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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 August 2020 (13.08.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/163631 A1

(51) International Patent Classification:
G01N 33/50 (2006.01) *C12N 15/113* (2010.01)
C12Q 1/6886 (2018.01) *G16H 50/20* (2018.01)

(21) International Application Number:
PCT/US2020/017062

(22) International Filing Date:
06 February 2020 (06.02.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/802,620 07 February 2019 (07.02.2019) US

(71) Applicant: **UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INCORPORATED** [US/US]; 223 Grinter Hall, Gainesville, Florida 32611-5500 (US).

(72) Inventors: **TRAN, David**; 2230 NW 104th Way, Gainesville, Florida 32606 (US). **LE, Son Bang**; PO Box 100265, Gainesville, Florida 32610 (US).

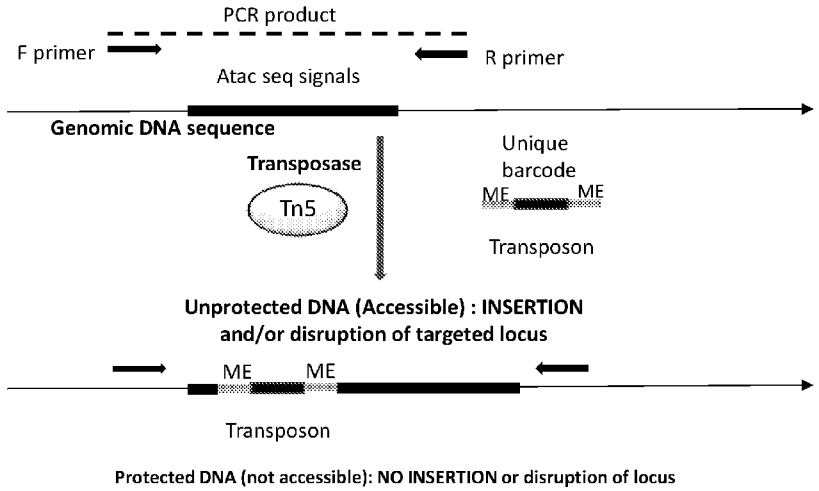
(74) Agent: **CHANG, Cindy** et al.; Alston & Bird LLP, 101 South Tryon Street, Bank of America Plaza, Suite 4000, Charlotte, North Carolina 28280-4000 (US).

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

WO 2020/163631 A1

(54) **Title:** METHODS FOR CANCER SCREENING AND MONITORING BY CANCER MASTER REGULATORS MARKERS IN LIQUID BIOPSY

FIG. 1



(57) **Abstract:** Provided herein are methods of assessing, detecting, monitoring the presence, or monitoring progression of cancer in a subject, or assessing or predicting prognosis or survival of a subject having cancer. Also provided are methods of measuring chromosomal accessibility of the chromosomal locus of cancer master regulator genes or downstream genes in the master regulator network, methods of measuring chromosomal DNA methylation at the chromosomal locus of cancer master regulator genes or downstream genes, and methods measuring chromosomal DNA methylation at the chromosomal locus of the promoter and/or regulatory regions of cancer master regulator genes or downstream genes.

[Continued on next page]

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

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

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

METHODS FOR CANCER SCREENING AND MONITORING BY CANCER MASTER REGULATORS MARKERS IN LIQUID BIOPSY

BACKGROUND

[0001] Current cancer screening using liquid biopsy relies on massive deep genomic sequencing to detect rare cancer-cell-derived genetic materials in bodily fluids. This process is costly and fraught with high false negative (FN) and false positive (FP) rates. The high FN rate is due to rare mutations accounting for a sizable share of cancer patients. The high FP rate is due to low penetrance of many cancer-associated mutations. Early detection of cancer in patients without obvious clinical evidence of cancer is important, and there remains a need for easy methods of cancer screening in blood.

SUMMARY

[0002] Provided are methods of assessing, detecting, monitoring the presence, or monitoring progression of cancer in a subject, or assessing or predicting prognosis or survival of a subject having cancer. In some embodiments, the methods comprise measuring chromosomal accessibility of the chromosomal locus or one or more cancer master regulator (master regulator) genes or one or more gene downstream of a master regulator in the master regulator network (downstream gene). In some embodiments, the methods comprise measuring chromosomal DNA methylation at the chromosomal locus of one or more master regulator genes or one or more downstream genes. In some embodiments, the methods comprise measuring chromosomal DNA methylation at the chromosomal locus of the promoter and/or regulatory regions of one or more master regulator genes or one or more downstream genes In some embodiments, the methods comprise measuring differential expression of one or more downstream genes.

[0003] Measuring chromosomal accessibility or chromosomal DNA methylation can comprise measuring the chromosomal accessibility or chromosomal DNA methylation of a master regulator gene or a gene downstream of the master regulator in the master regulator network in a sample from a subject and comparing the chromosomal accessibility or chromosomal DNA methylation with the chromosomal accessibility or chromosomal DNA methylation of a corresponding gene in a healthy reference sample. Similarly, measuring differential expression of a downstream gene can comprise measuring expression of the downstream gene in a sample from a subject and comparing the expression with the expression of a corresponding gene in a healthy reference sample. An

increase in the level of chromosomal accessibility of the master regulator, a decrease in DNA methylation of the master regulator, or differential (increased or decreased) expression of the downstream gene in the subject relative to the level of chromosomal accessibility, DNA methylation, or expression of the corresponding master regulator or downstream gene in the healthy reference sample indicates the possible presence of cancer in the subject, an increase or risk of increase in cancer progression in the subject, a possible increased risk of developing cancer in the subject, a poor prognosis, or decreased predicted survival time for the subject. In some embodiments, chromosomal accessibility or chromosomal DNA methylation is measure is the promotor or regulatory region of a master regulator gene or a gene downstream of the master regulator in the master regulator network. The methods can be used to guide or suggest treatments or changes in treatment of a subject.

[0004] In some embodiments, the methods are used to assess whether a subject has a poor survival prognosis for cancer comprising: analyzing the chromosomal accessibility or DNA methylation of least one master regulator in a sample from the subject, wherein an increase in the chromosomal accessibility or DNA methylation level of the at least one master regulator relative to the chromosomal accessibility or DNA methylation level of the corresponding master regulator in a healthy reference sample, is indicative that the subject has a poor survival prognosis for the cancer. The sample can be, but is not limited to, a liquid sample. The liquid sample can be, but is not limited to, a blood sample. Cancer cells, including cancer stem cells, can leave the primary tumor and spread. The methods described herein can be used to detect these migrating cancer cells in blood samples. The cancer can be, but is not limited to, glioblastoma (GBM) and glioblastoma-related cancers. Exemplary cancer master regulators are provided in FIG. 4-7.

[0005] In some embodiments, the methods are used to assess whether a subject has a poor survival prognosis for cancer comprising: analyzing expression of least one factor downstream in the master regulator network (downstream factor) in a sample from the subject, wherein an increase in the expression level of the at least one downstream factor relative to the expression level of the corresponding downstream factor in a healthy reference sample, is indicative that the subject has a poor survival prognosis for the cancer. The sample can be, but is not limited to, a liquid sample. The liquid sample can be, but is not limited to, a blood sample. The cancer can be, but is not limited to, glioblastoma and glioblastoma-related cancers. The downstream factor can

be a downstream factor differentially expressed in glioblastoma stem cells (GSCs). Exemplary downstream factors are provided in Example 4 and Table 1.

[0006] In some embodiments, one or more cancer therapies is administered to a subject identified as having a poor prognosis or decreased survival time.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S)

[0007] Having thus described the invention in general terms, reference will now be made to the accompanying drawings, which are not necessarily drawn to scale.

[0008] FIG. 1 illustrates the ATAC-seq (Assay for Transposase-Accessible Chromatin using sequence) approach to detect chromosomal accessibility at specific chromosomal locations. Presence or absence of insertion and/or disruption of a gene locus can be detected by loss of the locus as measured by reduced PCR amplification of the locus, or by Sanger sequencing of PCR amplified locus.

[0009] FIG. 2 illustrates the ATAC-seq (Assay for Transposase-Accessible Chromatin using sequence) approach to detect chromosomal accessibility at specific chromosomal locations. The transposon contains a phosphate group at the end of the transposon. Presence or absence of insertion and/or disruption of a gene locus can be detected by loss of the locus as measured by reduced PCR amplification of the locus, or by Sanger sequencing of PCR amplified locus.

[0010] FIG. 3 illustrates the top master regulators that are differentially expressed in glioblastoma stem cells (GSCs) v. astrocytes.

[0011] FIG. 4 illustrates the OTP methylation profile and ATAC-seq signals in GSCs and peripheral blood mononuclear cells (PBMCs). Examples of the correlation between methylation and accessibility: Methylated CpG islands = Gene not expressed = Not accessible, and vice versa.

[0012] FIG. 5 illustrates the OLIG2 methylation profile and ATAC-seq signals in GSCs and peripheral blood mononuclear cells (PBMCs). Examples of the correlation between methylation and accessibility: Methylated CpG islands = Gene not expressed = Not accessible, and vice versa.

[0013] FIG. 6 illustrates the BATF2 methylation profile and ATAC-seq signals in GSCs and peripheral blood mononuclear cells (PBMCs). Examples of the correlation between methylation and accessibility: Methylated CpG islands = Gene not expressed = Not accessible, and vice versa.

[0014] FIG. 7 illustrates the NKX2-2 methylation profile and ATAC-seq signals in GSCs and peripheral blood mononuclear cells (PBMCs). Examples of the correlation between methylation and accessibility: Methylated CpG islands = Gene not expressed = Not accessible, and vice versa.

[0015] FIG. 8 illustrates PCR products of nuclei preparations treated with transposon and transposase. Accessible regions of indicated GSC master regulator genes are disrupted by transposase specifically in GSCs but not PBMCs. The GAPDH locus was used as equal input control after transposase treatment.

[0016] FIG. 9 illustrates the results from using the gene expression approach on downstream target ID4.

[0017] FIG. 10 illustrates the results from using the gene expression approach on downstream target FREM2.

[0018] FIG. 11 illustrates the results from using the gene expression approach on downstream target NES.

[0019] FIG. 12 illustrates the results from using the gene expression approach on downstream target SALL1.

DEFINITIONS

[0020] A "sample" comprises any tissue or material isolated from a subject, such as a patient. The sample may contain cellular and/or non-cellular material from the subject, and may contain any biological material suitable for detecting a desired biomarker, such a DNA or RNA. The sample can be isolated from any suitable biological tissue or fluid such as, but not limited to, a tissue or blood. A sample may be treated physically, chemically, and/or mechanically to disrupt tissue or cell structure, thus releasing intracellular components into a solution which may further contain enzymes, buffers, salts, detergents and the like, which are used to prepare the sample for analysis.

[0021] A "master regulator" or "cancer master regulator" is a gene or protein that acts to drive one or more intermediary gene or proteins in a pathway or network important in initiating or maintaining a cancerous state or initiating or maintaining one or more deleterious cancerous behaviors. Some master regulators are involved in pathways in the transition to a cancer state. Some master regulators are involved in pathways of aggressive (bad) cancer behavior. Expression of master regulators is indicative of poor prognosis in subjects having cancer.

[0022] A "master regulator network" refers to a master regulator and one or more genes downstream of the master regulator whose transcription level is dependent on or affected by the master regulator.

