

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 March 2012 (15.03.2012)

PCT

(10) International Publication Number
WO 2012/031329 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C40B 50/06* (2006.01)

WANG, Qiao [CN/AU]; 15/164 Chapel Street, St. Kilda, Victoria 3182 (AU).

(21) International Application Number:
PCT/AU2011/001161

(74) Agents: HUGHES, E., John, L. et al.; Davies Collison Cave, 1 Nicholson Street, Victoria, Melbourne 3000 (AU).

(22) International Filing Date:
8 September 2011 (08.09.2011)

(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, 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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2010904085 10 September 2010 (10.09.2010) AU

(71) Applicants (for all designated States except US): MURDOCH CHILDRENS RESEARCH INSTITUTE [AU/AU]; Royal Children's Hospital, Flemington Road, Parkville, Victoria 3052 (AU). NATIONAL ICT AUSTRALIA (NICTA) [AU/AU]; Australian Technology Park, Level 5, 13 Garden Street, Eveleigh, New South Wales 2015 (AU).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): WONG, Nicholas Chau Lun [AU/AU]; 14/111 Victoria Road, Hawthorn East, Victoria 3123 (AU). CRAIG, Jeffrey Mark [AU/AU]; 32 Brentwood Drive, Avondale Heights, Victoria 3034 (AU). SAFFERY, Richard Eric [AU/AU]; 7 MacDonald Street, Glen Iris, Victoria 3146 (AU). ASHLEY, David [AU/AU]; Director of Medical Oncology, Geelong Hospital, Deakin University, The Andrew Love Cancer Centre, 70 Swanston Street Geelong, Victoria 3220 (AU). BEDO, Justin [AU/FR]; 4 rue Martel, F-75010 Paris (FR). KOWALCZYK, Adam [AU/AU]; 8 Cappella Court, Glen Waverley, Victoria 3150 (AU).

Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: ASSAY FOR DETECTION AND MONITORING OF CANCER

(57) Abstract: The present disclosure relates generally to an assay for the determination of epigenetic profiles including epigenetic profiles associated with a pathological condition. The present disclosure teaches an assay to detect epigenetic profiles associated with cancer including its form or type, state and minimum residual disease status. The assay enabled herein further identifies and monitors forms and sub-types of leukemia and other hematological malignancies. Kits and assays for medicaments are also taught herein. Methods for screening agents which modulate methylation are also enabled.



WO 2012/031329 A1

- 1 -

ASSAY FOR DETECTION AND MONITORING OF CANCER

FILING DATA

5

[0001] This application is associated with and claims priority from Australian Provisional Patent Application No. 2010904085, filed on 10 September 2010, entitled "Assay for detection and monitoring of cancer", the entire contents of which, are incorporated herein by reference.

10

FIELD

[0002] The present disclosure relates generally to an assay for the determination of epigenetic profiles including epigenetic profiles associated with a pathological condition.

15

The present disclosure teaches an assay to detect epigenetic profiles associated with cancer including its form or type, state and minimum residual disease status. The assay enabled herein further identifies and monitors forms and sub-types of leukemia and other hematological malignancies. Kits and assays for medicaments are also taught herein. Methods for screening agents which modulate methylation are also enabled.

20

BACKGROUND

[0003] Bibliographic details of the publications referred to by authors in this specification are collected alphabetically at the end of the description.

25

[0004] Reference to any prior art in this specification is not, and should not be taken as, an acknowledgment or any form of suggestion that this prior art forms part of the common general knowledge in any country.

- 2 -

[0005] Cancer is one of the leading causes of mortality and morbidity across all societies and ethnic groups. Cancer has a complex etiology influenced genetic and environmental pressures. In fact, cancer can be considered a broad spectrum of pathological conditions.

5 [0006] One type of cancer is the hematological form, leukemia. Leukemia is classified on the cell type which becomes malignant. Major types include acute lymphoblastic leukemia (ALL) [also known as acute lymphocytic or lymphatic leukemia], acute myeloid leukemia (AML) [also known as acute myelogenous leukemia], chronic myeloid leukemia (CML) and chronic lymphocytic leukemia (CLL). Within these forms there are sub-types
10 including leukemias which tend to have an onset in an age-dependent manner. For example, some types of leukemia appear more frequently in younger members of a population while others are more frequent in an aging population.

[0007] ALL is the most prevalent of the childhood malignancies in developed countries.
15 Although current treatment protocols give rise to a proportion of 5-year event free survival outcome, many patients still succumb to the disease, relapse or experience secondary effects of the treatment due to their non-specific nature. It is unclear how leukemia arises and the mechanisms involved in this hematological malignancy. It is also difficult if not impossible to select treatment protocols based on a prediction as to their likely efficacy.

20

[0008] One of the difficulties in the treatment and management of cancers is the sensitivity in determining the presence of malignant cells which constitutes the minimum residual disease (MRD). Considering the pervasiveness of cancer there is a need to be able to identify the MRD at frequencies $<10^{-5}$. This will be important for prognostic purposes and
25 for tailored, individual medicine.

[0009] It is apparent that DNA methylation plays a role in the regulation of gene expression in higher organisms. The importance of DNA methylation has been highlighted by its involvement in several human diseases. Methylation of cytosine at the 5' position is a
30 widely known covalent modification of human genomic DNA. In particular, methylation of CpG islands within regulatory regions of the genome appears to be highly tissue

- 3 -

specific. It is now apparent that methylation of cytosines outside CpG islands is also important. These regions, within 2kb of CpG islands, have been named "shores" or "island shores" (Irizarry *et al.*, *Nature Genetics* 41(2):178-186, 2009). Methylation modifications which are also potentially important include the generation of hydroxymethylcytosines and
5 other base methylations as well as RNA methylation.

[0010] Despite the availability of a range of methylation assays (see, for example, Laird, *Nature Reviews Genetics* 11:191-203, 2010; Hansen *et al.*, *Nature Genetics* 43(8):768-775, 2011; Rein *et al.*, *Nucleic Acids Res.* 26:2255, 1998 and Fraga and Estella, *Biotechniques*
10 33:632-649, 2002), selection of regions to amplify and screen is an important aspect of determining an epigenetic profile characteristic of a disease condition. There is a need to identify these crucial regions in the genome and to associate epigenetic profiles to pathological conditions such as cancer. There is also a need to develop better panels of assays to facilitate more precise diagnosis of cancer subtypes, to more accurately predict
15 treatment response and long term outcome, to facilitate the development of patient-specific tailored treatment regimens, and to allow monitoring of disease progression.

SUMMARY

[0011] Throughout this specification, unless the context requires otherwise, the word "comprise" or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or method step or group of elements or integers or method steps but not the exclusion of any element or integer or method step or group of elements or integers or method steps.

[0012] Nucleotide and amino acid sequences are referred to by a sequence identifier number (SEQ ID NO). The SEQ ID NOs correspond numerically to the sequence identifiers <400>1 (SEQ ID NO:1), <400>2 (SEQ ID NO:2), etc. A summary of the sequence identifiers is provided in Table 1. A sequence listing is provided after the claims.

TABLE 1

Summary of sequence identifiers

SEQUENCE ID NO:	DESCRIPTION
1	Oligonucleotide
2	Oligonucleotide
3	Primer sequence 5'3' - sq_FOXE3_F
4	Primer sequence 5'3' - sq_FOXE3_R
5	Primer sequence 5'3' - sq_TLX3_F
6	Primer sequence 5'3' - sq_TLX3_R
7	Primer sequence 5'3' - sqClorf76_F
8	Primer sequence 5'3' - sqClorf76_R
9	Primer sequence 5'3' - sqFBXO39_F
10	Primer sequence 5'3' - sqFBXO39_R
11	Primer sequence 5'3' - sqMYOD1_F
12	Primer sequence 5'3' - sqMYOD1_R
13	Primer sequence 5'3' - sqGUCY1A2_F

- 5 -

SEQUENCE ID NO:	DESCRIPTION
14	Primer sequence 5'3' - sqGUCY1A2_R
15	Primer sequence 5'3' - sqHRH3_F
16	Primer sequence 5'3' - sqHRH3_R
17	Primer sequence 5'3' - sqPKDREJ_F
18	Primer sequence 5'3' - sqPKDREJ_R
19	Primer sequence 5'3' - sqPNMA2_F
20	Primer sequence 5'3' - sqPNMA2_R
21	Primer sequence 5'3' - sqPPARG_F
22	Primer sequence 5'3' - sqPPARG_R
23	Primer sequence 5'3' - sqPRLHR_F
24	Primer sequence 5'3' - sqPRLHR_R
25	Primer sequence 5'3' - sqPTPRZ1_F
26	Primer sequence 5'3' - sqPTPRZ1_R
27	Primer sequence 5'3' - sqSCRN1_F
28	Primer sequence 5'3' - sqSCRN1_R
29	Primer sequence 5'3' - sqSOX17_F
30	Primer sequence 5'3' - sqSOX17_R
31	Primer sequence 5'3' - sqUnknown_F
32	Primer sequence 5'3' - sqUnknown_R

[0013] The present disclosure teaches the identification of genomic regions including loci, the extent of epigenetic change to which, is an indicator of a form or type, state or level of cancer or minimal residual disease (MRD) with respect to cancer. The term "epigenetic change" includes distinction of epigenetic sites. Reference to "loci" or "locus" includes genes associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity. Table 2 lists examples of genes in each of these classes. Reference to "cancer" includes hematological malignancies such as leukemias. Reference to "leukemias" include all forms, types and sub-types of leukemias including leukemias which have a high frequency of onset during childhood years (pediatric leukemias) or later onset leukemias (adult leukemias). Hence,

- 6 -

the present disclosure is instructional of an assay to detect an epigenetic profile indicative of a form or type (including sub-type), state or MRD status of a cancer including leukemia. Such leukemias include acute lymphatic leukemia (ALL), acute myeloid leukemia (AML), chronic myeloid leukemia (CML), small lymphocytic lymphoma (SLL), chronic lymphocytic leukemia (CLL), acute monocytic leukemia (AMOL), Hodgkin's lymphomas (all types), non-Hodgkin's lymphomas (all types) and other lymphoid and myeloid malignancies.

[0014] Furthermore, the epigenetic profile is also informative as to the spectrum of cancer disease conditions including its various forms, types, sub-types and is useful in monitoring treatment protocols or patients after treatment. Identification of the minimum residual disease (MRD) is important in the decision process in relation to undertaking more or less intensive or toxic therapy and hence is useful for prognosis and tailored therapy on a case by case basis.

[0015] An example of epigenetic change is the extent of change in methylation or change in distinction of methylation sites in one or more regions in one or more of the genetic loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity. Examples are listed in Tables 2 and 3. By "one or more" includes from 1 to a number which provides 100% confidence that the change in methylation or other epigenetic marker is associated with a disease condition. In an embodiment this range is from 1 to 16 including 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or 16 different genetic loci or 1 to 16 different regions in any one or more of the listed genetic loci. Greater than 16 loci or regions can nevertheless be measured such as from 16 to 1000.

[0016] Accordingly, enabled herein is a method for identifying an epigenetic profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for the extent of change in epigenetic profile relative to a control or normal/non-cancer tissue within or proximal to a locus selected from one or more loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA

- 7 -

binding; (iv) subcellular location; and (v) transcription regulator activity wherein a change in the epigenetic profile in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

5 [0017] Another aspect taught herein is a method for identifying a methylation profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of methylation within or proximal to a locus selected from one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA
10 binding; (iv) subcellular location; and (v) transcription regulator activity wherein a change in the extent of methylation in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

[0018] The present specification is further instructional of a method for identifying an
15 epigenetic profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of change in epigenetic profile within or proximal to from 1 to 16 regions in one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v)
20 transcription regulator activity wherein a change in the epigenetic profile in or proximal to one or more regions is indicative of the presence of a cancerous condition or a propensity to develop same.

[0019] A still further aspect enabled herein is a method for identifying a methylation
25 profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of methylation within or proximal to from 1 to 16 regions in one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator
30 activity wherein a change in the extent of methylation in or proximal to one or more

regions is indicative of the presence of a cancerous condition or a propensity to develop same.

[0020] Examples of genes or loci in each class of associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity; are provided in Table 2. Particular loci which are differentially subject to epigenetic change are listed in Table 3. By "proximal" is meant a region up to approximately 1000kb up- or down-stream from the 5' or 3' end of a locus and includes regulatory elements within this region.

[0021] By "epigenetic change" includes a change in epigenetic profile. By "epigenetic profile" includes epigenetic modifications such as methylation including hypermethylation and hypomethylation, RNA/DNA interactions, expression profiles of non-coding RNA, histone modification, changes in acetylation, ubiquitylation, phosphorylation and sumoylation, as well as chromatin altered transcription factor levels and the like leading to activation or deactivation of genetic locus expression. In an embodiment, the extent of methylation, RNA/DNA interaction and non-coding RNA expression are determined as well as any changes therein. In an aspect, the epigenetic modification is an increase or decrease in methylation or an alteration in distribution of methylation sites or other epigenetic sites.

[0022] "Methylation" includes methylation of any base in DNA or RNA including methylation of cytosine, hydroxymethylation of cytosine, 5-methylcytosine and methylation of adenine. The present disclosure teaches a method for detecting methylation in DNA or RNA associated with a spectrum of cancerous conditions including pediatric and adult leukemia or their various types and sub-types. Reference to a "cancerous condition" includes hematological malignancies and tissue cancers, such as leukemias, sarcomas, carcinomas and other tumors. A "leukemia" may, therefore, be of any hematological type or sub-type and extends to pediatric and adult leukemias.

- 9 -

[0023] The epigenetic profile is determined in the genome of a cell or sub-population of cells of a subject. Any cell may be tested including but not limited to blood cells, cerebrospinal fluid (CSF) cells, bone marrow cells, buccal cells and cells from rectal swabs or feces. Cells from pre-natal tissues and embryos may also be tested. In addition, cell free
5 DNA or RNA circulating in whole blood, serum or plasma may also be tested. Circulating DNA or RNA may also be tested in other fluids such as urine, pus, respiratory fluid, lymph fluid, feces, bile, saliva, sputum, semen, vaginal flow, cerebral spinal fluid, brain fluid, ascites, milk, secretions from the genitourinary tract and a lavage of a tissue or organ (e.g. a lung).

10

[0024] In an embodiment, DNA or RNA methylation is detected by chemical conversion or methylation specific or sensitive restriction enzymes.

15

[0025] Selective chemical conversion, such as using bisulfite treatment [e.g. WO 2005/038051] of unmethylated cytosines to uracil, leaving methylated cytosines unchanged. Uracil has the same base pairing behaviour as thymine. It therefore forms a base pair with adenine. 5-Methylcytosine, on the other hand, base pairs with guanine. Methylated and unmethylated cytosines can therefore be differentiated. Either approach typically employs a PCR step.

20

25

[0026] The present disclosure further enables a method for monitoring the treatment of a cancerous condition in which the treatment modulates the epigenetic profile of one or more loci listed in Table 2, the method comprising monitoring for a change relative to a control or a pre- and post-treatment sample in the epigenetic profile within or near the locus or
30 loci. In an embodiment, the extent of change in methylation profile is determined. The sensitivity of the subject assay allows for determination of MRD at a frequency down to about 10^{-4} to about 10^{-8} (i.e. one cell per 10^{-4} to about 10^{-8} cells) which enables better prognostic determinations. Furthermore, identification of malignant cells in treated patients or patients in remission allows for the selection of more intensive or less toxic therapies.

- 10 -

[0027] The present disclosure also teaches for the use of an epigenetic profile in one or more loci listed in Table 2 in the manufacture of an assay to identify an epigenetic profile of a cancerous pathological condition. The epigenetic profile includes the determination of extent of epigenetic change such as extent of methylation in any base in DNA or RNA including in cytosine bases. The cytosines include those in CpG and CpNpG islands and shores and in non-CpG and CpNpG islands and shores. By "locus" or "loci" includes coding and non-coding regions (e.g. promoter regions, 5' non-coding regions, exons, introns and 3' non-coding regions). The region encompassed by a locus includes its coding sequence, promoter and up- or down-stream regulatory elements typically within approximately 1000kb of a transcriptional start site or transcriptional termination signal. Hence, this 1000kb region is regarded as the region proximal to the locus.

[0028] The assay enabled herein may also be used alone or in combination with assays to detect gene expression/transcription of a locus associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity. Particular examples of loci are provided in Tables 2 and 3. The assay taught herein is also useful in epidemiological studies of different ethnic populations with cancer. Accordingly, the present disclosure further provides a method of identifying an epigenetic profile in populations of subjects indicative of a cancerous pathological condition, the method comprising screening for a change relative to a control in a statistically significant number of subjects the extent of epigenetic profile in one or more loci listed in Table 2 or 3 wherein a significant difference in the extent of epigenetic change compared to a control is indicative of the presence of the pathological condition or a propensity to develop same. In accordance with this method, the assay may comprise the further step of determining the extent of expression such as by quantitative reverse transcriptase PCR (qRT-PCR), TaqMan, gene expression micro arrays or by Northern or Western blot analysis. In an embodiment, the epigenetic change is extent of methylation or distribution of methylation sites.

[0029] The present disclosure further enables a method for screening for an agent which modulates to epigenetic profile of one or more loci listed in Table 2 or 3 in the presence or

- 11 -

absence of an agent to be tested, wherein an agent is selected if it induces a change in the epigenetic profile. An example of an agent which modulates methylation includes those which inhibit the DNA methyltransferase enzymes (DNMT1, DNMT2, DNMT3a, DNMT3b). For example 5-aza-C, 5-aza-dC (Decitabine) and Zebularine and other known
5 DNA demethylating agents (DMA's) as well as histone deacetylase inhibitors such as valproic acid, SAHA (Vorinostat), Trichostatin A, and combinations of these agents amongst many other agents.

[0030] A further embodiment taught herein is a kit for use in the above methods
10 comprising primers to amplify a region within a locus or loci listed in Table 2 or 3 for detection of epigenetic change such as methylation profile in DNA. The kit may also be adopted for use in a multiplex assay. In an embodiment, the kit contains components to amplify a genomic region and conduct a methylation assay.

15 [0031] Computer programs to monitor changes in epigenetic profile over time and/or in response to therapeutic intervention are also taught herein.

- 12 -

TABLE 2*Genetic loci listed by function differentially methylated between leukemia and controls*

Function	Genes
cell fate commitment	MYOD1, PPARG, SIX1, ONECUT2, PTF1A, SOX9, TLX3
transcription factor activity	MYOD1, L3MBTL4, SOX1, EGR4, DMRT3, ONECUT2, PPARG, NKX2-8, SOX9, TP73, MSX2, SIX1, TFAP2C, TLX3, FOXE3
sequence-specific DNA binding	MSX2, MYOD1, SOX1, NKX2-8, PPARG, SIX1, ONECUT2, PTF1A, SOX9, TLX3, FOXE3
dna-binding	MYOD1, SOX1, ZNF454, EGR4, DMRT3, NKX2-8, ONECUT2, PTF1A, PPARG, SOX9, TP73, SALL3, MSX2, SIX1, SOX17, TFAP2C, TLX3, FOXE3
subcellular location	ACADL, BGLAP, CDH22, CELSR1, CPXM2, CXCL1, CYR61, DHCR24, DMRT3, EGR4, ELOVL4, EPOR, FOXE3, GALR1, GUCY1A2, HS3ST2, KCNK5, KISS1, L3MBTL4, LAMA1, LGI2, LPL, LY86, MAG, MMP11, MPST, MS4A7, MSX2, MYO3A, MYOD1, NELL1, NINJ1, NKX2-8, NPTX2, OGDHL, ONECUT2, PDE10A, PKDREJ, PLD4, PNMA2, PPARG, PRLHR, PRSS12, PTF1A, RBP1, RIMS4, SALL3, SCARF1, SCRNI1, SFRP1, SH3GL2, SIX1, SLC18A3, SLC22A3, SLC5A7, SLC8A2, SNAP91, SOX1, SOX17, SOX9, SSTR4, TCL1A, TFAP2C, TLX3, TP73, ZNF454, CYP24A1
transcription activity regulator	MYOD1, L3MBTL4, BGLAP, SOX1, EGR4, DMRT3, ONECUT2, PTF1A, PPARG, NKX2-8, SOX9, TP73, MSX2, SIX1, TFAP2C, TLX3, FOXE3

TABLE 3

Genetic loci differentially methylated between leukemia and controls

Gene Symbol	
ACADL	
ACTL6B	
ARNT2	
ASNS	
BTG4	
CDH22	
CELSR1	
GALR1	cg00662556
MALL	cg01615704
KCNK5	cg02128567
PNMA2	cg02154186
ONECUT2	cg02250594
GRM7	cg02332525
TFAP2C	cg02536286
CYP24A1	cg02604290
FBXO39	cg02613386
SOX17	cg02919422
VGLL2	cg03160135
IRX4	cg03963198
COL4A3	cg04324308
CYR61	cg04453065
GPR126	cg04461802
GALR1	cg04534765
PPARG	cg04632671
FLJ41327	cg05020203
TRAM1L1	cg05807991
PGM5	cg06190053
MARK2	cg06204948
OGDHL	cg06222851
CELSR1	cg06268694
SOX9	cg06391468
SOX1	cg06675478
CNTNAP2	cg07028533
SLC22A3	cg07237939

- 14 -

MYOD1	cg07271264
PRSS12	cg07625840
LAMA1	cg07846220
ACTL6B	cg08572611
LPL	cg08918749
MGC52057	cg09500672
CPXM2	cg09619146
MTNR1A	cg10063179
GPR78	cg10189695
IGSF4	cg10193817
RHCG	cg10453365
SCRN1	cg10639440
PKDREJ	cg11377136
SLC18A3	cg11389172
EPHA10	cg12128017
RBP4	cg12331389
PAK7	cg12388309
DMRT3	cg12420104
RBP1	cg12497564
WWTR1	cg12507125
MGC52057	cg12594641
NPTX2	cg12799895
CNTNAP3	cg13059782
ARL4	cg13140740
ADPN	cg13184872
ELOVL4	cg13297865
LGI2	cg13359415
EGR4	cg13481359
FLJ46831	cg13929328
FAM84A	cg14225485
ASNS	cg14785449
ACADL	cg14795968
OTOP1	cg14882700
RAB32	cg14889768
O	cg15121304
SALL3	cg15191648
IRX2	cg15433631
ZNF365	cg15591678
SFRP1	cg15839448

- 15 -

MTNR1B	cg15842276
NMNAT3	cg15878317
PRLHR	cg16196812
SLC5A7	cg16232126
SCGN	cg16954341
UNQ9433	cg17162024
ID4	cg17252960
NELL1	cg17371081
PTF1A	cg17389519
OVOL2	cg17404915
NKX2-8	cg17685628
L3MBTL4	cg17688525
ARNT2	cg17788682
SH3GL2	cg17977409
FOXE3	cg18815943
HS3ST2	cg19064258
RIMS4	cg19332710
SLCO2A1	cg19591206
HAND1	cg19721889
RAB32	cg20098887
FBXO39	cg20723355
SNAP91	cg21688264
SLC8A2	cg22123464
SFRP1	cg22418909
SLC12A5	cg22752533
ZNF454	cg23037403
BTG4	cg23211240
NEF3	cg23290344
MYO3A	cg23771603
RAB32	cg23833452
GUCY1A2	cg23984434
CDH22	cg24101578
THRB	cg24120841
CHODL	cg24130010
PDE10A	cg24133115
FLJ90650	cg25044651
CRYBB3	cg25221625
PLOD2	cg25250358
SLC22A3	cg25313204

- 16 -

FLJ21511	cg25484904
DHCR24	cg25536676
TLX3	cg25720804
KITLG	cg25741452
CXCL1	cg25806808
EIF5A2	cg26209058
GSH1	cg26609631
MSX2	cg27096144
KCNC1	cg27409364
CTTNBP2	cg27603796
CHODL	
CNTNAP2	
COL4A3	
CPZ /// GPR78	
CRYBB3	
CXCL1	
CYP24A1	
CYR61	
DHCR24	
EGR4	
EIF5A2	
ELOVL4	
FOXE3	
GALR1	
GPR126	
GRM7	
GUCY1A2	
HAND1	
HS3ST2	
ID4	
IRX4	
KCNC1	
KCNK5	
KITLG	
LAMA1	
LGI2	
LPL	
LPL	
MALL	
MARK2	
MSX2	

- 17 -

MTNR1A
MTNR1B
MYO3A
MYOD1
NELL1
NKX2-8
NPTX2
OGDHL
ONECUT2
OVOL2
PAK7
PDE10A
PGM5
PKDREJ
PLOD2
PNMA2
PPARG
PRSS12
RAB32
RBP1
RBP4
RHCG
SCGN
SCRN1
SFRP1
SH3GL2
SLC12A5
SLC18A3
SLC22A3
SLC5A7
SLC8A2
SLCO2A1
SNAP91
SOX1
SOX17
SOX9
TFAP2C
THRB
TLX3
WWTR1
ZNF365

- 18 -

[0032] Table 2 lists genetic loci classed by function. The class of function is (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity differentially methylated between leukemia and controls. Table 3 lists all the genes identified herein as being subject to differential epigenetic regulation. Loci are selected using algorithmic models to separate leukemia bone marrow from normal bone marrow, of which any combination of 16 genes can determine this 100% of the time. Algorithmic models include the Linear Models for Microarray Analysis (LIMMA) model, the Recursive Feature Elimination- Support Vector Machine (RFE-SVM) model and Centroid model. Top candidate loci are those which overlapped all three models. These models could also be used to separate leukemia cases with poor outcome from those with a good outcome.

[0033] A list of abbreviations used herein is provided in Table 4.

TABLE 4
Abbreviations

ABBREVIATION	DESCRIPTION
A	adenine
Ab	Antibody
ALL	Acute lymphatic leukemia
AML	Acute myeloid leukemia
AMOL	Acute monocytic leukemia
AROC	Average area under the receiver operating characteristic curve
AvgLeuk	Average Beta value (methylation) of Leukemia Samples
AvgRemFol	Average Beta value (methylation) of remission and followup samples
C	Cytosine
CLL	Chronic lymphocytic leukemia
CML	Chronic myeloid leukemia
COMPARE-MS	Combination of methylated-DNA precipitation and methylation-sensitive restriction enzymes
CpG	Cytosine and guanine separated by phosphate (C-phosphate-G), which links the two nucleotides together in DNA
CpNpG	Cytosine and guanine separated by a nucleotide (N) where N is any nucleotide but guanine. The cytosine and N nucleotide are phosphorylated.
DiffBeta	Difference between AvgLeuk and AvgRemFol
DNA	Deoxyribonucleic acid
G	Guanine
LIMMA	Linear Models for Microarray Analysis
Lymphatic	Lymphoblastic or Lymphocytic
Lymphoblastic	Lymphatic
Lymphocyte	Lymphatic or Lymphoblastic

- 20 -

ABBREVIATION	DESCRIPTION
MALDI-TOF	Matrix assisted laser desorption ionisation time-of-flight
MLPA	Multiplex ligation-dependent probe amplification
MRD	Minimal/Minimum residual disease
MS-HRM	methylation specific high resolution melting
MS-MCA	methylation specific melting curve analysis
Myelogenous	Myeloid
Myeloid	Myelogenous
PCR	Polymerase Chain Reaction
ROC	Receiver operating characteristics
RFE	Recursive Feature Elimination
SLL	Small lymphocytic lymphoma
SVM	Support Vector Machine
T	Thymine
U	uracil

BRIEF DESCRIPTION OF THE FIGURES

5 [0034] Some figures contain color representations or entities. Color photographs are available from the Patentee upon request or from an appropriate Patent Office. A fee may be imposed if obtained from a Patent Office.

10 [0035] Figure 1 is a graphical representation showing correlation between methylation values of CpG units from bone marrow samples flash frozen or subject to smearing, staining with Giemsa and mounting with DPX on a microscopic slide.

15 [0036] Figure 2 is a schematic representation of hierarchical clustering of genes associated with leukemia and non-leukemia cases. This heatmap is provided in black and white. Color prints of the heatmaps are available from the patentee upon request (black represents methylated DNA 50-100% while white represents 0-50% and gradients thereof).

[0037] Figure 3 is a graphical representation showing correlation between methylation of FOXE3 and model cancer cell lines.

20 [0038] Figure 4 is a graphical representation showing correlation between methylation of TLX3 and cancer cell lines.

[0039] The cancer cell lines referred to in Figures 3 and 4 are defined in Table 5.

25 [0040] Figure 5 is a diagrammatic representation showing the single allele base extension reaction (SABER): (A) Primers are initially designed to bisulphite converted DNA (not shown), amplifying a region of interest (approximately 100-300bp). (B) An extension primer is then designed to hybridize within the amplified region. The extension primer is designed up to the CpG site of interest.

- 22 -

[0041] The SABER approach differs from normal base extension (C) as only one nucleotide is added in the extension reaction. The primer then extends over the CpG site and terminates following the incorporation of a single nucleotide.

5 [0042] The extended primers are then assayed by Mass spectrometry (D). This method has been shown to be more sensitive than standard base extension when used for SNP interrogation (Ding *et al.*, *J. Invest. Dermatol.* 123(3):470-473, 2004) and could be applied as a multiplex assay to test multiple loci for DNA methylation.

