEQ ID NO: 757 |
| SLC1A5-16-t6    | ggttgagggtcggt      | SEQ ID NO: 758 |
| RPL18-17-t6     | gggggaaaggataaagat  | SEQ ID NO: 759 |
| PPP1R15A-18-t6  | tgattcagaagcagcct   | SEQ ID NO: 760 |
| LINC00085-19-t6 | tggctatgaagtctgggt  | SEQ ID NO: 761 |
| RPS9-20-t6      | cttcgggaagtgggtc    | SEQ ID NO: 762 |
| ICAM1-21-t6     | ctattgggtatgctgag   | SEQ ID NO: 763 |
| SUPT5H-22-t6    | gaagcaggccattg      | SEQ ID NO: 764 |
| PVR-23-t6       | tttcgtatgctgccag    | SEQ ID NO: 765 |
| PPP1R15A-24-t6  | ccagattcttgagtcttag | SEQ ID NO: 766 |
| EGLN2-25-t6     | cccgtcgggctcg       | SEQ ID NO: 767 |

The disclosures of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entirety.

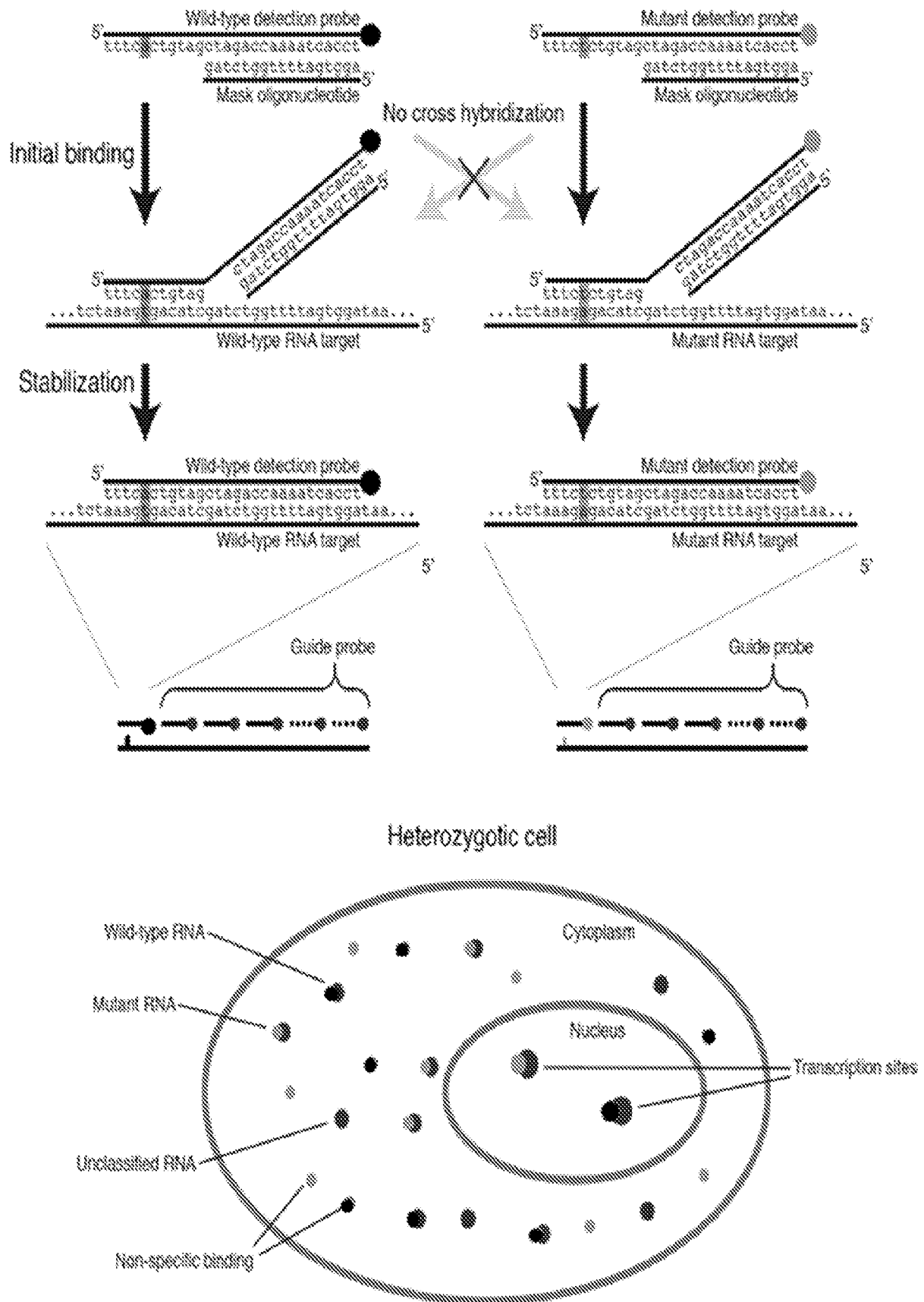
- While this invention has been disclosed with reference to specific
- 5   embodiments, it is apparent that other embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations.