[0023] Compositions or methods "comprising" or "including" one or more recited elements may include other elements not specifically recited. For example, a composition that "comprises" or "includes" a protein may contain the protein alone or in combination with other ingredients.

[0024] Designation of a range of values includes all integers within or defining the range, and all subranges defined by integers within the range.

[0025] Unless otherwise apparent from the context, the term "about" encompasses values within a standard margin of error of measurement (e.g., SEM) of a stated value or variations ± 0.5%, 1%, 5%, or 10% from a specified value.

[0026] The singular forms of the articles "a," "an," and "the" include plural references unless the context clearly dictates otherwise. For example, the term "an antigen" or "at least one antigen" can include a plurality of antigens, including mixtures thereof.

Statistically significant means $p \leq 0.05$.

DETAILED DESCRIPTION

[0027] Various embodiments of the inventions now will be described more fully hereinafter, in which some, but not all embodiments of the inventions are shown. Indeed, these inventions may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will satisfy applicable legal requirements. The term "or" is used herein in both the alternative and conjunctive sense, unless otherwise indicated. The terms "illustrative" and "exemplary" are used to be examples with no indication of quality level.

[0028] Provided are methods for screening for cancer in a subject.. The methods take advantage of the finding that cancer master regulators are often not mutated but used by cancer causative mutated genes to establish the cancer state. Many cancer master regulators are developmentally restricted and repressed epigenetically in normal adult tissues. In the cancer or pre-cancer state, the master regulator genes become unrestricted. Thus, measuring or determining the status of epigenetic markers in cancer master regulator genes can be used to detect cancer. Epigenetic markers include DNA methylation, chromosomal accessibility, and differential expression of factors downstream of a master regulator in the master regulator network.

[0029] In some embodiments, chromosomal accessibility can be measured either by disruption of a target gene locus or by insertion of exogenous DNA barcodes in the target gene locus. In

some embodiments, the target gene is a master regulator. In some embodiments, master regulators of a particular cancer are identified using GeneRep/nSCORE as described in WO2018/069891, which is incorporated by reference in its entirety.

[0030] In some embodiments, ATAC-seq can be used to measure chromosomal accessibility. In some embodiments, ATAC-seq, is used to digest hypomethylated and accessible regions in DNA present in a sample. In some embodiments, ATAC-seq, is used to insert exogenous DNA barcodes into accessible regions in DNA present in a sample. The level of digestion or insertion in the area of a target region can be measured using PCR and primers designed to amplify the target region DNA. In some embodiments, the target region is a region of a master regulator, such as, but not limited to, a promoter of a master regulator.

[0031] Determining or measuring DNA accessibility may be done using methods known in the art. Exemplary methods of determining or measuring DNA accessibility include, but are not limited to, ATAC-seq, CRISPR, DNAse-seq, and MNase-seq

[0032] Determining or measuring DNA methylation may be done using methods known in the art. Exemplary methods of determining or measuring DNA methylation include, but are not limited to, ATAC-seq, digestion based assay followed by PCR, methylation specific PCR, E-ice-COLD-PCR, bead array analysis, pyrosequencing, PCR with high resolution melting, and bisulfite sequencing.

[0033] In some embodiments, ATAC-seq can be used to detect GSC in patient blood. Using ATAC-seq, accessible regions in master regulators of GCS are identified.

[0034] Described are methods assessing, detecting, monitoring the presence, or monitoring progression of cancer in a subject, or assessing or predicting prognosis or survival of a subject having cancer. The methods comprise

a) obtaining or having obtained a sample from a subject

b) measuring or having measured the chromosomal accessibility level of at least one master regulator in the sample; and

c) comparing the chromosomal accessibility level with the chromosomal accessibility level of a corresponding master regulator gene in a healthy reference sample;

wherein an increase in the chromosomal accessibility level of the at least one master regulator in the subject relative to the chromosomal accessibility level of the corresponding master regulator in the healthy reference sample indicates the possible presence of cancer in the subject, an increase

or risk of increase in cancer progression in the subject, an increased risk of developing cancer in the subject, a poor prognosis, or decreased predicted survival time for the subject. The methods can be used to suggest treatments or changes in treatment of the subject. The sample can be, but is not limited to, a liquid sample. A liquid sample can be, but is not limited to, a blood sample.

[0035] Described are methods assessing, detecting, monitoring the presence, or monitoring progression of cancer in a subject, or assessing or predicting prognosis or survival of a subject having cancer. The methods comprise

a) obtaining or having obtained a sample from a subject

b) measuring or having measured the chromosomal DNA methylation level of at least one master regulator in the sample; and

c) comparing the chromosomal DNA methylation level with the chromosomal DNA methylation level of a corresponding master regulator gene in a healthy reference sample;

wherein a decrease in the chromosomal DNA methylation level of the at least one master regulator in the subject relative to the chromosomal DNA methylation level of the corresponding master regulator in the healthy reference sample indicates the possible presence of cancer in the subject, an increase or risk of increase in cancer progression in the subject, an increased risk of developing cancer in the subject, a poor prognosis, or decreased predicted survival time for the subject. The methods can be used to suggest treatments or changes in treatment of the subject. The sample can be, but is not limited to, a liquid sample. A liquid sample can be, but is not limited to, a blood sample. In some embodiments, chromosomal DNA methylation is measure in a promoter or regulator region of at least one master regulator gene.

[0036] Described are methods assessing, detecting, monitoring the presence, or monitoring progression of cancer in a subject, or assessing or predicting prognosis or survival of a subject having cancer. The methods comprise

a) obtaining or having obtained a sample from a subject

b) measuring or having measured the expression level of at least gene downstream of a master regulator in the master regulator network in the sample (downstream gene); and

c) comparing the expression level with the expression level of a corresponding downstream gene in a healthy reference sample;

wherein an increase or decrease in expression level of the at least one downstream gene in the subject relative to the expression level of the corresponding gene in the healthy reference sample

indicates the possible presence of cancer in the subject, an increase or risk of increase in cancer progression in the subject, an increased risk of developing cancer in the subject, a poor prognosis, or decreased predicted survival time for the subject. The methods can be used to suggest treatments or changes in treatment of the subject. The sample can be, but is not limited to, a liquid sample. A liquid sample can be, but is not limited to, a blood sample.

[0037] Methods of determining gene expression in a sample can be performed using methods known in the art. Such methods included, but are not limited to, nucleotide amplification assays (including but not limited to PCR, RT-PCR, serial analysis of gene expression, and differential display), microarray technologies, proteomics, HPLC, and Western electrophoresis.

[0038] In some embodiments, the methods are used to assess whether a subject has a decreased predicted survival time for cancer comprising: determining a level of chromosomal accessibility or chromosomal DNA methylation of a cancer master regulator in a sample from the subject, wherein an increase chromosomal accessibility level of the at least one master regulator relative to the chromosomal accessibility level of the corresponding master regulator in a healthy reference sample, or a decrease chromosomal DNA methylation level of the at least one master regulator relative to the chromosomal DNA methylation level of the corresponding master regulator in a healthy reference sample is indicative that the subject has a poor survival prognosis for the cancer.

[0039] In some embodiments, chromosomal accessibility or chromosomal DNA methylation levels of 2, 3, 4, 5, 6, 7, 8, 9, 10. 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 master regulators in a subject sample are measured and compared with the chromosomal accessibility or chromosomal DNA methylation level of the corresponding master regulators in a healthy reference (control) sample.

[0040] In some embodiments, the methods are used to assess whether a subject has a decreased predicted survival time for cancer comprising: determining an expression level of a least one gene downstream of a cancer master regulator in the cancer master regulator network (downstream gene) in a sample from the subject, wherein an change in expression level of the at least one downstream gene relative to the expression level of the corresponding gene in a healthy reference sample is indicative that the subject has a poor survival prognosis for the cancer.

[0041] In some embodiments, a change in expression levels of 2, 3, 4, 5, 6, 7, 8, 9, 10. 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 downstream gene in a subject sample are measured and compared

with the expression levels of the corresponding downstream genes in a healthy reference (control) sample.

[0042] In some embodiments, glioblastoma is detected in a patient by analyzing differential expression of 1, 2, 3, or 4 of ID4, FREM2, NES, and SALL1 in a blood sample.

[0043] In some embodiments, chromosomal accessibility for one or more master regulators and chromosomal DNA methylation for one or more master regulators is measured. The master regulators can be for the same cancer type. Chromosomal accessibility and chromosomal DNA methylation can be measured for the same master regulators, different master regulator or overlapping sets of master regulators.

[0044] In some embodiments, measurement of chromosomal accessibility and/or chromosomal DNA methylation for one or more master regulators is combined with measurement of differential expression of one or more downstream genes.

[0045] In some embodiments, master regulators are selected based on the cancer type.

[0046] Expression of master regulator genes in cancer drive bad cancer behavior or poor prognosis of the cancer. Poor prognosis can include, but is not limited to, poor response to typical cancer treatment, aggressive cancer growth, increased metastasis, and/or decreased survival time. Identification of poor prognosis in a patient can be used to diagnose and/or prescribe treatment. Such treatment can include, but is not limited to, master regulator-specific treatment and/or more aggressive treatment. Master regulator-specific treatment includes treatments, including adjuvants, known to be effective in treating similar cancers in other patients expressing the same master regulator gene(s). As an example, patients having increased expression of VDR or VDR-related genes may be given vitamin D.

EXAMPLES

Example 1: Cancer screening method: chromosomal accessibility approach

[0047] The regulatory chromosomal elements of master regulators of GSCs, especially those that are developmentally restricted, are accessible in GSCs and not accessible in adult normal cells. Developmentally restricted genes are highly methylated and thus highly coiled and folded in normal adult cells and are therefore not accessible to transcription factors so that they are not expressed haphazardly. In contrast, these master regulators in cancer cells are unmethylated and therefore accessible. Additionally, cancer stem cells tend to leave the primary tumor and spread

which leads to metastases. Therefore, migrating GSCs can be detected by measuring and amplifying the accessibility of these regulatory chromosomal elements among other normal adult cells, *e.g.*, blood.

[0048] Methods such as ATAC-seq (FIG. 1) or CRISPR in combination with qPCR or nanostring/multiplexing can be used to access and detect the unmethylated DNA of the master regulators of cancer in the blood of a subject. Additionally, technologies that can access DNA, cleave DNA, and insert a barcode can also be used.

[0049] FIGs. 4-7 show methylation profiles and ATAC-seq signals in GSCs and four master regulators of GSCs, OTP, OLIG2, BATF2, and NKX2-2. FIG. 8 shows PCR products of nuclei preparations treated with transposon and transposase and shows that accessible regions of indicated GSC master regulator genes are disrupted by transposase specifically in GSCs but not PBMCs. Thus, ATAT-seq mediated degradation of, disruption of, or insertions into OTP, OLIG2, BATF2, and/or NKX2-2 can be used in the diagnosis of GSC. Further, the ATAC-seq assay can be used to detect and diagnose GCS using blood samples.

[0050] The same approach can be used to detect cancer stem cells from other cancers by identifying the master regulators of cancer.

Example 2. Cancer screening method: chromosomal accessibility approach.