10 [0043] Figure 6 is a representation of an unsupervised clustering heatmap plot of the DNA methylation beta-values from the Illumina Infinium HumanMethylation27 BeadArray of 115 ALL-specific probes identified using three supervised learning methods (x-axis). Three distinct clusters comprising the leukemia, remission/non-leukemic and the cell lines are apparent (y-axis). The majority of these probes are hypermethylated in all leukemic
15 samples analyzed. These clusters remained when all 14,876 probes, retained after p-detection cutoff, were taken into account (Figure 9).

[0044] Figure 7 is a representation of heatmap plot of SEQUENOM EpiTYPER DNA methylation results generated from 85 cases of B-Cell ALL with matching leukemic and
20 remission bone marrow samples, controls (DONOR) and cancer cell line (REH). DNA methylation data for a total of 103 CpG sites encompassing 16 probes we selected by Infinium analysis are shown here (x-axis). DNA hypermethylation of 77/85 (91%) leukemic bone marrow samples was observed regardless of ALL subtype confirming the existence of a DNA methylation signature associated with leukemia.

25

[0045] Figure 8 is a representation showing ALL is associated with an increase in average promoter methylation. Average global promoter DNA methylation levels of bone marrow according to sample group. The overall average beta-value of probes that passed stringent quality control for all samples within a sample group are displayed. Bars represent
30 standard deviation.

- 23 -

[0046] Figure 9 shows unsupervised clustering of Infinium beta-values of 14,876 probes accurately delineates disease free tissue from leukemia. Heatmap plot of unsupervised hierarchical clustering of the beta-values of 14,876 probes passing stringent quality control.

5

[0047] Figure 10 is a representation showing the performance and the average area under the receiver operating characteristics (ROC) curve (AROC) curves of the supervised learning methods applied to the data for DNA methylation profiling. (a) Centroid, (b) RFE-SVM and (c) LIMMA plots are depicted. The average accuracy of classification (accuracy) and the average area under the receiver operating characteristics (ROC) curve (AROC) are plotted against the number of features included in the classification. The error bars represent estimated 95% confidence intervals for the obtained results. Accuracy and AROC reached 100% after 16 features for the Centroid method. Whilst the same accuracy and AROC level was achieved in less than 4 features using RFE-SVM and LIMMA.

15

[0048] Figure 11 is a representation of unsupervised hierarchical clustering heatmap of SEQUENOM EpiTYPER results. 163 paediatric leukemia cases were analyzed at three loci *Corf76*, *FBXO39* and *MYOD1* for DNA methylation. Matching leukaemic and remission bone marrow samples were analyzed from each case. Leukemia subtype for each diagnosed case is also depicted and illustrates a hypermethylation signature associated with leukaemic bone marrow.

20

[0049] Figure 12 is a graphical representation showing results from the modified multiplex methylation MALDI-TOF technique. Methylation is a measure of the of amount of extension primer that has been extended (extension occurring only on the methylated template) over the total extension primer input into each reaction.

25

- 24 -

TABLE 5
Cell Line/Sample Cell Type

CELL LINE	DESCRIPTION
MCF7	Breast Cancer
DR75	Lymphoblastoid Cell Line
DG75	Lymphoblastoid Cell Line
DR149	Unknown
40424	Normal Fibroblast
20299	Normal Fibroblast
980342	Normal Fibroblast
23970cyto	Normal Lymphoblast
25046cyto	Normal Lymphoblast
29390cyto	Normal Lymphoblast
GM10926D(N10)	Somatic Hybrid Normal
NTERA	Embryonal Carcinoma
CHANG	HeLa Derivative
K562	Myelogenous Leukemia
HELA	Cervical Cancer
PaJu	Neuroblastoma Cell Line
CCL-129 IMR32	Spleen Tumor Cell Line
Jurkat	Leukemia
GM05837A	Somatic Hybrid Normal
DR-10-22	Unknown
REH	Non-T, Non-B cell leukaemia
CEM-CCRF	Leukkaemia
SW48	Colon Cancer
SKNAS-HOXC4	Neuroblastoma Cell Line
JAM	Unknown
NCCIF	Germ Cell Tumor (NCCIT)

DETAILED DESCRIPTION

[0050] Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or method step or group of elements or integers or method steps but not the exclusion of any other element or integer or method step or group of elements or integers or method steps.

[0051] The present disclosure teaches a method for identifying an epigenetic profile associated with indicative, instructive or informative of a pathological condition associated with cancer. The epigenetic profile is in loci associated with one or more of (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity. Particular examples are listed by function in Table 2. The pathological condition may also be described as a cancerous condition or cancer such as a hematological or blood born cancer (e.g. leukemia) or solid cancerous tumors or various types and sub-types thereof. Hence, leukemias, sarcomas, carcinomas and the like including hematological malignancies of the blood, bone marrow and lymph nodes are encompassed by a "cancer" or "cancerous condition" as are there various types and sub-types. By "epigenetic change" includes a change in epigenetic profile. By "epigenetic profile" includes epigenetic modifications such as methylation including hypermethylation, hypomethylation and hydroxymethylation, RNA/DNA interactions, expression profiles of non-coding RNA, histone modification, changes in acetylation, ubiquitylation, phosphorylation and sumoylation, as well as chromatin altered transcription factor levels and the like leading to activation or deactivation of genetic locus expression. In an embodiment, the extent of methylation, RNA/DNA interaction and non-coding RNA expression are determined as well as any changes therein. In an aspect, the epigenetic modification is an elevation in methylation, an increase or decrease in methylation or an alteration in distribution of methylation sites. An example of types of leukemia include acute lymphatic leukemia (ALL), acute myeloid leukemia (AML), chronic myeloid leukemia (CML) and chronic lymphocytic leukemia (CLL), small lymphocytic lymphoma (SLL), acute monocytic leukemia (AMOL), Hodgkin's lymphomas (all types) and non-

- 26 -

Hodgkin's lymphomas (all types). Sub-types include forms of leukemia having an age-related onset bias such as pediatric leukemias and adult leukemias.

[0052] Hematological malignancies include malignancies of either myeloid or lymphoid lineages including lymphomas, lymphocytic leukemias and myelomas (lymphoid lineage malignancies) and acute and chronic myelogenous leukemia, myelodysplastic syndromes and myeloproliferative diseases (myeloid lineage malignancies).

[0053] The epigenetic profile is in one or more regions in one or more loci listed in Table 2 or 3. These loci are grouped into five classes based on function: i.e. (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity. Reference to a "locus" or "loci" includes promoter, intron, exon and non-encoding 3' and 5' regions proximal to the locus which regions are up to approximately 1000kb upstream and downstream of the locus as well as known and yet-to-be-defined regulatory elements associated with the gene. The 1000kb region is referred to herein as being "proximal" to a locus. The epigenetic profile enables determination of minimal residual disease (MRD) which comprises malignant cells at a frequency down to 10^{-3} to about 10^{-8} including frequencies of 10^{-5} , 10^{-6} and 10^{-7} and frequencies inbetween (i.e. one cancer cell per 10^3 to 10^8 cells). The epigenetic profile is determined in the genome of a cell or sub-population of cells of a subject. Any cell may be tested including but not limited to blood cells, cerebrospinal fluid (CSF) cells, bone marrow cells, buccal cells and cells from rectal swabs or feces. Cells from pre-natal tissues and embryos may also be tested. In addition, cell free DNA or RNA circulating in whole blood, serum or plasma may also be tested. Circulating DNA or RNA may also be tested in other fluids such as urine, pus, respiratory fluid, lymph fluid, feces, bile, saliva, sputum, semen, vaginal flow, cerebral spinal fluid, brain fluid, ascites, milk, secretions from the genitourinary tract and a lavage of a tissue or organ (e.g. a lung).

[0054] In an embodiment, the epigenetic profile is extent of methylation or distribution in methylation sites including an increase or decrease in this extent. Methylation includes

- 27 -

methylation of any base in DNA or RNA including methylation of cytosine, hydroxymethylation of cytosine and methylation of adenine.

[0055] Accordingly, enabled herein is a method for identifying an epigenetic profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for the extent of change in epigenetic profile relative to a control or normal/non-cancer tissue within or proximal to a locus selected from one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity wherein a change in epigenetic profile in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

[0056] Another aspect taught herein is a method for identifying a methylation profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of methylation within or proximal to a locus selected from one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity wherein a change in the extent of methylation in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

[0057] The present specification is further instructional of a method for identifying an epigenetic profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of change in epigenetic profile within or proximal to a locus selected from 1 to 16 regions in one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity wherein a change in the epigenetic profile in or proximal to one or more regions is indicative of the presence of a cancerous condition or a propensity to develop same.

- 28 -

[0058] A still further aspect enabled herein is a method for identifying a methylation profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of methylation within or proximal to from 1 to 16 regions in one or more of the loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity wherein a change in the extent of methylation in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

10 [0059] Examples of loci grouped by function are listed in Table 2. All loci identified herein are listed in Table 3.

[0060] Accordingly, enabled herein is a method for identifying an epigenetic profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for the extent of change in epigenetic profile relative to a control or normal/non-cancer tissue within or proximal to a locus selected from one or more of the loci listed in Table 2 or 3 wherein a change in epigenetic profile in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

20

[0061] Another aspect taught herein is a method for identifying a methylation profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of methylation within or proximal to a locus selected from one or more of the loci listed in Table 2 or 3 wherein a change in the extent of methylation in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

[0062] The present specification is further instructional of a method for identifying an epigenetic profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control

30

- 29 -

or normal/non-cancer tissue in the extent of change in epigenetic profile within or proximal to a locus selected from 1 to 16 regions in one or more of the loci listed in Table 2 or 3 wherein a change in the epigenetic profile in or proximal to one or more regions is indicative of the presence of a cancerous condition or a propensity to develop same.

5

[0063] A still further aspect enabled herein is a method for identifying a methylation profile in the genome of a cell indicative of a cancerous condition or a predisposition thereto, the method comprising screening for a change relative to a control or normal/non-cancer tissue in the extent of methylation within or proximal to from 1 to 16 regions in one or more of the loci listed in Table 2 or 3 wherein a change in the extent of methylation in or proximal to one or more loci is indicative of the presence of a cancerous condition or a propensity to develop same.

10

[0064] Reference can be made to either Table 2 or Table 3 to identify loci which are subject to differential epigenetic regulation.

15

[0065] The present disclosure also teaches the use of an epigenetic profile within or proximal to a locus selected from one or more of the list in Table 2 or 3 in the manufacture of an epigenetic assay to provide data which are indicative of the presence of a cancerous condition or a propensity to develop a cancerous condition. In an embodiment, the epigenetic assay determines extent of methylation and/or distribution of methylation sites within or proximal to the locus. In an embodiment, from 1 to the number of loci or regions within a locus required to achieve a 100% confidence level that a subject has or does not have cancer are selected. In an embodiment, this number is 16. Hence, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or 16 different loci, or 1 to 16 different regions within a single locus, or a combination of both are assayed for any epigenetic change such as change in methylation profile. By function, the loci are associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity.

20

25

30

- 30 -

[0066] The epigenetic profile enables the determination of MRD with high sensitivity (down to 10^{-3} to 10^{-8}) in a patient. The identification of MRD facilitates the decision process on whether more intensive therapy is required, such as more toxic therapy, a bone marrow transplant or stem cell therapy or whether an alternative type of therapy is required. Hence, the MRD influences the prognosis of a cancer in a subject.

[0067] Hence, taught herein is a method for determining MRD in a subject, the method comprising determining an epigenetic profile in or proximal to one or more loci listed in Table 2 or 3 wherein an elevation or a decrease in epigenetic change of a specific gene or genes, relative to disease free control tissue is an indication of the MRD.

[0068] The present disclosure further enables a method for determining MRD in a subject, the method comprising determining a change in methylation profile in or proximal to one or more loci or in one or more regions within the one or more loci listed in Table 2 or 3 wherein an elevation or a decrease in extent of methylation is an indication of the MRD. By "methylation profile" means extent of and/or distribution of methylation sites within multiple loci, a single locus or a combination of both.

[0069] Further taught herein is a method for determining MRD in a subject, the method comprising determining an epigenetic profile in or proximal to from 1 to 16 loci listed in Table 2 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 wherein an elevation or a decrease in extent of epigenetic change is an indication of the MRD.

[0070] Yet a further aspect enabled herein is a method for determining MRD in a subject, the method comprising determining a methylation profile in or proximal from 1 to 16 loci listed in Table 2 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 wherein an elevation or a decrease in extent of methylation is an indication of the MRD.

[0071] As indicated above, 16 different loci or 16 different regions within one or more loci provides a 100% confidence level of the presence or absence of cancer in a subject. The present disclosure contemplates confidence controls of from about 40% to 100% including

- 31 -

40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 and 100%. Any number of loci or regions within one or more loci may be selected to provide from about 40% to about 100% confidence level of the presence or absence of cancer. More than 16 loci or regions may be assayed such as from 16 to 1000 or greater. This may occur when an array is used such as in solid phase amplification using from 1 to 1000 immobilized primers on, for example, a chip. Reference to "1 to 1000" include "16 to 1000" means 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449,

- 32 -

450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467,
468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485,
486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503,
504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521,
5 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539,
540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557,
558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575,
576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593,
594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611,
10 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629,
630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647,
648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665,
666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683,
684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701,
15 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719,
720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737,
738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755,
756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773,
774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791,
20 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809,
810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827,
828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845,
846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863,
864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881,
25 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899,
900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917,
918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935,
936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953,
954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971,
30 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989,
990, 991, 992, 993, 994, 995, 996, 997, 998, 999 or 1000.

- 33 -

[0072] In accordance with the present disclosure, a method is provided wherein the extent of methylation provides a quantitative or semi-quantitative or qualitative indication of malignant cells at a frequency down to 10^{-3} to 10^{-8} , including 10^{-5} , 10^{-6} and 10^{-7} and frequencies inbetween. The method described herein may also be used in conjunction with other assays such as quantitative Reverse transcriptase PCR (qRT-PCR), TaqMan, gene expression micro arrays, Northern or Western blot procedures to measure gene expression and transcription. Other assays contemplated herein include quantitative fluorescent PCR (QF-PCR), multiplex fluorescent PCR (MF-PCR), real time PCR (RT-PCR), single cell PCR, restriction fragment length PCR (PCR-RFLP), PCR-RFLP/RT-PCR-RFLP, hot start PCR, nested PCR, *in situ* polonony PCR, *in situ* rolling circle amplification, bridge PCR, picotiter PCR, emulsion PCR, next generation/massively parallel sequencing and single molecule sequencing. Other amplification methods include selective amplification of target polynucleotide sequences, consensus sequence primed PCR (CP-PCR), arbitrarily primed PCR (AP-PCR), degenerate oligonucleotides-primed PCR (DOP-PCR) and nucleic acid based sequence amplification (NABSA).

[0073] Following amplification, DNA methylation status can be determined in any number of ways such as by electrophoresis which includes capillary, capillary zone, capillary isoelectric focusing and capillary gel electrophoresis as well as capillary electrochromatography, micellar electrokinetic capillary chromatography and transient isotachopheresis by use of arrays, beads, gas chromatography, supercritical fluid chromatography, liquid chromatography (which encompasses partition, adsorption, ion exchange, size exclusion, thin-layer and affinity chromatography). Other techniques include comparative genomic hybridization, microarrays, bead arrays, high-throughput genotyping such as with the use of a molecular inversion probe, MALDI-TOF and sequencing by synthesis (SBS).

[0074] Within the meaning of a "pathological condition" or "disease condition" includes an abnormal malignancy as defined by objective or subjective manifestations of cancer diseases. The assay of the present disclosure enables a molecular (i.e. genetic or

- 34 -

epigenetic) determination to be made to complement other symptom-based diagnoses such as based on physiological or medical studies or may be made in its own right. The assay may be part of a suite of diagnostic or prognostic genetic assays of embryos, pre- and post-natal subjects. The terms "method", "assay", "system", "test", "determination",
5 "prognostic", "diagnostic", "report" and the like may all be used to describe an epigenetic assay including a methylation assay of selected regions of one or more loci listed in Table 2, the number selected on the basis of the required level of confidence. The methylation assay determines the epigenetic profile or extent of methylation change compared to a control which suggests or indicates or is instructive of a cancerous condition. The present
10 disclosure also teaches epidemiological studies of populations including studies of different ethnic populations with cancer.

[0075] Accordingly, the present disclosure further provides a method of identifying an epigenetic profile in a population of subjects indicative of a cancerous pathological
15 condition the method comprising screening for a change relative to a control in a statistically significant number of subjects of the extent of epigenetic change within or proximal to one or more loci listed in Table 2 or 3 wherein the extent of epigenetic change is indicative of the presence of the pathological condition or a propensity to develop same.

20 [0076] Also taught herein is a method of identifying a methylation profile in a population of subjects indicative of a cancerous pathological condition the method comprising screening for a change relative to a control in a statistically significant number of subjects the extent of methylation within or proximal to one or more loci listed in Table 2 or 3 wherein an increase or decrease in extent of methylation is indicative of the presence of the
25 pathological condition or a propensity to develop same.

[0077] Another aspect of the present disclosure enabled herein is a method of identifying epigenetic profile in a population of subjects indicative of a cancerous pathological condition the method comprising screening for a change relative to a control in a
30 statistically significant number of subjects the extent of epigenetic change within or proximal to from 1 to about 16 loci listed in Table 2 or 3 or in 1 to 16 regions within one or

- 35 -

more loci listed in Table 2 or 3 wherein an elevation or decrease in extent of epigenetic change is indicative of the presence of the pathological condition or a propensity to develop same.

- 5 [0078] Yet a further taught herein is a method of identifying a methylation profile in a population of subjects indicative of a cancerous pathological condition the method comprising screening for a change relative to a control in a statistically significant number of subjects the extent of methylation within or proximal to one or more loci listed in Table 2 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 wherein an increase or
10 decrease in extent of methylation is indicative of the presence of the pathological condition or a propensity to develop same.

[0079] As indicated above, the total number of sites selected depends on the level of confidence required for a diagnosis. For example, from about 40% to 100% confidence
15 levels. A total of 16 different loci, or 16 different regions within one or more loci provides a 100% confidence level of the presence or absence of cancer.

[0080] These methods may also be conducted wherein no change or the presence of a change in extent of methylation is indicative of a non-pathological condition or at least an
20 amelioration in the pathogenicity of the condition.

[0081] The term "proximal" includes the region up to approximately 1000kb upstream or downstream (5' or 3' terminal regions) of a locus.

25 [0082] In an embodiment, a genome wide methylation map has been constructed in accordance with the present disclosure using standard techniques such as and DNA hybridisation microarray and high throughput mass spectrometry of various cells. Any cell type cell may be assayed. These cells include blood cells, CSF cells, bone marrow cells, buccal cells and cells from rectal swabs or faeces.

- 36 -

5 [0083] The present disclosure contemplates that the extent of epigenetic change in within or proximal to one or more loci listed in Table 2 or 3 which corresponds to a healthy condition or a level of disease within the spectrum of cancer. The extent of change in epigenetic profile, as indicated above, permits the sensitive determination of MRD in a subject and its use in prognosis of cancer.

[0084] The cancerous condition includes leukemias and its various types and sub-types such as pediatric and adult leukemias.

10 [0085] Insofar as the epigenetic profile relates to methylation, methylation may occur anywhere within the DNA or RNA of or proximal to a locus including of any cytosine whether in islands or shores or other areas of the nucleic acid. As used herein, the terms "cytosine" or "C", "CpG islands", "CpNpG islands", "island shores" and "shores" all include these basis in a locus or in a region up to approximately 1000kb in distance
15 upstream or downstream from a locus listed in Table 2 or 3 or more particularly Table 2 or 3 as well as within the locus and upstream and downstream enhancer elements. Multiple loci may be screened or multiple regions within one or more loci screened.

[0086] As used herein, the terms "subject", "case", "patient", "individual", "target" and the
20 like refer to any organism or cell of the organism on which an assay described herein is performed whether for experimental, diagnostic, prophylactic, and/or therapeutic purposes. Typical subjects include both male and female humans but the present disclosure extends to experimental animals such as non-human mammals, (e.g., monkeys, chimpanzees, orangutangs, gorillas, mice, rats, rabbits, sheep, pigs, cows, horses and guinea
25 pigs/hamsters). Other test cells may be used such as yeast, microorganisms, single cell organisms to test fundatmental gene function which is conserved throughout phyla. The "subject" may also be referred to as a population since the present disclosure is useful in epidemiological studies or assays of an ethnic population.

30 [0087] The term "genomic DNA" includes all DNA in a cell, group of cells, or in an organelle of a cell and includes exogenous DNA such a transgenes introduced into a cell.

- 37 -

[0088] Hence, the present disclosure teaches a method for identifying an epigenetic profile in the genome of a cell indicative of a cancerous pathological condition, the method comprising screening for a change relative to the control in the extent of epigenetic change within or proximal to one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 wherein the extent of epigenetic change is indicative of the presence of the cancerous pathological condition or a propensity to develop same. This method is particularly useful in the screening of a change in extent of methylation or distribution of methylation sites. The change may be indicative of health, remission or cancer.

[0089] As indicated above, the loci may be grouped based on function, i.e. loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity.

[0090] Any methylation assay may be employed such as bisulfite sequencing (see, for example, WO 2005/038051), methylation specific melting curve analysis (MS-MCA), high resolution melting (MS-HRM) [Dahl *et al.*, *Clin Chem* 53(4):790-793, 2007; Wojdacz *et al.*, *Nucleic Acids Res.* 35(6):e41, 2007], MALDI-TOF MS (Tost *et al.*, *Nucleic Acids Res.* 31(9):e50, 2003), methylation specific MLPA (Nygren *et al.*, *Nucleic Acids Res.* 33(14):e128, 2005); methylated-DNA precipitation/enrichment and methylation-sensitive restriction enzymes (COMPARE-MS) [Yegnasubramanian *et al.*, *Nucleic Acids Res.* 34(3):e19, 2006] or methylation sensitive oligonucleotide microarray (Gitan *et al.*, *Genome Res.* 12(1):158-164, 2002), Infinium (Bibikova *et al.*, *Genome Research* 16:383-393, 2006) and MethyLight (Trinh *et al.*, 2001 *supra*; Shivapunkar *et al.*, 2005 *supra*; WO 00/70090; US Patent No. 6,331,393), via antibodies and protein binding domains targeted to methylated DNA as well as single molecule real time sequencing (Pac Biosciences; Flushberg *et al.*, *Nat Methods* 7(6):461-465, 2010). Multiplex methylation based PCR assays and multiplex kits are also contemplated for use herein. See also Laird 2010 *supra*; and Hansen *et al.* 2011 *supra*. The primers used to amplify DNA prior to bisulfite treatment may need to be modified if the bisulfite treated DNA is to be amplified.

- 38 -

[0091] Insofar as a methylation assay involves any amplification, an amplification methodology may be employed. Amplification methodologies contemplated herein include the polymerase chain reaction (PCR) such as disclosed in U.S. Patent Nos. 4,683,202 and 5 4,683,195; the ligase chain reaction (LCR) such as disclosed in European Patent Application No. EP-A-320 308 and gap filling LCR (GLCR) or variations thereof such as disclosed in International Patent Publication No. WO 90/01069, European Patent Application EP-A-439 182, British Patent No. GB 2,225,112A and International Patent Publication No. WO 93/00447. Other amplification techniques include Q β replicase such 10 as described in the literature; Strand Displacement Amplification (SDA) such as described in European Patent Application Nos. EP-A-497 272 and EP-A-500 224; Self-Sustained Sequence Replication (3SR) such as described in Fahy *et al.*, *PCR Methods Appl.* 1(1):25-33, 1991) and Nucleic Acid Sequence-Based Amplification (NASBA) such as described in the literature.

15

[0092] A PCR amplification process is useful in the practice of the assay enabled herein.

[0093] Another aspect of the present disclosure contemplates a method for determining the epigenetic profile within the genome of a eukaryotic cell or group of cells, the method 20 comprising obtaining a sample of genomic DNA or RNA from the cell or group of cells and subjecting the genomic DNA or RNA to bisulfite treatment/conversion followed by primer-specific amplification within one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 and assaying for extent of change in epigenetic profile relative to a control.

25

[0094] The present disclosure further provides a method for determining the methylation profile within the genome of a eukaryotic cell or group of cells, the method comprising obtaining a sample of genomic DNA or RNA from the cell or group of cells and subjecting the genomic DNA or RNA to bisulfite treatment/conversion followed by primer-specific 30 amplification within one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one

- 39 -

or more loci listed in Table 2 or 3 and assaying for extent of change in methylation relative to a control.

5 [0095] The present disclosure still further enables a method for determining the epigenetic profile within the genome of a eukaryotic cell or group of cells, the method comprising obtaining a sample of genomic DNA or RNA from the cell or group of cells and subjecting the genomic DNA or RNA to bisulfite treatment/conversion followed by primer-specific amplification within from 1 to 16 loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 and assaying for extent of change in epigenetic profile.

10

[0096] Yet a further aspect taught herein is a method for determining the methylation profile within the genome of a eukaryotic cell or group of cells, the method comprising obtaining a sample of genomic DNA or RNA from the cell or group of cells and subjecting the genomic DNA or RNA to bisulfite treatment/conversion followed by primer-specific amplification within from 1 to 16 loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 and assaying for extent of methylation relative to a control.

15

[0097] From 1 to 16 different loci may be assayed or from 1 to 16 different sites within from 1 to 16 loci may be assayed.

20

[0098] A change in epigenetic profile such as a change in extent of methylation or distribution of methylation sites compared to a control (i.e. a subject with known disease status) determines if a subject has cancer, is healthy or is in remission.

25

[0099] Similarly, genomic DNA or RNA could be subjected to methylation sensitive restriction digestion and amplification for similar assaying for the extent of methylation. Furthermore, the above methods can be used to detect a decrease in extent of methylation.

30 [0100] A "nucleic acid" as used herein, is a covalently linked sequence of nucleotides in which the 3' position of the phosphorylated pentose of one nucleotide is joined by a

- 40 -

phosphodiester group to the 5' position of the pentose of the next nucleotide and in which the nucleotide residues are linked in specific sequence; i.e. a linear order of nucleotides. A "polynucleotide" as used herein, is a nucleic acid containing a sequence that is greater than about 100 nucleotides in length. An "oligonucleotide" as used herein, is a short polynucleotide or a portion of a polynucleotide. An oligonucleotide typically contains a sequence of about two to about one hundred bases. The word "oligo" is sometimes used in place of the word "oligonucleotide". The term "oligo" also includes a particularly useful primer length in the practice of the present invention of up to about 10 nucleotides.

10 [0101] As used herein, the term "primer" refers to an oligonucleotide or polynucleotide that is capable of hybridizing to another nucleic acid of interest under particular stringency conditions. A primer may occur naturally as in a purified restriction digest or be produced synthetically, by recombinant means or by PCR amplification. The terms "probe" and "primers" may be used interchangeably, although to the extent that an oligonucleotide is
15 used in a PCR or other amplification reaction, the term is generally "primer". The ability to hybridize is dependent in part on the degree of complementarity between the nucleotide sequence of the primer and complementary sequence on the target DNA.

[0102] The terms "complementary" or "complementarity" are used in reference to nucleic
20 acids (i.e. a sequence of nucleotides) related by the well-known base-pairing rules that A pairs with T or U and C pairs with G. For example, the sequence 5'-A-G-T-3' is complementary to the sequence 3'-T-C-A-5' in DNA and 3'-U-C-A-5' in RNA or bisulfite treated genomic DNA. Complementarity can be "partial" in which only some of the nucleotide bases are matched according to the base pairing rules. On the other hand, there
25 may be "complete" or "total" complementarity between the nucleic acid strands when all of the bases are matched according to base-pairing rules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands as known well in the art. This is of particular importance in detection methods that depend upon binding between nucleic acids, such as
30 those of the invention. The term "substantially complementary" is used to describe any primer that can hybridize to either or both strands of the target nucleic acid sequence under

- 41 -

conditions of low stringency as described below or, preferably, in polymerase reaction buffer heated to 95°C and then cooled to room temperature. As used herein, when the primer is referred to as partially or totally complementary to the target nucleic acid, that refers to the 3'-terminal region of the probe (i.e. within about 10 nucleotides of the 3'-
5 terminal nucleotide position).

[0103] The present disclosure teaches, therefore, a methylation profile of the sites within one or more loci listed in Table 2 or 3 in a genome of a eukaryotic cell or group of cells, the methylation profile comprising the extent or level of methylation, the method
10 comprising obtaining a sample of genomic DNA or RNA from the cell or group of cells, subjecting the restriction digested DNA or RNA or bisulfite treated DNA or RNA to an amplification reaction using primers selected to amplify one or more regions in one or more loci listed in Table 2 or 3 or modified primers which can amplify bisulfite treated DNA or RNA and then subjecting the amplified DNA or RNA to methylation detection
15 means to determine relative to control the extent of methylation wherein a change in methylation relative to the control is indicative of a cancerous pathological condition. The control may be a normal subject or a subject with known disease status.

[0104] The present disclosure also contemplates kits for determining the epigenetic profile
20 of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells. The kits may comprise many different forms but in one embodiment, the kits comprise reagents for the bisulfite methylation assay and primers for amplification.