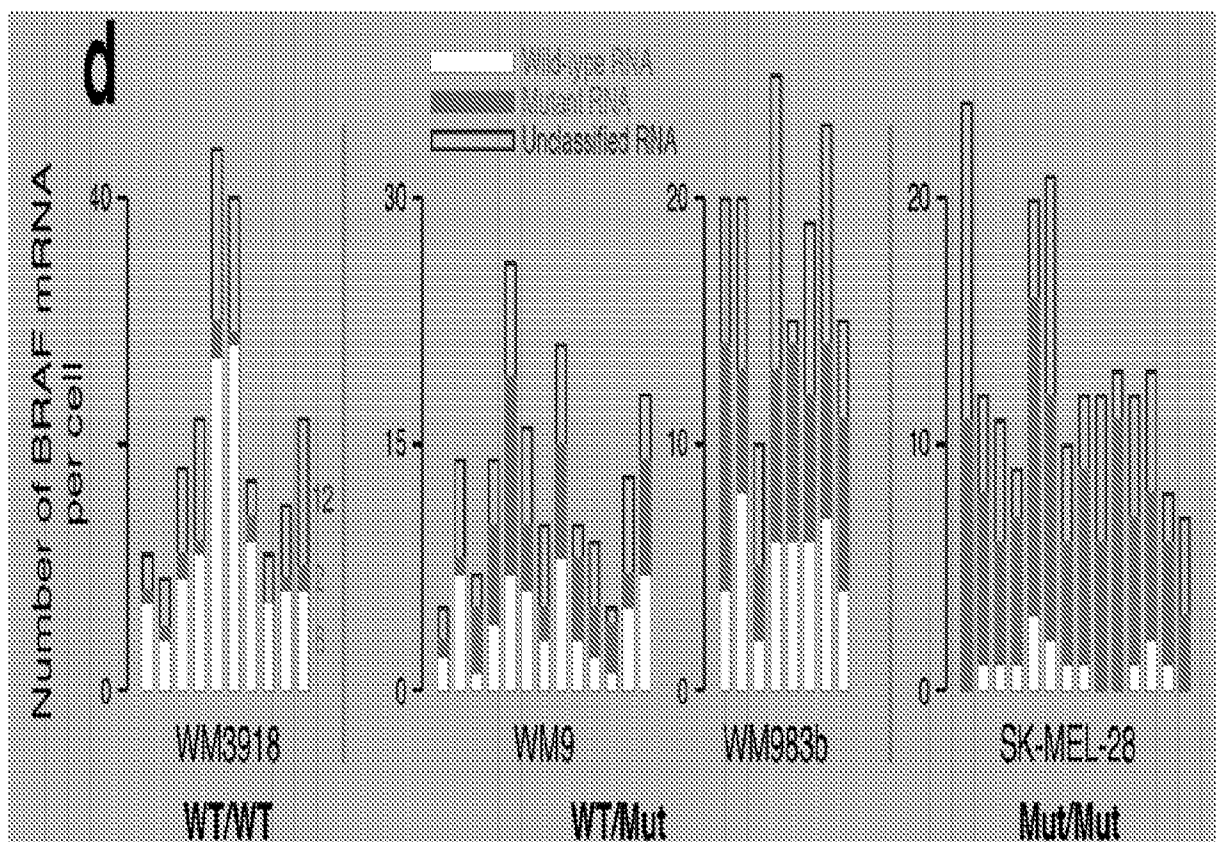
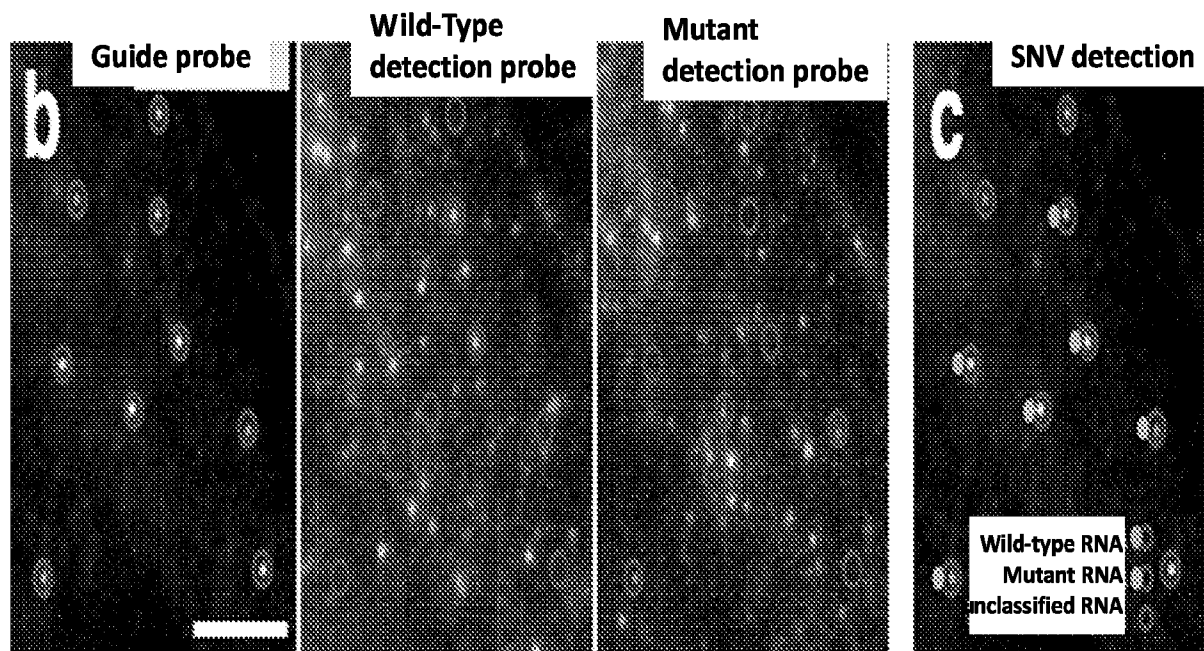
## CLAIMS