[0051] The data in example 1 was generated using a transposon without a phosphate (PO_4) at the end of the transposon. In the absence of the terminal PO_4 (FIG. 1) the transposase can destroy the target promoter loci since ligation can't proceed. Addition of the PO_4 at the end of the transposon (FIG. 2) permits testing of the insertion of the unique transposon specifically at the accessible regions in master regulator gene promoters. This results in a larger amplified product (either as a ladder or a smear) extending upward from the native product, using the primer F and R pair.

Example 3. Master regulators

[0052] Using genetic analyses, such as GeneRep/nSCORE, the top 5-20 genes in the GSC regulatory network with the highest differential expression between GSC and PBMC are identified. Analysis of master regulators is then combined with downstream factors, such as OTP, OLIG2, BATF2, and NKX2-2, to detect both upstream master regulators and downstream

pathways at the same time to simultaneously increase sensitivity and specificity for the screening technology.

Example 4: Cancer screening method: gene expression approach

[0053] Master regulators are expressed at a low level and therefore directly measuring the expression level of the master regulators in the blood can be unreliable. However, target genes (and their expression protein products) downstream of GSC master regulators amplify GBM state signals to maintain the GBM state. The expression of these downstream factors can reach hundreds of fold higher than in normal healthy adult cells and can be used to detect rare migratory GSCs, *e.g.*, in blood. The same approach can be used to detect cancer stem cells from other cancers by identifying the downstream genes.

[0054] We identified the following downstream factors with the highest levels of expression in GSCs compared to PBMCs: ENSG00000277459, SALL1, ID4, FOXG1, FJX1, FREM2, NES, TUBB2B, MSI1, FKBP10, CBARP, GNG12, CD276, FAT1, CTXN1, DPYSL5, TYRO3, GJC1, CHST3, LRP4, PHF21B, ADAMTS9, CSPG5, TEAD1, BCAR1, EML1, RBFOX2, MPDZ, MSX1, EFNB3, ENAH, LYPD6, GTF2IRD1, TBX2, ANK2, C1QL1, ZIC1, EMP2, TMEM132A, CX3CL1, SYDE1, SLC16A2, SOX9, RND3, LARP6, CNN3, SPRY4, DPF1, PCGF2, BOC, OBSL1, SOX2, EPB41L1, MEX3A, NCS1, SMO, TMC7, SEZ6L2, POU3F2, FAM171A2, DENND2A, TANC1, PROX1, ENSG00000268592, PNPLA3, TSKU, DLX1, KIAA1549, MTSS1L, PTPRF, IRX5, DZIP1, MAGI1, ADCY6, BCHE, CXADR, TEAD4, PTPN21, CDR2L, ADGRL3, REEP2, SHROOM3, ARC, EEF1A2, ETV4, EGFR, UCHL1, KAZALD1, TJP1, ENSG00000237004, ETV5, CKB, KCNF1, MAP1B, S100A16, COL27A1, VGF, ALDH7A1, GAS1, LOC101927480, CITED1, ETV1, NOVA1, JPH1, FBXO17, CNKSR3, PDXP, PLEKHH3, PYGO1, SCARA3, RTL8B, LAMB1, MYH14, CASKIN1, NLGN2, PACSIN3, CA12, ARHGAP39, LAMA4, ZNF462, NR2F1, CSRP2, ARHGEF25, GLI3, TMEFF1, MID1, B4GALNT1, SH3D19, SV2A, VAX2, CASC10, CSPG4, SNAI2, ARSJ, FSCN1, KHDRBS3, RASSF8, SMARCA1, TNC, PIR, ANTXR1, PHLDB1, CASKIN2, LAMC1, PAX6, ASPHD1, MAPK12, CAVIN1, TSPAN6, SEMA6D, NDRG4, PRR36, PFN2, SOX21, SPTBN2, GPC1, NRSN2, AADAT, GNA11, TCEAL9, DDAH1, KIAA1549L, LGR4, MEX3D, TNKS1BP1, PIMREG, SCD, FRMD6, HUNK, TMEM136, LRRC49, ARNT2, ENSG00000259495, DOCK1, PTPRS, TTC23, PPFIBP1, LHX2, SIX4, MAPK8IP1, IGDCC3, DMRTA2, STXBP1, PTPN14,

SLC2A10, ARMC9, TBC1D16, FLNC, RHPN2, RHBDF1, P3H4, ENSG00000261578, DTNA, CELSR2, NOVA2, GPR176, VPS37D, SLC26A10, DNAJC22, ZNRF3. Using a series of these factors, one can build a biomarker structure highly specific to GSCs compared to blood cells.

Table 1. Downstream factors with higher levels of expression in GSCs compared to PBMCs.

Gene	min_GSC_vs_max_blood
ENSG00000277459	832.3084162
SALL1	458.4426045
ID4	337.4452973
FREM2	161.4199241
NES	155.6845922
TUBB2B	149.6286058
MSI1	145.1664194

[0055] For this study we used PCR to look at gene expression of four of the downstream factors: ID4, FREM2, NES, and SALL1 (FIGs. 9-12, Tables 2-9). The CT (cycle threshold) is defined as the number of cycles required for the fluorescent signal to cross the threshold, *i.e.*, exceed the background level. CT levels are inversely proportional to the amount of target nucleic acid in the sample, *i.e.*, the lower the CT level the greater the amount of target nucleic acid in the sample. Here, we have shown that expression of downstream factors ID4, FREM2, NES, and SALL1 is higher in GBM patient samples than in control healthy patient samples (Figures 9-12). The blue panel in Figures 9-12 shows results from control healthy patient blood from Red Cross, and each sample represents a pooled sample of three donors. The yellow/orange panels show results from an experiment testing the detection threshold for the qPCR method used where control samples are spiked with GSCs. The pink panels show results from four GBM patients. It is therefore possible to measure expression of master regulators indirectly by measuring expression of downstream factors in patient blood.

Table 2. PCR analysis of ID4.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3	ID4	Undetermined	Undetermined	
fPBMC3	ID4	Undetermined	Undetermined	
fPBMC3	ID4	Undetermined	Undetermined	21.80076
fPBMC1	ID4	38.158	36.485	
fPBMC1	ID4	35.667	36.485	
fPBMC1	ID4	35.629	36.485	16.243

fPBMC2	ID4	35.408	34.231	
fPBMC2	ID4	34.347	34.231	
fPBMC2	ID4	32.939	34.231	15.557
fPBMC4	ID4	34.974	35.172	
fPBMC4	ID4	34.728	35.172	
fPBMC4	ID4	35.816	35.172	16.288
fPBMC5	ID4	33.462	33.686	
fPBMC5	ID4	33.689	33.686	
fPBMC5	ID4	33.906	33.686	16.989

Table 3. PCR analysis of ID4.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3 GSC25	ID4	37.307	37.380	
fPBMC3 GSC25	ID4	37.454	37.380	
fPBMC3 GSC25	ID4	Undetermined	37.380	21.163
fPBMC3 GSC125	ID4	36.327	35.782	
fPBMC3 GSC125	ID4	35.864	35.782	
fPBMC3 GSC125	ID4	35.155	35.782	19.422
fPBMC3 GSC1250	ID4	Undetermined	33.750	
fPBMC3 GSC1250	ID4	34.705	33.750	
fPBMC3 GSC1250	ID4	32.795	33.750	19.744
fPBMC3 GSC25000	ID4	27.774	28.606	
fPBMC3 GSC25000	ID4	28.845	28.606	
fPBMC3 GSC25000	ID4	29.200	28.606	19.380
CA7 2000	ID4	23.076	22.892	
CA7 2000	ID4	22.783	22.892	
CA7 2000	ID4	22.816	22.892	17.411
Pt1				
10475	ID4	29.354	29.007	
10475	ID4	28.367	29.007	
10475	ID4	29.301	29.007	17.601
Pt2				
32235	ID4	29.774	31.233	
32235	ID4	32.338	31.233	
32235	ID4	31.588	31.233	14.911
Pt3				
19907	ID4	27.033	27.599	
19907	ID4	27.813	27.599	
19907	ID4	27.952	27.599	19.393
Pt4				
18921	ID4	28.937	28.987	
18921	ID4	28.776	28.987	
18921	ID4	29.247	28.987	15.936

Table 4. PCR analysis of FREM2.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3	FREM2	Undetermined	Undetermined	
fPBMC3	FREM2	Undetermined	Undetermined	
fPBMC3	FREM2	Undetermined	Undetermined	22.451
fPBMC1	FREM2	Undetermined	Undetermined	
fPBMC1	FREM2	Undetermined	Undetermined	17.402
fPBMC2	FREM2	Undetermined	Undetermined	
fPBMC2	FREM2	Undetermined	Undetermined	16.042
fPBMC4	FREM2	Undetermined	Undetermined	
fPBMC4	FREM2	Undetermined	Undetermined	17.197
fPBMC5	FREM2	35.393	35.914	
fPBMC5	FREM2	36.434	35.914	17.821

Table 5. PCR analysis of FREM2.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3 GSC25	FREM2	Undetermined	Undetermined	
fPBMC3 GSC25	FREM2	Undetermined	Undetermined	
fPBMC3 GSC25	FREM2	Undetermined	Undetermined	22.354
fPBMC3 GSC125	FREM2	36.170	36.296	
fPBMC3 GSC125	FREM2	38.124	36.296	
fPBMC3 GSC125	FREM2	34.593	36.296	20.173
fPBMC3 GSC1250	FREM2	Undetermined	39.730	
fPBMC3 GSC1250	FREM2	Undetermined	39.730	
fPBMC3 GSC1250	FREM2	39.730	39.730	21.953
fPBMC3 GSC25000	FREM2	30.768	30.699	
fPBMC3 GSC25000	FREM2	31.001	30.699	
fPBMC3 GSC25000	FREM2	30.328	30.699	19.789
CA7 2000	FREM2	23.897	24.108	
CA7 2000	FREM2	23.915	24.108	
CA7 2000	FREM2	24.510	24.108	18.294
Pt1				
10475	FREM2	34.720	34.095	
10475	FREM2	32.916	34.095	
10475	FREM2	34.650	34.095	18.272
Pt2				
32235	FREM2	31.853	32.483	
32235	FREM2	32.611	32.483	
32235	FREM2	32.986	32.483	15.748
Pt3				
19097	FREM2	31.443	31.167	
19097	FREM2	30.647	31.167	
19097	FREM2	31.412	31.167	19.934
Pt4				
18921	FREM2	34.565	33.675	

18921	FREM2	33.358	33.675	
18921	FREM2	33.100	33.675	16.627

Table 6. PCR analysis of NES.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3	NES	Undetermined	Undetermined	
fPBMC3	NES	Undetermined	Undetermined	
fPBMC3	NES	Undetermined	Undetermined	22.451
fPBMC1	NES	38.107	38.107	
fPBMC1	NES	Undetermined	38.107	17.402
fPBMC2	NES	36.521	36.223	
fPBMC2	NES	35.926	36.223	16.042
fPBMC4	NES	34.633	36.677	
fPBMC4	NES	38.721	36.677	17.197
fPBMC5	NES	Undetermined	34.639	
fPBMC5	NES	34.639	34.639	17.821