[0105] A further embodiment taught by the present disclosure is a kit for the use in the
25 above methods comprising primers to amplify a particular site within one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 or which amplify a strand after bisulfite conversion.

[0106] The kit may also comprise instructions for use.

- 42 -

[0107] Conveniently, the kits are adapted to contain compartments for two or more of the above-listed components. Furthermore, buffers, nucleotides and/or enzymes may be combined into a single compartment.

5 [0108] As stated above, instructions optionally present in such kits instruct the user on how to use the components of the kit to perform the various methods taught herein. It is contemplated that these instructions include a description of the detection methods enabled herein.

10 [0109] The present disclosure further contemplates kits which contain a primer for a nucleic acid target of interest with the primer being complementary to a predetermined nucleic acid target or a bisulfite treated template. In another embodiment, the kit contains multiple primers or probes, each of which contains a different base at an interrogation position or which is designed to interrogate different target DNA or RNA sequences. In a
15 contemplated embodiment, multiple probes are provided for a set of nucleic acid target sequences that give rise to analytical results which are distinguishable for the various probes. The multiple probes may be in microarray format for ease of use.

[0110] It is contemplated that a kit comprises a vessel containing a purified and isolated
20 enzyme whose activity is to release one or more nucleotides from the 3' terminus of a hybridized nucleic acid probe and a vessel containing pyrophosphate. In one embodiment, these items are combined in a single vessel. It is contemplated that the enzyme is either in solution or provided as a solid (e.g. as a lyophilized powder); the same is true for the pyrophosphate. Preferably, the enzyme is provided in solution. Some contemplated kits
25 contain labeled nucleic acid probes. Other contemplated kits further comprise vessels containing labels and vessels containing reagents for attaching the labels. Microtiter trays are particularly useful and these may comprise from two to 100,000 wells or from about six to about 10,000 wells or from about six to about 1,000 wells.

30 [0111] The present disclosure also teaches genome wide screening for epigenetic profiles including methylation profiles.

[0112] Another important application is in the high throughput screening of agents which are capable of demethylation genomes. This may be important, for example, in de-differentiating cells and cancer therapies.

5

[0113] The present disclosure further enables a method for screening for an agent which modulates epigenetic regulation of one or more loci or a region within one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising screening for a change relative to a control in the epigenetic profile within or proximal to a locus in the presence or absence of an agent to be tested, wherein an agent is selected if it induces a change in the extent of epigenetic.

[0114] A method is also provided for screening for an agent which modulates methylation of one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising screening for a change relative to a control in the extent of methylation within or proximal to a locus in the presence or absence of an agent to be tested, wherein an agent is selected if it induces a change in the extent of methylation.

[0115] A further method is provided for screening for an agent which modulates epigenesis of one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising screening for a change relative to a control in the extent of change in the epigenetic profile within or proximal to a locus in the presence or absence of an agent to be tested, wherein an agent is selected if it induces a change in the extent of epigenesis.

[0116] A still further method is provided for screening for an agent which modulates methylation of one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising screening for a change relative to a control in the extent of methylation within or proximal to a locus in the presence or

- 44 -

absence of an agent to be tested, wherein an agent is selected if it induces a change in the extent of methylation.

5 [0117] The present disclosure further enables a method for monitoring the treatment of a cancer in which the treatment modulates the epigenetic profile of one or more loci or a region within one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising monitoring for a change relative to a control or a pre- and post-treatment sample in the epigenetic profile within or proximal to one or more loci.

10

[0118] A method for monitoring the treatment of a cancer is provided in which the treatment modulates the methylation of one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising monitoring for a change relative to a control or a pre- and post-treatment sample in the extent of
15 methylation within or proximal to one or more loci.

[0119] A further method for monitoring the treatment of a cancer is provided in which the treatment modulates the epigenetic profile of one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising
20 monitoring for a change relative to a control or a pre- and post-treatment sample in the extent of the epigenetic profile within or proximal to one or more loci.

[0120] A still further method for monitoring the treatment of a cancer is provided in which the treatment modulates the methylation of one or more loci or a region within one or more
25 loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 the method comprising monitoring for a change relative to a control or a pre- and post-treatment sample in the extent of methylation within or proximal to one or more loci.

[0121] Reference to "one or more loci listed in Table 2 or 3" includes from 1 to 16 loci
30 listed in Table 2 or 3. By function, the 1 to 16 loci may be associated with (i) cell fate

- 45 -

commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity (see Table 2).

5 [0122] In another embodiment, the identification of MRD is provided such as a malignant cell at a frequency down to 10^{-4} to 10^{-8} (i.e. one cell in from 10^4 to 10^8 cells) which may be instructive in the decision as to whether to apply more intensive therapy such as more toxic therapy, a bone marrow transplant or stem cell therapy.

10 [0123] In a particular embodiment, methylation is detected in any C such as a C in CpG or CpNpG islands and/or island shores. In another particular embodiment, the cancer is a leukemia or its various types or sub-types.

15 [0124] In cases where the gene is methylated and silenced in affected individuals or tissues, compounds are screened in high throughput fashion in stable cell lines or individuals to identify drugs that result in demethylation and reactivation of the affected gene. Alternatively, a normal active copy of the affected gene is transfected as a transgene into cells to correct the defect. Such transgenes are introduced with modulating sequences that protect the transgene from methylation and keep it unmethylated and transcriptionally active.

20

[0125] In cases where the gene is unmethylated and transcriptionally active or transcriptionally over-active in affected individuals or tissues, compounds are screened in high throughput fashion in stable cell lines to identify drugs that result in methylation and silencing of the affected gene. Alternatively, a transgene encoding a double stranded RNA homologous to the affected sequences or homologs thereof, are transfected as a transgene
25 into cells to methylate the gene, silence it and thereby correct the defect. Such double stranded RNA-encoding transgenes are introduced with modulating sequences which protect it from methylation, keep it transcriptionally active and producing double stranded RNA.

30

- 46 -

[0126] The present disclosure further teaches a computer program and hardware which monitors the changing state, if any, of extent of methylation over time or in response to therapeutic and/or behavioral modification. This includes for monitoring MRD. Such a computer program has important utility in monitoring disease progression, response to
5 intervention and may guide modification of therapy or treatment. The computer program is also useful in understanding the association between increasing methylation and disease progression.

[0127] Thus, in accordance with the present disclosure, index values are assigned to levels
10 of methylation in selected loci or proximal thereto which are stored in a machine-readable storage medium, which is capable of processing the data to provide an extent of disease progression or change in methylation for a subject.

[0128] Thus, in another aspect, the disclosure teaches a computer program product for
15 assessing progression of a pathological condition associated with cancer in a subject, the product comprising:

- (1) assigning values to extent of methylation or other epigenetic change in one or more regions of one or more loci listed in Table 2 or 3;
- 20 (2) means to converting the value to a code; and
- (3) means to store the code in a computer readable medium.

[0129] In a related aspect, the present disclosure levels a computer for assessing an association between extent of methylation or other epigenetic change within one or more
25 regions within one or more loci listed in Table 2 or 3 and progression of a cancerous condition wherein the computer comprises:

- (1) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein the machine-readable data comprise values
30 associated with extent of methylation or other epigenetic change in one or more regions of one or more loci listed in Table 2 or 3;

- 47 -

- (2) means to converting the value to a code; and
- (3) means to store the code in a computer readable medium.

[0130] In a related aspect, the present disclosure extends to a computer for assessing an association between extent of methylation or other epigenetic change within one or more loci listed in Table 2 or 3 or in 1 to 16 regions within one or more loci listed in Table 2 or 3 and progression of a cancer disease condition wherein the computer comprises a machine-readable data storage medium comprising a data storage material encoded within machine-readable data, wherein the machine-readable data comprise values associated with the features extent of methylation or other epigenetic change in one or more regions of one or more loci listed in Table 2 or 3.

[0131] The present disclosure also teaches a computer comprising:

- (1) a working memory for storing instructional codes for processing the machine-readable data;
- (2) a central-processing unit coupled to the working memory and to the machine-readable data storage medium, for processing the machine-readable data to provide data instructional or informative of changing methylation or other epigenetic patterns or disease progression; and
- (3) an output hardware coupled to the central processing unit, for receiving the data.

[0132] The computer system herein may also be linked to detection systems such as MALDI-TOF machines.

[0133] Reference to one or more loci herein includes from 1 to 100 such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 and 100. In an

- 48 -

embodiment, from 1 to 50 or 1 to 30 or 1 to 20 loci are screened including 1 to 16 such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 and 16. The present disclosure further encompasses from 1 to 100 sites or regions in 1 or more loci such as 16 sites from one or more loci.

5

[0134] Aspects taught herein are now further described by the following non-limiting Examples.

EXAMPLE 1***Analysis of genetic loci*****Methods**

- 5 [0135] Let $\vec{x} \in \mathbb{R}^n$ be a lipid profile (with or without clinical features) and $y \in \{-1, 1\}$ be the corresponding binary label (e.g. SU or CD). A linear least-squares support vector machine was used (Bernhard and Smola, *Learning with kernels* MIT Press, 2002) model of the form $f_{\vec{\beta}, b}(\vec{x}) = \langle \vec{x}, \vec{\beta} \rangle$. The coefficients $\vec{\beta} = (\beta_i) \in \mathbb{R}^n$ are found by minimizing the following functional:

10

$$\arg \max_{\vec{\beta}} \Xi(\vec{\beta}) \triangleq \sum_i \max^2 \left(1 - \bar{y}_i (\langle \vec{x}_i, \vec{\beta} \rangle), 0 \right) + \lambda \|\vec{\beta}\|^2, \quad (1)$$

where λ , is the regularization constant.

- [0136] After initial assessment of few different techniques for feature selection, focus was
15 on recursive feature elimination technique (Guyon *et al.*, *Machine Learning* 46:389-4122, 2002):

1. Let $\theta_1 = \{1, \dots, n\}$ be the set of active features.
2. Build an initial model $\vec{\beta}^{(1)}$ using all features $(x_i \mid i \in \theta_1)$ by minimizing Equation
- 20 1. Then iterate the following "for" loop.
3. For $t = 1, \dots, n$.
 - (a) Estimate the least important feature $i_t := \arg \min_i |\beta_i^{(t)}|$.
 - (b) Discard feature i_t , set $\theta_{t+1} = \theta_t \setminus \{i_t\}$.
 - 25 (c) Build the $t + 1$ -th model

$$\vec{\beta}^{(t+1)} = \beta_i^{(t+1)} \in \mathbb{R}^{n-t-1}$$

using the features set with i_t feature discarded, i.e. by minimizing Equation
1 for the features $(x_i \mid i \in \theta_{t+1})$, where $\theta_{t+1} = \theta_t \setminus \{i_t\}$.

- 50 -

[0137] Feature importance was determined depending on when the feature was removed from the model. The first feature eliminated is assigned the lowest rank; the next feature removed is assigned the next lowest rank, etc. This provides a ranking from the best feature (last removed) to the worst (first removed). The overall ranking of the features across all bootstrap samples is defined as the average rank across all subsampled data sets.

Evaluation

[0138] To estimate the generalization performance of the models, $\epsilon=0$ bootstrap repeated 50 times (Hastie *et al.*, *The Elements of Statistical Learning* Springer, 2001). One bootstrap sample consists of constructing a training set of the same size as the original data set by sampling the data with replacement. A corresponding independent test set is then formed using all samples not present in the training set.

15

[0139] Two performance metrics were used. The area under the receiver operating characteristic (AROC) and the error rate.

[0140] The AROC is the chance of the classifier correctly ranking samples, i.e. the chance that the model would take a random sample from the positive class and a random sample from the negative class and rank the positive class higher than the negative class. This metric is popular in machine learning and biomedical research.

[0141] More formally, given a model $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the AROC is defined as the probability of pairs being correctly ordered, or more formally given two pairs of test points with $y > y'$, the AROC is defined as:

$$AROC(f) \triangleq P[f(\bar{x}) > f(\bar{x}') \mid y > y'] + \frac{1}{2} P[f(\bar{x}) = f(\bar{x}') \mid y > y'].$$

- 51 -

assuming above that $(\vec{x}, y)$ and $(\vec{x}', y')$ belong to the space of labeled sample $\mathcal{X} \subset \mathbb{R}^n \times \{-1, 1\}$. In particular, given a finite set of test samples X , the AROC can be estimated as:

$$\frac{1}{|\mathcal{X}^+||\mathcal{X}^-|} \sum_{\vec{x} \in \mathcal{X}^+} \sum_{\vec{x}' \in \mathcal{X}^-} \mathbb{I}(f(\vec{x}) > f(\vec{x}')) + \frac{1}{2} \mathbb{I}(f(\vec{x}) = f(\vec{x}')),$$

5

where $|\cdot|$ denotes set cardinality; $\mathbb{I}(\cdot)$ denotes the indicator function, defined as 1 if its argument is true, and 0, otherwise; and $\mathcal{X}^+ := \{\vec{x} | (\vec{x}, 1) \in \mathcal{X}\}$ and $\mathcal{X}^- := \{\vec{x} | (\vec{x}, -1) \in \mathcal{X}\}$.

EXAMPLE 2

10

Selection of samples

[0142] Leukemia cases were all diagnosed within the Children's Cancer Centre, Royal Children's Hospital, Melbourne, Victoria, Australia. Primary tumor, 28-day post induction and follow up bone marrow smears on microscope slides (6, 12, 24 months post induction) were used to extract genomic DNA for DNA methylation analysis. A subset of primary tumor bone marrow was used to determine if DNA methylation changes occur as a result of the microscope slide process.

15

EXAMPLE 3

20

Illumina Infinium human methylation 27 DNA microarray

[0143] Thirteen cases of TEL/AML(ETV6/RUNX1) positive leukemia (7 male, 6 female) were chosen for genome-wide DNA methylation analysis. Matched 28-day remission and subsequent follow up bone marrow samples were also analyzed along with non-leukemic bone marrow samples as controls to compare to. The cases from which the samples were analyzed are listed in Table 6 below.

25

- 52 -

TABLE 6
Leukemia samples

Case ID	Age Dx (years)	Sex	Cytogenetics	Leuk	Rem	Follow
235	5.13	F	46,XX,del(12)(p13)[3]/46,XX[14] TEL/AML1 Fusion	Yes	Yes	Yes
088	9.32	M	46,XY,l(21)(q10)[15]/46,XY[1]	No	Yes	Yes
102	2.97	M	47,XY,+21c	Yes	No	No
084	7.37	M	46,XY Normal FISH	No	Yes	Yes
089	5.37	M	46,XY Normal FISH	Yes	No	No
094	7.24	M	46,XY,del(7)(q32),-11,+r[3]/46,XY[7]	Yes	No	No
114	6.96	F	46,XX Normal FISH	Yes	Yes	Yes
123	5.99	F	46,XX Normal FISH	Yes	No	No
127	4.76	F	46,XX,del(11)(q23)[2]/45,X,-X,del(11)(q23),inc[5]/46,XX [5], TEL/AML1 positive	Yes	Yes	No
133	2.37	M	46,XY, TEL/AML Positive	Yes	No	Yes
143	13.05	F	45,X,-X,del(6)(q21),add(12)(p13),-13,+16[3]/46,idem,+21[2]/46,XX[7] TEL/AML1 positive	Yes	No	Yes
175	2.81	M	46,XY	Yes	Yes	No
216	4.21	M	46,XY nuc ish(TELx2),(AML1x3), (TEL con AML1x1)[185/199]/ (MLLx1) [39/198]	Yes	Yes	Yes
344	2.57	F	47,X,- X, ?add(9)(q34), +10,add(12)(p13),+mar[6]/46,XX[9] TEL/AML1 positive	Yes	Yes	Yes

- 53 -

[0144] The genomic DNA samples from the selected cases were processed by the Australian Genome Research Facility (AGRF) for analysis. Results from Infinium analysis were cleaned to remove poor performing probes across all samples analyzed (with a p-detection value of great than 0.05). R was used for statistical analysis (<http://www.r-project.org>) and generation of heatmap plots using the gplots library.

EXAMPLE 4

Bisulphite conversion of genomic DNA

[0112] Genomic DNA was subjected to bisulphite conversion using either MethylEasy Xceed (Human Genetic Signatures) or EZ-96 DNA Methylation gold (Zymo Research) kits according to manufacturer's instructions. Converted DNA was eluted with sufficient volume of elution buffer to give a final sample concentration of 20ng/μL.

EXAMPLE 5

SEQUENOM MassArray EpiTYPER chemistry

[0113] PCR amplification of bisulphite converted genomic DNA was performed using gene specific primers containing the necessary tags for SEQUENOM analysis (Table 7) and Fast Start PCR master mix (Roche Applied Sciences) according to manufacturer's instructions. Typical PCR cycling conditions performed are: 95°C for 10 minutes denaturation followed by five cycles of 95°C for 10 seconds denaturation, 56°C annealing for 20 seconds, 72°C extension for two minutes and 40 cycles of 95°C for 10 seconds denaturation, 56°C annealing for 20 seconds, 72°C extension for 1.5 minutes. Successful amplicons were then subject to SEQUENOM EpiTYPER chemistry consisting of SAP treatment, RNA transcription and cleavage prior to mass spectrophotometry analysis as outlined in the manufacturer's instructions. Mass spectra were processed using EpiTYPER viewer software v1.0.5 (SEQUENOM Inc.) and cleaned using an in-house R-script to remove poor quality CpG units and samples. Heatmaps of SEQUENOM data were also drawn using heatmap.2 of the gplots library (<http://www.r-project.org>).

- 54 -

TABLE 7

Genomic loci selection for SEQUENOM DNA methylation analysis

Gene Name	Location of Assay (hg18)	Primer Length
SYT9_rep01	chr11:722931+7230180	21,23
LAMA1_rep01	chr18:7107420+7107714	28,31
TLX3rep01	chr5:170668817+17066912	30,24
THEM4_rep01	chr1:150148496+15014884	28,25
WNT2_rep01	chr7:116750654+11675092	28,35
FOXE3_rep01	chr1:47654798+47655020	27,27
CYR61_rep01	chr1:85819176+85819330	28,28
KCNC3_rep01	chr19:55523578+55523941	32,24
THRB_rep01	chr3:24510979+24511199	34,31
ACTL6B_rep01	chr7:00091691+10009212	24,35
GALR1_rep01	chr18:73091207+73091513	23,23
EPB41L3_rep01	chr18:5533582+5533933	25,23
SH3GL3_rep01	chr15:81907023+81907340	27,24
EXT1_rep01	chr8:119193240+11919366	32,23
PLOD2_rep01	chr3:147361674+14736192	24,27
L3MBTL4_rep01	chr18:6404859+6405148	26,26
SLC18A3_rep01	chr10:50488165+50488452	30,27
CELSR1_rep01	chr22:45311123+45311476	25,25
RIMS4_rep01	chr20:42872176+42872407	23,25
PTF1A_rep01	chr10:23521471+23521781	24,20
5' Forward Tag	AGGAAGAGAG	
5' Reverse Tag	CAGTAATACGACTCACTATAGGGAGAAGGCT	

- 55 -

EXAMPLE 6***Sample justification***

[0114] DNA methylation was determined using SEQUENOM MassArray EpiTYPER of a select number of gene loci on identical bone marrow samples that were either flash frozen whole or subjected to smearing on a microscope stained with Giemsa and mounted in DPX for microscopy. There was good correlation between the methylation values of CpG units from both samples ($R^2=0.97$, Figure 1) suggesting that the archival process of making the smear slides does not affect DNA methylation. This allowed the performance of further analysis on samples derived from archived bone marrow aspirates.

EXAMPLE 7***Illumina Infinium analysis***

[0115] Thirteen TEL/AML positive cases were selected for genome-wide interrogation using Illumina Infinium HumanMethylation27 Bead Chip DNA arrays. DNA methylation of gene promoters of interest were identified by calculating the average beta-value between the two groups leukemia and remission samples and only retaining promoters that were at least 40% differential DNA methylation between the two groups. After data cleaning to applying this rule to on the data, 936 gene promoters were identified as differentially methylated between leukemia and controls (Table 8 lists genes ranked on average methylation difference). Unsupervised hierarchical clustering on these genes was able to separate the samples into two groups consisting of leukemic bone marrow and non-leukemic bone marrow by DNA methylation levels (Figure 2).

[0116] Along with leukemia samples, matched 28-day remission and post-treatment follow up samples were also analyzed on nine out of the 13 cases. The methylation profile of the leukemic sample was distinct from the remission and follow up samples obtained from these cases.

- 56 -

[0117] 74 genes showed on average a greater than 80% difference in DNA methylation between leukemic (methylated) and non-leukemic (unmethylated) samples while only 34 genes showed a greater than 40% difference in DNA methylation between leukemic (unmethylated) and non-leukemic (methylated) samples (Table 8).

5

TABLE 8
Differentially methylated genes in dataset

SYMBOL	AvgLeuk	AvgRemFol	DiffBeta
LAMA1	0.95	0.02	0.93
FOXE3	0.92	0.03	0.90
GALR1	0.93	0.04	0.89
CYR61	0.94	0.05	0.88
TLX3	0.94	0.06	0.88
CELSR1	0.94	0.07	0.88
FBXO39	0.94	0.06	0.87
RIMS4	0.91	0.04	0.87
SLC18A3	0.90	0.03	0.87
RBP1	0.89	0.02	0.87
NELL1	0.90	0.03	0.87
MYOD1	0.94	0.07	0.86
L3MBTL4	0.89	0.03	0.86
LGI2	0.88	0.02	0.86
ELOVL4	0.88	0.02	0.86
SH3GL2	0.90	0.04	0.86
SALL3	0.90	0.04	0.86
EPB41L3	0.88	0.03	0.85
NPTX2	0.90	0.04	0.85
GSH1	0.90	0.04	0.85
PPARG	0.92	0.06	0.85
MYO3A	0.89	0.03	0.85
SNAP91	0.89	0.04	0.85
TFAP2C	0.89	0.04	0.85
PKDREJ	0.95	0.10	0.85
SLC22A3	0.89	0.04	0.85
ACADL	0.87	0.02	0.85
UNQ9433	0.91	0.06	0.85
SFRP1	0.90	0.05	0.85
ARNT2	0.86	0.02	0.84
TRIM58	0.87	0.02	0.84
MGC52057	0.87	0.03	0.84
ADPN	0.89	0.05	0.84
SCRN1	0.91	0.07	0.84
PDE10A	0.89	0.05	0.84
PLOD2	0.87	0.03	0.84

- 57 -

SYMBOL	AvgLeuk	AvgRemFol	DiffBeta
KCNK5	0.88	0.04	0.84
COL4A3	0.88	0.04	0.84
FLJ90650	0.88	0.04	0.83
IRX2	0.88	0.05	0.83
NKX2-8	0.89	0.05	0.83
TP73	0.88	0.05	0.83
IRX4	0.88	0.05	0.83
NEF3	0.89	0.06	0.83
ADRB3	0.86	0.02	0.83
PAK7	0.86	0.03	0.83
KITLG	0.86	0.03	0.83
IGSF4	0.88	0.05	0.83
CXCL1	0.87	0.04	0.83
SLC6A2	0.86	0.03	0.83
ACTL6B	0.90	0.07	0.83
SLC16A12	0.86	0.03	0.83
FBXO39	0.89	0.07	0.83
BTG4	0.88	0.05	0.83
OGDHL	0.88	0.06	0.83
SOX9	0.91	0.09	0.82
SOX17	0.90	0.07	0.82
AMPH	0.87	0.04	0.82
XKR6	0.85	0.03	0.82
MSX2	0.87	0.05	0.82
CDCP1	0.85	0.03	0.82
ONECUT2	0.91	0.09	0.82
SOX21	0.83	0.01	0.82
MALL	0.85	0.04	0.81
FLJ21511	0.86	0.05	0.81
RHPN2	0.84	0.03	0.81
DMRT3	0.86	0.05	0.81
WWTR1	0.86	0.06	0.81
ZNF365	0.83	0.03	0.80
FLJ46831	0.83	0.03	0.80
THRB	0.86	0.05	0.80
SLCO2A1	0.84	0.03	0.80
ZNF454	0.87	0.07	0.80
FLJ46831	0.85	0.05	0.80
NINJ1	0.09	0.49	-0.40
CACNA1H	0.24	0.65	-0.41
SFTPB	0.18	0.59	-0.41
SORBS3	0.10	0.51	-0.41
SCT	0.41	0.83	-0.42
MPP7	0.21	0.63	-0.42
GNB3	0.21	0.64	-0.43
DNAI1	0.31	0.75	-0.44
CKM	0.26	0.69	-0.44
RAG1	0.20	0.64	-0.44

- 58 -

SYMBOL	AvgLeuk	AvgRemFol	DiffBeta
MMP11	0.21	0.65	-0.44
BZRAP1	0.14	0.59	-0.45
BLK	0.15	0.60	-0.45
LCK	0.16	0.61	-0.45
KAZALD1	0.23	0.68	-0.45
LRP8	0.19	0.64	-0.45
TBC1D14	0.15	0.61	-0.46
MAG	0.10	0.57	-0.47
MPST	0.07	0.55	-0.48
BGLAP	0.12	0.61	-0.49
PACAP	0.14	0.63	-0.49
PISD	0.37	0.86	-0.50
NPPA	0.24	0.75	-0.51
LY86	0.06	0.57	-0.51
PLD4	0.20	0.72	-0.52
EPOR	0.10	0.62	-0.52
KISS1	0.24	0.79	-0.55
	0.13	0.72	-0.58
KCNMB3	0.16	0.74	-0.59
FLJ41327	0.12	0.70	-0.59
TCF3	0.17	0.77	-0.59
SCARF1	0.19	0.79	-0.59
DDR1	0.19	0.80	-0.61
APCDD1	0.10	0.74	-0.63
CRYBB3	0.19	0.83	-0.64

[0118] Genes showing 60-60.99% as shown in Table 9.