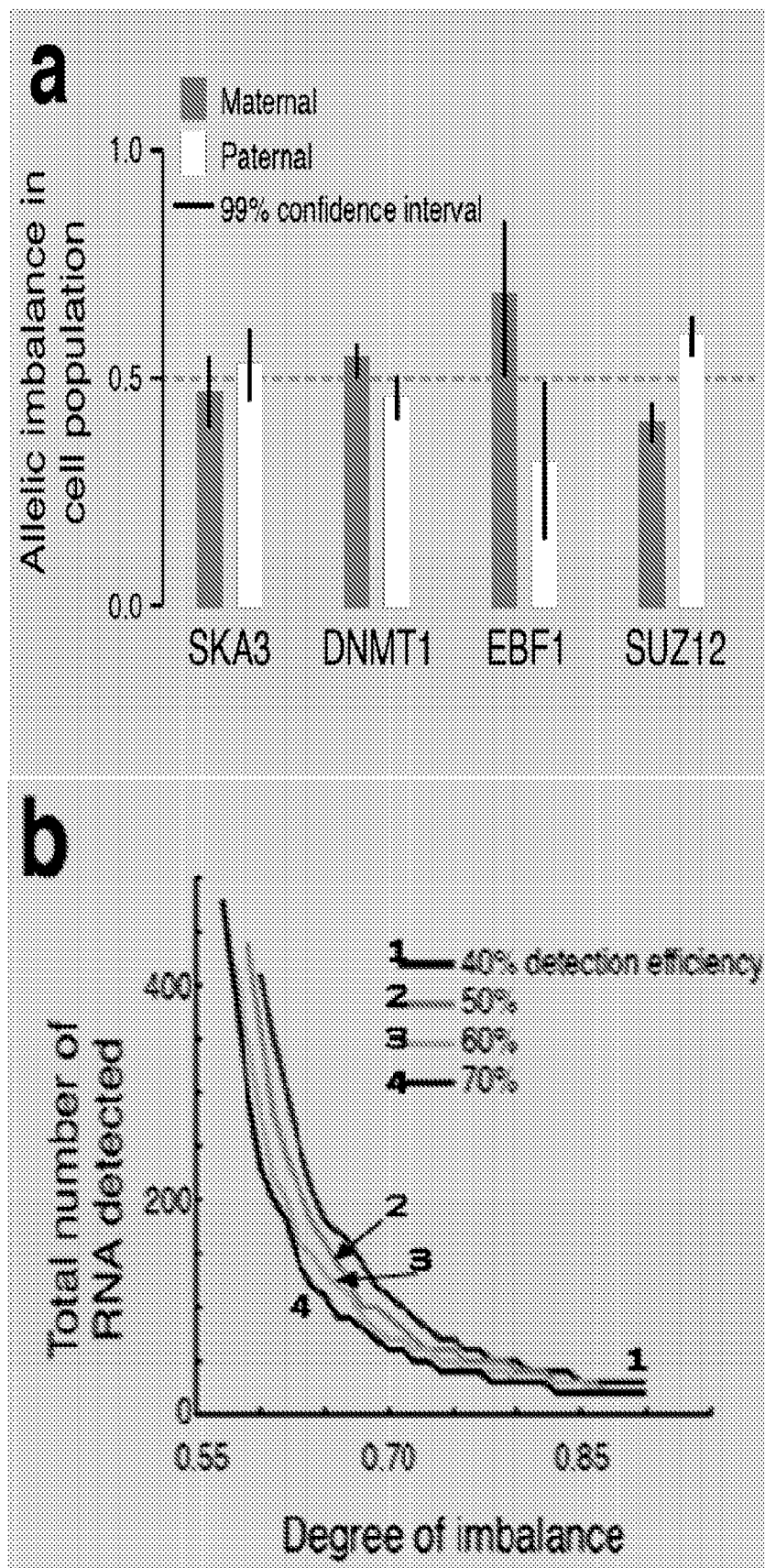
What is claimed is:

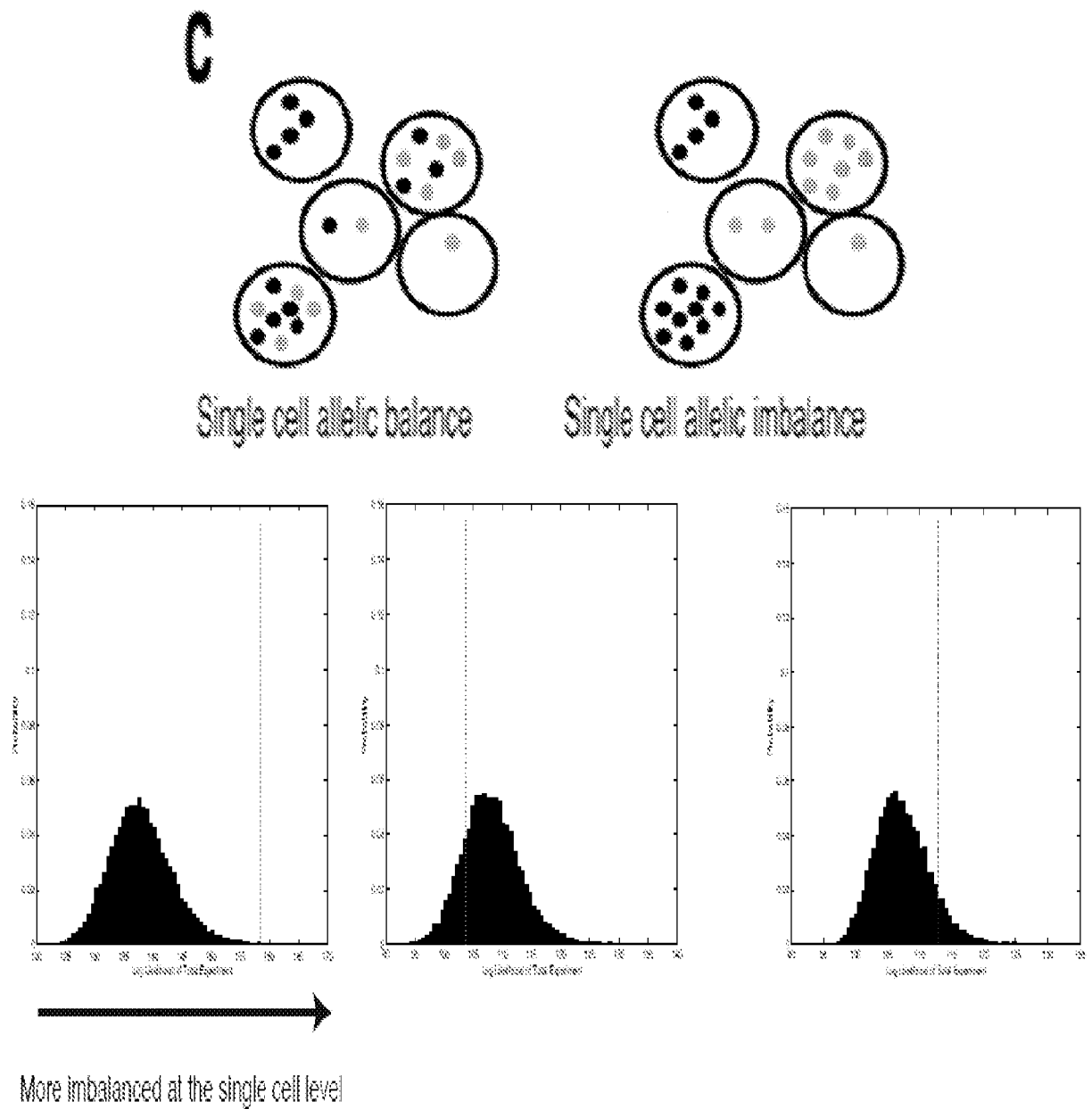
1. A method for detecting a mutation in a target nucleic acid, comprising:  
hybridizing at least one labeled, masked detection probe to a target nucleic acid having at least one mutation;  
wherein the hybridization region of the detection probe and target nucleic acid includes the mutation.
2. The method of claim 1, wherein the mutation is a small-scale mutation selected from the group consisting of a single nucleotide variant (SNV), insertion, and deletion.
3. The method of claim 1, further comprising hybridizing the target nucleic acid with at least one labeled guide probe.
4. The method of claim 3, wherein the hybridization region of the at least one guide probe and target nucleic acid does not include the mutation.
5. The method of claim 3, wherein the at least one guide probe and the detection probe are each labeled with a fluorophore.
6. The method of claim 5, wherein the fluorophore for the at least one guide probe and the fluorophore for the detection probe are distinguishable when visualized.
7. The method of claim 6, further comprising detecting the at least one guide probe fluorophore in a first image;  
detecting the detection probe fluorophore in a second image; and  
detecting the mutation in the nucleotide sequence of the target nucleic acid by comparing said first image with said second image to determine a colocation of the at least one guide probe fluorophore and the detection probe fluorophore.
8. The method of claim 1, wherein the detection probe is masked by an oligonucleotide that hybridizes to a portion of a nucleotide sequence of the detection probe.