Table 7. PCR analysis of NES.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3 GSC25	NES	Undetermined	Undetermined	
fPBMC3 GSC25	NES	Undetermined	Undetermined	
fPBMC3 GSC25	NES	Undetermined	Undetermined	22.354
fPBMC3 GSC125	NES	Undetermined	34.619	
fPBMC3 GSC125	NES	34.619	34.619	
fPBMC3 GSC125	NES	Undetermined	34.619	20.173
fPBMC3 GSC1250	NES	Undetermined	Undetermined	
fPBMC3 GSC1250	NES	Undetermined	Undetermined	
fPBMC3 GSC1250	NES	Undetermined	Undetermined	21.953
fPBMC3 GSC25000	NES	33.184	33.567	
fPBMC3 GSC25000	NES	32.671	33.567	
fPBMC3 GSC25000	NES	34.845	33.567	19.789
CA7 2000	NES	27.880	27.279	
CA7 2000	NES	27.529	27.279	
CA7 2000	NES	26.428	27.279	18.294
Pt1				
10475	NES	Undetermined	Undetermined	
10475	NES	Undetermined	Undetermined	
10475	NES	Undetermined	Undetermined	18.272
Pt2				
32235	NES	32.337	32.466	
32235	NES	32.595	32.466	
32235	NES	Undetermined	32.466	15.748
Pt3				
19097	NES	33.737	37.104	
19097	NES	38.399	37.104	

19097	NES	39.175	37.104	19.934
Pt4				
18921	NES	Undetermined	33.572	
18921	NES	34.457	33.572	
18921	NES	32.686	33.572	16.627

Table 8. PCR analysis of SALL1.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3	SALL1	Undetermined	Undetermined	
fPBMC3	SALL1	Undetermined	Undetermined	16.920
fPBMC1	SALL1	Undetermined	36.385	
fPBMC1	SALL1	36.385	36.385	
fPBMC1	SALL1	Undetermined	36.385	16.920
fPBMC2	SALL1	34.407	35.259	
fPBMC2	SALL1	35.803	35.259	
fPBMC2	SALL1	35.566	35.259	15.757
fPBMC4	SALL1	34.385	33.410	
fPBMC4	SALL1	32.931	33.410	
fPBMC4	SALL1	32.915	33.410	16.288
fPBMC5	SALL1	30.987	31.534	
fPBMC5	SALL1	31.947	31.534	
fPBMC5	SALL1	31.668	31.534	16.989

Table 9. PCR analysis of SALL1.

Sample Name	Target Name	CT	Ct Mean	ACTB
fPBMC3 GSC25	SALL1	34.921	34.204	
fPBMC3 GSC25	SALL1	33.487	34.204	22.287
fPBMC3 GSC125	SALL1	33.680	33.331	
fPBMC3 GSC125	SALL1	32.982	33.331	19.857
fPBMC3 GSC1250	SALL1	Undetermined	34.954	
fPBMC3 GSC1250	SALL1	34.954	34.954	21.335
fPBMC3 GSC25000	SALL1	30.514	30.508	
fPBMC3 GSC25000	SALL1	30.502	30.508	19.993
CA7 9000	SALL1	24.335	24.114	
CA7 9000	SALL1	23.894	24.114	18.154
Pt1				
10475	SALL1	28.261	28.139	
10475	SALL1	28.210	28.139	
10475	SALL1	27.946	28.139	18.058
Pt2				
32235	SALL1	30.978	31.177	
32235	SALL1	30.973	31.177	
32235	SALL1	31.580	31.177	15.398
Pt3				
19097	SALL1	30.222	29.933	

19097	SALL1	29.643	29.933	19.509
Pt4				
18921	SALL1	33.406	33.460	
18921	SALL1	33.515	33.460	16.743

[0056] The same concept can be used for other cancers in which expression profiles of their cancer stem cells are available.

Example 5. PCR detection Optimization.

[0057] We showed that using qPCR there is a limit of sensitivity due to the rarity of circulating GBM cells in the blood. In the above example, 5 GBM per 1 million blood cells was reliably detected. Reliable detection at a sensitivity of 1 GBM cell per 1 million of blood cells to 1 GBM cell for 5 million blood cells is desired.

[0058] A. Blood cell depletion: In order to enhance sensitivity, blood cells were depleted from the samples to enrich for cancer cells. We used a magnetic method to deplete CD45+ cells or immune cells (FIG. 1B) right). To test the method, normal peripheral blood mononuclear cells (PBMC) were spiked with GBM stem cells (GSC) at different frequencies, CD45+ cells were depleted, and sensitivity of detection was analyzed. For these tests, the FREM2 target was used.

[0059] Without enrichment, a frequency of 5 cancer cells per 1 million blood cells (PBMC+GSC 5) was the lowest reliable limit of detection. Surprisingly, depletion of CD545+ cells did not improve sensitivity. On the contrary, depletion of CD45+ cells, decreased sensitivity of FREM2 detection. One likely explanation was that cancer cells bound to the magnetic beads non-specifically, thereby removing them from the sample with the CD45+ cells (FIG. 2B).

Table 10. Input 50 ng RNA/well

Sample Name	Target Name	CT	GAPDH
L2	FREM2	25.961	30.648
PBMC+GSC 100	FREM2	32.844	31.512
PBMC+GSC 50	FREM2	35.426	32.926
PBMC+GSC 5	FREM2	38.964	32.623
PBMC+GSC 1	FREM2	undetermined	32.651
PBMC	FREM2	35.543	38.596

[0060] B. RNA level: We next tested increasing the RNA input in the no enrichment sample. With a high input level, non-specific amplification became a problem. With 100 ng of RNA per

well we could no longer detect less than 50 cancer cells per 1 million blood cells. Non-specific binding of the PCR primers to the excess RNA template limited the level of detection.

Table 11. Input 100 ng/well

Sample Name	Target Name	CT	GAPDH
L2	FREM2	28.079	32.931
PBMC+GSC 100	FREM2	34.308	33.524
PBMC+GSC 50	FREM2	36.965	34.697
PBMC+GSC 5	FREM2	38.321	33.708
PBMC+GSC 1	FREM2	38.615	31.675
PBMC	FREM2	38.746	32.918

[0061] C. Two-step headstart PCR: 10 cycles of PCR were used to amplify and molecularly enrich rare RNA species. The amplification product of this first round of PCR was then used as input for the next 30 cycles of PCR amplification with the same primer pair. Two-step headstart PCR, increased sensitivity top reliably detect 1 cancer cell per 1 million blood cells (FIG. 4B).

Table 12. Two-step headstart PCR, input 50 ng RNA /well.

Sample Name	Target Name	CT	GAPDH
L2	FREM2	17.941	23.159
PBMC+GSC 100	FREM2	23.738	24.006
PBMC+GSC 50	FREM2	26.867	24.794
PBMC+GSC 5	FREM2	28.191	22.680
PBMC+GSC 1	FREM2	28.984	24.356
PBMC	FREM2	undetermined	23.260

[0062] D. Nested two-step headstart PCR. A non-enriched sample is PCR amplified for 10 cycles with a first primer pair, primer pair A. The product from this first round of PCR is then used as input for 30-40 more cycles in a second round of PCR amplification using a second, nested primer pair, primer pair B. Primer pair B amplifies a region of DNA contained within the primer pair A amplification product. In other embodiments, the first PCR reaction is 5-15 cycles and the second PCR reaction is 20-40 cycles.

[0063] Using nested two-step headstart PCR and 1-5 tumor cells per 1 million blood cells, reliable detection of cancer cells was observed. The data in Table 13 show that the nested method separated the lower end frequencies, 1 tumor cell vs 5 tumor cells, with CT difference of almost 4. Thus, nested two-step headstart PCR was able to detect low tumor cell frequencies. Based on

these results, it is expected that nested two-step headstart PCR will work for detection of 1 tumor cell in 5 million blood cells.

Table 13. Nested two-step headstart PCR detection of 1, 5, 50, or 1000 cancer cells per 1 million blood cells.

Sample Name	Target Name	CT	GAPDH
L2	FREM2	24.859	24.500
PBMC + GSC 1000	FREM2	30.984	24.734
PBMC + GSC 50	FREM2	35.298	25.676
PBMC + GSC 5	FREM2	33.655	25.487
PBMC + GSC 1	FREM2	37.227	23.361
PMBC	FREM2	undetermined	24.330

[0064] In some embodiments, a third, nested primer pair, primer pair C, is used, nested three step headstart PCR. Primer pair C amplifies a region of DNA contained within the primer pair B amplification product. Various number of cycles can be used in each round. An exemplary protocol, for use of three primer pairs is: 10 cycles PCR amplification using primer pair A (first PCR reaction), followed by 10 cycles PCR amplification using primer pair B (second PCR reaction), followed by 30 cycles PCR amplification using primer pair C (third PCR reaction). The amplification product using primer pair A is used as input for amplification using primer pair B and the amplification product using primer pair B is used as input for amplification using primer pair C. In other embodiments, the first PCR reaction is 5-15 cycles, the second PCR reaction is 5-15 cycles, and the third PCR reaction is 20-40 cycles.

[0065] Nested multistep PCR can be used to detect GCSs in blood. Nested multistep PCR can be used to detect FREM2 as described above. Nested multistep PCR can also be used to detect SALL1, ID4, or NES using nested primers specific for these target genes. In some embodiments, nested multistep PCR is used to detect two or more of FREM2, SALL1, ID4, and NES. In some embodiments, nested multistep PCR is used to detect three or more of FREM2, SALL1, ID4, and NES. In some embodiments, nested multistep PCR is used to detect FREM2, SALL1, ID4, and NES. Nested multistep PCR to detect two or more genes can be performed in a single multiplex reaction or in separate uniplex reactions.

[0066] In addition to the four genes described above, nested multistep PCR can be used to detect other genes in the GCS regulatory network or downstream genes in other cancer master

regulator networks. Using genetic analyses, the top 20 genes in a cancer of interest regulatory network, such as the GSC regulatory network, with the highest differential expression between the cancer cell, such as GSC, and PBMC are identified. Detection of the top 20 genes is performed in subjects with known GBM and healthy control subjects. Statistical methods are then used to assign weight to each of the 20 factors and develop an algorithm of scoring. A correlation between the strength of the scoring system and tumor burden and survival in patients is used as a GBM disease assessment and treatment response monitoring method. The described tests can be used to screen for cancer in apparently healthy or at risk subjects. The described tests can also be used to monitor the disease in cancer patients.

[0067] In some embodiments are described methods of assessing cancer risk comprising: using a statistical analysis to identify the top 20 genes differentially expressed in subjects having a cancer of interest; obtaining or having obtained a sample from a patient having, suspected of having, or at risk of developing in the cancer of interest; measuring, or having measured, expression of a plurality of the top 20 genes in the sample; assessing tumor burden or cancer risk based on the expression of the plurality of the top 20 genes in the sample.

[0068] Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which the inventions pertain having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the inventions are not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

WHAT IS CLAIMED IS:

1. A method of diagnosing a subject with a cancer, measuring in a liquid biopsy of said subject, the level of unmethylated DNA in a promoter and/or regulatory region of one or more master regulator genes of the cancer and comparing the results with a healthy reference sample, wherein a higher amount of unmethylated DNA of the master regulator genes detected in the liquid biopsy from the subject relative to the healthy reference sample is an indication of the cancer.

2. The method of claim 1, further comprising identifying the one or more master regulator genes of the cancer before the detecting step.