TABLE 9

5

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta2	MethRatio
OTOP1	NM_177998.1	0.88	0.18	0.70	0.70	4.79
EDN3	NM_000114.2	0.74	0.05	0.70	0.70	16.32
SULT4A1	NM_014351.2	0.72	0.03	0.70	0.70	28.15
POU4F3	NM_002700.1	0.77	0.08	0.70	0.70	10.02
CART1	NM_006982.1	0.77	0.07	0.70	0.70	10.68
FOXQ1	NM_033260.2	0.74	0.05	0.70	0.70	16.00
TFCP2L1	NM_014553.1	0.75	0.06	0.70	0.70	13.61
CYP1B1	NM_000104.2	0.75	0.05	0.69	0.69	13.85
TBX4	NM_018488.2	0.76	0.07	0.69	0.69	11.45
TRPA1	NM_007332.1	0.76	0.07	0.69	0.69	11.33
GCM2	NM_004752.1	0.83	0.13	0.69	0.69	6.15
CDC14B	NM_003671.2	0.75	0.06	0.69	0.69	12.91
GULP1	NM_016315.2	0.73	0.04	0.69	0.69	19.23
CHAT	NM_020986.1	0.73	0.04	0.69	0.69	19.83
CDC42BPB	NM_006035.2	0.73	0.04	0.69	0.69	19.18
AP3B2	NM_004644.3	0.75	0.06	0.69	0.69	13.30
GSCL	NM_005315.1	0.73	0.04	0.69	0.69	18.45
ID4	NM_001546.2	0.74	0.05	0.69	0.69	14.31
CHODL	NM_024944.2	0.84	0.15	0.69	0.69	5.59
CYP24A1	NM_000782.3	0.83	0.14	0.69	0.69	5.99
KCNJ8	NM_004982.2	0.74	0.05	0.69	0.69	14.47
DCC	NT_010966.13	0.74	0.06	0.69	0.69	12.79
TMEFF2	NM_016192.2	0.73	0.04	0.68	0.68	17.32
HRH3	NM_007232.1	0.71	0.02	0.68	0.68	32.17
LHX5	NM_022363.2	0.73	0.04	0.68	0.68	16.56
SYN2	NM_133625.2	0.76	0.08	0.68	0.68	9.83
PDE10A	NM_006661.1	0.74	0.06	0.68	0.68	13.37
SYTL4	NM_080737.1	0.83	0.15	0.68	0.68	5.67
KCNB2	NM_004770.2	0.75	0.07	0.68	0.68	10.62
PPP1R3C	NM_005398.3	0.72	0.04	0.68	0.68	20.44
SLC5A8	NM_145913.2	0.77	0.09	0.68	0.68	8.42
BMPR1A	NM_004329.2	0.71	0.03	0.68	0.68	27.13
STEAP2	NM_152999.2	0.70	0.02	0.68	0.68	31.64
GALNT11	NM_022087.1	0.76	0.08	0.68	0.68	9.63
AGTR1	NM_000685.3	0.76	0.08	0.68	0.68	9.00
ABO	NM_020469.2	0.72	0.04	0.68	0.68	16.50
INA	NM_032727.2	0.71	0.03	0.68	0.68	23.17
TRPA1	NM_007332.1	0.72	0.04	0.68	0.68	16.99
NPR3	NM_000908.2	0.72	0.04	0.68	0.68	18.37
CPXM2	NM_198148.1	0.83	0.16	0.68	0.68	5.33

- 60 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBe ta2	MethR atlo
MGC16291	NM_032770.3	0.80	0.12	0.68	0.68	6.41
ADCY2	NM_020546.1	0.72	0.04	0.68	0.68	16.23
NAALAD2	NM_005467.2	0.74	0.07	0.68	0.68	11.13
POU4F2	NM_004575.1	0.79	0.12	0.68	0.68	6.81
PTGS2	NT_004487.18	0.73	0.06	0.67	0.67	13.22
PRLHR	NM_004248.1	0.80	0.13	0.67	0.67	6.26
PTF1A	NM_178161.1	0.80	0.13	0.67	0.67	6.29
GRB14	NM_004490.2	0.72	0.04	0.67	0.67	16.09
LRIG3	NM_153377.3	0.75	0.08	0.67	0.67	9.46
WNT2	NM_003391.1	0.72	0.04	0.67	0.67	16.81
IRX2	NM_033267.2	0.75	0.08	0.67	0.67	9.49
HTRA1	NM_002775.3	0.71	0.04	0.67	0.67	19.59
NR2F2	NM_021005.2	0.73	0.06	0.67	0.67	12.44
SEMA3C	NM_006379.2	0.74	0.07	0.67	0.67	11.24
NKX3-1	NM_006167.2	0.72	0.05	0.67	0.67	15.56
LOC63928	NM_022097.1	0.72	0.05	0.67	0.67	14.32
EVX1	NM_001989.2	0.73	0.06	0.67	0.67	11.84
MGC13040	NM_032930.1	0.73	0.06	0.67	0.67	12.12
EVX1	NM_001989.2	0.71	0.05	0.67	0.67	15.81
YAP1	NM_006106.2	0.74	0.08	0.67	0.67	9.49
PDZRN3	NM_015009.1	0.80	0.14	0.67	0.67	5.77
IHH	NM_002181.1	0.73	0.06	0.67	0.67	11.83
HOXC10	NM_017409.2	0.69	0.02	0.67	0.67	29.17
VMP	NM_080723.2	0.73	0.07	0.66	0.66	10.58
PCDH8	NM_032949.1	0.78	0.11	0.66	0.66	7.01
GDA	NM_004293.2	0.70	0.04	0.66	0.66	18.82
LOC401498	NM_212558.1	0.72	0.06	0.66	0.66	13.05
SH3GL2	NM_003026.1	0.79	0.12	0.66	0.66	6.43
EPHX2	NM_001979.4	0.76	0.10	0.66	0.66	7.53
CDH1	NM_004360.2	0.69	0.03	0.66	0.66	22.37
PCDHAC2	NM_031883.2	0.76	0.09	0.66	0.66	8.04
DCC	NM_005215.1	0.71	0.05	0.66	0.66	15.38
DMRT1	NM_021951.2	0.70	0.03	0.66	0.66	20.44
FLJ46380	NM_207396.1	0.74	0.08	0.66	0.66	9.33
COMP	NM_000095.2	0.73	0.07	0.66	0.66	10.32
NKX2-2	NM_002509.2	0.74	0.08	0.66	0.66	9.50
TTC22	NM_017904.1	0.70	0.04	0.66	0.66	18.30
HNT	NM_016522.2	0.70	0.04	0.66	0.66	16.21
SIX1	NM_005982.1	0.72	0.06	0.66	0.66	11.89
PRG-3	NM_017753.2	0.73	0.07	0.66	0.66	11.04
P2RY1	NM_002563.2	0.72	0.06	0.66	0.66	11.77
TMEM98	NM_015544.2	0.69	0.03	0.66	0.66	22.19
MYF5	NM_005593.1	0.82	0.16	0.66	0.66	5.11
FGF10	NM_004465.1	0.72	0.06	0.66	0.66	11.16
WDR17	NM_181265.2	0.71	0.05	0.66	0.66	13.35
CTNND2	NM_001332.2	0.70	0.04	0.66	0.66	17.94
EYA4	NT_025741.14	0.72	0.07	0.66	0.66	11.06

- 61 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBe ta2	MethR atio
ZNF300	NM_052860.1	0.77	0.12	0.65	0.65	6.50
GPR26	NM_153442.1	0.79	0.14	0.65	0.65	5.67
KCNB1	NM_004975.2	0.77	0.12	0.65	0.65	6.64
NPTX2	NM_002523.1	0.87	0.22	0.65	0.65	4.01
ATP6V0A1	NM_005177.2	0.68	0.02	0.65	0.65	27.87
CNTNAP3	NM_033655.2	0.73	0.08	0.65	0.65	9.30
SH3GL3	NM_003027.2	0.69	0.03	0.65	0.65	20.08
C10orf82	NM_144661.2	0.77	0.12	0.65	0.65	6.58
GLRB	NM_000824.2	0.68	0.03	0.65	0.65	25.96
GPR88	NM_022049.1	0.72	0.07	0.65	0.65	9.98
DKKL1	NM_014419.3	0.72	0.07	0.65	0.65	10.15
RASEF	NM_152573.2	0.69	0.03	0.65	0.65	20.66
SCGN	NM_006998.3	0.79	0.14	0.65	0.65	5.75
LPA	XM_926329.1	0.83	0.18	0.65	0.65	4.59
ARHGAP8	NM_181335.2	0.73	0.08	0.65	0.65	9.59
ADAMTS5	NM_007038.2	0.69	0.04	0.65	0.65	18.08
SNAP25	NM_003081.2	0.71	0.06	0.65	0.65	12.35
CNTNAP2	NM_014141.3	0.75	0.10	0.65	0.65	7.62
CIDEA	NM_001279.2	0.70	0.05	0.65	0.65	14.18
NEFH	NM_021076.2	0.76	0.11	0.65	0.65	7.17
SLC18A2	NM_003054.2	0.69	0.04	0.65	0.65	16.45
FAM80A	NM_173642.1	0.72	0.07	0.65	0.65	10.25
IGFBP3	NT_007819.16	0.68	0.03	0.65	0.65	20.17
FLJ21159	NM_024826.1	0.68	0.04	0.65	0.65	18.62
FZD9	NM_003508.2	0.88	0.23	0.65	0.65	3.78
GENX-3414	NM_003943.1	0.69	0.05	0.65	0.65	15.26
LRRTM1	NM_178839.3	0.71	0.06	0.65	0.65	11.86
C1orf188	NM_173795.2	0.76	0.11	0.65	0.65	6.73
NMBR	NM_002511.1	0.80	0.16	0.65	0.65	5.07
PITX3	NM_005029.3	0.68	0.04	0.64	0.64	16.91
PRMT8	NM_019854.3	0.67	0.03	0.64	0.64	23.40
ZFP41	NM_173832.3	0.73	0.09	0.64	0.64	8.53
FBXL2	NM_012157.2	0.68	0.04	0.64	0.64	18.89
PTPRO	NM_002848.2	0.67	0.02	0.64	0.64	27.60
NR2F2	NM_021005.2	0.69	0.05	0.64	0.64	13.84
FLJ45717	NM_207401.1	0.69	0.05	0.64	0.64	14.85
MTNR1B	NM_005959.3	0.89	0.25	0.64	0.64	3.61
ADRA2A	NM_000681.2	0.69	0.05	0.64	0.64	13.87
EYA2	NM_005244.3	0.71	0.07	0.64	0.64	10.55
SLC8A2	NM_015063.1	0.85	0.21	0.64	0.64	4.10
FLJ10916	NM_018271.2	0.70	0.06	0.64	0.64	11.34
EYA4	NM_004100.2	0.68	0.04	0.64	0.64	16.57
DES	NM_001927.3	0.68	0.04	0.64	0.64	15.27
WBSCR14	NM_032951.1	0.70	0.06	0.64	0.64	11.17
ACTC	NM_005159.3	0.72	0.08	0.64	0.64	9.05
CRABP1	NM_004378.1	0.68	0.04	0.64	0.64	15.68
EYA4	NM_004100.2	0.67	0.03	0.64	0.64	23.61

- 62 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBe ta2	MethR atio
SCGN	NM_006998.3	0.69	0.05	0.64	0.64	14.15
GPM6A	NM_201592.1	0.69	0.06	0.64	0.64	12.15
KCNK5	NM_003740.2	0.70	0.06	0.64	0.64	11.53
CTNNA2	NM_004389.2	0.70	0.06	0.64	0.64	11.25
ERRFI1	NM_018948.2	0.67	0.04	0.64	0.64	17.39
ADCY2	NM_020546.1	0.71	0.07	0.64	0.64	9.53
KIF6	NM_145027.3	0.67	0.03	0.64	0.64	21.70
DOCK5	NM_024940.4	0.68	0.05	0.63	0.63	14.74
SOX14	NM_004189.2	0.69	0.05	0.63	0.63	13.40
PRRX2	NM_016307.3	0.70	0.06	0.63	0.63	10.91
CLSTN2	NM_022131.1	0.68	0.05	0.63	0.63	14.20
CXCL5	NM_002994.3	0.72	0.09	0.63	0.63	8.02
PAWR	NM_002583.2	0.67	0.04	0.63	0.63	17.37
GALR1	NM_001480.2	0.73	0.10	0.63	0.63	7.31
LRIG3	NM_153377.3	0.66	0.03	0.63	0.63	21.90
DOK5	NM_018431.3	0.72	0.09	0.63	0.63	8.18
DKK3	NM_015881.5	0.68	0.05	0.63	0.63	14.14
HOXB4	NM_024015.3	0.66	0.03	0.63	0.63	24.69
CART	NM_004291.2	0.71	0.08	0.63	0.63	9.12
MARK2	NM_017490.1	0.79	0.16	0.63	0.63	4.85
LYPD5	NM_182573.1	0.72	0.10	0.63	0.63	7.57
DCC	NT_010966.13	0.73	0.10	0.63	0.63	7.32
TMEM67	NM_153704.3	0.68	0.05	0.63	0.63	13.15
FAM19A4	NM_182522.3	0.68	0.05	0.63	0.63	13.57
ACTN2	NM_001103.1	0.75	0.13	0.62	0.62	5.96
MSC	NM_005098.2	0.79	0.17	0.62	0.62	4.68
ADCY1	NM_021116.1	0.69	0.06	0.62	0.62	10.83
DCC	NT_010966.13	0.71	0.09	0.62	0.62	8.16
LANCL3	NM_198511.1	0.82	0.20	0.62	0.62	4.16
GPR103	NM_198179.1	0.71	0.09	0.62	0.62	8.01
RELN	NM_005045.2	0.69	0.07	0.62	0.62	10.45
SKIP	NM_030623.1	0.78	0.16	0.62	0.62	4.93
DRD2	NM_016574.2	0.74	0.12	0.62	0.62	6.24
GLRA1	NM_000171.1	0.75	0.13	0.62	0.62	5.81
FLJ30277	NM_153008.3	0.65	0.03	0.62	0.62	20.49
TCERG1L	NM_174937.1	0.72	0.10	0.62	0.62	7.22
ADAMTS1	NM_006988.3	0.66	0.04	0.62	0.62	17.17
UTF1	NM_003577.2	0.69	0.07	0.62	0.62	9.49
GATA6	NM_005257.3	0.71	0.09	0.62	0.62	7.93
LOC401498	NM_212558.1	0.68	0.07	0.62	0.62	9.99
HOXC13	NM_017410.2	0.67	0.05	0.62	0.62	12.84
MEGF10	NM_032446.1	0.73	0.12	0.62	0.62	6.19
GBX2	NM_001485.2	0.70	0.09	0.62	0.62	8.23
TRHDE	NM_013381.1	0.69	0.07	0.62	0.62	9.58
FOXG1B	NM_005249.3	0.68	0.07	0.61	0.61	9.84
SLC22A17	NM_020372.2	0.64	0.02	0.61	0.61	28.11
BARHL2	NM_020063.1	0.78	0.17	0.61	0.61	4.64

- 63 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta2	MethRatio
INADL	NM_005799.2	0.68	0.06	0.61	0.61	10.92
ACSL6	NM_015256.2	0.64	0.03	0.61	0.61	22.97
ZNF285	NM_152354.2	0.75	0.13	0.61	0.61	5.61
PDE1C	NM_005020.1	0.76	0.15	0.61	0.61	5.08
TSPAN2	NM_005725.3	0.67	0.06	0.61	0.61	11.88
MARVELD3	NM_052858.3	0.66	0.05	0.61	0.61	13.49
RAB32	NM_006834.2	0.66	0.05	0.61	0.61	13.87
G0S2	NM_015714.2	0.66	0.05	0.61	0.61	12.12
DKFZP566N034	NM_030923.2	0.64	0.03	0.61	0.61	21.97
MIPOL1	NM_138731.2	0.64	0.03	0.61	0.61	20.65
RIC3	NM_024557.2	0.65	0.04	0.61	0.61	15.90
HAND2	NM_021973.1	0.71	0.10	0.61	0.61	6.81
MITF	NM_000248.2	0.69	0.08	0.61	0.61	8.69
SLC6A3	NM_001044.2	0.72	0.12	0.61	0.61	6.27
PRG-3	NM_017753.2	0.65	0.04	0.61	0.61	14.69
TM6SF1	NM_023003.1	0.63	0.03	0.61	0.61	23.28
MINPP1	NM_004897.2	0.63	0.03	0.61	0.61	23.32
CCDC37	NM_182628.1	0.68	0.07	0.61	0.61	9.44
ALDH1L1	NM_012190.2	0.65	0.04	0.61	0.61	16.13
RP11-450P7.3	NM_153270.1	0.85	0.25	0.60	0.60	3.42
LOC348840	NM_182631.1	0.85	0.25	0.60	0.60	3.43
SIX6	NM_007374.1	0.65	0.05	0.60	0.60	12.76
RBP1	NM_002899.2	0.66	0.06	0.60	0.60	11.16
KIF6	NM_145027.3	0.65	0.05	0.60	0.60	14.35
MT1G	NM_005950.1	0.65	0.05	0.60	0.60	12.99
GPR39	NM_001508.1	0.65	0.04	0.60	0.60	14.44
GRM7	NM_181874.1	0.71	0.11	0.60	0.60	6.33
SMOC2	NM_022138.1	0.69	0.09	0.60	0.60	7.88
IRS4	NM_003604.1	0.85	0.25	0.60	0.60	3.45
BNC1	NT_077661.2	0.72	0.12	0.60	0.60	5.98
UNQ9433	NM_207413.1	0.68	0.08	0.60	0.60	8.96
SCARF1	NM_145349.1	0.19	0.79	0.60	-0.60	0.24
APCDD1	NM_153000.3	0.10	0.71	0.60	-0.60	0.14
KCNMB3	NM_014407.3	0.16	0.76	0.61	-0.61	0.20
CRYBB3	NM_004076.3	0.19	0.82	0.62	-0.62	0.24

[0119] Genes showing 70-70.99% as shown in Table 10.

TABLE 10

5

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta2	MethRatio
UTF1	NM_003577.2	0.82	0.02	0.80	0.80	35.02
PNMA2	NM_007257.4	0.89	0.09	0.80	0.80	9.42
BNC1	NT_077661.2	0.85	0.05	0.80	0.80	15.67
SIX1	NM_005982.1	0.85	0.05	0.80	0.80	17.78
TDRD5	NM_173533.2	0.84	0.04	0.80	0.80	18.92
GRM6	NM_000843.2	0.86	0.06	0.80	0.80	14.04
CAMK2B	NM_001220.3	0.82	0.03	0.80	0.80	30.92
SOX1	NM_005986.2	0.89	0.09	0.80	0.80	9.82
SLC16A12	NM_213606.1	0.83	0.03	0.79	0.79	25.93
C1orf76	NM_173509.2	0.82	0.02	0.79	0.79	32.80
EFCAB1	NM_024593.2	0.85	0.05	0.79	0.79	16.08
TJP1	NM_003257.2	0.84	0.05	0.79	0.79	18.50
HCN4	NM_005477.1	0.83	0.03	0.79	0.79	23.65
SLC6A2	NM_001043.2	0.84	0.05	0.79	0.79	16.78
SLC34A2	NM_006424.1	0.82	0.03	0.79	0.79	27.02
CYYR1	NM_052954.2	0.83	0.04	0.79	0.79	21.42
THRB	NM_000461.2	0.83	0.04	0.79	0.79	19.49
PHF21B	NM_138415.1	0.83	0.05	0.79	0.79	17.65
MYOD1	NM_002478.3	0.86	0.07	0.79	0.79	11.62
PRSS12	NM_003619.2	0.87	0.09	0.78	0.78	9.98
RBP4	NM_006744.3	0.82	0.03	0.78	0.78	25.70
NKX6-2	NM_177400.1	0.83	0.04	0.78	0.78	19.37
FBN1	NM_000138.2	0.81	0.03	0.78	0.78	26.91
LPL	NM_000237.1	0.84	0.06	0.78	0.78	13.43
SFRP1	NM_003012.3	0.82	0.04	0.78	0.78	19.76
SLC1A1	NM_004170.4	0.80	0.03	0.78	0.78	31.82
ADRA1D	NM_000678.2	0.82	0.04	0.78	0.78	20.68
SNCAIP	NM_005460.2	0.81	0.04	0.78	0.78	21.25
VAX1	NM_199131.1	0.84	0.06	0.78	0.78	13.12
T	NM_003181.2	0.84	0.06	0.77	0.77	13.36
RHCG	NM_016321.1	0.81	0.04	0.77	0.77	22.53
DHCR24	NM_014762.2	0.85	0.07	0.77	0.77	11.35
ID4	NM_001546.2	0.84	0.06	0.77	0.77	13.38
GPR6	NM_005284.2	0.82	0.04	0.77	0.77	18.89
RAB32	NT_025741.14	0.85	0.08	0.77	0.77	11.08
KLK10	NT_011109.15	0.80	0.03	0.77	0.77	25.92
CPLX2	NM_001008220.1	0.84	0.07	0.77	0.77	11.31
HS3ST2	NM_006043.1	0.86	0.09	0.77	0.77	9.73
ADRA2C	NM_000683.3	0.87	0.10	0.77	0.77	8.74
GPR78	NM_080819.2	0.82	0.05	0.77	0.77	17.05

- 65 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta2	MethRatio
NELL1	NM_006157.2	0.80	0.03	0.77	0.77	27.03
FLJ21511	NM_025087.1	0.90	0.13	0.77	0.77	6.90
HAND1	NM_004821.1	0.80	0.03	0.77	0.77	23.91
ANKRD38	NM_181712.2	0.80	0.03	0.76	0.76	24.74
DOCK5	NM_024940.4	0.80	0.04	0.76	0.76	20.92
RASGRF2	NM_006909.1	0.79	0.03	0.76	0.76	27.04
GALR1	NM_001480.2	0.82	0.06	0.76	0.76	14.35
PARVA	NM_018222.2	0.80	0.04	0.76	0.76	19.88
GUCY1A2	NM_000855.1	0.80	0.04	0.76	0.76	18.58
MTNR1A	NM_005958.3	0.83	0.07	0.76	0.76	12.18
PLAC2	NM_153375.1	0.83	0.07	0.76	0.76	11.89
SPSB4	NM_080862.1	0.80	0.04	0.76	0.76	19.05
WNT7A	NM_004625.3	0.81	0.05	0.76	0.76	17.02
MOSC1	NM_022746.2	0.79	0.03	0.76	0.76	23.60
PTHR2	NM_005048.2	0.81	0.05	0.76	0.76	16.64
L3MBTL4	NM_173464.1	0.79	0.03	0.76	0.76	24.86
BRUNOL6	NM_052840.3	0.78	0.02	0.76	0.76	32.00
FA2H	NM_024306.2	0.78	0.02	0.76	0.76	38.14
AFAR3	NM_201252.1	0.87	0.12	0.76	0.76	7.47
KCNC1	NM_004976.2	0.92	0.17	0.76	0.76	5.47
TLX3	NM_021025.2	0.82	0.06	0.75	0.75	12.68
FAM19A4	NM_182522.3	0.82	0.06	0.75	0.75	12.96
ARL4	NM_005738.2	0.82	0.07	0.75	0.75	12.08
CNTNAP2	NM_014141.3	0.81	0.06	0.75	0.75	14.20
GRM7	NM_181874.1	0.81	0.06	0.75	0.75	13.65
GPR126	NM_198569.1	0.84	0.08	0.75	0.75	9.94
VGLL2	NM_153453.1	0.82	0.07	0.75	0.75	11.97
HBQ1	NM_005331.3	0.81	0.06	0.75	0.75	14.13
NEUROG1	NT_034772.5	0.80	0.05	0.75	0.75	16.69
EDIL3	NM_005711.3	0.80	0.05	0.75	0.75	15.85
KRT18	NM_000224.2	0.78	0.03	0.75	0.75	29.36
GDNF	NM_000514.2	0.82	0.07	0.75	0.75	11.83
GREM1	NM_013372.5	0.80	0.05	0.75	0.75	15.81
RAB32	NM_006834.2	0.80	0.05	0.75	0.75	15.40
PDE4C	NM_000923.1	0.86	0.11	0.75	0.75	7.56
SLC24A3	NM_020689.3	0.81	0.06	0.75	0.75	13.77
MOSC2	NM_017898.3	0.79	0.04	0.75	0.75	18.69
SST	NM_001048.3	0.83	0.08	0.75	0.75	10.16
PDZRN3	NM_015009.1	0.81	0.07	0.75	0.75	12.32
KCNG1	NM_002237.2	0.81	0.06	0.74	0.74	12.87
CDH22	NM_021248.1	0.85	0.11	0.74	0.74	7.78
G0S2	NM_015714.2	0.79	0.04	0.74	0.74	18.12
SLC27A6	NM_014031.3	0.80	0.06	0.74	0.74	13.55
TRPC1	NM_003304.3	0.82	0.07	0.74	0.74	11.18
NPY5R	NM_006174.2	0.86	0.11	0.74	0.74	7.55
SEMA3C	NM_006379.2	0.81	0.06	0.74	0.74	12.69
THEM4	NM_053055.3	0.78	0.04	0.74	0.74	20.93

- 66 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta2	MethRatio
CD164L2	NM_207397.1	0.82	0.08	0.74	0.74	10.04
KAL1	NM_000216.1	0.78	0.04	0.74	0.74	19.11
ADAM12	NM_003474.3	0.81	0.07	0.74	0.74	12.21
MSC	NM_005098.2	0.84	0.11	0.74	0.74	7.98
CYP24A1	NM_000782.3	0.83	0.09	0.74	0.74	8.78
EIF5A2	NM_020390.5	0.78	0.04	0.74	0.74	17.81
GALR1	NM_001480.2	0.80	0.06	0.74	0.74	12.49
SLC5A7	NM_021815.2	0.83	0.09	0.74	0.74	8.93
ADAMTS17	NM_139057.1	0.78	0.04	0.74	0.74	19.90
PKP2	NM_004572.2	0.78	0.04	0.74	0.74	19.30
ARNT2	NM_014862.3	0.79	0.06	0.74	0.74	14.27
CTTNBP2	NM_033427.1	0.79	0.06	0.73	0.73	13.73
KCNN2	NM_021614.2	0.80	0.07	0.73	0.73	11.67
SLIT3	NM_003062.1	0.78	0.05	0.73	0.73	16.81
FLJ42486	NM_207379.1	0.80	0.07	0.73	0.73	11.48
PRP2	NM_173490.4	0.77	0.04	0.73	0.73	19.02
FLJ14503	NM_152780.2	0.84	0.11	0.73	0.73	7.62
TRAM1L1	NM_152402.1	0.88	0.15	0.73	0.73	5.93
SLC18A3	NM_003055.1	0.82	0.08	0.73	0.73	9.66
MLF1	NM_022443.2	0.76	0.03	0.73	0.73	25.30
NEFH	NM_021076.2	0.88	0.15	0.73	0.73	5.77
DSCAML1	NM_020693.2	0.80	0.07	0.73	0.73	11.24
ASNS	NM_001673.2	0.81	0.08	0.73	0.73	10.40
EIF5A2	NM_020390.5	0.80	0.07	0.73	0.73	11.32
RAB32	NM_006834.2	0.79	0.06	0.73	0.73	13.94
SRD5A2	NM_000348.2	0.80	0.07	0.73	0.73	11.00
OVOL2	NM_021220.2	0.84	0.11	0.73	0.73	7.62
CRYBA2	NM_057093.1	0.75	0.02	0.73	0.73	33.33
KIAA1944	NM_133448.1	0.79	0.06	0.73	0.73	12.74
ATP8A2	NM_016529.3	0.82	0.09	0.73	0.73	8.81
EDIL3	NM_005711.3	0.77	0.04	0.73	0.73	18.72
KCNN2	NM_021614.2	0.79	0.06	0.73	0.73	13.47
RGS7	NM_002924.2	0.79	0.06	0.73	0.73	13.08
DMN	NM_145728.1	0.80	0.07	0.73	0.73	10.99
COL14A1	NM_021110.1	0.75	0.03	0.72	0.72	27.04
HLF	NM_002126.3	0.78	0.06	0.72	0.72	13.47
GALR1	NT_025004.13	0.85	0.12	0.72	0.72	6.98
NAV2	NM_182964.4	0.78	0.06	0.72	0.72	13.99
LARP6	NM_197958.1	0.77	0.05	0.72	0.72	15.82
ESX1	NM_153448.2	0.88	0.16	0.72	0.72	5.63
KIF17	NM_020816.1	0.74	0.02	0.72	0.72	40.60
GATA4	NT_077531.3	0.82	0.10	0.72	0.72	8.04
CBLN4	NM_080617.4	0.79	0.07	0.72	0.72	11.68
NMNAT3	NM_178177.2	0.78	0.06	0.72	0.72	13.65
FNDC4	NM_022823.1	0.76	0.04	0.72	0.72	18.83
FOXF2	NM_001452.1	0.74	0.03	0.71	0.71	26.90
RASGRF1	NM_002891.3	0.80	0.08	0.71	0.71	9.71

- 67 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta2	MethRatio
PGM5	NM_021965.3	0.87	0.15	0.71	0.71	5.68
EPHA10	NM_173641.1	0.77	0.06	0.71	0.71	12.37
EPHA10	NM_173641.1	0.76	0.05	0.71	0.71	16.45
PRRX1	NM_022716.2	0.77	0.06	0.71	0.71	13.17
HOXD11	NM_021192.2	0.77	0.06	0.71	0.71	13.14
MT1A	NM_005946.2	0.78	0.07	0.71	0.71	10.95
SNX7	NM_015976.2	0.78	0.07	0.71	0.71	11.13
PRICKLE1	NM_153026.1	0.74	0.03	0.71	0.71	28.51
EFHA2	NM_181723.1	0.75	0.04	0.71	0.71	20.58
ME1	NM_002395.2	0.75	0.04	0.71	0.71	17.07
GRIA2	NM_000826.2	0.80	0.09	0.71	0.71	9.05
GPC6	NM_005708.2	0.81	0.10	0.71	0.71	7.89
DMRT2	NM_006557.3	0.76	0.05	0.71	0.71	14.90
PLA2G7	NM_005084.2	0.74	0.03	0.71	0.71	22.16
CTSL2	NM_001333.2	0.74	0.03	0.71	0.71	22.32
SSTR4	NM_001052.1	0.81	0.10	0.71	0.71	7.78
BMP7	NM_001719.1	0.77	0.06	0.70	0.70	12.05
FOXA2	NM_021784.3	0.77	0.07	0.70	0.70	11.68
SLC12A5	NM_020708.3	0.85	0.15	0.70	0.70	5.74
TSCOT	NM_033051.2	0.77	0.07	0.70	0.70	11.74
BNC1	NT_077661.2	0.76	0.06	0.70	0.70	13.18
CACNA1A	NM_023035.1	0.78	0.07	0.70	0.70	10.37
DLX5	NT_007933.14	0.79	0.09	0.70	0.70	8.53
ADAMTS14	NM_080722.2	0.73	0.03	0.70	0.70	21.30
PENK	NM_006211.2	0.77	0.07	0.70	0.70	11.00

[0120] Genes showing greater than 80% as shown in Table 11. Differentially methylated genes in dataset are shown in Table 12.