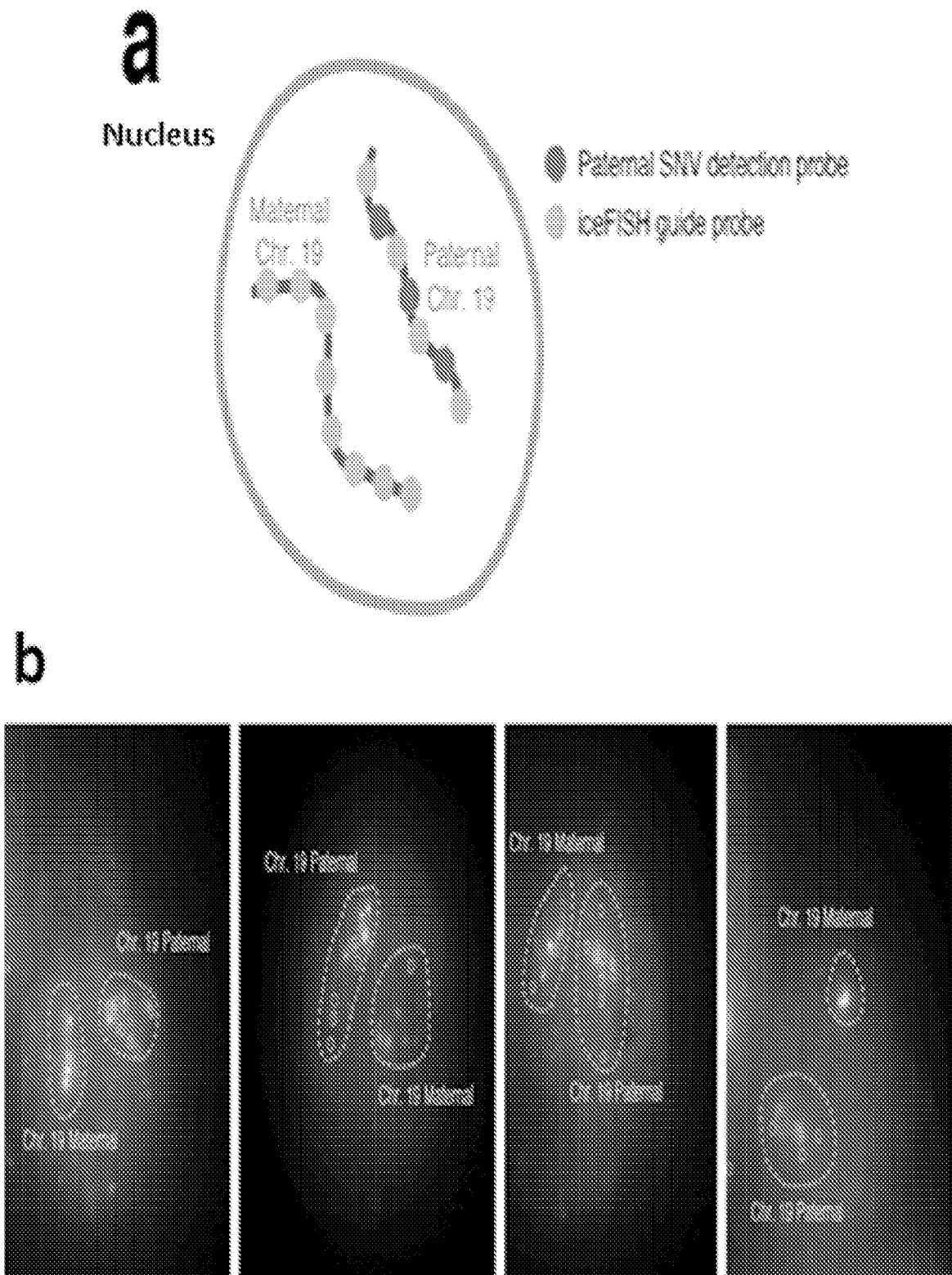
9. The method of claim 8, wherein the detection probe hybridizes to the target nucleic acid on the unmasked region of the detection probe.
10. The method of claim 9, wherein the oligonucleotide masking a portion of the nucleotide sequence of the detection probe denatures from the detection probe when the detection probe hybridizes to the target nucleic acid.
11. The method of claim 10, wherein the previously masked portion of the detection probe hybridizes to the target nucleic acid subsequent to the oligonucleotide denaturing from the detection probe.
12. The method of claim 1, wherein the target nucleic acid is a ribonucleic acid (RNA).
13. The method of claim 12, wherein the RNA is selected from the group consisting of messenger RNA, intronic RNA, exonic DNA, and non-coding RNA.
14. The method of claim 1, wherein the detection probe comprises a nucleotide sequence of at least 10 bases.
15. The method of claim 1, wherein the at least one guide probe comprises a nucleotide sequence of at least 10 bases.
16. The method of claim 1, wherein the oligonucleotide masking the detection probe has a nucleotide sequence of at least 5 bases.
17. The method of claim 1, wherein the detection probe comprises an unmasked nucleotide sequence of at least 2 bases when the oligonucleotide mask is hybridized to the detection probe.