3. The method of claim 2, wherein the master regulator genes of the cancer are identified using GeneRep/nSCORE.

4. The method of any preceding claim, wherein the unmethylated DNA is detected using ATAC-seq.

5. The method of any preceding claim, wherein the unmethylated DNA is detected using a method comprising: CRISPR, digestion based assay followed by PCR, methylation specific PCR, E-ice-COLD-PCR, bead array analysis, pyrosequencing, PCR with high resolution melting, bisulfite sequencing.

6. A method of diagnosing a subject with a cancer, comprising measuring the expression level of one or more downstream targets of one or more master regulator genes of the cancer in a liquid biopsy of the subject and comparing the results with a healthy reference sample, wherein a higher expression level of downstream targets in the liquid biopsy from the subject relative to the healthy reference sample is an indication of the cancer.

7. The method of claim 6, further comprising identifying the one or more master regulators genes of the cancer and/or the one or more downstream targets of the master regulator genes before the measuring step.

8. The method of claim 6, wherein the cancer is GBM and the downstream targets are select from the group consisting of SALL1, ID4, FREM2, NES, MLXIPL, NKX2-2, KCNIP3, HLF, DDN, BATF2, MEOX2, OLIG2, PARGC1B, ACTN2, OTP, PRKCB, HOXA13, MNX1, ATOH7, RXRG, HOXA11, HOXD13, PEG3, RPH3A, HOXD3, CEBPB, ZNF248, BHLHE40, NMI, POU4F1, THRB, ENSG00000277459, FOXG1, FJX1, TUBB2B, MSI1,

FKBP10, CBARP, GNG12, CD276, FAT1, CTXN1, DPYSL5, TYRO3, GJC1, CHST3, LRP4, PHF21B, ADAMTS9, CSPG5, TEAD1, BCAR1, EML1, RBFOX2, MPDZ, MSX1, EFNB3, ENAH, LYPD6, GTF2IRD1, TBX2, ANK2, C1QL1, ZIC1, EMP2, TMEM132A, CX3CL1, SYDE1, SLC16A2, SOX9, RND3, LARP6, CNN3, SPRY4, DPF1, PCGF2, BOC, OBSL1, SOX2, EPB41L1, MEX3A, NCS1, SMO, TMC7, SEZ6L2, POU3F2, FAM171A2, DENND2A, TANC1, PROX1, ENSG00000268592, PNPLA3, TSKU, DLX1, KIAA1549, MTSS1L, PTPRF, IRX5, DZIP1, MAGI1, ADCY6, BCHE, CXADR, TEAD4, PTPN21, CDR2L, ADGRL3, REEP2, SHROOM3, ARC, EEF1A2, ETV4, EGFR, UCHL1, KAZALD1, TJP1, ENSG00000237004, ETV5, CKB, KCNF1, MAP1B, S100A16, COL27A1, VGF, ALDH7A1, GAS1, LOC101927480, CITED1, ETV1, NOVA1, JPH1, FBXO17, CNKSR3, PDXP, PLEKHH3, PYGO1, SCARA3, RTL8B, LAMB1, MYH14, CASKIN1, NLGN2, PACSIN3, CA12, ARHGAP39, LAMA4, ZNF462, NR2F1, CSRP2, ARHGEF25, GLI3, TMEFF1, MID1, B4GALNT1, SH3D19, SV2A, VAX2, CASC10, CSPG4, SNAI2, ARSJ, FSCN1, KHDRBS3, RASSF8, SMARCA1, TNC, PIR, ANTXR1, PHLDB1, CASKIN2, LAMC1, PAX6, ASPHD1, MAPK12, CAVIN1, TSPAN6, SEMA6D, NDRG4, PRR36, PFN2, SOX21, SPTBN2, GPC1, NRSN2, AADAT, GNA11, TCEAL9, DDAH1, KIAA1549L, LGR4, MEX3D, TNKS1BP1, PIMREG, SCD, FRMD6, HUNK, TMEM136, LRRC49, ARNT2, ENSG00000259495, DOCK1, PTPRS, TTC23, PPFIBP1, LHX2, SIX4, MAPK8IP1, IGDCC3, DMRTA2, STXBP1, PTPN14, SLC2A10, ARMC9, TBC1D16, FLNC, RHPN2, RHBDF1, P3H4, ENSG00000261578, DTNA, CELSR2, NOVA2, GPR176, VPS37D, SLC26A10, DNAJC22, and ZNRF3.

9. A method of diagnosing cancer in a subject with a cancer comprising

(a) obtaining or having obtained a sample from the subject

(b) analyzing or having analyzed chromosomal accessibility and/or DNA methylation of one or more master regulator genes of the cancer; and

(c) comparing the chromosomal accessibility and/or DNA methylation of the one or more master regulator genes of the cancer in the sample with the chromosomal accessibility and/or DNA methylation of the one or more master regulator genes of the cancer of a healthy reference sample, wherein an increase in chromosomal accessibility or a decrease in DNA methylation of a promotor and/or regulatory region of the master regulator genes detected in the sample relative to the healthy reference sample is an indication of the cancer.

10. The method of claim 9, wherein the sample is a liquid sample.

11. The method of claim 10, wherein the liquid sample is a blood sample.

12. The method of any one of claims 9-11, wherein chromosomal accessibility is analyzed by analyzing DNA methylation.

13. The method of claim 12, wherein analyzing chromosomal accessibility and/or DNA methylation comprises ATAC-seq, CRISPR, DNAse-seq, MNase-seq, a digestion based assay followed by PCR, methylation specific PCR, E-ice-COLD-PCR, Bead array analysis, pyrosequencing, PCR with high resolution melting, bisulfite sequencing.

14. The method any one of claims 9-13, wherein the cancer is GBM and the master regulator genes are selected from the group consisting of: TBX2, NKX2-2, BATF2, OLIG2, OTP, SALL1, ID4, FREM2, NES, MLXIPL, KCNIP3, HLF, DDN, MEOX2, PARGC1B, ACTN2, PRKCB, HOXA13, MNX1, ATOH7, RXRG, HOXA11, HOXD13, PEG3, RPH3A, HOXD3, CEBPB, ZNF248, BHLHE40, NMI, POU4F1, THRB, ENSG00000277459, FOXG1, FJX1, TUBB2B, MSI1, FKBP10, CBARP, GNG12, CD276, FAT1, CTXN1, DPYSL5, TYRO3, GJC1, CHST3, LRP4, PHF21B, ADAMTS9, CSPG5, TEAD1, BCAR1, EML1, RBFOX2, MPDZ, MSX1, EFNB3, ENAH, LYPD6, GTF2IRD1, ANK2, C1QL1, ZIC1, EMP2, TMEM132A, CX3CL1, SYDE1, SLC16A2, SOX9, RND3, LARP6, CNN3, SPRY4, DPF1, PCGF2, BOC, OBSL1, SOX2, EPB41L1, MEX3A, NCS1, SMO, TMC7, SEZ6L2, POU3F2, FAM171A2, DENND2A, TANC1, PROX1, ENSG00000268592, PNPLA3, TSKU, DLX1, KIAA1549, MTSS1L, PTPRF, IRX5, DZIP1, MAGI1, ADCY6, BCHE, CXADR, TEAD4, PTPN21, CDR2L, ADGRL3, REEP2, SHROOM3, ARC, EEF1A2, ETV4, EGFR, UCHL1, KAZALD1, TJP1, ENSG00000237004, ETV5, CKB, KCNF1, MAP1B, S100A16, COL27A1, VGF, ALDH7A1, GAS1, LOC101927480, CITED1, ETV1, NOVA1, JPH1, FBXO17, CNKSR3, PDXP, PLEKHH3, PYGO1, SCARA3, RTL8B, LAMB1, MYH14, CASKIN1, NLGN2, PACSIN3, CA12, ARHGAP39, LAMA4, ZNF462, NR2F1, CSRP2, ARHGEF25, GLI3, TMEFF1, MID1, B4GALNT1, SH3D19, SV2A, VAX2, CASC10, CSPG4, SNAI2, ARSJ, FSCN1, KHDRBS3, RASSF8, SMARCA1, TNC, PIR, ANTXR1, PHLDB1, CASKIN2, LAMC1, PAX6, ASPHD1, MAPK12, CAVIN1, TSPAN6, SEMA6D, NDRG4, PRR36, PFN2, SOX21, SPTBN2, GPC1, NRSN2, AADAT, GNA11, TCEAL9, DDAH1, KIAA1549L, LGR4, MEX3D, TNKS1BP1, PIMREG, SCD, FRMD6, HUNK, TMEM136, LRRC49, ARNT2, ENSG00000259495, DOCK1, PTPRS, TTC23, PPFIBP1, LHX2, SIX4, MAPK8IP1, IGDCC3, DMRTA2, STXBP1, PTPN14, SLC2A10, ARMC9, TBC1D16, FLNC, RHPN2, RHBDF1,

P3H4, ENSG00000261578, DTNA, CELSR2, NOVA2, GPR176, VPS37D, SLC26A10, DNAJC22, and ZNRF3

15. A method of diagnosing cancer in a subject with a cancer comprising

(a) identifying or having identified one or more genes in a master regulator network downstream of a master regulator of the cancer,

(b) obtaining or having obtained a sample from the subject,

(c) analyzing or having analyzed expression of the one or more genes; and

(d) comparing the expression level of the one or more genes in the sample with the expression level of the one or more genes of a healthy reference sample, wherein a differential expression of the one or more genes in the sample relative to the healthy reference sample is an indication of the cancer.

16. The method of claim 15, further comprising identifying or having identified at least one master regulator of the cancer before step (a).

17. The method of claim 16, where identifying one or more master regulators of the cancer comprises using GeneRep/nSCORE.

18. The method of any one of claims 15-17, wherein the sample is a liquid sample.

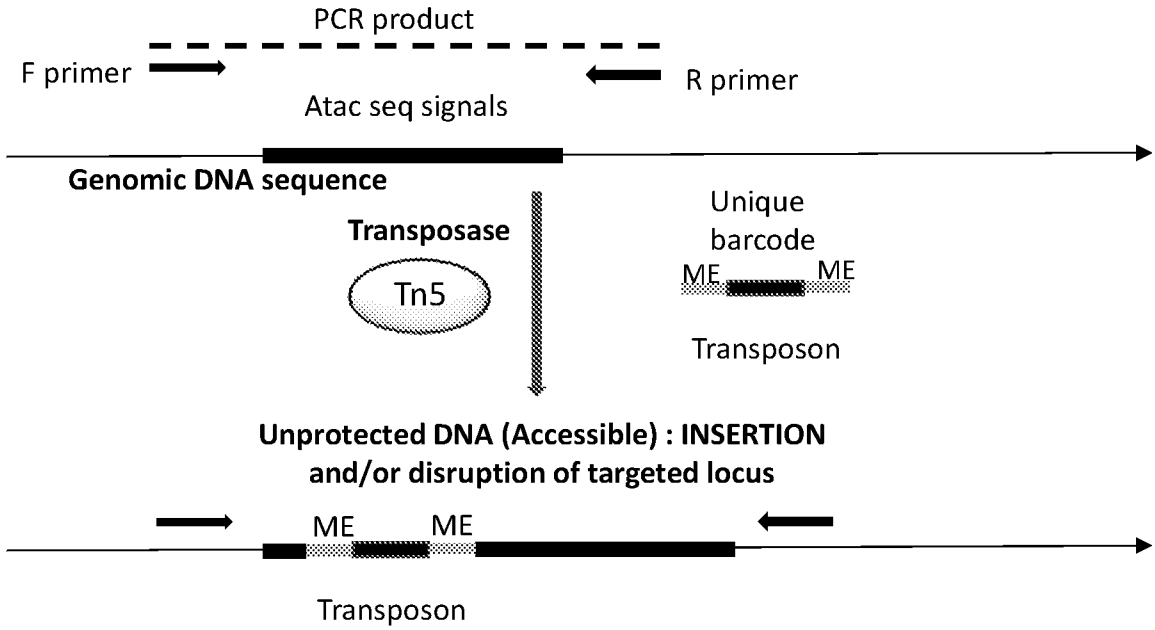
19. The method of claim 18, wherein the liquid sample is a blood sample.