- 68 -

TABLE 11

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta 2	MethR atio
LAMA1	NM_005559.2	0.95	0.02	0.93	0.93	38.69
FOX E3	NM_012186.1	0.92	0.03	0.90	0.90	35.76
GALR1	NM_001480.2	0.93	0.04	0.89	0.89	22.02
CELSR1	NM_014246.1	0.94	0.06	0.88	0.88	15.10
CYR61	NM_001554.3	0.94	0.06	0.88	0.88	16.56
FBXO39	NM_153230.1	0.94	0.06	0.88	0.88	15.30
TLX3	NM_021025.2	0.94	0.06	0.88	0.88	15.04
MYOD1	NM_002478.3	0.94	0.07	0.87	0.87	13.50
RBP1	NM_002899.2	0.89	0.02	0.87	0.87	36.35
SLC18A3	NM_003055.1	0.90	0.03	0.87	0.87	30.68
RIMS4	NM_182970.2	0.91	0.04	0.87	0.87	21.77
NELL1	NM_006157.2	0.90	0.03	0.87	0.87	26.31
NPTX2	NM_002523.1	0.90	0.03	0.86	0.86	27.36
LGI2	NM_018176.2	0.88	0.02	0.86	0.86	41.97
ELOVL4	NM_022726.2	0.88	0.02	0.86	0.86	45.54
SALL3	NM_171999.1	0.90	0.04	0.86	0.86	22.10
SH3GL2	NM_003026.1	0.90	0.04	0.86	0.86	21.42
L3MBTL4	NM_173464.1	0.89	0.04	0.86	0.86	24.28
EPB41L3	NM_012307.2	0.88	0.03	0.85	0.85	32.86
PKDREJ	NM_006071.1	0.95	0.09	0.85	0.85	10.17
GSH1	NM_145657.1	0.90	0.04	0.85	0.85	19.99
TFAP2C	NM_003222.3	0.89	0.04	0.85	0.85	23.58
MYO3A	NM_017433.3	0.89	0.04	0.85	0.85	24.77
SNAP91	NM_014841.1	0.89	0.04	0.85	0.85	23.53
PPARG	NM_138712.2	0.92	0.07	0.85	0.85	13.54
SLC22A3	NM_021977.2	0.89	0.04	0.85	0.85	23.62
SFRP1	NM_003012.3	0.90	0.05	0.85	0.85	18.20
ACADL	NM_001608.2	0.87	0.02	0.85	0.85	38.21
TRIM58	NM_015431.2	0.87	0.02	0.84	0.84	39.63
ARNT2	NM_014862.3	0.86	0.02	0.84	0.84	42.85
UNQ9433	NM_207413.1	0.91	0.07	0.84	0.84	13.56
ADPN	NM_025225.2	0.89	0.05	0.84	0.84	18.66
MGC52057	NM_194317.2	0.87	0.02	0.84	0.84	34.87
PDE10A	NM_006661.1	0.89	0.05	0.84	0.84	18.65
SCRN1	NM_014766.2	0.91	0.07	0.84	0.84	13.22
KCNK5	NM_003740.2	0.88	0.03	0.84	0.84	25.02
PLOD2	NM_182943.2	0.87	0.03	0.84	0.84	25.56
FLJ90650	NM_173800.3	0.88	0.04	0.84	0.84	22.47
TP73	NM_005427.1	0.88	0.04	0.84	0.84	20.99
NKX2-8	NM_014360.2	0.89	0.05	0.84	0.84	17.79
COL4A3	NM_000091.2	0.88	0.04	0.84	0.84	21.69
CXCL1	NM_001511.1	0.87	0.03	0.83	0.83	28.02
IRX2	NM_033267.2	0.88	0.05	0.83	0.83	18.27
SOX9	NM_000346.2	0.91	0.08	0.83	0.83	11.92

- 69 -

SYMBOL	ACCESSION	AvgLeuk	AvgRem Fol	DiffBeta	DiffBeta 2	MethR atio
SLC6A2	NM_001043.2	0.86	0.03	0.83	0.83	27.60
ADRB3	NM_000025.1	0.86	0.02	0.83	0.83	37.45
PAK7	NM_177990.1	0.86	0.03	0.83	0.83	26.83
FBXO39	NM_153230.1	0.89	0.06	0.83	0.83	14.44
IRX4	NM_016358.1	0.88	0.05	0.83	0.83	17.76
IGSF4	NM_014333.2	0.88	0.05	0.83	0.83	18.05
ACTL6B	NM_016188.3	0.90	0.07	0.83	0.83	12.38
NEF3	NM_005382.1	0.89	0.06	0.83	0.83	15.34
KITLG	NM_000899.3	0.86	0.04	0.83	0.83	24.55
SOX17	NM_022454.2	0.90	0.07	0.83	0.83	12.92
SLC16A12	NM_213606.1	0.86	0.03	0.83	0.83	24.69
BTG4	NM_017589.2	0.88	0.05	0.82	0.82	17.23
MSX2	NM_002449.3	0.87	0.04	0.82	0.82	20.35
OGDHL	NM_018245.1	0.88	0.06	0.82	0.82	14.85
XKR6	NM_173683.2	0.85	0.03	0.82	0.82	27.53
AMPH	NM_001635.2	0.87	0.05	0.82	0.82	19.10
ONECUT2	NM_004852.1	0.91	0.09	0.82	0.82	10.31
CDCP1	NM_178181.1	0.85	0.03	0.82	0.82	27.97
MALL	NM_005434.3	0.85	0.04	0.82	0.82	23.42
SOX21	NM_007084.2	0.83	0.01	0.82	0.82	58.10
FLJ21511	NM_025087.1	0.86	0.05	0.81	0.81	17.09
RHPN2	NM_033103.3	0.84	0.03	0.81	0.81	26.41
FLJ46831	NM_207426.1	0.85	0.04	0.81	0.81	19.47
ZNF454	NM_182594.1	0.87	0.06	0.81	0.81	13.54
DMRT3	NM_021240.2	0.86	0.05	0.81	0.81	16.86
WWTR1	NM_015472.3	0.86	0.06	0.81	0.81	15.51
ZNF365	NM_199451.1	0.83	0.03	0.81	0.81	30.39
FLJ46831	NM_207426.1	0.83	0.03	0.80	0.80	28.23
THRB	NM_000461.2	0.86	0.06	0.80	0.80	15.24
SLCO2A1	NM_005630.1	0.84	0.04	0.80	0.80	23.10
MGC52057	NM_194317.2	0.85	0.05	0.80	0.80	18.43

TABLE 12

Example of Differentially methylated genes in dataset

SYMBOL	AvgLeuk	AvgRemFol	DiffBeta
LAMA1	0.95	0.02	0.93
FOXE3	0.92	0.03	0.90

- 5 [0121] Tables 13a through 13e show genetic loci with greater than 90% differential methylation between leukemia and controls. Table 13a lists the top 100 genes identified using the centroid model (PMID:8917796) to separate leukemia bone marrow from normal bone marrow, of which any combination of 16 genes can determine this 100% of the time; Table 13b lists the top 100 genes identified using the Linear Models for Microarray
- 10 Analysis (LIMMA) [De Hertogh *et al.*, *BMC Bioinformatics* 11:17, 2020 (PMID 20064233)] model to separate leukemia bone marrow from normal bone marrow, of which any combination of 16 genes can determine this 100% of the time; Table 13c lists the top 100 genes identified using the Support Vector Machine (SVM) [Varewyck and Martens, *IEEE Trans. Syusk. Man. Cybern B Cybern* 41(2):330-340, 2011 (PMID:20699214)]
- 15 model to separate leukemia bone marrow from normal bone marrow, of which any combination of 16 genes can determine this 100% of the time; Table 13d lists the top 100 genes identified using SVM to separate leukemia cases that will relapse (poort outcome) to those with a good outcome, of which any combination of four genes can determine this 90% of the time; and Table 13e lists the top 100 genes identified using Centroid to separate
- 20 leukemia cases that will relapse (poor outcome) to those with a good outcome, of which any combination of four genes can determine this 90% of the time.

- 71 -

TABLE 13a
Centroid Top 100

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	CFS Centroid	CFSCentroid AVG RNK
cg16196812	GeneID:2834	GR3; GPR10; PrRPR; MGC126539; MGC126541;	NM_004248.1	PRLHR	1	19.5
cg05020203	GeneID:401045		NM_207485.1	FLJ41327	2	20.43
cg07271264	GeneID:4654	PUM; MYF3; MYOD;	NM_002478.3	MYOD1	3	21.93
cg24133115	GeneID:10846	HSPDE10A;	NM_006661.1	PDE10A	4	20.43
cg13481359	GeneID:1961	NGFIC; NGFI-C; PAT133;	NM_001965.1	EGR4	5	29.23
cg17389519	GeneID:256297	PTF1-p48;	NM_178161.1	PTF1A	6	32.63
cg04534765	GeneID:2587	GALNR; GALNR1;	NM_001480.2	GALR1	7	33.8
cg17977409	GeneID:6456	CNSA2; SH3P4; EEN-B1; SH3D2A; FLJ20276;	NM_003026.1	SH3GL2	8	32.93
cg24101578	GeneID:64405	C20orf25; MGC39564; dJ998H6.1;	NM_021248.1	CDH22	9	40.63
cg20723355	GeneID:162517	Fbx39; MGC35179;	NM_153230.1	FBXO39	10	40.3
cg25484904	GeneID:80157		NM_025087.1	FLJ21511	11	41.87
cg08918749	GeneID:4023	LIPD;	NM_000237.1	LPL	12	44.93
cg11389172	GeneID:6572	VACHT;	NM_003055.1	SLC18A3	13	41.93
cg17688525	GeneID:91133	HsT1031;	NM_173464.1	L3MBTL4	14	40.47
cg07846220	GeneID:284217	LAMA;	NM_005559.2	LAMA1	15	44.83
cg23037403	GeneID:285676	FLJ37444;	NM_182594.1	ZNF454	16	49.57
cg02919422	GeneID:64321	FLJ22252;	NM_022454.2	SOX17	17	58.13
cg11377136	GeneID:10343		NM_006071.1	PKDREJ	18	60.63
cg06675478	GeneID:6656		NM_005986.2	SOX1	19	58.07
cg10639440	GeneID:9805	SES1; KIAA0193;	NM_014766.2	SCRN1	20	53.37
cg02613386	GeneID:162517	Fbx39; MGC35179;	NM_153230.1	FBXO39	21	61.63
cg18815943	GeneID:2301	ASMD; FKHL12; FREAC8;	NM_012186.1	FOXE3	22	63.73
cg12497564	GeneID:5947	CRBP; RBPC; CRBP1; CRABP-I;	NM_002899.2	RBP1	23	61.1
cg06391468	GeneID:6662	CMD1; SRA1; CMPD1;	NM_000346.2	SOX9	24	67.4
cg06268694	GeneID:9620	ME2; FMI2; CDHF9; HFMI2; DKFZp434P0729;	NM_014246.1	CELSR1	25	75.73
cg10193817	GeneID:23705	BL2; ST17; NECL2; TSLC1; IGSF4A; SYNCAM; sTSLC-1; syncAM1; DKFZp686F1789;	NM_014333.2	IGSF4	26	67.63
cg27603796	GeneID:83992	Orf4; C7orf8; CORTBP2; KIAA1758; MGC104579;	NM_033427.1	CTTNBP2	27	69.77

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	CFS Centroid	CFS Centroid AVG RNK
cg17162024	GeneID:389658		NM_207413.1	UNQ9433	28	72.47
cg22418909	GeneID:6422	FRP; FRP1; FrzA; FRP-1; SARP2;	NM_003012.3	SFRP1	29	76.07
cg19064258	GeneID:9956	3OST2; 3OST2;	NM_006043.1	HS3ST2	30	68.17
cg21688264	GeneID:9892	AP180; KIAA0656;	NM_014841.1	SNAP91	31	71.1
cg04632671	GeneID:5468	NR1C3; PPARG1; PPARG2; HUMPPARG;	NM_138712.2	PPARG	32	74.37
cg25806808	GeneID:2919	GRO1; GROa; MGSA; NAP-3; SCYB1; MGSA-a; MGSA alpha;	NM_001511.1	CXCL1	33	74.7
cg14795968	GeneID:33	LCAD; ACAD4;	NM_001608.2	ACADL	34	84.17
cg25720804	GeneID:30012	RNX; HOX11L2; MGC29804;	NM_021025.2	TLX3	35	82
cg10019507	GeneID:6754		NM_001052.1	SSTR4	36	100.83
cg12420104	GeneID:58524	DMRTA3;	NM_021240.2	DMRT3	37	86.33
cg12799895	GeneID:4885	NP2; NARP; NP-II;	NM_002523.1	NPTX2	38	96.1
cg27096144	GeneID:4488	FPP; MSH; PFM; CRS2; HOX8; PFM1;	NM_002449.3	MSX2	39	95.93
cg07625840	GeneID:8492	MRT1; BSSP3; BSSP-3; MGC12722; MOTOPSIN;	NM_003619.2	PRSS12	40	79.43
cg07237939	GeneID:6581	EMT; EMTH; OCT3;	NM_021977.2	SLC22A3	41	91.73
cg26609631	GeneID:219409	Gsh-1;	NM_145657.1	GSH1	42	83.97
cg17371081	GeneID:4745	NRP1; IDH3GL;	NM_006157.2	NELL1	43	86.63
cg25044651	GeneID:206338	MGC125378; MGC125379;	NM_173800.3	FLJ90650	44	92.3
cg04453065	GeneID:3491	CCN1; GIG1; IGFBP10;	NM_001554.3	CYR61	45	94.23
cg02604290	GeneID:1591	CP24; CYP24; MGC126273; MGC126274; P450-CC24;	NM_000782.3	CYP24A1	46	90.8
cg17685628	GeneID:26257	NKX2H; NKX2.8; Nkx2-9;	NM_014360.2	NKX2-8	47	84.43
cg02154186	GeneID:10687	MA2; MM2; RGAG2; KIAA0883;	NM_007257.4	PNMA2	48	99.9
cg12839593	GeneID:6495		NM_005982.1	SIX1	49	93.03
cg02250594	GeneID:9480	OC2; OC-2; MGC120377; MGC120378;	NM_004852.1	ONECUT2	50	99.77
cg03160135	GeneID:245806	VGL2; VITO1;	NM_153453.1	VGLL2	51	97.87
cg23771603	GeneID:53904	DFNB30;	NM_017433.3	MYO3A	52	99.57
cg10063179	GeneID:4543	MEL-1A-R;	NM_005958.3	MTNR1A	53	108
cg13359415	GeneID:55203	LGIL2; FLJ10675; KIAA1916; MGC126808; MGC126810;	NM_018176.2	LGI2	54	103.8
cg25313204	GeneID:6581	EMT; EMTH; OCT3;	NM_021977.2	SLC22A3	55	88.43

- 73 -

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	CFS Centroid	CFS Centroid AVG RNK
cg13297865	GeneID:6785	ADMD; STGD2; STGD3;	NM_022726.2	ELOVL4	56	106.47
cg15191648	GeneID:27164		NM_171999.1	SALL3	57	118.17
cg23290344	GeneID:4741	NFM; NEFM; NF-M;	NM_005382.1	NEF3	58	99.2
cg04005707	GeneID:80157		NM_025087.1	FLJ21511	59	104.93
cg02141570	GeneID:348840		NM_182631.1	LOC348840	60	139.67
cg15842276	GeneID:4544	MEL-1B-R;	NM_005959.3	MTNR1B	61	114.97
cg19721889	GeneID:9421	Hxt; eHand; Thing1;	NM_004821.1	HAND1	62	128.83
cg15839448	GeneID:6422	FRP; FRP1; FrzA; FRP-1; SARP2;	NM_003012.3	SFRP1	63	118.37
cg18191162	GeneID:221458	C6orf102; MGC33317; dJ137F1.4; dJ188D3.1; dJ1043E3.1;	NM_145027.3	KIF6	64	99.27
cg25250358	GeneID:5352	LH2; TLH;	NM_182943.2	PLOD2	65	107.97
cg18555440	GeneID:4654	PUM; MYF3; MYOD;	NM_002478.3	MYOD1	66	122.53
cg19332710	GeneID:140730	RIM4; C20orf190; dJ781B1.3;	NM_182970.2	RIMS4	67	114.83
cg23211240	GeneID:54766	PC3B; MGC33003;	NM_017589.2	BTG4	68	126.83
cg19885761	GeneID:10814	CPX2; 921-L; CPX-2;	NM_001008220.1	CPLX2	69	117.53
cg03693099	GeneID:1056	BAL; FAP; BSDL; BSSL; CELL; FAPP; LIPA; CEase;	NM_001807.2	CEL	70	126.47
cg11530960	GeneID:10655		NM_006557.3	DMRT2	71	149.73
cg10862587	GeneID:79411	MGC10771;	NM_024506.3	GLB1L	72	130.83
cg13929328	GeneID:399823	FOXI2;	NM_207426.1	FLJ46831	73	119.87
cg17252960	GeneID:3400		NM_001546.2	ID4	74	123.03
cg27409364	GeneID:3746	KV4; NGK2; KV3.1; MGC129855;	NM_004976.2	KCNC1	75	153
cg22123464	GeneID:6543	NCX2;	NM_015063.1	SLC8A2	76	130.23
cg14785449	GeneID:440	TS11;	NM_001673.2	ASNS	77	125.6
cg03963198	GeneID:50805		NM_016358.1	IRX4	78	138.8
cg14882700	GeneID:133060		NM_177998.1	OTOP1	79	129.77
cg07028533	GeneID:26047	NRXN4; CASPR2; DKFZp781D1846;	NM_014141.3	CNTNAP2	80	129.8
cg12388309	GeneID:57144	PAK5; KIAA1264; MGC26232;	NM_177990.1	PAK7	81	137.5
cg10009830	GeneID:256714	MGC104944;	NM_152780.2	FLJ14503	82	149
cg00546491	GeneID:5137	Hcam3;	NM_005020.1	PDE1C	83	125.8
cg02536286	GeneID:7022	ERF1; TFAP2G; hAP-2g; AP2-GAMMA;	NM_003222.3	TFAP2C	84	138.7
cg14889768	GeneID:10981		NM_006834.2	RAB32	85	147.7
cg22752533	GeneID:57468	KCC2; KIAA1176;	NM_020708.3	SLC12A5	86	164.33
cg23759710	GeneID:165140	GPCR; GPR170; TG1019;	NM_148962.3	OXER1	87	142.4

- 74 -

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	CFS Centroid	CFSCentroid AVG RNK
cg13059782	GeneID:79937	CASPR3; CNTNAP3A; RP11-290L7.1; RP11-138L21.1;	NM_033655.2	CNTNAP3	88	147.93
cg15433631	GeneID:153572		NM_033267.2	IRX2	89	146.13
cg02994956	GeneID:4744	NFH;	NM_021076.2	NEFH	90	156.13
cg09261015	GeneID:80712	ESX1L; ESXR1;	NM_153448.2	ESX1	91	156.33
cg11376198	GeneID:246181	AKR7A4;	NM_201252.1	AFAR3	92	165.37
cg24130010	GeneID:140578	MT75; PRED12; C21orf68; FLJ12627;	NM_024944.2	CHODL	93	134
cg13334990	GeneID:2026	NSE;	NM_001975.2	ENO2	94	141.17
cg13899108	GeneID:5143	DPDE1; MGC126222; PDE4C-791;	NM_000923.1	PDE4C	95	133.77
cg00662556	GeneID:2587	GALNR; GALNR1;	NT_025004.13	GALR1	96	153.8
cg19703610	GeneID:57453	KIAA1132;	NM_020693.2	DSCAML1	97	146.9
cg00263760	GeneID:11023	MGC126743; MGC126745;	NM_199131.1	VAX1	98	155.57
cg10235817	GeneID:152	ADRA2L2; ADRARL2; ADRA2RL2; ALPHA2CAR;	NM_000683.3	ADRA2C	99	165.57
cg00565688	GeneID:7161	P73;	NM_005427.1	TP73	100	155.53

- 75 -

TABLE 13b
LIMMA Top 100

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	LIMMA	LIMMA AVG RNK
cg07846220	GeneID:284217	LAMA;	NM_005559.2	LAMA1	1	2.066667
cg07271264	GeneID:4654	PUM; MYF3; MYOD;	NM_002478.3	MYOD1	2	4.033333
cg20723355	GeneID:162517	Fbx39; MGC35179;	NM_153230.1	FBXO39	3	5.1
cg11377136	GeneID:10343		NM_006071.1	PKDREJ	4	7.233333
cg06391468	GeneID:6662	CMD1; SRA1; CMPD1;	NM_000346.2	SOX9	5	9.566667
cg18815943	GeneID:2301	ASMD; FKHL12; FREAC8;	NM_012186.1	FOXE3	6	10.83333
cg04534765	GeneID:2587	GALNR; GALNR1;	NM_001480.2	GALR1	7	12.86667
cg06268694	GeneID:9620	ME2; FMI2; CDHF9; HFM12; DKFZp434P0729;	NM_014246.1	CELSR1	8	13.6
cg25720804	GeneID:30012	RNX; HOX11L2; MGC29804;	NM_021025.2	TLX3	9	14.56667
cg04453065	GeneID:3491	CCN1; GIG1; IGFBP10;	NM_001554.3	CYR61	10	16.56667
cg02919422	GeneID:64321	FLJ22252;	NM_022454.2	SOX17	11	17.83333
cg25484904	GeneID:80157		NM_025087.1	FLJ21511	12	18.03333
cg23037403	GeneID:285676	FLJ37444;	NM_182594.1	ZNF454	13	20.3
cg10639440	GeneID:9805	SES1; KIAA0193;	NM_014766.2	SCRN1	14	21.6
cg02613386	GeneID:162517	Fbx39; MGC35179;	NM_153230.1	FBXO39	15	24.1
cg22418909	GeneID:6422	FRP; FRP1; FrzA; FRP-1; SARP2;	NM_003012.3	SFRP1	16	24.5
cg11389172	GeneID:6572	VACHT;	NM_003055.1	SLC18A3	17	27.63333
cg25044651	GeneID:206338	MGC125378; MGC125379;	NM_173800.3	FLJ90650	18	27.66667
cg27603796	GeneID:83992	Orf4; C7orf8; CORTBP2; KIAA1758; MGC104579;	NM_033427.1	CTTNBP2	19	28.23333
cg23290344	GeneID:4741	NFM; NEFM; NF-M;	NM_005382.1	NEF3	20	32.46667
cg26609631	GeneID:219409	Gsh-1;	NM_145657.1	GSH1	21	33.26667
cg17977409	GeneID:6456	CNSA2; SH3P4; EEN-B1; SH3D2A; FLJ20276;	NM_003026.1	SH3GL2	22	36.06667
cg02250594	GeneID:9480	OC2; OC-2; MGC120377; MGC120378;	NM_004852.1	ONECUT2	23	39.03333
cg23211240	GeneID:54766	PC3B; MGC33003;	NM_017589.2	BTG4	24	40.53333
cg04632671	GeneID:5468	NR1C3; PPARG1; PPARG2; HUMPPARG;	NM_138712.2	PPARG	25	42.43333
cg00565688	GeneID:7161	P73;	NM_005427.1	TP73	26	44.1
cg05020203	GeneID:401045		NM_207485.1	FLJ41327	27	45.6
cg02154186	GeneID:10687	MA2; MM2; RGAG2; KIAA0883;	NM_007257.4	PNMA2	28	46.46667
cg17162024	GeneID:389658		NM_207413.1	UNQ9433	29	46.63333
cg17371081	GeneID:4745	NRP1; IDH3GL;	NM_006157.2	NELL1	30	46.66667
cg15191648	GeneID:27164		NM_171999.1	SALL3	31	46.76667
cg19064258	GeneID:9956	30ST2; 30ST2;	NM_006043.1	HS3ST2	32	47.16667

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	LIMMA	LIMMA AVG RNK
cg13359415	GeneID:55203	LGIL2; FLJ10675; KIAA1916; MGC126808; MGC126810;	NM_018176.2	LGIL2	33	47.93333
cg21688264	GeneID:9892	AP180; KIAA0656; CRBP; RBPC; CRBP1; CRABP-I;	NM_014841.1	SNAP91	34	49.8
cg12497564	GeneID:5947	ADMD; STGD2; STGD3;	NM_002899.2	RBP1	35	50.36667
cg13297865	GeneID:6785	CHT; CHT1; hCHT;	NM_022726.2	ELOVL4	36	51.2
cg16232126	GeneID:60482	RIM4; C20orf190; dJ781B1.3;	NM_021815.2	SLC5A7	37	52.56667
cg19332710	GeneID:140730	LCAD; ACAD4;	NM_182970.2	RIMS4	38	54.5
cg14795968	GeneID:33	NP2; NARP; NP-II;	NM_001608.2	ACADL	39	54.9
cg12799895	GeneID:4885	LIPD;	NM_002523.1	NPTX2	40	54.93333
cg08918749	GeneID:4023	MRT1; BSSP3; BSSP-3; MGC12722; MOTOPSIN;	NM_000237.1	LPL	41	57.8
cg07625840	GeneID:8492	BL2; ST17; NECL2; TSLC1; IGSF4A; SYNCAM; sTSLC-1; synCAM1; DKFZp686F1789;	NM_003619.2	PRSS12	42	57.86667
cg10193817	GeneID:23705	HSPDE10A;	NM_014333.2	IGSF4	43	58.6
cg24133115	GeneID:10846	EMT; EMT3;	NM_006661.1	PDE10A	44	58.8
cg07237939	GeneID:6581	DFNB30;	NM_021977.2	SLC22A3	45	62.33333
cg23771603	GeneID:53904	GRO1; GROa; MGSA; NAP-3; SCYB1; MGSA-a; MGSA alpha;	NM_017433.3	MYO3A	46	62.43333
cg25806808	GeneID:2919	UNQ676;	NM_001511.1	CXCL1	47	63.7
cg09619146	GeneID:119587	NCX2;	NM_198148.1	CPXM2	48	63.73333
cg22123464	GeneID:6543	PAK5; KIAA1264; MGC26232;	NM_015063.1	SLC8A2	49	64.13333
cg12388309	GeneID:57144	CASPR3; CNTNAP3A; RP11- 290L7.1; RP11- 138L21.1;	NM_177990.1	PAK7	50	64.16667
cg06675478	GeneID:6656		NM_005986.2	SOX1	51	65.2
cg00662556	GeneID:2587	GALNR; GALNR1;	NT_025004.13	GALR1	52	66.7
cg13059782	GeneID:79937	VGL2; VITO1;	NM_033655.2	CNTNAP3	53	67.53333
cg03963198	GeneID:50805	PTF1-p48;	NM_016358.1	IRX4	54	68.86667
cg03160135	GeneID:245806	PNPLA3; C22orf20; IPLA(2)epsilon;	NM_153453.1	VGLL2	55	72.23333
cg17389519	GeneID:256297	TFT; MGC104817;	NM_178161.1	PTF1A	56	72.66667
cg13184872	GeneID:80339	FPP; MSH; PFM; CRS2; HOX8; PFM1;	NM_025225.2	ADPN	57	73.06667
cg17188046	GeneID:6862	PUM; MYF3; MYOD;	NM_003181.2	T	58	74.4
cg27096144	GeneID:4488	FRP; FRP1; FrZA; FRP-1; SARP2;	NM_002449.3	SFRP1	59	75.36667
cg18555440	GeneID:4654	NKX2H; NKX2.8; Nkx2-9;	NM_002478.3	MSX2	60	77.73333
cg15839448	GeneID:6422	CP24; CYP24; MGC126273; MGC126274; P450- CC24;	NM_003012.3	MYOD1	61	79.96667
cg17685628	GeneID:26257		NM_014360.2	NKX2-8	62	80.23333
cg02604290	GeneID:1591		NM_000782.3	CYP24A1	63	85.16667

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	LIMMA	LIMMA AVG RNK
cg00027083	GeneID:23136	4.1B; DAL1; DAL-1; KIAA0987;	NM_012307.2	EPB41L3	64	85.2
cg25942450	GeneID:30012	RNX; HOX11L2; MGC29804;	NM_021025.2	TLX3	65	85.53333
cg02536286	GeneID:7022	ERF1; TFAP2G; hAP- 2g; AP2-GAMMA;	NM_003222.3	TFAP2C	66	86.56667
cg10063179	GeneID:4543	MEL-1A-R;	NM_005958.3	MTNR1A	67	86.83333
cg16196812	GeneID:2834	GR3; GPR10; PrRPR; MGC126539; MGC126541;	NM_004248.1	PRLHR	68	87.5
cg27409364	GeneID:3746	KV4; NGK2; KV3.1; MGC129855;	NM_004976.2	KCNC1	69	89.03333
cg15591678	GeneID:22891	UAN; KIAA0844; MGC41821; MGC87345;	NM_199451.1	ZNF365	70	90.2
cg10947146	GeneID:286046	XRG6; C8orf7; C8orf21;	NM_173683.2	XKR6	71	90.33333
cg18671950	GeneID:2200	FBN; SGS; WMS; MASS; MFS1; OCTD; fibrillin;	NM_000138.2	FBN1	72	94.6
cg00263760	GeneID:11023	MGC126743; MGC126745;	NM_199131.1	VAX1	73	95.26667
cg19721889	GeneID:9421	Hxt; eHand; Thing1;	NM_004821.1	HAND1	74	97.8
cg25250358	GeneID:5352	LH2; TLH;	NM_182943.2	PLOD2	75	98.06667
cg17688525	GeneID:91133	HsT1031;	NM_173464.1	L3MBTL4	76	98.13333
cg12331389	GeneID:5950		NM_006744.3	RBP4	77	99.3
cg16954341	GeneID:10590	SEGN; SECRET; setagin; DJ501N12.8;	NM_006998.3	SCGN	78	99.83333
cg06222851	GeneID:55753		NM_018245.1	OGDHL	79	102.0333
cg17788682	GeneID:9915	KIAA0307;	NM_014862.3	ARNT2	80	103.1
cg02844545	GeneID:9247	GCMB; hGCMB;	NM_004752.1	GCM2	81	103.5333
cg24101578	GeneID:64405	C20orf25; MGC39564; dJ998H6.1;	NM_021248.1	CDH22	82	105.1333
cg12594641	GeneID:130574		NM_194317.2	MGC52057	83	108.3
cg15433631	GeneID:153572		NM_033267.2	IRX2	84	109
cg13929328	GeneID:399823	FOXI2;	NM_207426.1	FLJ46831	85	109.3
cg17051321	GeneID:646	BNC; BSN1; HsT19447;	NT_077661.2	BNC1	86	111.9
cg19591206	GeneID:6578	PGT; OATP2A1; SLC21A2;	NM_005630.1	SLC02A1	87	112.3667
cg22752533	GeneID:57468	KCC2; KIAA1176;	NM_020708.3	SLC12A5	88	113.2333
cg14473924	GeneID:23024	LNK3; SEMACAP3;	NM_015009.1	PDZRN3	89	114.8333
cg12699371	GeneID:2587	GALNR; GALNR1;	NM_001480.2	GALR1	90	116.0333
cg17252960	GeneID:3400		NM_001546.2	ID4	91	116.4
cg10453365	GeneID:51458	RHGK; PDRC2; C15orf6;	NM_016321.1	RHCG	92	116.8667
cg10235817	GeneID:152	ADRA2L2; ADRARL2; ADRA2RL2; ALPHA2CAR;	NM_000683.3	ADRA2C	93	118.6333
cg17619823	GeneID:155	BETA3AR;	NM_000025.1	ADRB3	94	119.7