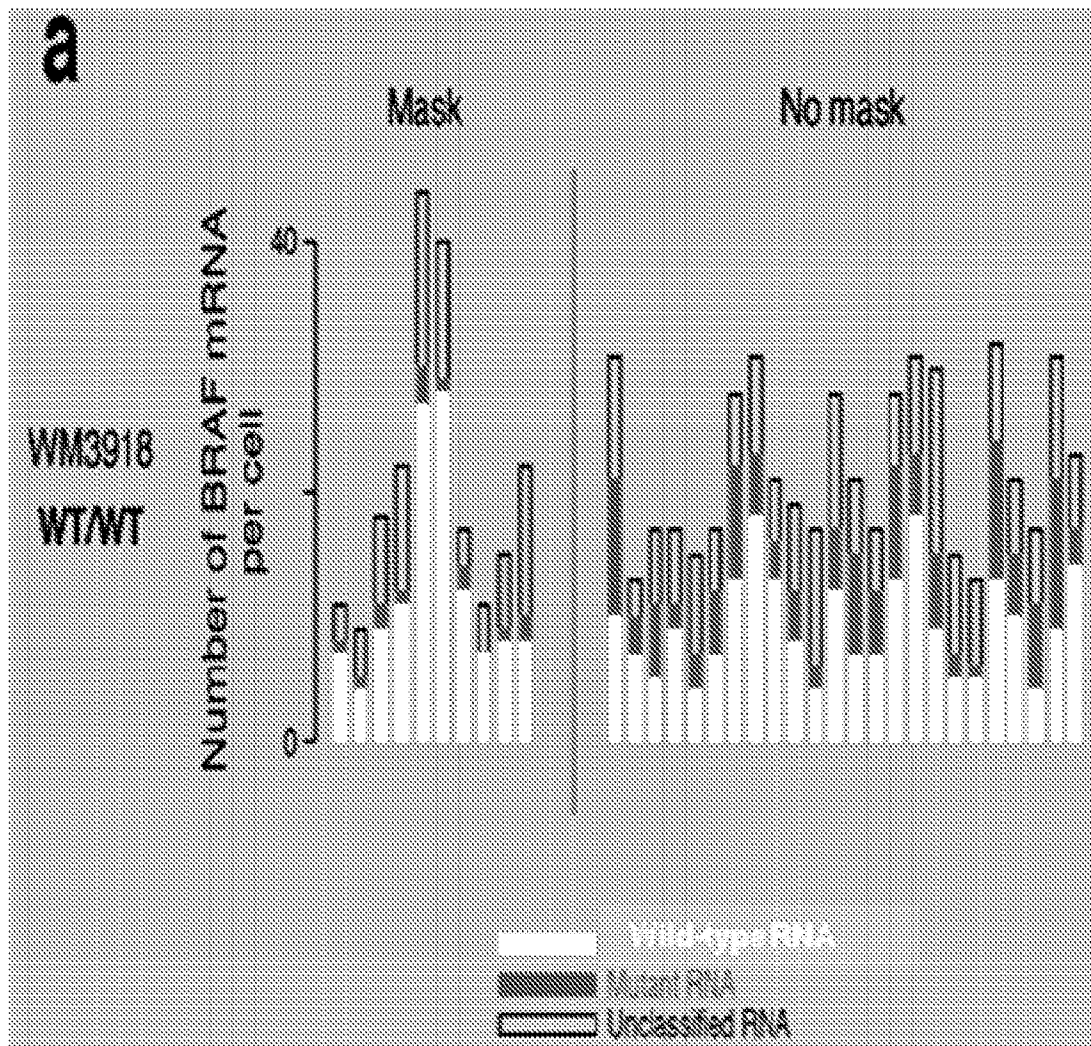
**FIG. 1A**

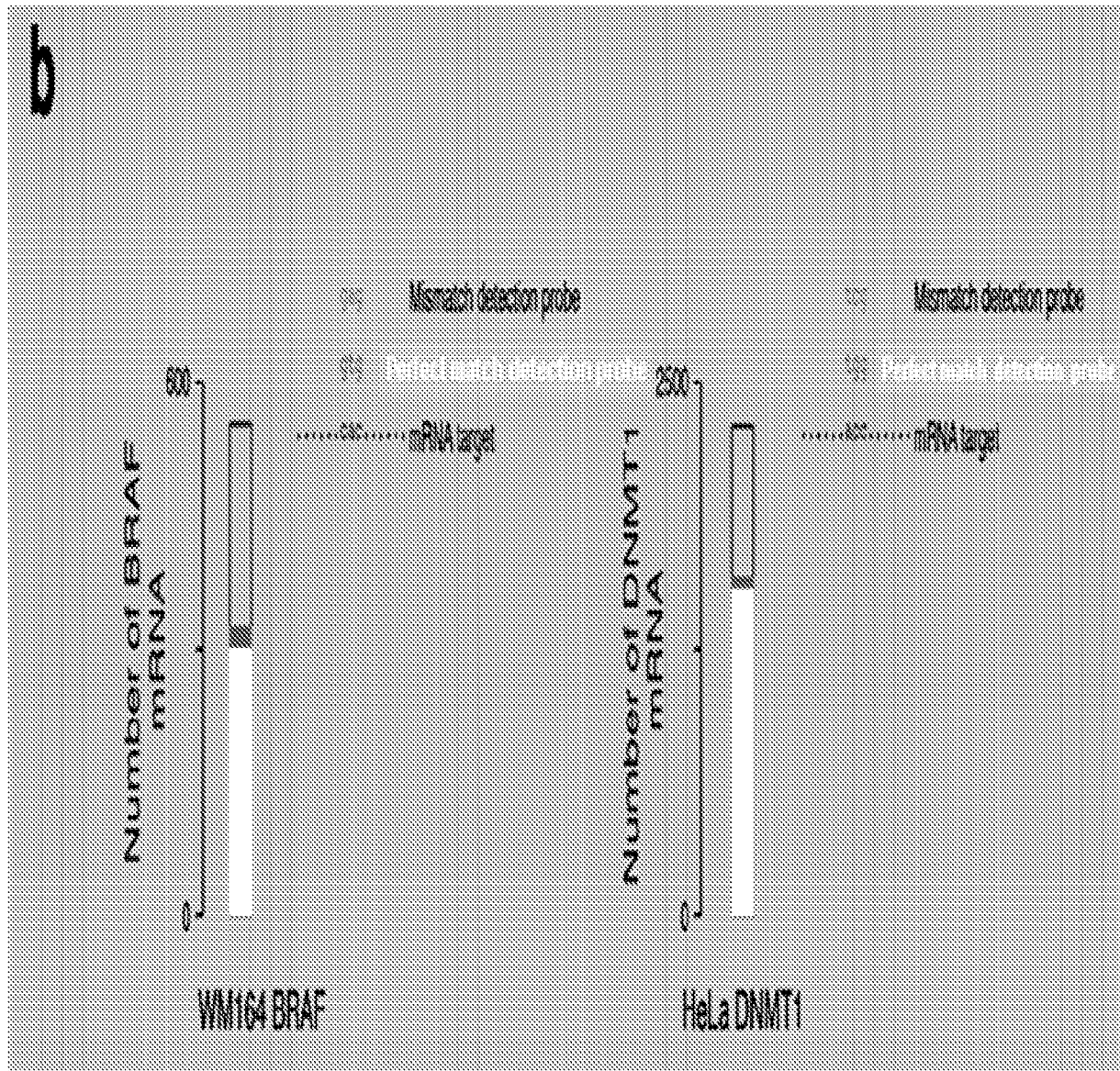
FIGs. 1B-1D

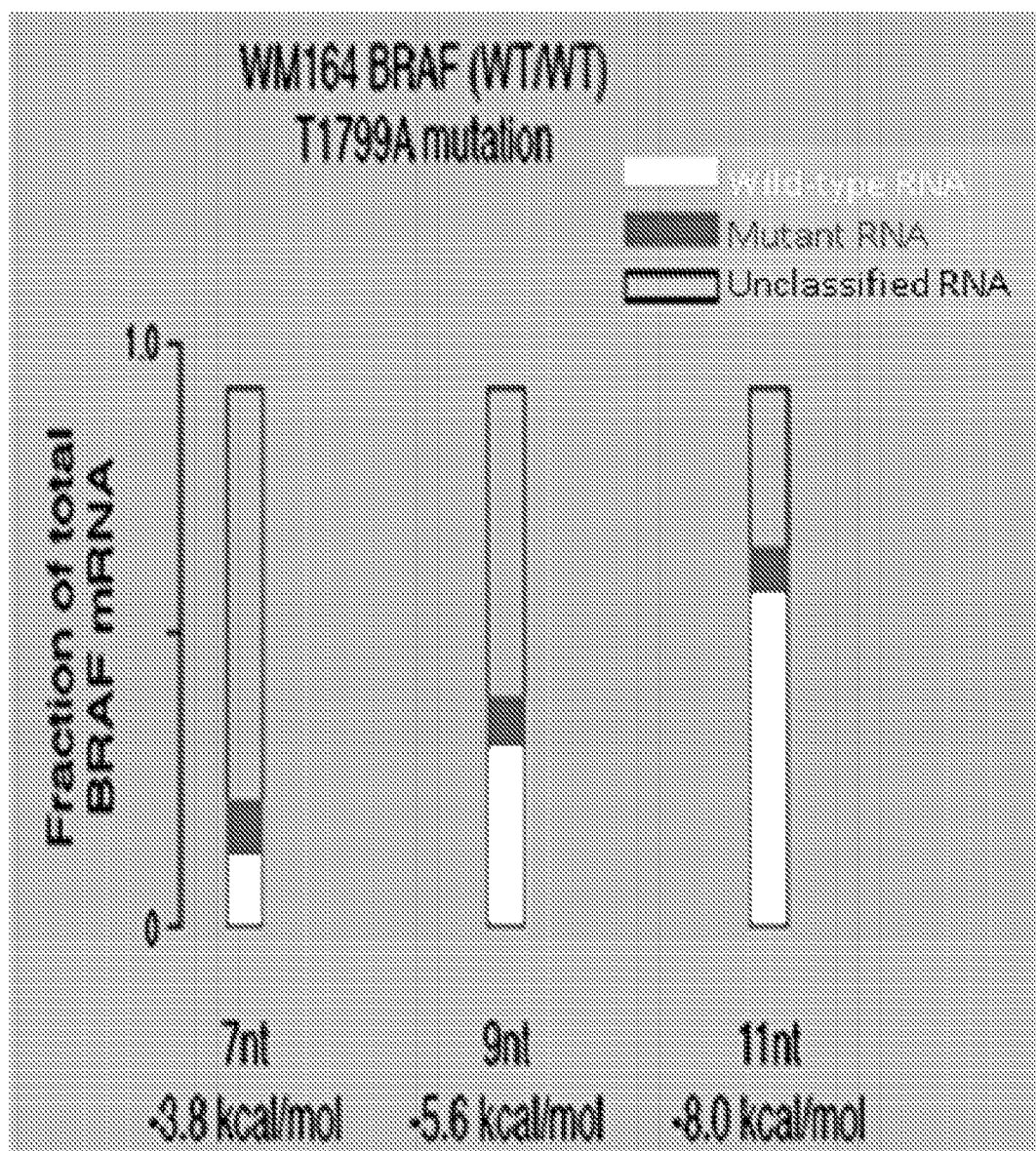
**FIGs. 2A-2B**

**FIG. 2C**

**FIGs. 3A-3B**

**FIG. 4A**

**FIG. 4B**

**FIG. 5**

## INTERNATIONAL SEARCH REPORT

014/025/03 US 07.2014

International application No.

PCT/US 14/25703

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12Q 1/68, C07H 21/02 (2014.01)

USPC - 435/6.11, 536/24.3

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - C12Q 1/68, C07H 21/02 (2014.01)

USPC - 435/6.11, 536/24.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
USPC - 435/6.11, 536/24.3 (keyword limited; terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

ProQuest Dialog, Google Scholar