20. The method of any one of claims 15-19, wherein the expression level is analyzed by PCR .

21. The method of claim 20, wherein the PCT comprises multi-step headstart PCR.

22. The method of any one of claims 15-21, wherein identifying or having identified one or more genes in a master regulator network downstream of a master regulator of the cancer comprises identifying or having identified the top 20 gene that are differentially expressed between the cancer and peripheral blood mononuclear cells.

FIG. 1



Protected DNA (not accessible): NO INSERTION or disruption of locus

FIG. 2

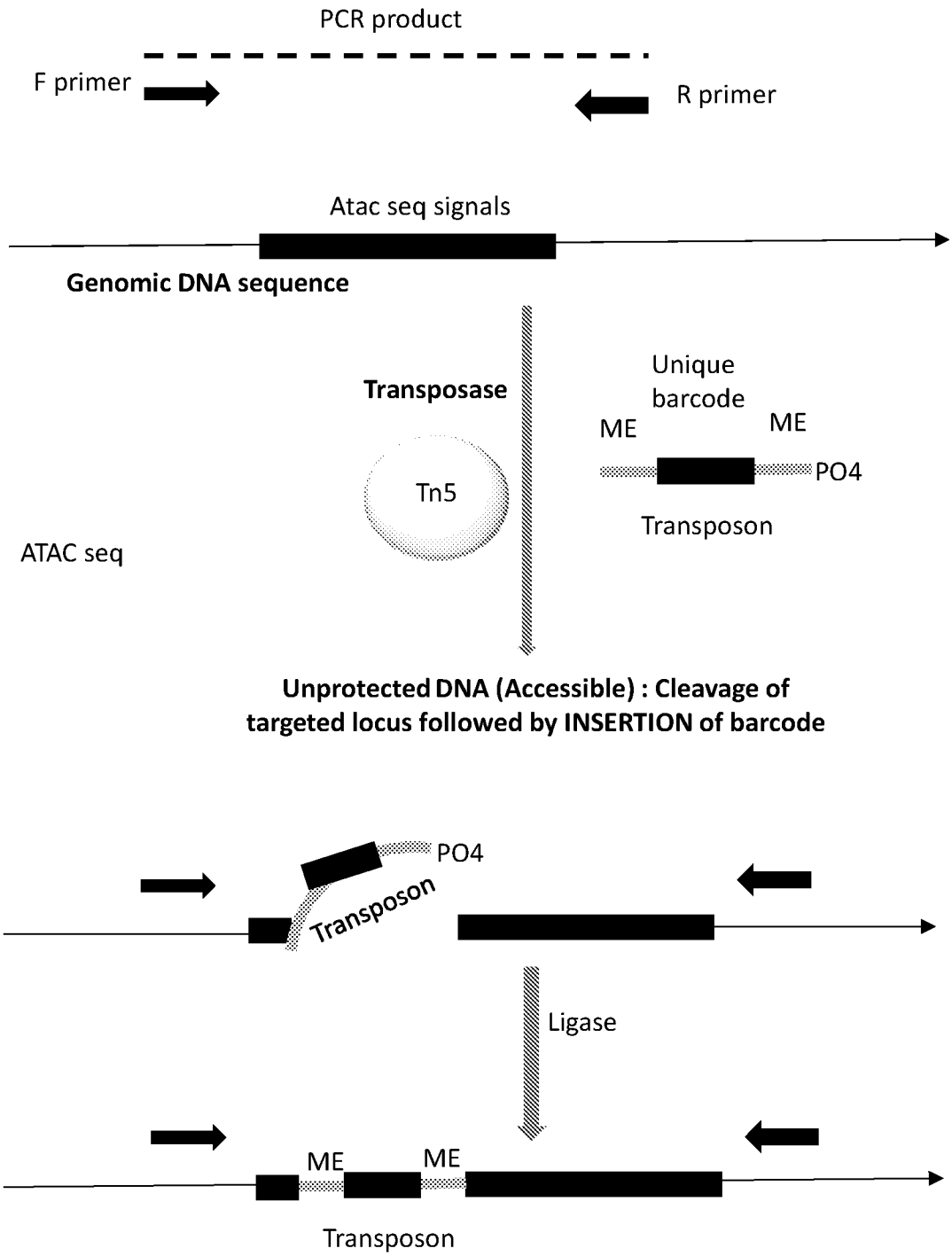

PCR product
F primer
R primer
Atac seq signals
Genomic DNA sequence
Transposase
Tn5
ATAC seq
Unique barcode
ME
ME
PO4
Transposon
Unprotected DNA (Accessible) : Cleavage of targeted locus followed by INSERTION of barcode
PO4
Transposon
Ligase
ME
ME
Transposon

FIG. 3

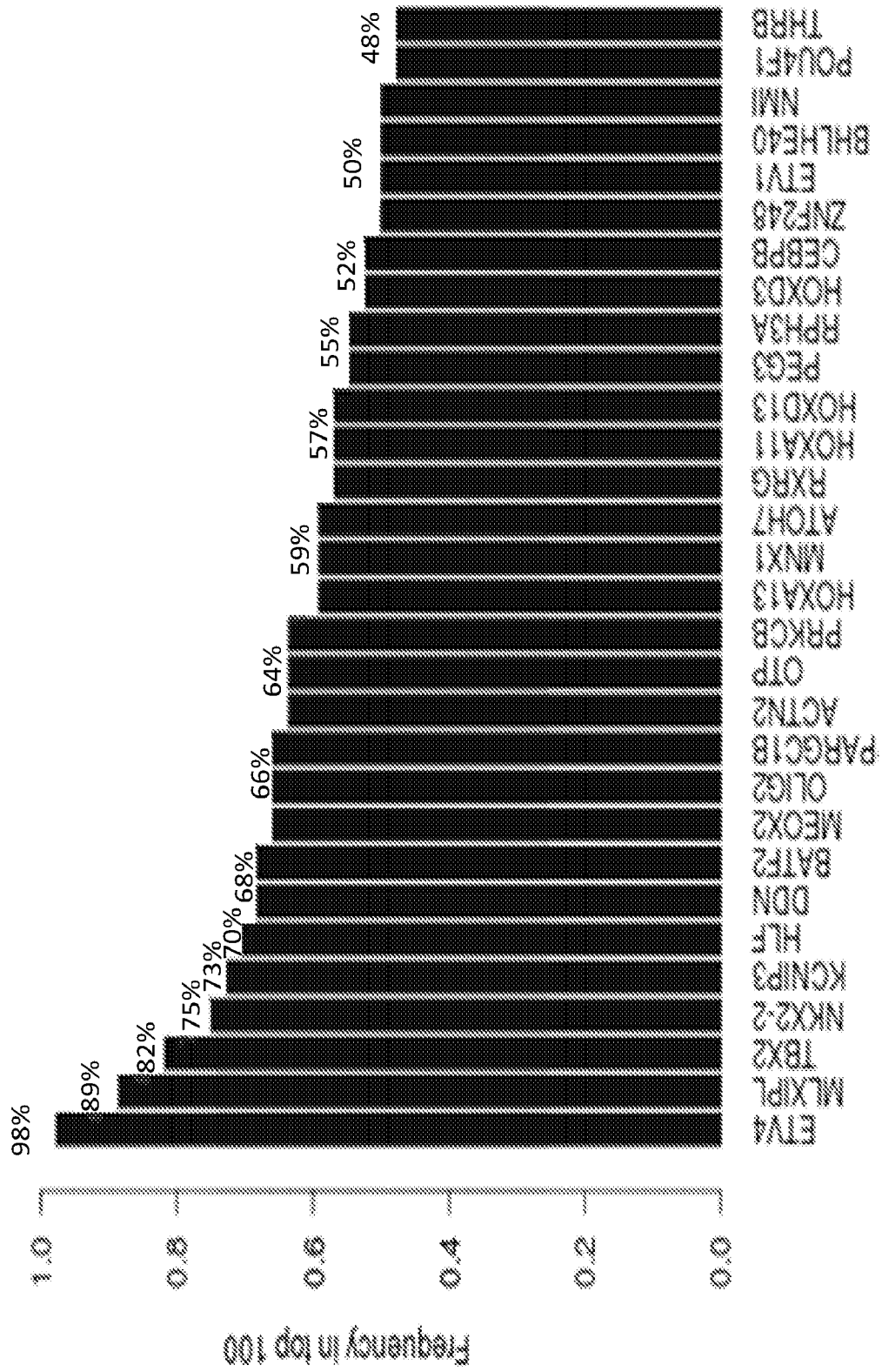

AST TO GSC top genes ranking consistency
Frequency in bp 100
1.0
0.8
0.6
0.4
0.2
0.0
ETV4
MLXIPL
TBX2
NKX2-2
KCNIP3
HLF
DDN
BATF2
MEOX2
OLIG2
PARGC1B
ACTN2
OTP
PRKCB
HOXA13
MNX1
ATOH7
RXRG
HOXA11
HOXD13
PEG3
RPH3A
HOXD3
CEBPB
ZNF243
ETV1
BHLHE40
NMI
POU4F1
THRB
98%
89%
82%
75%
73%
70%
68%
66%
64%
59%
57%
55%
52%
50%
48%