- 78 -

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	LIMMA	LIMMA AVG RNK
cg24120841	GeneID:7068	GRTH; THR1; ERBA2; NR1A2; THRB1; THRB2; ERBA-BETA; MGC126109; MGC126110;	NM_000461.2	THRB	95	120.5333
cg14882700	GeneID:133060		NM_177998.1	OTOP1	96	123.1
cg25536676	GeneID:1718	KIAA0018; SELADIN1; Nbla03646; seladin- 1;	NM_014762.2	DHCR24	97	123.5333
cg12839593	GeneID:6495		NM_005982.1	SIX1	98	124
cg05807991	GeneID:133022	MGC26568;	NM_152402.1	TRAM1L1	99	124.4333
cg26721264	GeneID:2587	GALNR; GALNR1;	NM_001480.2	GALR1	100	127.9333

TABLE 13c
RFE-SVM Top 100

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	SVM+ TTEST	SVM+ TTEST AVG RNK
cg12969595	GeneID:26022	DKFZP564K1964;	NM_015544.2	TMEM98	1	8.5
cg12594641	GeneID:130574		NM_194317.2	MGC52057	2	9.87
cg21801378	GeneID:60677	CELF6;	NM_052840.3	BRUNOL6	3	14.43
cg07846220	GeneID:284217	LAMA;	NM_005559.2	LAMA1	4	16.03
cg16041660	GeneID:144165	RILP; FLJ31627; FLJ31937;	NM_153026.1	PRICKLE1	5	17.57
cg13168820	GeneID:11122	RPTPrho; KIAA0283;	NM_133170.2	PTPRT	6	18.13
cg07533148	GeneID:25893	BIA2; DKFZp434C091;	NM_015431.2	TRIM58	7	22.6
cg13297865	GeneID:6785	ADMD; STGD2; STGD3;	NM_022726.2	ELOVL4	8	23.43
cg25381844	GeneID:6457	CNSA3; EEN-B2; SH3D2C; SH3P13; HsT19371; EEN- 2B-L3;	NM_003027.2	SH3GL3	9	25.23
cg01775414	GeneID:112885	PHF4; BHC80L; FLJ34161;	NM_138415.1	PHF21B	10	28.53
cg19063972	GeneID:11166	SOX25;	NM_007084.2	SOX21	11	29.83
cg08186362	GeneID:11255	HH3R; GPCR97;	NM_007232.1	HRH3	12	30.6
cg09053680	GeneID:8433		NM_003577.2	UTF1	13	31.03
cg17688525	GeneID:91133	HsT1031;	NM_173464.1	L3MBTL4	14	31.4
cg10227191	GeneID:3875	K18; CYK18;	NM_000224.2	KRT18	15	31.67
cg13191808	GeneID:255738	FH3; NARC1; NARC-1; HCHOLA3;	NM_174936.2	PCSK9	16	32
cg17788682	GeneID:9915	KIAA0307;	NM_014862.3	ARNT2	17	32.3
cg13449778	GeneID:148753	NDSP; MGC16664; RP11-12M5.2;	NM_173509.2	C1orf76	18	34.67
cg02090283	GeneID:79937	CASPR3; CNTNAP3A; RP11- 290L7.1; RP11- 138L21.1;	NM_033655.2	CNTNAP3	19	35.63
cg17619823	GeneID:155	BETA3AR;	NM_000025.1	ADRB3	20	41.8
cg17776353	GeneID:657	ALK3; CD292; ACVRLK3;	NM_004329.2	BMPRI1A	21	42.03
cg09787254	GeneID:79152	FAAH; FAXDC1; FLJ25287;	NM_024306.2	FA2H	22	42.53
cg19884262	GeneID:399823	FOXI2;	NM_207426.1	FLJ46831	23	43.7
cg09186006	GeneID:387700		NM_213606.1	SLC16A12	24	44.73
cg00826384	GeneID:5803	PTPZ; HPTPZ; PTP18; PTPRZ; RPTPB;	NM_002851.1	PTPRZ1	25	45.9
cg10646402	GeneID:5800	PTPU2; GLEPP1; PTP-U2;	NM_002848.2	PTPRO	26	50.6
cg09688546	GeneID:2743		NM_000824.2	GLRB	27	52.87
cg20113732	GeneID:4745	NRP1; IDH3GL;	NM_006157.2	NELL1	28	52.9
cg09450020	GeneID:261729	STMP; IPCA1; IPCA-1; STAMP1; PCANAP1;	NM_152999.2	STEAP2	29	53.7

- 80 -

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	SVM+ TTEST	SVM+ TTEST AVG RNK
cg25806808	GeneID:2919	GRO1; GROa; MGSA; NAP-3; SCYB1; MGSA-a; MGSA alpha;	NM_001511.1	CXCL1	30	54.03
cg06572160	GeneID:3748	KV3.3; KSHIID;	NM_004977.2	KCNC3	31	55.83
cg25764191	GeneID:9118	NEF5; NF-66; TXBP-1; MGC12702;	NM_032727.2	INA	32	64.3
cg14795968	GeneID:33	LCAD; ACAD4;	NM_001608.2	ACADL	33	64.63
cg20082357	GeneID:535	a1; Stv1; VPP1; Vph1; ATP6N1; ATP6N1A; DKFZp781J1951;	NM_005177.2	ATP6V0A1	34	65.87
cg09500672	GeneID:130574		NM_194317.2	MGC52057	35	66.43
cg09775312	GeneID:5507	PPP1R5;	NM_005398.3	PPP1R3C	36	67.93
cg20182358	GeneID:4291		NM_022443.2	MLF1	37	68.9
cg02128567	GeneID:8645	TASK2; TASK-2; FLJ11035;	NM_003740.2	KCNK5	38	68.97
cg04711050	GeneID:6505	EAAC1; EAAT3;	NM_004170.4	SLC1A1	39	72.97
cg25250358	GeneID:5352	LH2; TLH;	NM_182943.2	PLOD2	40	73.6
cg19721889	GeneID:9421	Hxt; eHand; Thing1;	NM_004821.1	HAND1	41	75.2
cg12297221	GeneID:116969	MGC22848;	NM_053017.2	ART5	42	75.47
cg02144298	GeneID:1050	CEBP; C/EBP- alpha;	NM_004364.2	CEBPA	43	77.2
cg23984434	GeneID:2977	GC-SA2; GUC1A2;	NM_000855.1	GUCY1A2	44	77.47
cg12128017	GeneID:284656	FLJ16103; FLJ33655; MGC43817;	NM_173641.1	EPHA10	45	82.83
cg23433607	GeneID:1114	SCG1;	NM_001819.1	CHGB	46	87.37
cg04786857	GeneID:245711	SPY1; MGC57218; MGC110856;	NM_001008779.1	SPDY1	47	88.07
cg12388309	GeneID:57144	PAK5; KIAA1264; MGC26232;	NM_177990.1	PAK7	48	89
cg11625946	GeneID:121227	FLJ90440; KIAA3016;	NM_153377.3	LRIG3	49	90.27
cg18815943	GeneID:2301	ASMD; FKHL12; FREAC8;	NM_012186.1	FOXO3	50	90.7
cg17371081	GeneID:4745	NRP1; IDH3GL;	NM_006157.2	NELL1	51	91.3
cg13184872	GeneID:80339	PNPLA3; C22orf20; iPLA(2)epsilon;	NM_025225.2	ADPN	52	92.97
cg08876932	GeneID:401	ARIX; FEOM2; NCAM2; PMX2A; CFEOM2; MGC52227;	NM_005169.2	PHOX2A	53	94
cg06151165	GeneID:30813	PPD; KTCN; PPCD; RINX;	NM_014588.4	VSX1	54	94.8
cg23739862	GeneID:56341	HRMT1L3; HRMT1L4;	NM_019854.3	PRMT8	55	94.8
cg18191162	GeneID:221458	C6orf102; MGC33317; dJ137F1.4; dJ188D3.1; dJ1043E3.1;	NM_145027.3	KIF6	56	95.37
cg15839448	GeneID:6422	FRP; FRP1; FrzA; FRP-1; SARP2;	NM_003012.3	SFRP1	57	96.3

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	SVM+ TTEST	SVM+ TTEST AVG RNK
cg10639440	GeneID:9805	SES1; KIAA0193;	NM_014766.2	SCRN1	58	96.63
cg18070061	GeneID:1846	TYP; HVH2; MKP2;	NM_057158.2	DUSP4	59	97.17
cg01805282	GeneID:2070	CMD1J; DFNA10;	NT_025741.14	EYA4	60	97.37
cg24659201	GeneID:2863		NM_001508.1	GPR39	61	98.5
cg08999486	GeneID:9578	MRCKB;	NM_006035.2	CDC42BPB	62	99.63
cg22734085	GeneID:152641	KIAA1124;	NM_153008.3	FLJ30277	63	101.67
cg10644361	GeneID:145282	DKFZp313M2036;	NM_138731.2	MIPOL1	64	102.2
cg02265318	GeneID:64757	FLJ22390; RP11- 295M18.1;	NM_022746.2	MOSC1	65	105.7
cg23196831	GeneID:7373	UND;	NM_021110.1	COL14A1	66	106
cg04180868	GeneID:56952	HHGP; FLJ11888;	NM_020200.5	PRTFDC1	67	106.43
cg10453365	GeneID:51458	RHGK; PDRC2;	NM_016321.1	RHCG	68	108.37
cg26963271	GeneID:5142	C15orf6;				
		DPDE4; PDEIVB;				
		MGC126529;	NM_002600.2	PDE4B	69	110.3
		DKFZp686F2182;				
cg15613048	GeneID:57576	KIF3X; KIF17B;	NM_020816.1	KIF17	70	111.13
		KIAA1405;				
cg00756058	GeneID:22873	DZIP; DZIPT1;	NM_198968.1	DZIP1	71	112.63
cg26831415	GeneID:286097	KIAA0996; RP11- 23E3.3;	NM_181723.1	EFHA2	72	113.43
cg07271264	GeneID:4654	PUM; MYF3;	NM_002478.3	MYOD1	73	113.5
		MYOD;				
cg16534233	GeneID:284716	MGC47816; RP11- 157D18.1;	NM_173642.1	FAM80A	74	113.93
cg11965370	GeneID:50863	NTM; MGC60329;	NM_016522.2	HNT	75	115.47
cg06222851	GeneID:55753		NM_018245.1	OGDHL	76	118.7
cg08900043	GeneID:3226	HOX3I; MGC5259;	NM_017409.2	HOXC10	77	118.9
cg26252167	GeneID:2830		NM_005284.2	GPR6	78	122.1
cg04534765	GeneID:2587	GALNR; GALNR1;	NM_001480.2	GALR1	79	123.9
cg23967169	GeneID:151647	TAF4; TAF4-4;	NM_182522.3	FAM19A4	80	126.07
		FLJ25161;				
cg09892203	GeneID:27092	MGC11138;	NM_014405.2	CACNG4	81	126.1
		MGC24983;				
cg09381701	GeneID:23328	KIAA0790;	NM_015278.3	SASH1	82	126.6
		dJ323M4.1; RP3- 323M4.1;				
cg22040627	GeneID:284111	NACT;	NM_177550.2	SLC13A5	83	127.67
cg21688264	GeneID:9892	MGC138356;	NM_014841.1	SNAP91	84	128.67
cg11846956	GeneID:5655	DKFZp686E17257;	NT_011109.15	KLK10	85	130.4
		AP180; KIAA0656;				
cg07237939	GeneID:6581	NES1; PRSSL1;	NM_021977.2	SLC22A3	86	130.6
		EMT; EMTH;				
		OCT3;				
cg20052718	GeneID:7291	SCS; ACS3;	NT_007819.16	TWIST1	87	133.57
		BPES2; BPES3;				
		TWIST;				
cg12087643	GeneID:2674	GDNFR; RET1L;	NM_005264.2	GFRA1	88	134.23
cg22594309	GeneID:127833	RETL1; TRNR1;	NM_177402.3	SYT2	89	136.03
		GDNFRA;				
		MGC23045; GFR- ALPHA-1;				
		SytII;				

- 82 -

TargetID	GENE_ID	SYNONYM	ACCESSION	SYMBOL	SVM+ TTEST	SVM+ TTEST AVG RNK
cg08768421	GeneID:9615	CYPIN; GUANASE; MGC9982; NEDASIN; KIAA1258;	NM_004293.2	GDA	90	137.97
cg01557297	GeneID:51310	BOCT; BOIT; hBOIT;	NM_020372.2	SLC22A17	91	139.57
cg23752923	GeneID:55283	FLJ11006; MGC71509;	NM_018298.9	MCOLN3	92	142.13
cg02755525	GeneID:81831	NEOT2; FLJ10430; FLJ14724; FLJ90456;	NM_018092.3	NETO2	93	143.37
cg06187947	GeneID:51454	CED6; GULP; CED- 6;	NM_016315.2	GULP1	94	143.6
cg19709625	GeneID:26577	PCPE2;	NM_013363.2	PCOLCE2	95	145
cg08016383	GeneID:170691	FLJ32769;	NM_139057.1	ADAMTS17	96	145.23
cg22415432	GeneID:83259	PCDHY; PCDH22; PCDH11X;	NM_032971.1	PCDH11Y	97	145.87
cg18671950	GeneID:2200	FBN; SGS; WMS; MASS; MFS1; OCTD; fibrillin;	NM_000138.2	FBN1	98	145.97
cg07380496	GeneID:4131	MAP5; FUTSCH; DKFZp686F1345;	NM_032010.1	MAP1B	99	146.03
cg27603796	GeneID:83992	Orf4; C7orf8; CORTBP2; KIAA1758; MGC104579;	NM_033427.1	CTTNBP2	100	147.6

TABLE 13d
Relapsed Leukaemia RFE-SVM Top 100

Rank	RFE	Symbol	Accession
1	cg22903370	SSR4	NM_006280.1
2	cg13060154	DAB2IP	NT_008470.18
3	cg00915289	CKS1B	NM_001826.1
4	cg07643942	LACRT	NM_033277.1
5	cg13494498	TUBB3	NM_006086.2
6	cg04601137	ADAMTSL5	NM_213604.1
7	cg17704839	UBL5	NM_024292.2
8	cg04576021	HLA-DOB	NM_002120.2
9	cg19784477	TFPI2	NT_007933.14
10	cg02525396	TOMM40L	NM_032174.3
11	cg00112517	PPP1R1B	NM_032192.2
12	cg05944800	INSIG2	NM_016133.2
13	cg05781767	PODXL2	NM_015720.1
14	cg19948393	ANKRD33	NM_182608.2
15	cg15540820	EOMES	NM_005442.2
16	cg00231140	ACAA2	NM_006111.1
17	cg06585893	CAPN9	NM_006615.2
18	cg07679836	BAK1	NM_001188.2
19	cg18611122	LASS2	NM_013384.3
20	cg09276451	SLTL2	NM_138440.1
21	cg09352789	XPNPEP1	NM_020383.2
22	cg11523020	UPK3A	NM_006953.2
23	cg02915544	PSMAL	NM_153696.2
24	cg26134665	STX1B2	NM_052874.1
25	cg09907395	KIAA1618	NM_020954.2
26	cg22933847	MRGPRF	NM_145015.2
27	cg06836736	ME1	NM_002395.2

Rank	RFE	Symbol	Accession
28	cg24745738	EDNRB	NT_024524.13
29	cg24924779	KCNG1	NM_002237.2
30	cg23022999	FLJ45909	NM_198445.1
31	cg06600447	MGC2749	NM_024069.2
32	cg23509869	LST1	NM_205838.1
33	cg01172972	ZFYVE9	NM_004799.2
34	cg16075940	FLJ37396	NM_173671.1
35	cg20903926	C1orf177	NM_152607.1
36	cg27390220	RIS1	NM_015444.1
37	cg09076012	NOV	NM_002514.2
38	cg12167564	LYST	NM_000081.2
39	cg15823100	PISD	NM_014338.3
40	cg08013810	PHF20	NM_016436.3
41	cg17983307	DKFZp434B1231	NM_178275.3
42	cg10602543	H19	NT_009237.17
43	cg23850212	ZFP28	NM_020828.1
44	cg00727590	PLA2G3	NM_015715.2
45	cg20849549	PTEN	NM_000314.3
46	cg26271255	ARHGAP12	NM_018287.4
47	cg15198335	MGC4562	NM_133375.2
48	cg11714502	AK1	NM_000476.1
49	cg06615667	GPR107	NM_020960.2
50	cg04311964	LYPD2	NM_205545.1
51	cg02829654	LYST	NM_000081.2
52	cg21616088	FBXL4	NM_012160.3
53	cg04794887	LETM1	NM_012318.1
54	cg13277939	CTAGE5	NM_203357.1
55	cg12108912	MGC10993	NM_030577.1
56	cg25488547	CHM	NM_000390.1
57	cg23282559	KL	NM_004795.2

Rank	RFE	Symbol	Accession
58	cg10139846	INADL	NM_005799.2
59	cg04713352	ATP4A	NM_000704.2
60	cg06034822	USF1	NM_007122.3
61	cg04528819	KLF14	NM_138693.1
62	cg05473677	OSTalpha	NM_152672.3
63	cg07533148	TRIM58	NM_015431.2
64	cg15819171	FLJ23588	NM_022785.2
65	cg20644981	RPS3A	NM_001006.3
66	cg23640903	GFOD1	NM_018988.1
67	cg05225390	ZW10	NM_004724.2
68	cg14986420	YIPF6	NM_173834.2
69	cg22578204	TIMP3	NT_011520.11
70	cg01612158	FCRL4	NM_031282.1
71	cg07816439	C1QTNF2	NM_031908.3
72	cg08578641	DNAI1	NM_012144.2
73	cg17687282	A4GNT	NM_016161.1
74	cg01929065	CD63	NM_001780.3
75	cg09554443	CD3Z	NM_000734.2
76	cg22320000	HNRPU	NM_031844.2
77	cg19154438	CKM	NM_001824.2
78	cg06521852	HRIHFB2122	NM_138632.1
79	cg21096399	MCAM	NM_006500.2
80	cg12359315	SLC6A20	NM_020208.2
81	cg25511429	NRN1	NM_016588.2
82	cg15129294	CCL4L2	NM_207007.2
83	cg02579736	FANCD2	NM_033084.3
84	cg15534366	CDH4	NM_001794.2
85	cg05409038	TEX261	NM_144582.2
86	cg04498679	C2	NM_000063.3
87	cg02489552	FLJ40365	NM_173482.1

- 86 -

Rank	RFE	Symbol	Accession
88	cg26145228	NTNG1	NM_014917.2
89	cg27342122	PHF20L1	NM_024878.1
90	cg07753583	LRRC61	NM_023942.1
91	cg19722847	IPO8	NM_006390.2
92	cg00840516	HYAL2	NM_003773.2
93	cg01816336	MINPP1	NM_004897.2
94	cg14904464	NELL2	NM_006159.1
95	cg15415507	EPHA2	NM_004431.2
96	cg13156411	PTHR1	NM_000316.2
97	cg11402505	FBXO27	NM_178820.3
98	cg26135716	GLUL	NM_001033056.1
99	cg01875775	CHEK2	NM_007194.3
100	cg19103609	PKN1	NM_002741.3

TABLE 13e
Relapse Leukaemia Centroid Top 100

Rank	Centroid	Symbol	Accession
1	cg00915289	CKS1B	NM_001826.1
2	cg22903370	SSR4	NM_006280.1
3	cg13060154	DAB2IP	NT_008470.18
4	cg19784477	TFPI2	NT_007933.14
5	cg15540820	EOMES	NM_005442.2
6	cg02915544	PSMAL	NM_153696.2
7	cg04601137	ADAMTSL5	NM_213604.1
8	cg04576021	HLA-DOB	NM_002120.2
9	cg01172972	ZFYVE9	NM_004799.2
10	cg07679836	BAK1	NM_001188.2
11	cg24697031	EVPL	NM_001988.1
12	cg13156411	PTHR1	NM_000316.2
13	cg16902509	ITGA8	NM_003638.1
14	cg18384097	PTPN7	NM_002832.2
15	cg15823100	PISD	NM_014338.3
16	cg25619607	IMPACT	NM_018439.1
17	cg13494498	TUBB3	NM_006086.2
18	cg10602543	H19	NT_009237.17
19	cg06912252	C9orf125	NM_032342.1
20	cg24924779	KCNG1	NM_002237.2
21	cg26619317	CNN3	NM_001839.2
22	cg07643942	LACRT	NM_033277.1
23	cg19154438	CKM	NM_001824.2
24	cg22954818	APOBEC3A	NM_145699.2
25	cg23022999	FLJ45909	NM_198445.1
26	cg08602689	KDELC1	NM_024089.1
27	cg15958424	ACPP	NM_001099.2

- 88 -

Rank	Centroid	Symbol	Accession
28	cg11812218	GHSR	NM_198407.1
29	cg12167564	LYST	NM_000081.2
30	cg14611174	SIX6	NM_007374.1
31	cg08578641	DNAI1	NM_012144.2
32	cg12108912	MGC10993	NM_030577.1
33	cg12291552	GOLPH4	NM_014498.2
34	cg13277939	CTAGE5	NM_203357.1
35	cg15383392	CTTN	NM_005231.2
36	cg11714502	AK1	NM_000476.1
37	cg15819171	FLJ23588	NM_022785.2
38	cg15534366	CDH4	NM_001794.2
39	cg14155416	L3MBTL4	NM_173464.1
40	cg20352798	ZP3	NM_007155.4
41	cg21096399	MCAM	NM_006500.2
42	cg18755783	SPG20	NM_015087.3
43	cg01816336	MINPP1	NM_004897.2
44	cg22231902	EN1	NM_001426.2
45	cg25956985	MAP3K7IP1	NM_153497.2
46	cg15669092	TUBB3	NM_006086.2
47	cg06585893	CAPN9	NM_006615.2
48	cg08013810	PHF20	NM_016436.3
49	cg07909265	ITGAV	NM_002210.2
50	cg27360560	TMEM67	NM_153704.3
51	cg14904464	NELL2	NM_006159.1
52	cg11956146	CABYR	NM_012189.2
53	cg09085198	ANKRD29	NM_173505.2
54	cg08132931	ADD2	NM_017482.1
55	cg13365761	SCNN1G	NM_001039.2
56	cg00727590	PLA2G3	NM_015715.2
57	cg22584138	SLC6A4	NM_001045.2

Rank	Centroid	Symbol	Accession
58	cg03297731	GDPD3	NM_024307.1
59	cg11027570	RBP1	NM_002899.2
60	cg08990057	PLS3	NM_005032.3
61	cg23509869	LST1	NM_205838.1
62	cg09907395	KIAA1618	NM_020954.2
63	cg00243313	IRX4	NM_016358.1
64	cg16652063	SLC13A5	NM_177550.2
65	cg16422907	CCNA1	NM_003914.2
66	cg17688525	L3MBTL4	NM_173464.1
67	cg07533148	TRIM58	NM_015431.2
68	cg05781767	PODXL2	NM_015720.1
69	cg10139846	INADL	NM_005799.2
70	cg22917487	CX3CR1	NM_001337.3
71	cg07815386	RNF149	NM_173647.2
72	cg02265318	MOSC1	NM_022746.2
73	cg22040627	SLC13A5	NM_177550.2
74	cg01401376	EYA4	NT_025741.14
75	cg17471019	TMEM67	NM_153704.3
76	cg02525396	TOMM40L	NM_032174.3
77	cg09076012	NOV	NM_002514.2
78	cg07908874	TUBGCP2	NM_006659.1
79	cg03626672	CMKOR1	NM_020311.1
80	cg07730329	PCDHGA12	NM_003735.2
81	cg09276451	SLITL2	NM_138440.1
82	cg05944800	INSIG2	NM_016133.2
83	cg05473677	OSTalpha	NM_152672.3
84	cg09352789	XPNPEP1	NM_020383.2
85	cg02724472	LRRC17	NM_005824.1
86	cg24745738	EDNRB	NT_024524.13
87	cg01029592	SOX15	NM_006942.1

- 90 -

Rank	Centroid	Symbol	Accession
88	cg10006582	BBS5	NM_152384.2
89	cg11647681	PCDHGA12	NM_003735.2
90	cg23719124	GSTM5	NM_000851.2
91	cg16731240	ZNF577	NM_032679.1
92	cg07136254	HFE2	NM_145277.3
93	cg23850212	ZFP28	NM_020828.1
94	cg22970435	SPATS1	NM_145026.2
95	cg23493704	SLC10A3	NM_019848.2
96	cg06615667	GPR107	NM_020960.2
97	cg05938725	NUAK2	NM_030952.1
98	cg06836736	ME1	NM_002395.2
99	cg05368341	SYT6	NM_205848.1
100	cg10517312	RNF36	NM_080745.2

EXAMPLE 8***Generation of primers***

[0145] The nature of bisulphite treatment is that it results in a homogenous mixture of highly converted templates to ones that are not converted at all. To favor amplification of the fully converted templates, primers are chosen which this taken into account.

[0146] The + and – strands are not complementary and thus primers are designed for one strand and it is usually the plus strand as given by the text editor.

10

[0147] The criteria for primer selection includes:

- Primer sets amplifying no more than a maximum of 450bp.
- At least 30% C within the primer.
- 15 • A primer that selectively amplifies bisulfite converted templates.
- An even distribution of bases prior to bisulfite conversion.
- As high GA content as possible.
- A lack of CpG residues and if this can not be avoided then degeneracy at the possibly methylated cytosine residue.
- 20 • A long primer (25-30mer) to ensure uniqueness of the primer.
- Have a nested set to do the second round of PCR. Nested or hemi-nested primer is fine.
- T_m calculations to be equivalent for primer pairs. Input primer sequence into Primer Express to check the T_m . Primer express uses nearest neighbour method to calculate
- 25 T_m . Match primer T_m . If they can not be matched then that can not be helped. Set T_m of the PCR reaction to be 2°C lower than the lowest T_m of the primer set. T_m calculation between Primer Express and NetPrimer are different.

- 92 -

EXAMPLE 9

MRD assay development

[0148] Assays for MRD of methylation markers FOXE3, TLX3 (and others) identified by
5 Illumina Infinium Methylation Array (e.g. RIMS4, OGDHL, LVRN, NK2-X2, THRB, MSX2 and GSX1) are developed.

[0149] Robust and sensitive for MRD assays are developed using MS-PCR, MethyLight, HRM in addition to SEQUENOM for DNA methylation detection at the promoter regions
10 of these genes.

[0150] Methods are tested for methylation detection which include enrichment strategies for methylated DNA and depletion studies excluding the normal non-methylated DNA.

15 [0151] The limits of detection and sensitivity of these assays are assessed and evaluated.

EXAMPLE 10

TEL/AML subtype

20 [0152] Genes are analyzed by SEQUENOM assay in all TEL/AML(ETV6/RUNX1) samples (tumor and remission) available in the cohort.

[0153] The assays are extended to other subtypes of leukemia using the top candidates found in TEL/AML study, beginning with normal karyotype and hyperdiploidy.

25

[0154] For prognosis, methylation changes are identified in samples that predict relapse in leukemia and/or poor long term outcome.

[0155] Other cohorts of TEL/AML samples are also tested.

30

- 93 -

EXAMPLE 11

Normal baseline controls

5 [0156] CD34+ bone marrow cells (stem cell lineages) are obtained and DNA Methylation profiles are determined by Illumina Infinium. This is a more appropriate control to use and false positive hits can be removed for DNA methylation associated with normal development.

10 [0157] *In vitro* culture methods of the CD34+ cells extracted from donors are also developed.

[0158] Non-leukemia bone marrow samples are sourced to obtain controls to compare against the tumor samples.

15 EXAMPLE 12

Subtypes of leukemia

[0159] Other subtypes of childhood leukemia defined by genetic aberrations including hyperdiploidy, normal karyotype and T-cell leukemias in the cohort are profiled by
20 Infinium. Results are validated by SEQUENOM.

[0160] The Infinium datasets are combined to truly identify pan-leukemia DNA methylation markers that could be used for MRD.

25 [0161] In addition, cases are profiled which had a poor outcome, poor response to therapy, died from disease and/or relapsed to identify prognostic markers of relapse (in addition to MRD) in paediatric leukemia.