Search terms: single nucleotide variant, single nucleotide polymorphism, SNV, SNP, probe, mask, protect, toe-hold, toe hold, toehold, dissociate, denature, fluorophore, fluorescent, label, image, detect, in situ, FISH, co-locate

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category*     | Citation of document, with indication, where appropriate, of the relevant passages                                                                                              | Relevant to claim No.                |
|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|
| X<br>---<br>Y | WO 2012/058488 A1 (ZHANG et al.) 03 May 2012 (03.05.2012) p 2, para 4-5; p 3, para 1-3; p 7, para 3; p 16, para 3; p 19, para 1; p 25, para 3; p 57, para 4; Fig. 1A            | 1-2, 8-14, 16-17<br>-----<br>3-7, 15 |
| Y             | WO 2011/038403 A1 (LOU et al.) 31 March 2011 (31.03.2011) para [0093], [00698], [00705], [00706], [00708]                                                                       | 3-7, 15                              |
| X, P          | LEVESQUE et al., Visualizing SNVs to quantify allele-specific expression in single cells. Nat Methods. September 2013 (published online 4 August 2013), Vol 10, No 9, pp 865-7. | 1-17                                 |

☐ Further documents are listed in the continuation of Box C.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

29 May 2014 (29.05.2014)

Date of mailing of the international search report

03 JUL 2014

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents  
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774