FIG. 4
OTP methylation and Atac-seq signals

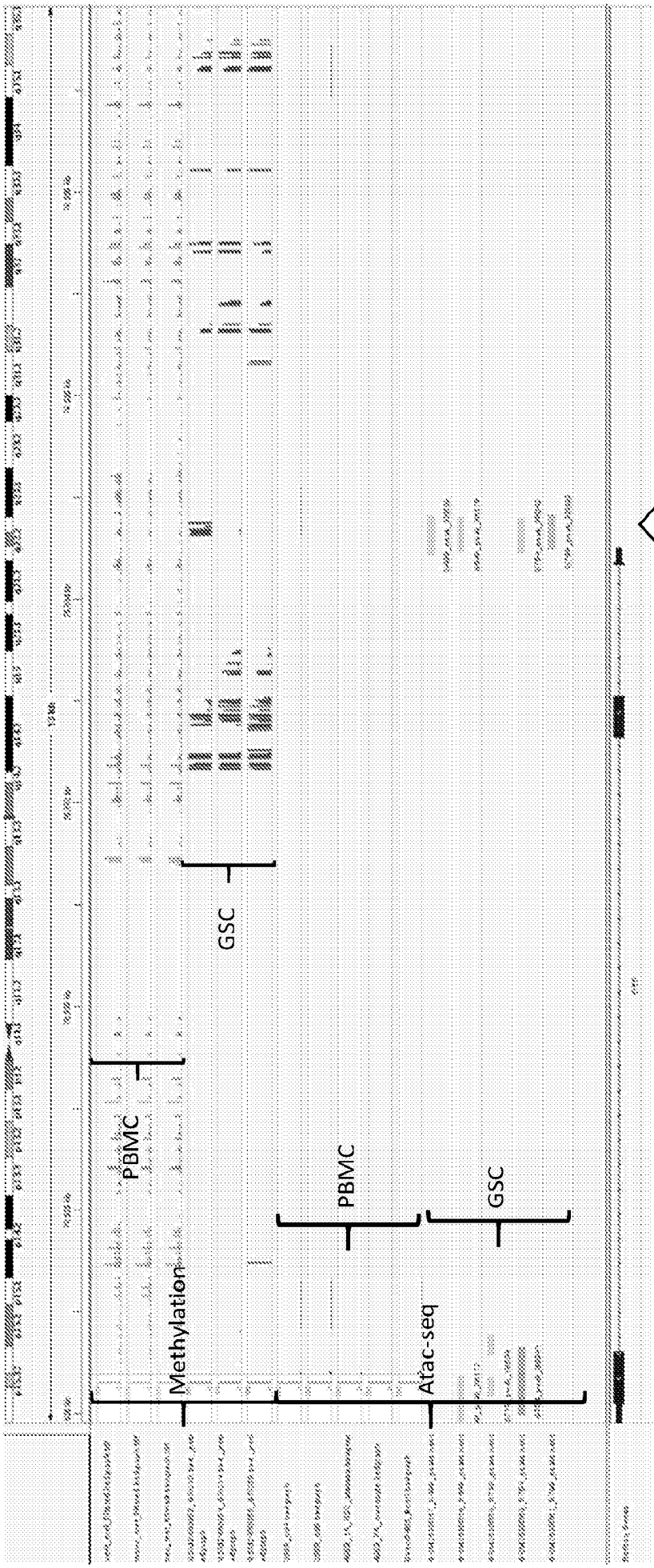


FIG. 5

OLIG2 methylation and Atac-seq signals

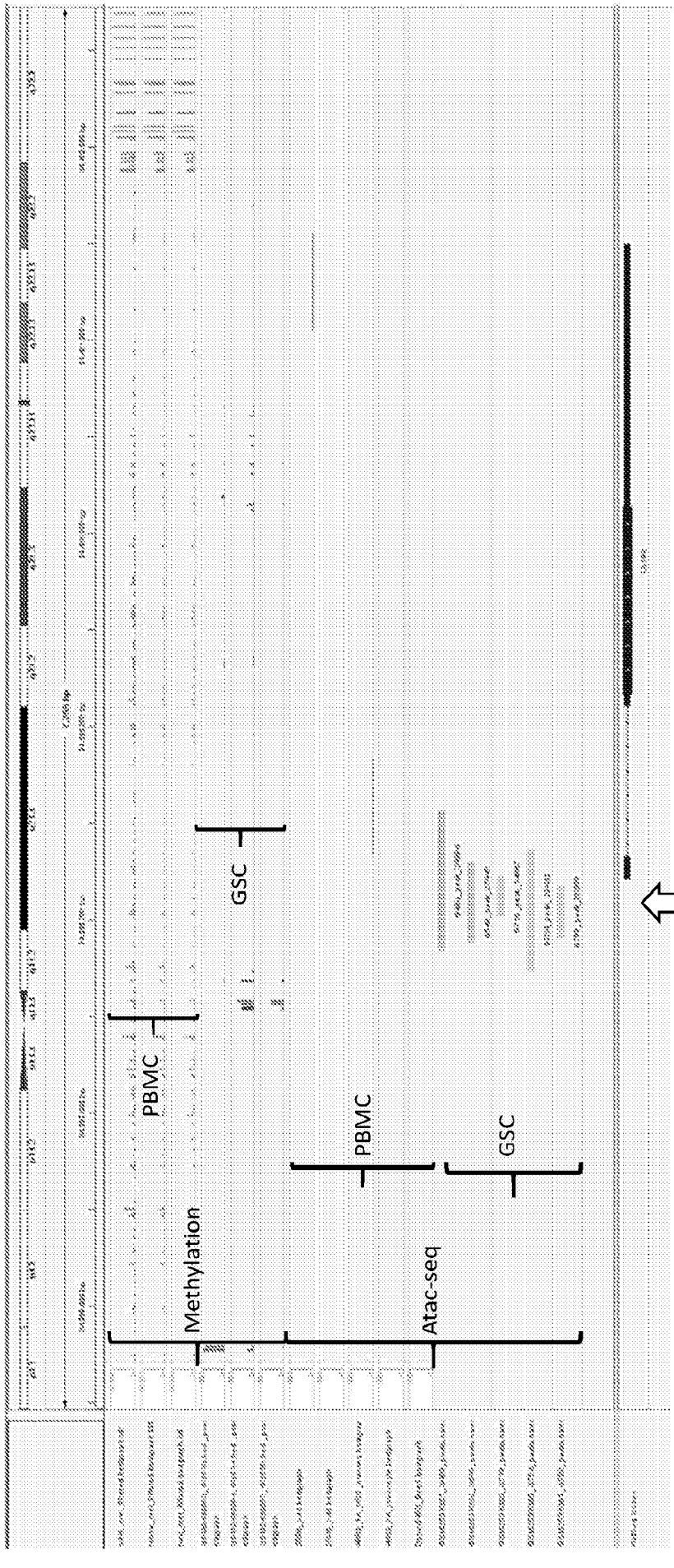


FIG. 6

BATF2 methylation and Atac-seq signals

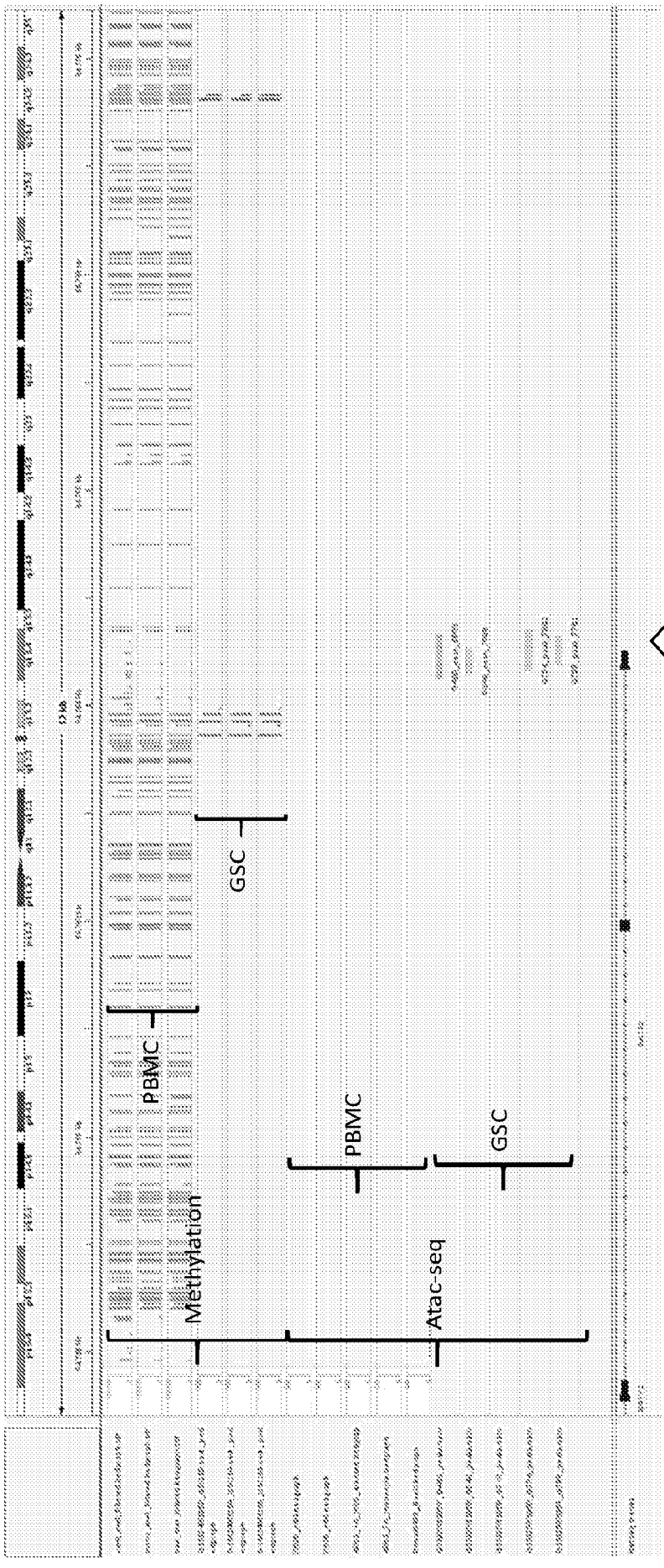


FIG. 7
NKX2-2 methylation and Atac-seq signals

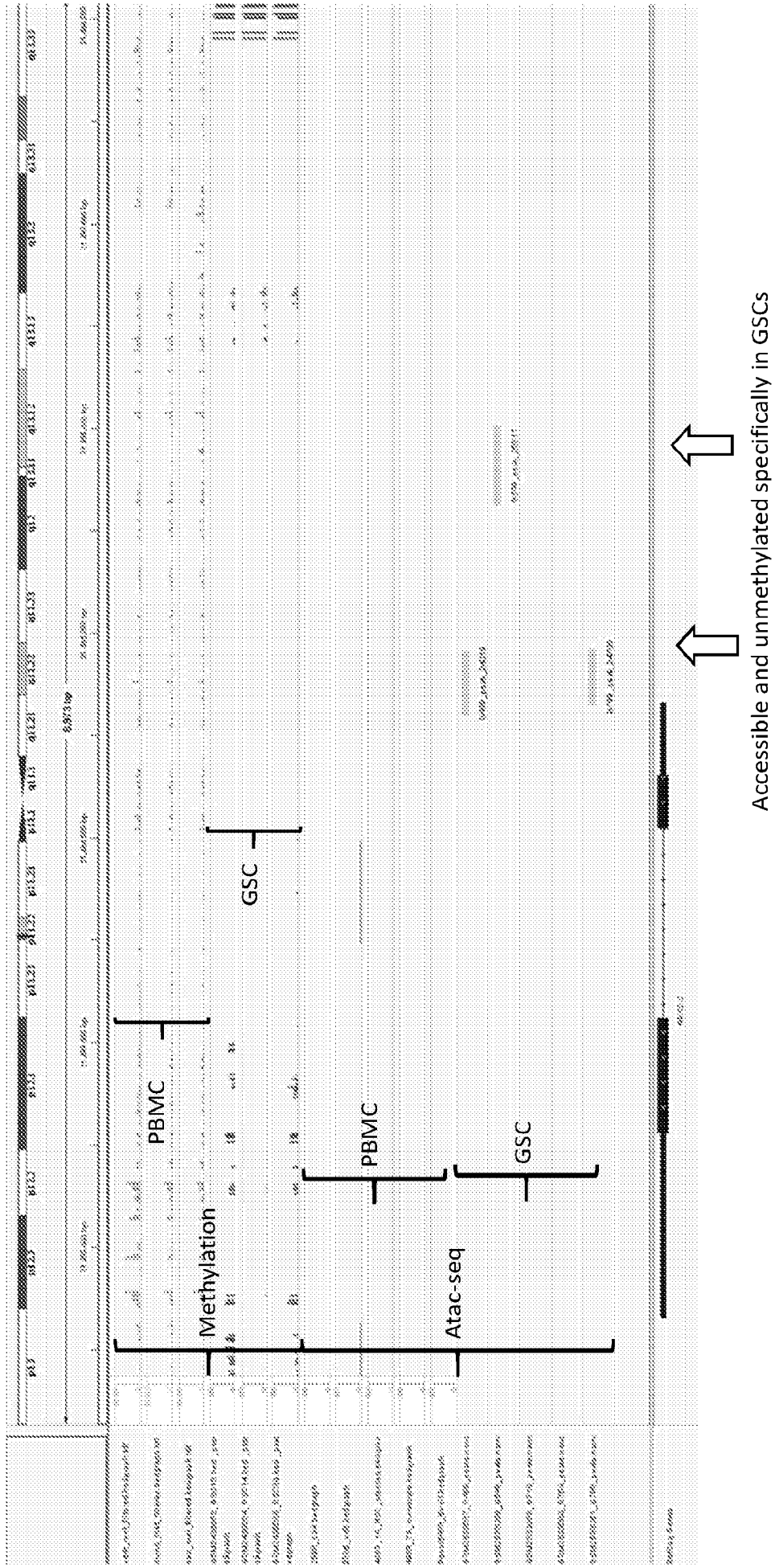


FIG. 8

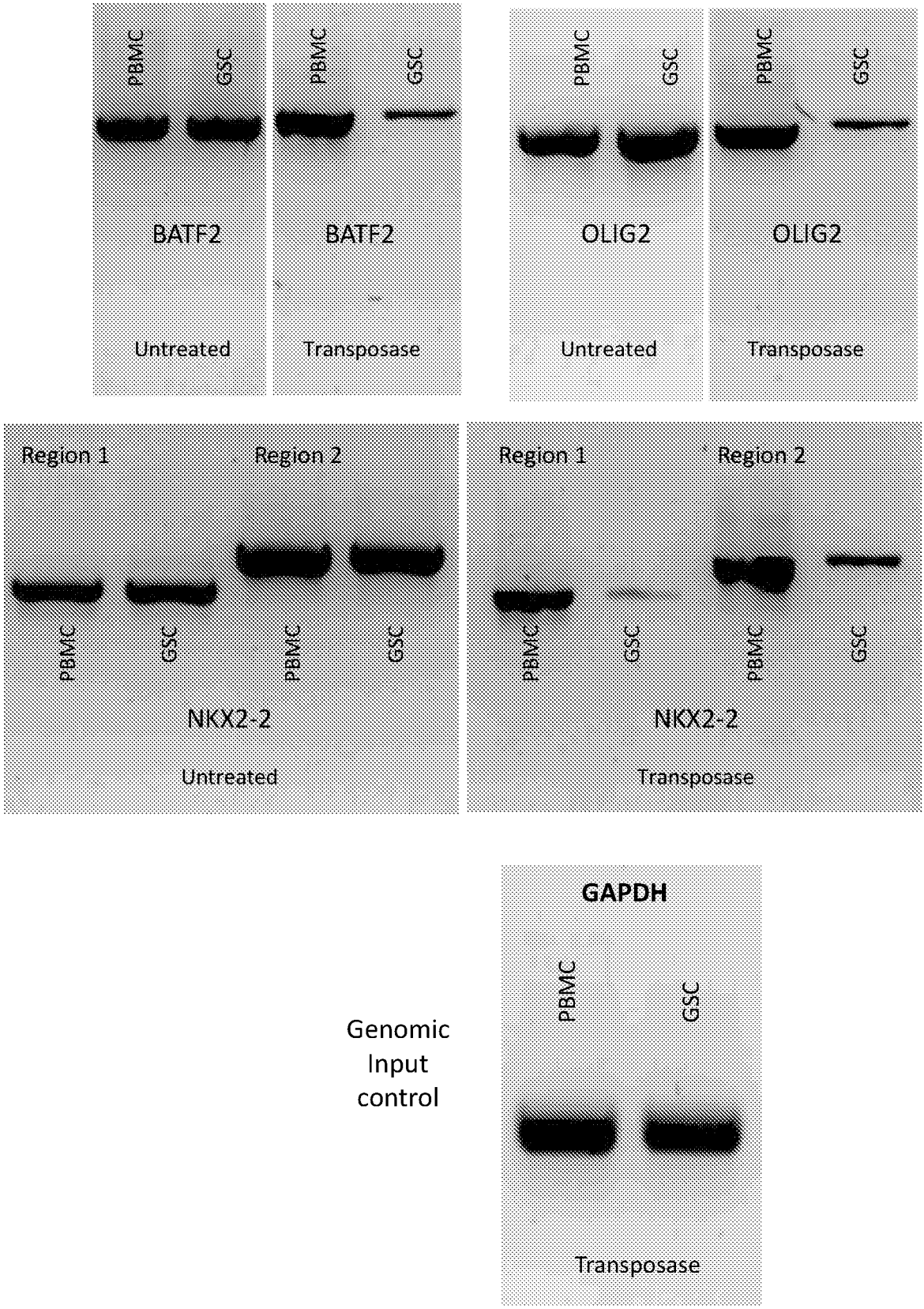