- 94 -

EXAMPLE 13

Cancer types

5 [0162] DNA Methylation is profiled by Illumina Infinium from samples from other adult leukemia including AML and MDS and compared back to childhood leukemia dataset.

[0163] Methylation markers identified in TEL/AML dataset across model cancer cell lines of various tissue types including leukemia, colon cancer, breast cancer and the like are screened. Non-cancer cell lines are used as controls. Primary samples from patients with
10 other cancers are then screened.

EXAMPLE 14

Functional Biology

15 [0164] CD34+ cells are cultures *in vitro* and allowed to differentiate towards the two lineages (T-cell and B-cell), cells are captured at different stages of development and DNA Methylation profiled by Illumina Infinium.

[0165] Cells with gene promoters identified here as being methylated are treated with
20 epigenetic modifying drugs including decitabine, zebularine in conjunction with other drugs currently used as well as those which modulate histone modifications to determine if these drugs can alter the aberrant DNA methylation.

[0166] DNA methylation markers are backtracked to identify markers present in samples
25 taken at birth.

EXAMPLE 15***Diagnostic capabilities of DNA methylation******Methylation diagnostics***

5 [0167] To establish the utility of DNA methylation signatures as diagnostic targets, DNA from patient bone marrow aspirates throughout their therapeutic regime are initially bisulfite converted to establish if DNA methylation targets can correctly track disease within these patients. DNA methylation targets have potential in diagnosis and disease monitoring especially within patients, such as those with normal karyotype or which lack
10 PCR amplifiable targets which are patient specific, thus requiring a large amount of work up. TLX3 and FOXE3 are defined and validated to be exclusively hypermethylated in ALL. Methylation of these genes is associated with poor outcome by normal karyotype patients.

15 [0168] Results indicate DNA methylation can diagnose and track disease throughout therapy.

[0169] Completely methylated and unmethylated artificial oligonucleotides are created, designed from the TLX3 and FOXE3 sequences. The SEQUENOM EpiTYPER
20 methylation results for the standards were comparable to the restriction digest and Liquid Chromatography Mass Spec (LC-MS) results from NMI. Gravimetric dilution (accurate dilution method where samples are weighed between each dilution) of the standards highlighted the analytical sensitivity of the EpiTYPER cleavage chemistry (approximately 5%).

25 [0170] In an attempt to increase the analytical sensitivity of the MALDI-TOF system to make it applicable to disease detection such as MRD detection where current techniques are capable of detecting down to 1 in 10,000 cells, SEQUENOM iPLEX primer extension methodology was investigated. However, the approach was modified, biasing the
30 extension over the methylated template through incorporation of only three nucleotides, absent was the nucleotide required for extension of the unmethylated template (shown in

- 96 -

Figure 5). This is known as single allele base extension reaction (SABER), developed for SNP detection.

[0171] Using the SABER approach methylation was detected in 0.01% methylated template in an unmethylated background above that of completely unmethylated template representing an analytical sensitivity comparable to current MRD techniques.

EXAMPLE 16

Detection of leukemia in patients

10

[0172] An unsupervised clustering heatmap plot of the DNA methylation beta-values from the Illumina Infinium HumanMethylation27 BeadArray of 115 ALL-specific probes identified using three supervised learning methods (x-axis) is shown in Figure 6. Three distinct clusters comprising the leukemia, remission/non-leukemic and the cell lines are apparent (y-axis). The majority of these probes are hypermethylated in all leukemic samples analyzed. These clusters remained when all 14,876 probes, retained after p-detection cutoff, were taken into account (Figure 7).

[0173] A heatmap plot of SEQUENOM EpiTYPER DNA methylation results generated from 85 cases of B-Cell ALL with matching leukemic and remission bone marrow samples, controls (DONOR) and cancer cell line (REH) is shown in Figure 7. DNA methylation data for a total of 103 CpG sites encompassing 16 probes were selected by Infinium analysis are shown here (x-axis). DNA hypermethylation of 77/85 (91%) leukemic bone marrow samples was observed regardless of ALL subtype confirming the existence of a DNA methylation signature associated with leukemia.

[0174] Figure 8 shows that ALL is associated with an increase in average promoter methylation. Average global promoter DNA methylation levels of bone marrow according to sample group. The overall average b-value of probes that passed stringent quality control for all samples within a sample group are displayed.

30

- 97 -

[0175] Unsupervised clustering of Infinium β -values of 14,876 probes accurately delineating disease free tissue from leukemia is shown in Figure 9. Heatmap plot of unsupervised hierarchical clustering of the beta-values of 14,876 probes passing stringent quality control.

5

[0176] Figure 10 is a representation showing the performance and the average area under the receiver operating characteristics (ROC) curve (AROC) curves of the supervised learning methods applied to the data for DNA methylation profiling. Centroid, RFE-SVM and LIMMA plots are depicted (Figures 10a through c, respectively). The average accuracy of classification (accuracy) and the average area under the receiver operating characteristics (ROC) curve (AROC) are plotted against the number of features included in the classification. Accuracy and aroc reached 100% after 16 features for the Centroid method. Whilst the same accuracy and AROC level was achieved in less than 4 features using RFE-SVM and LIMMA.

15

[0177] Unsupervised hierarchical clustering heatmap of SEQUENOM EpiTYPER results is shown in Figure 11. A total of 163 paediatric leukemia cases were analyzed at three loci, *Corf76*, *FBXO39* and *MYOD1*, for DNA methylation. Matching leukaemic and remission bone marrow samples were analyzed from each case. Leukemia subtype for each diagnosed case is also depicted and illustrates a hypermethylation signature associated with leukaemic bone marrow.

[0178] The results from the modified MALDI-TOF technique is shown in Figure 12. Methylation is a measure of the of amount of extension primer that has been extended (extension occurring on methylated template) over the total extension primer input into each reaction.

25

- 98 -

EXAMPLE 17

Expression association and demethylating agent interrogation

[0179] Bone marrow is processed from ALL diagnosed patients forming a cohort. These
5 cells are stored viably and create the basis for cell line culturing, and subsequent
expression interrogation and demethylating agent analysis.

[0180] Those skilled in the art will appreciate that aspects of aspects described herein are
susceptible to variations and modifications other than those specifically described. It is to
10 be understood that these aspects include all such variations and modifications. These
aspects also include all of the steps, features, compositions and compounds referred to or
indicated in this specification, individually or collectively, and any and all combinations of
any two or more of the steps or features.

BIBLIOGRAPHY

Bernhard and Smola, *Learning with kernels* MIT Press, 2002

Bibikova *et al.*, *Genome Research* 16:383-393, 2006

Dahl *et al.*, *Clin Chem* 53(4):790-793, 2007

De Hertogh *et al.*, *BMC Bioinformatics* 11:17, 2020 (PMID 20064233)

Ding *et al.*, *J. Invest. Dermatol.* 123(3):470-473, 2004

Fahy *et al.*, *PCR Methods Appl.* 1(1):25-33, 1991

Flushberg *et al.*, *Nat Methods* 7(6):461-465, 2010

Fraga and Estella, *Biotechniques* 33:632-649, 2002

Gitan *et al.*, *Genome Res.* 12(1):158-164, 2002

Guyon *et al.*, *Machine Learning* 46:389-4122, 2002

Hansen *et al.*, *Nature Genetics* 43(8):768-775, 2011

Hastie *et al.*, *The Elements of Statistical Learning* Springer, 2001

Irizarry *et al.*, *Nature Genetics* 41(2):178-186, 2009

Laird, *Nature Reviews Genetics* 11:191-203, 2010

Nygren *et al.*, *Nucleic Acids Res.* 33(14):e128, 2005

- 100 -

Rein *et al.*, *Nucleic Acids Res.* 26:2255, 1998

Tost *et al.*, *Nucleic Acids Res* 31(9):e50, 2003

Wojdacz *et al.*, *Nucleic Acids Res.* 35(6):e41, 2007

Varewyck and Martens, *IEEE Trans. Syusk. Man. Cybern B Cybern* 41(2):330-340, 2011
(PMID 20699214)

Yegnasubramanian *et al.*, *Nucleic Acids Res.* 34(3):e19, 2006

CLAIMS:

1. A method for identifying an epigenetic profile in the genome of a cell in a subject indicative of a cancerous condition or a predisposition thereto, said method comprising screening for a change relative to a control in the epigenetic profile within or proximal to a locus selected from one or more loci associated with (i) cell fate commitment; (ii) transcription factor activity; (iii) DNA binding; (iv) subcellular location; and (v) transcription regulator activity or in 1 to 16 regions within one or more of these loci wherein a change in epigenetic profile in or proximal to one or more loci or regions therein is indicative of the presence of a cancerous condition or a propensity to develop same.
2. The method of Claim 1 wherein the loci are selected from those listed in Table 2.
3. The method of Claim 1 wherein the loci are selected from those listed in Table 3.
4. The method of Claim 1 wherein the epigenetic profile is extent of methylation or distribution of methylation sites.
5. The method of Claim 1 wherein the number of loci or regions within loci selected is determined by the percentage level of confidence required to determine which subject has or does not have cancer.
6. The method of Claim 5 wherein if a 100% confidence level is required, the epigenetic profile is determined in 16 different loci or in 16 regions within one or more loci.
7. The method of Claim 1 wherein the cancerous condition is a hematological cancer.
8. The method of Claim 1 wherein the cancerous conditions is a leukemia or sub-type thereof.

- 102 -

9. The method of Claim 8 wherein the leukemia is an early age onset leukemia.
10. The method of Claim 8 wherein the leukemia is selected from acute lymphatic leukemia (ALL), acute myeloid leukemia (AML), chronic myeloid leukemia (CML), small lymphocytic lymphoma (SLL), chronic lymphocytic leukemia (CLL), acute monocytic leukemia (AMOL), Hodgkin's lymphoma and non-Hodgkin's lymphoma.
11. The method of any one of Claims 1 to 10 wherein the epigenetic profile is determined in DNA or RNA from a cell extract.
12. The method of any one of Claims 1 to 10 wherein the epigenetic profile is determined in DNA or RNA present in a sample selected from the list consisting of urine, pus, respiratory fluid, lymph fluid, feces, bile, saliva, sputum, semen, vaginal flow, cerebral spinal fluid, brain fluid, ascites, milk, secretions from the genitourinary tract and a lavage of a tissue or organ (e.g. a lung).
13. The method of any one of Claims 1 to 12 wherein the method identifies a malignant cell at a frequency down from one cell per 10^8 to one cell per 10^4 cells.
14. The method of Claim 13 wherein the frequency is the minimal residual disease (MRD) value.
15. The method of any one of Claims 1 to 14 wherein the epigenetic profile is used to monitor treatment of the cancerous condition.
16. A method for determining the epigenetic profile within the genome of a eukaryotic cell or group of cells, the method comprising obtaining a sample of genomic DNA or RNA from the cell or group of cells and subjecting the genomic DNA or RNA to bisulphite treatment and then primer-specific amplification within one or more regions of one or more loci listed in Table 2 or 3 and assaying for extent of methylation relative to a control.

- 103 -

17. The method of Claim 16 wherein the epigenetic profile is extent of methylation or distribution of methylation sites.
18. The method of Claim 16 wherein the epigenetic profile is extent of methylation or distribution of methylation sites.
19. The method of Claim 16 or 17 or 18 wherein the number of loci or regions within loci selected is determined by the percentage level of confidence required to determine which subject has or does not have cancer.
20. The method of Claim 16 wherein the cancerous condition is a hematological cancer.
21. The method of Claim 16 wherein the cancerous condition is a leukemia or sub-type thereof.
22. The method of Claim 21 wherein the leukemia is an early age onset leukemia.
23. The method of Claim 16 wherein the leukemia is ALL, AML, CLL, SLL, CML, AMOL, Hodgkin's lymphoma and non-Hodgkin's lymphoma.
24. The method of Claim 16 wherein the epigenetic profile is determined in DNA or RNA from a cell extract.
25. The method of Claim 16 wherein the epigenetic profile is determined in DNA or RNA present in a sample selected from the list consisting of urine, pus, respiratory fluid, lymph fluid, feces, bile, saliva, sputum, semen, vaginal flow, cerebral spinal fluid, brain fluid, ascites, milk, secretions from the genitourinary tract and a lavage of a tissue or organ (e.g. a lung).

- 104 -

26. The method of Claim 16 wherein the epigenetic profile is used to monitor treatment of the cancerous condition.

27. A kit for the use in the method of any one of Claims 1 to 26 comprising bisulphite treatment of genomic DNA and primers to amplify a site within one or more loci before or after bisulfite treatment listed in Table 2 or 3 for DNA methylation analysis.

1/14

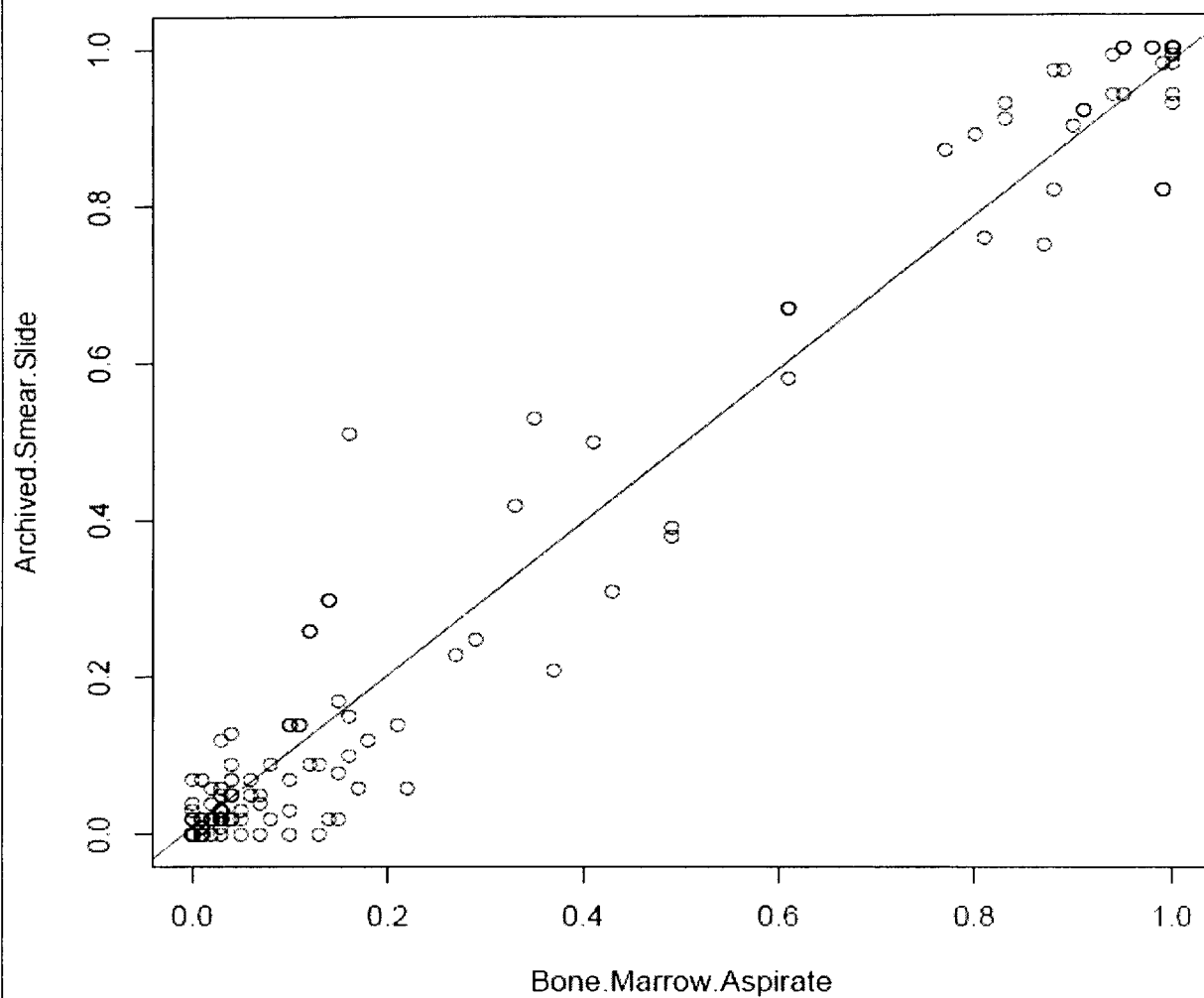
**Figure 1**

Figure 2

3/14

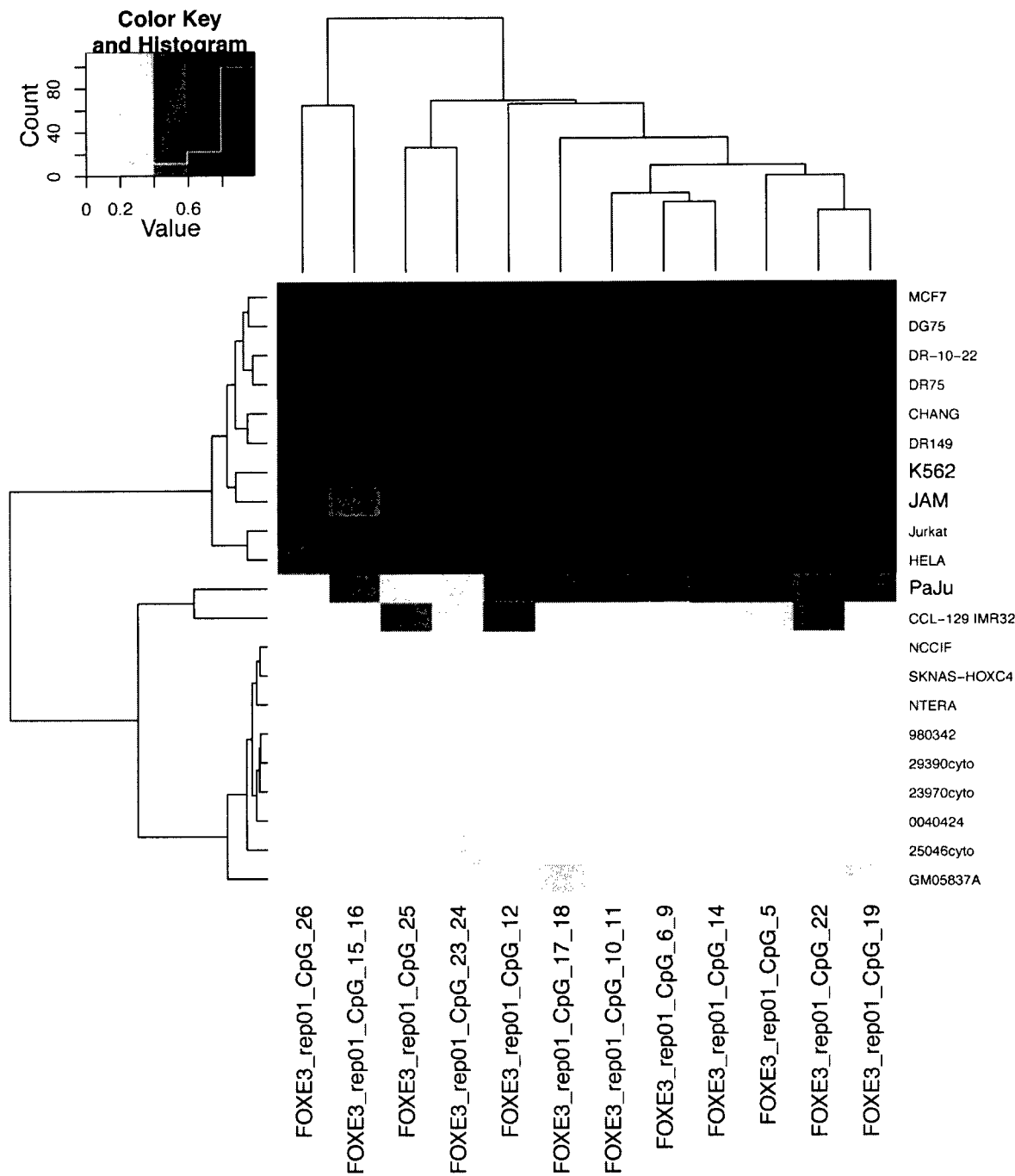


Figure 3

4/14

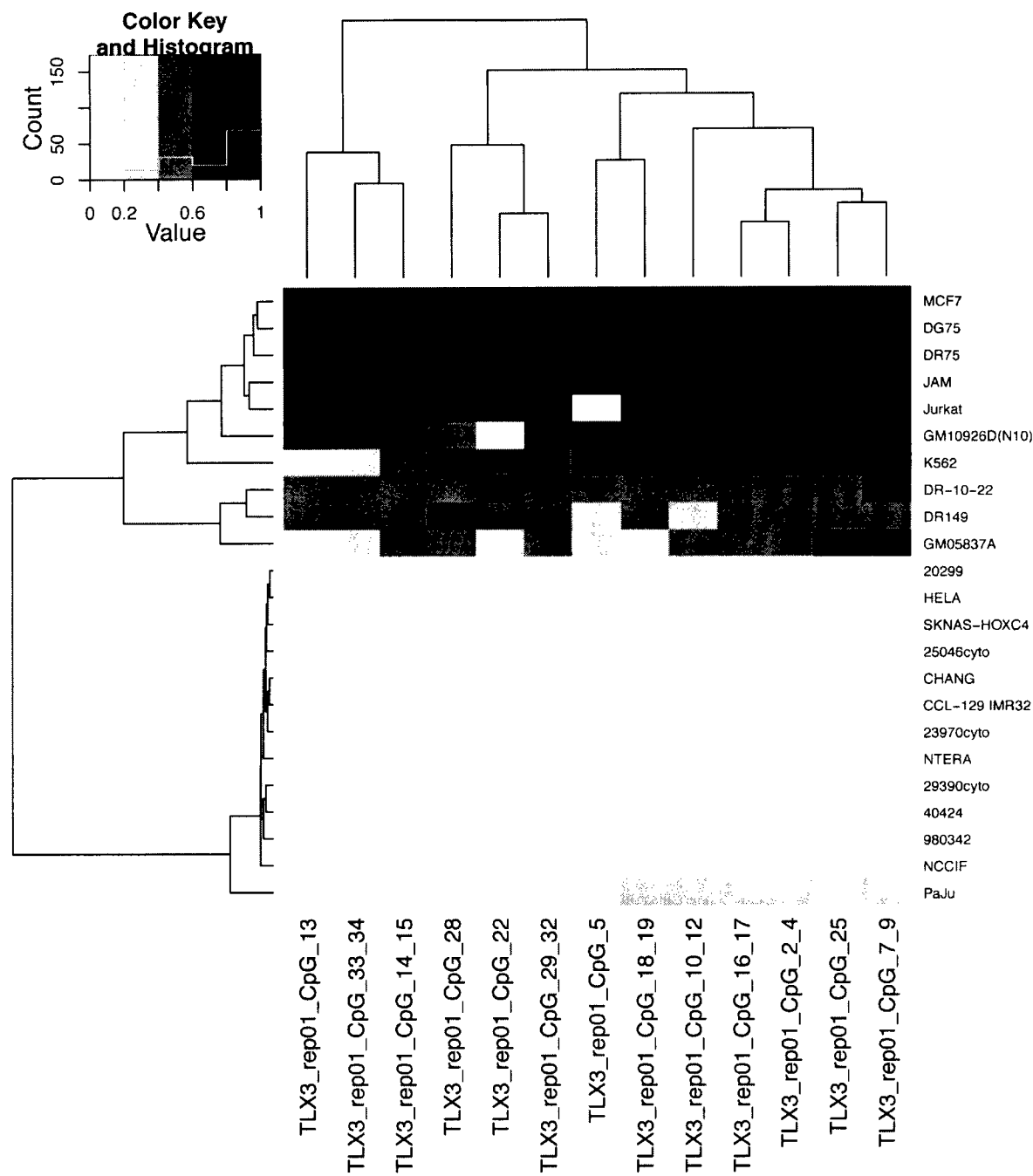


Figure 4

5/14

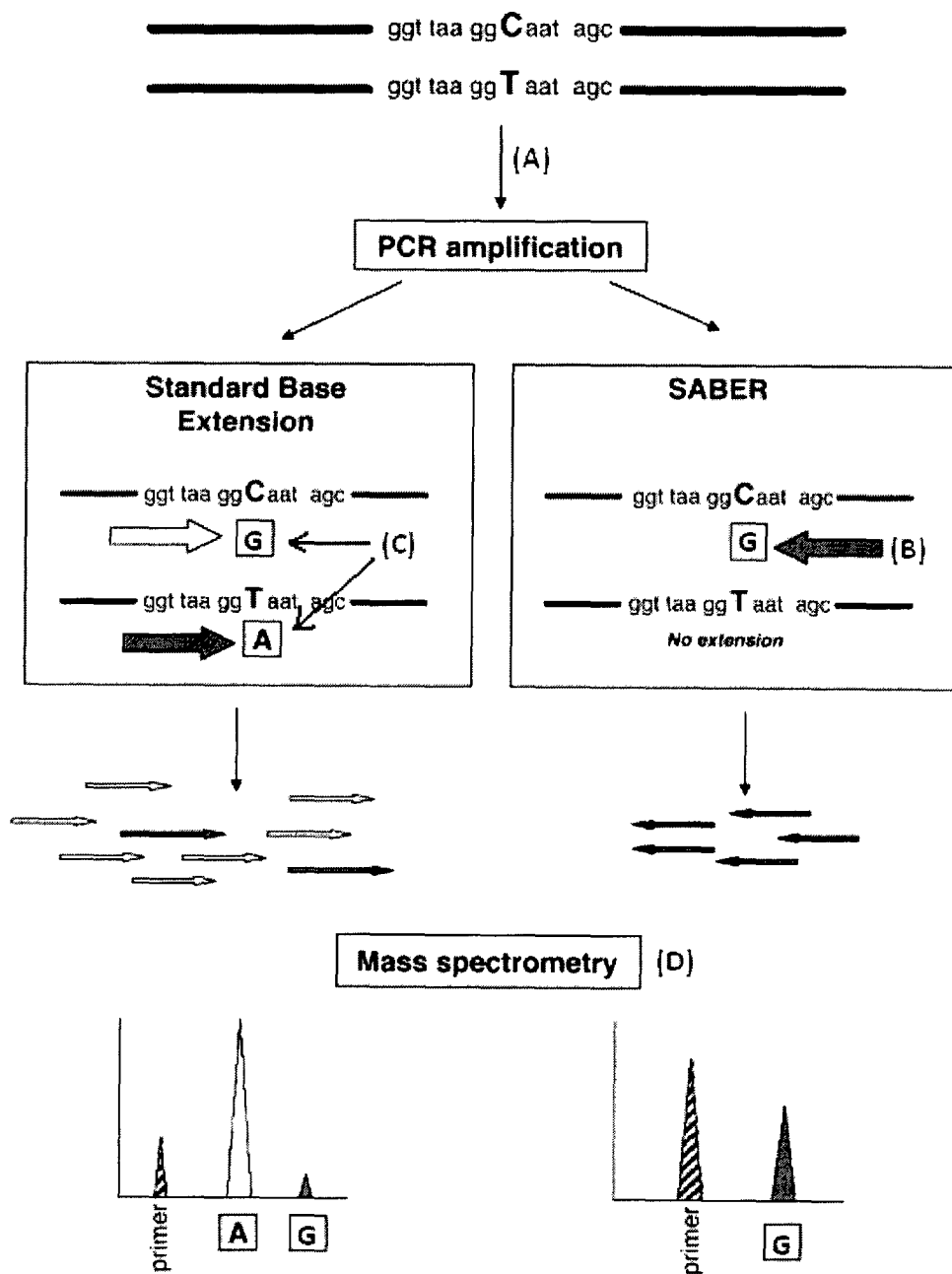


Figure 5

Figure 6

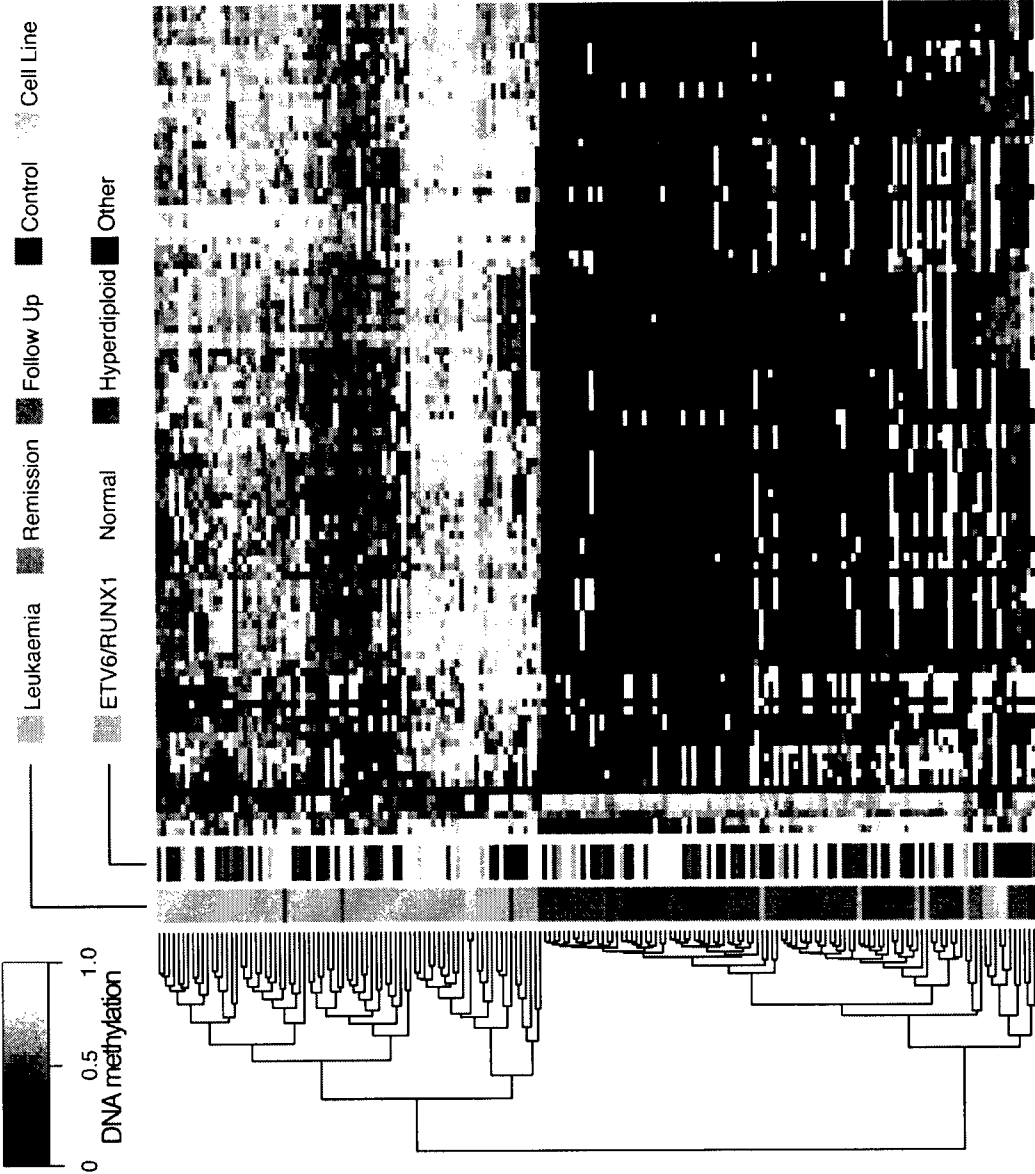


Figure 7

8/14

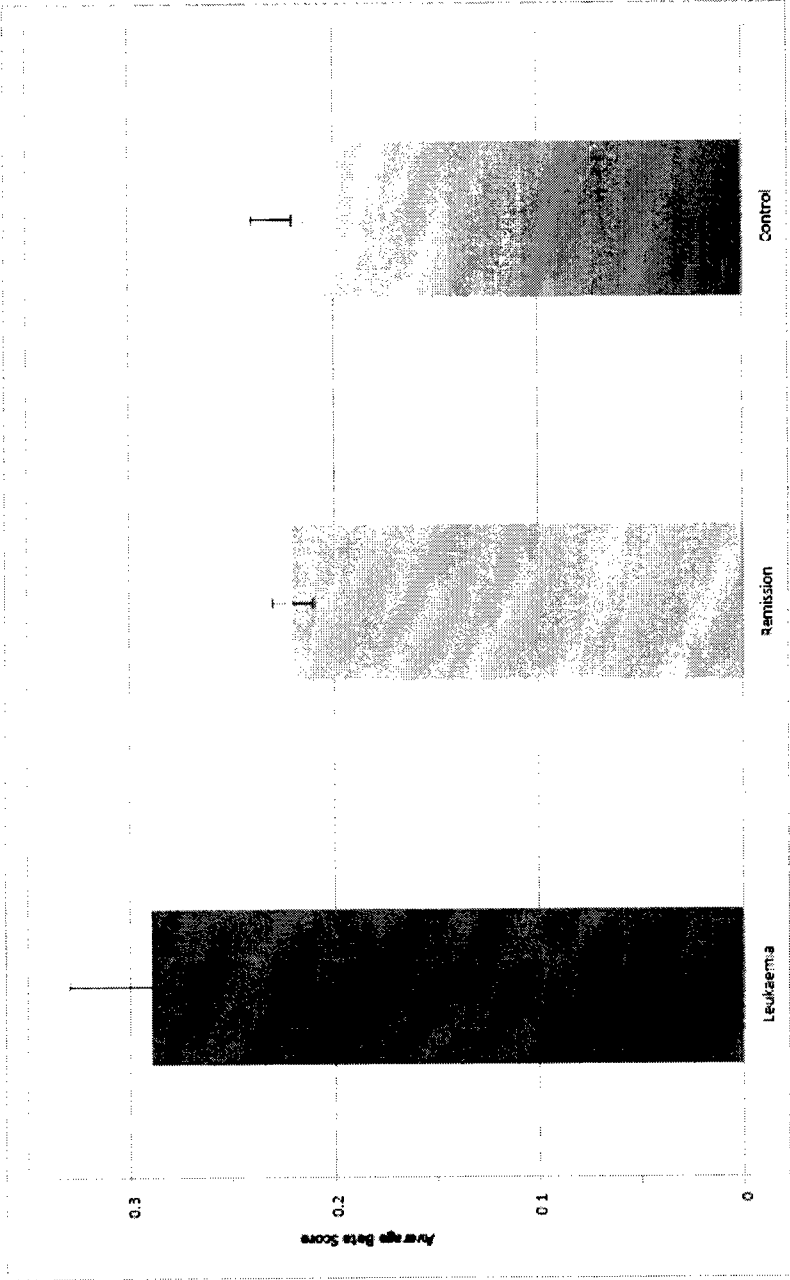


Figure 8

9/14

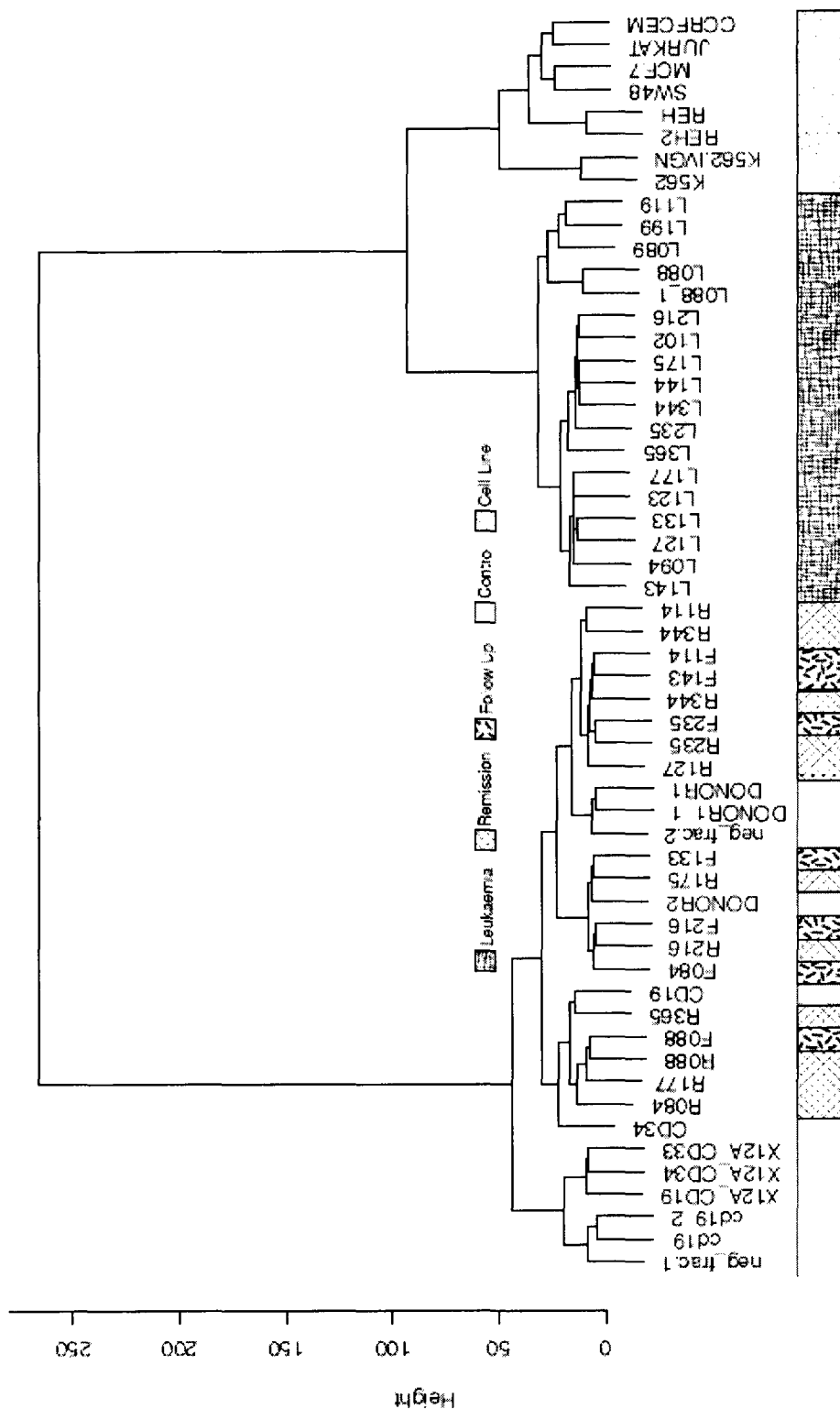


Figure 9

10/14

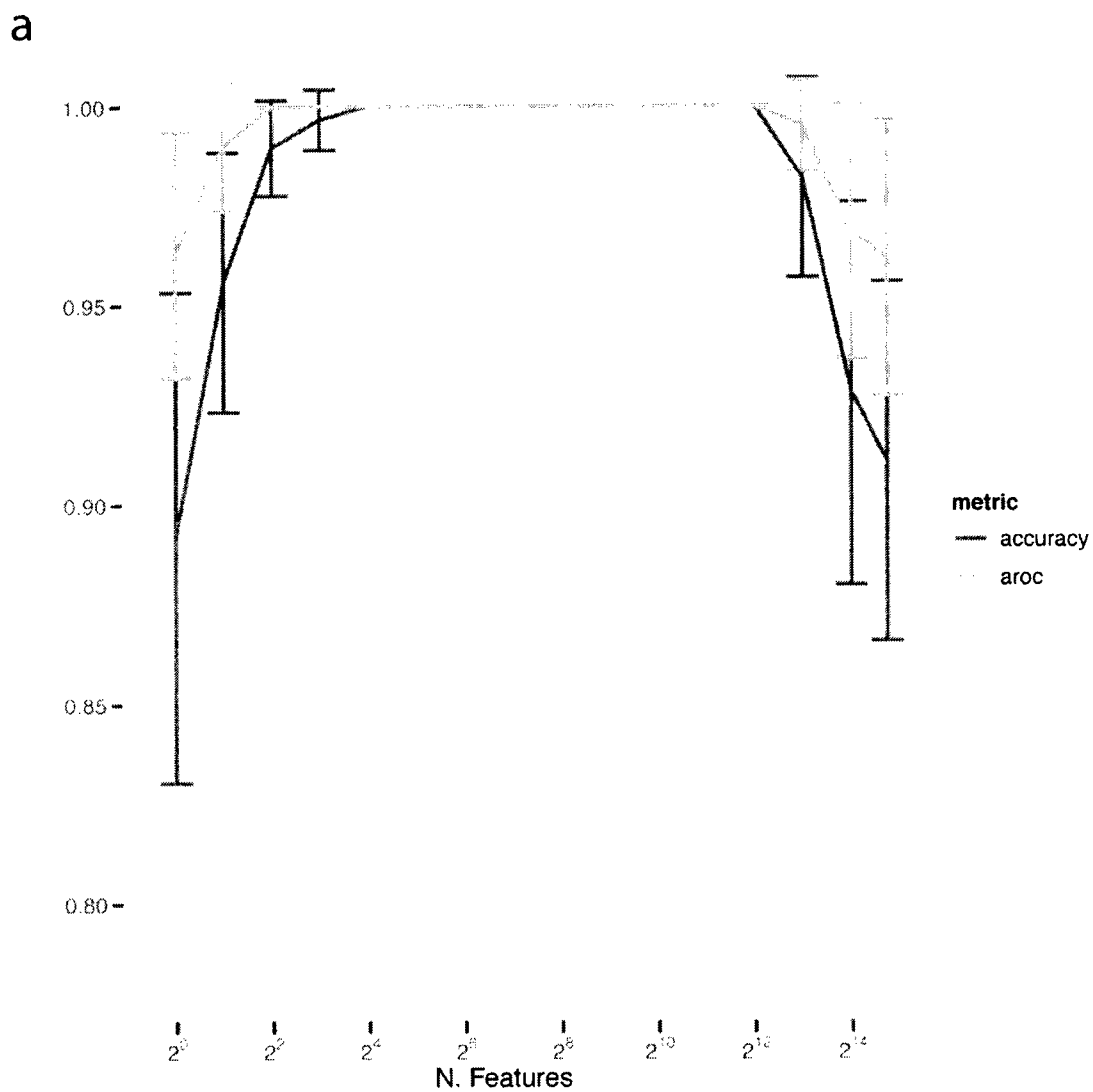


Figure 10a

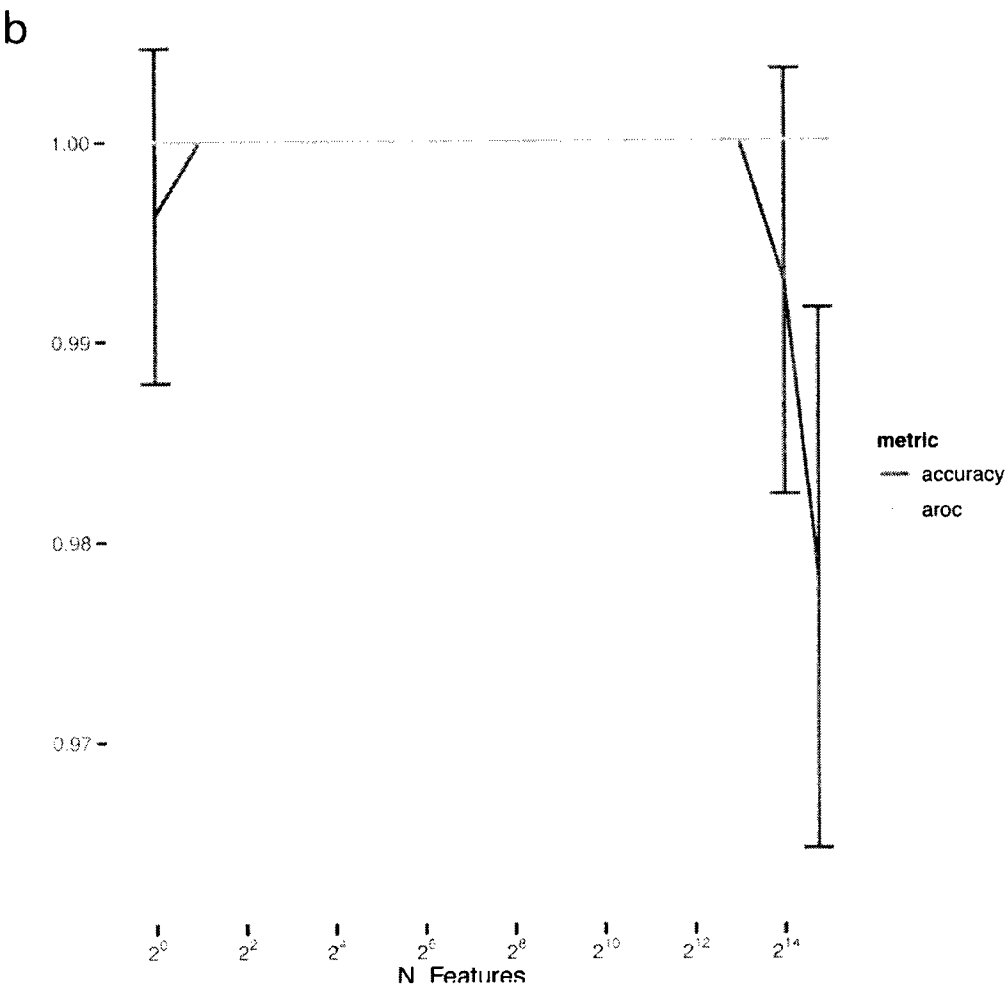


Figure 10b

12/14

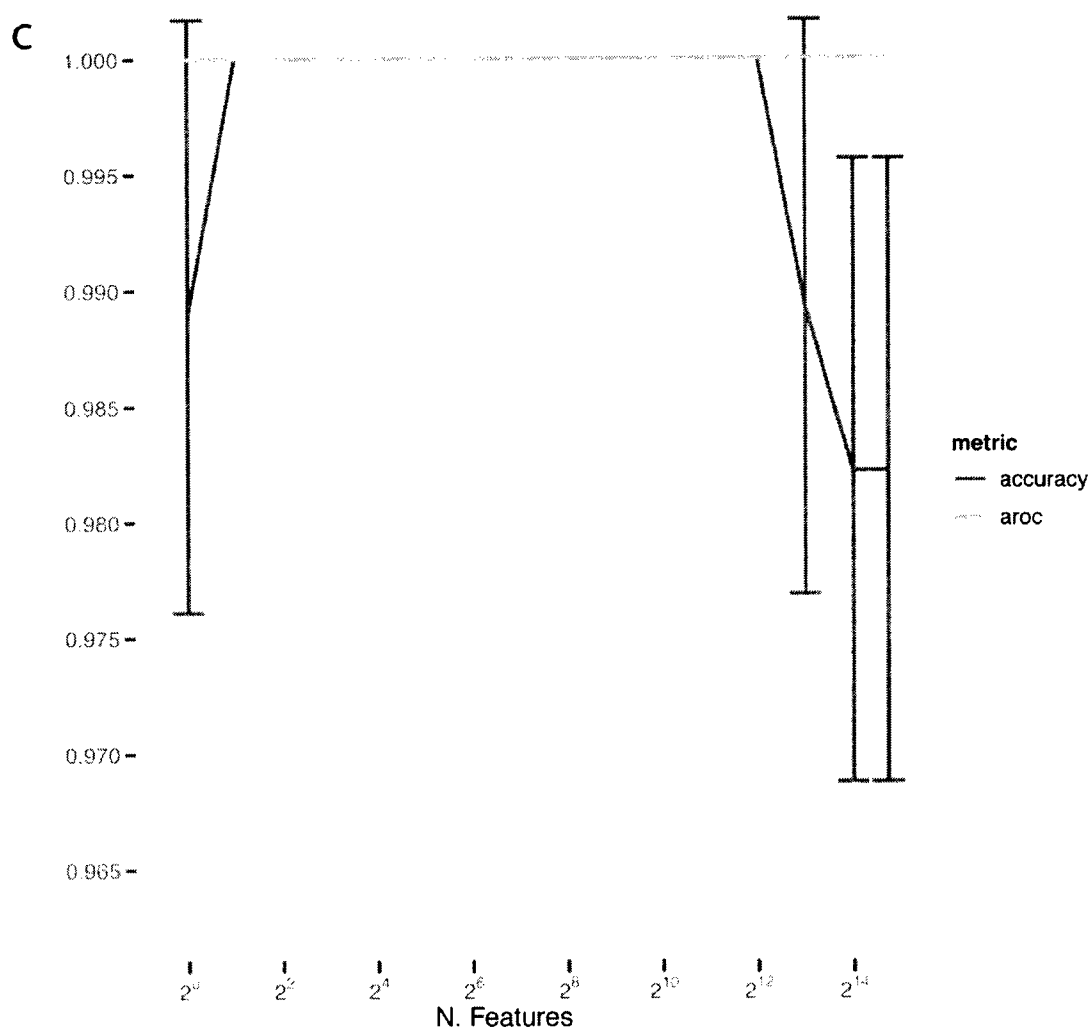


Figure 10c

13/14

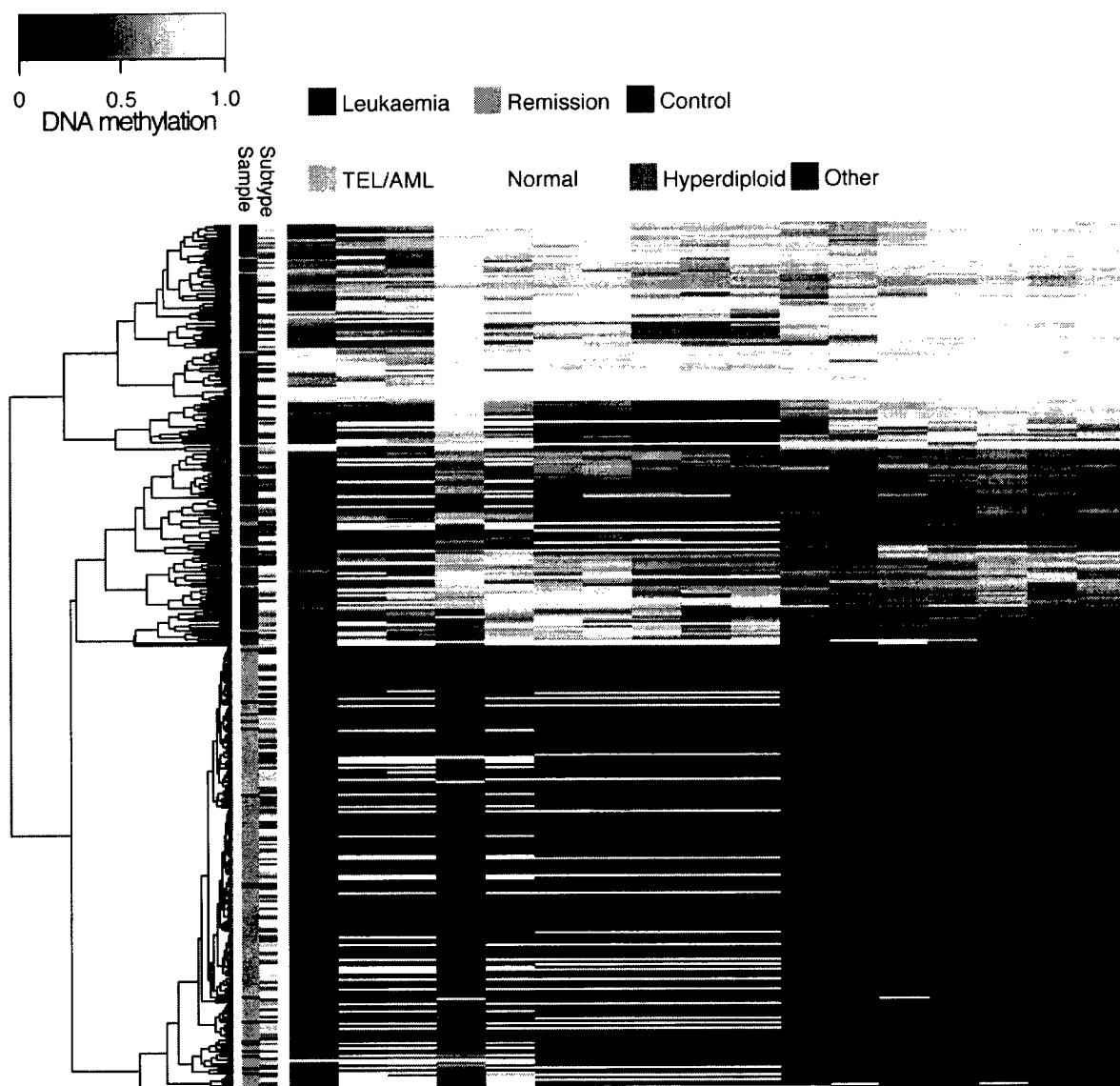
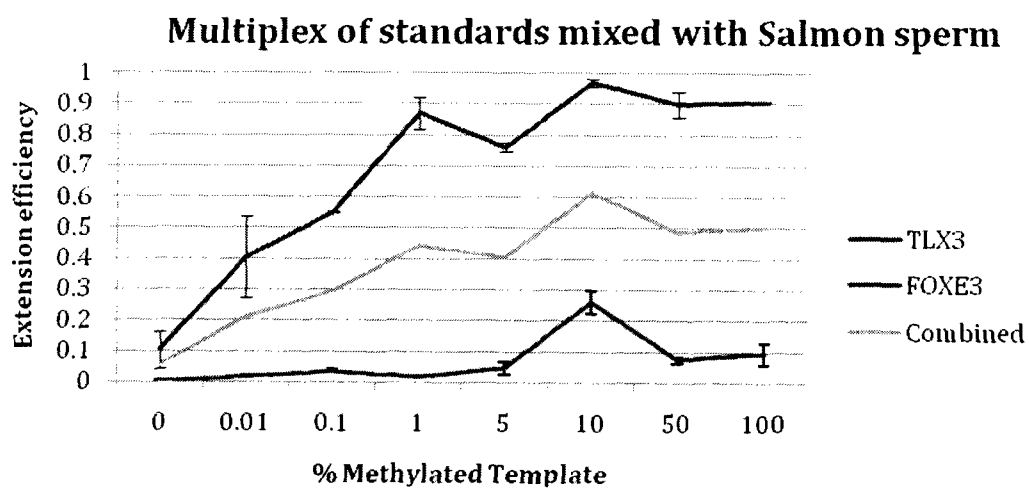


Figure 11

14/14

**Figure 12**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2011/001161

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12Q 1/68 (2006.01)

C40B 50/06 (2006.01)

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)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

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

WPI, EPODOC, MEDLINE, HCA, BIOSIS: epigenetic, methylate, CpG, hypermethylate, leukemia, lymphoma, hodgkin's, myeloma, cancer, blood, haematological, profile, diagnose, confidence, correlation, risk, sensitive and like terms

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2010/086388 A1 (AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH) 5 August 2010 See: abstract; pages 5-7; page 15, paragraph 2; page 42, paragraphs 2 and 4; pages 47-49	1-6, 11-15, 16-19 and 24-27
X	WO 2010/086389 A1 (AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH) 5 August 2010 See: abstract; table 1; claim 15; page 14, paragraph 3; page 15, paragraph 2; pages 51-52	1-5 and 7-27
X	ADORJÁN, P. et al. 'Tumour class prediction and discovery by microarray-based DNA methylation analysis' Nucleic Acid Research 2002, Vol. 30, e21 See: Abstract; table 1; page 1, column 2, paragraph 3; page 2, column 1, paragraph 2 and column 2, paragraph 2; page 5, column 2, paragraph 1	1-27



Further documents are listed in the continuation of Box C



See patent family annex

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
25 November 2011Date of mailing of the international search report
02/12/2011Name and mailing address of the ISA/AU
AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustalia.gov.au
Facsimile No. +61 2 6283 7999Authorized officer
ELLA JUNGERTH
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : +61 3 9935 9641

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2011/001161

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

a. (means)

☐ on paper

☐ in electronic form

b. (time)

☐ in the international application as filed

☐ together with the international application in electronic form

☐ subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

There was a sequence listing originally filed but it was not used for the purposes of this search and opinion.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2011/001161

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

[See Supplemental Box]

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2011/001161

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MILANI, L. et al. 'DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia' Blood 2010, Vol. 115, pages 1214-1225 See: abstract; figure 2; page 1215, column 1, paragraphs 2-5; page 1217, column 2, paragraph 2; page 1221, column 1	1-5 and 7-27
X	US 2009/0005268 A1 (BERLIN) 1 January 2009 See: abstract; tables 1 and 2; [0040]; [0184]; [0243]-[0247]; [0260]-[0268]	1-5 and 7-27

Supplemental Box

(To be used when the space in any of Boxes I to IV is not sufficient)

Continuation of Box No: III

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

This Authority has found that there are different inventions based on the following features that separate the claims into distinct groups:

- Claims 1-27 (all in part). The feature of a method for identifying an epigenetic profile in the genome of a cell in a subject indicative of a cancerous condition, in which the expression of MYOD1 is compared relative to a control is specific to this group of claims.
- Claims 1-27 (all in part). The feature of a method for identifying an epigenetic profile in the genome of a cell in a subject indicative of a cancerous condition, in which the expression of each of the remaining 123 loci listed in tables 2 and 3 is compared relative to a control is specific to each of these groups of claims.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

When there is no special technical feature common to all the claimed inventions there is no unity of invention. In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. The attorney has stated that it is the identification of 16 genes which provides 100% confidence that a particular subject has or does not have cancer and provides a technical relationship among the claims. However this feature does not make a contribution over the prior art because it is disclosed in:

WO 2010/086388 A1 (AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH) 5 August 2010

This document discloses that using 16 methylation markers, they were able to achieve 100% correct classification between tumour and normal tissue samples (page 49).

Therefore in the light of this document this common feature cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a posteriori*.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU2011/001161

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report				Patent Family Member			
WO	2010086388	CA	2750978	CA	2750979	EP	2233590
		WO	2010086389				
US	2009005268	AU	2006271906	CA	2615858	EP	1769086
		EP	1904649	EP	1934369	EP	1961827
		EP	2281902	EP	2298932	EP	2302073
		JP	2008506407	US	2009191548	US	2009298054
		US	2011269126	WO	2006008128	WO	2007009755
		WO	2007039234				
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.							
END OF ANNEX							