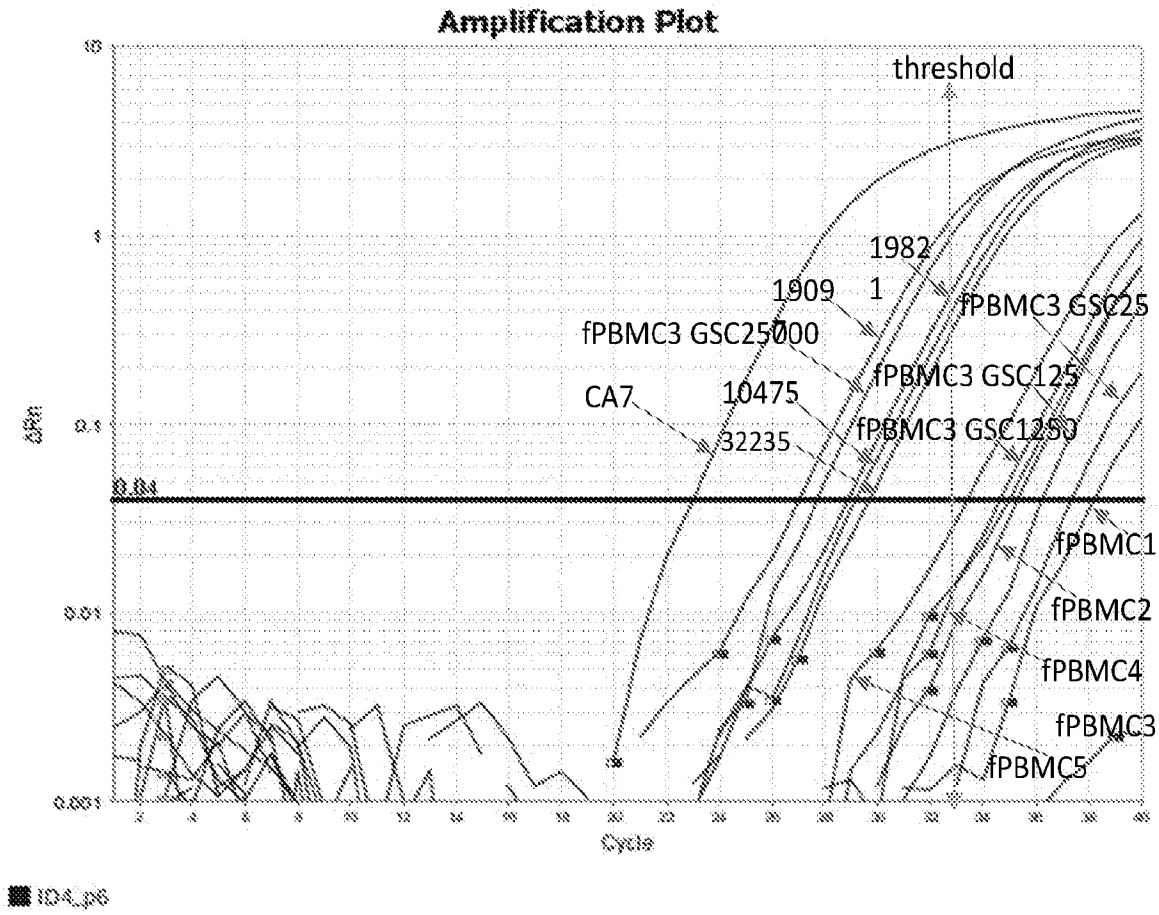

PBMC
GSC
BATF2
Untreated
PBMC
GSC
BATF2
Transposase
PBMC
GSC
OLIG2
Untreated
PBMC
GSC
OLIG2
Transposase
Region 1
Region 2
PBMC
GSC
PBMC
GSC
NKX2-2
Untreated
Region 1
Region 2
PBMC
GSC
PBMC
GSC
NKX2-2
Transposase
Genomic
Input
control
GAPDH
PBMC
GSC
Transposase

FIG. 9
Gene expression approach: ID4



40 cycle samples

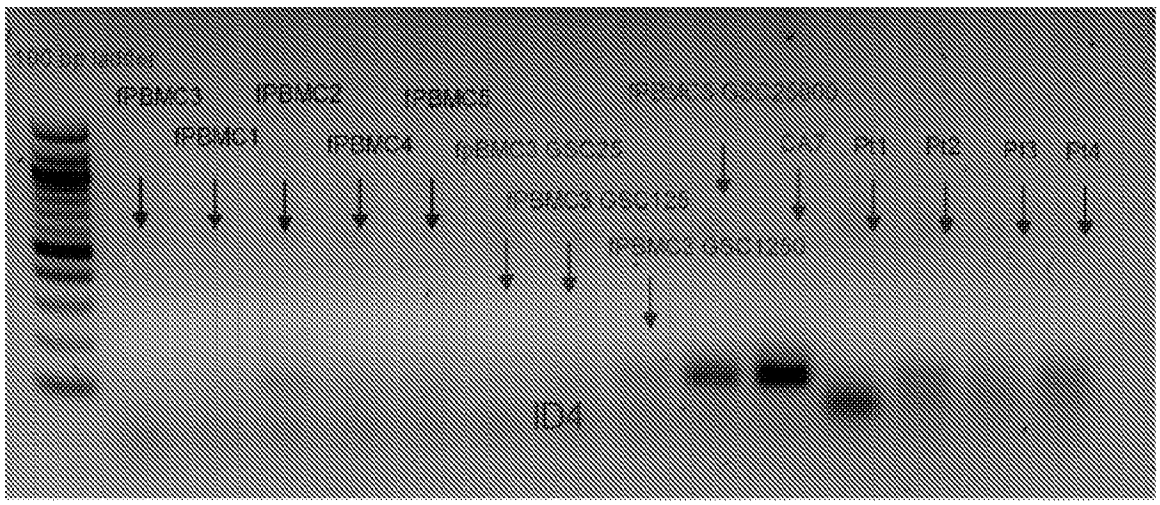
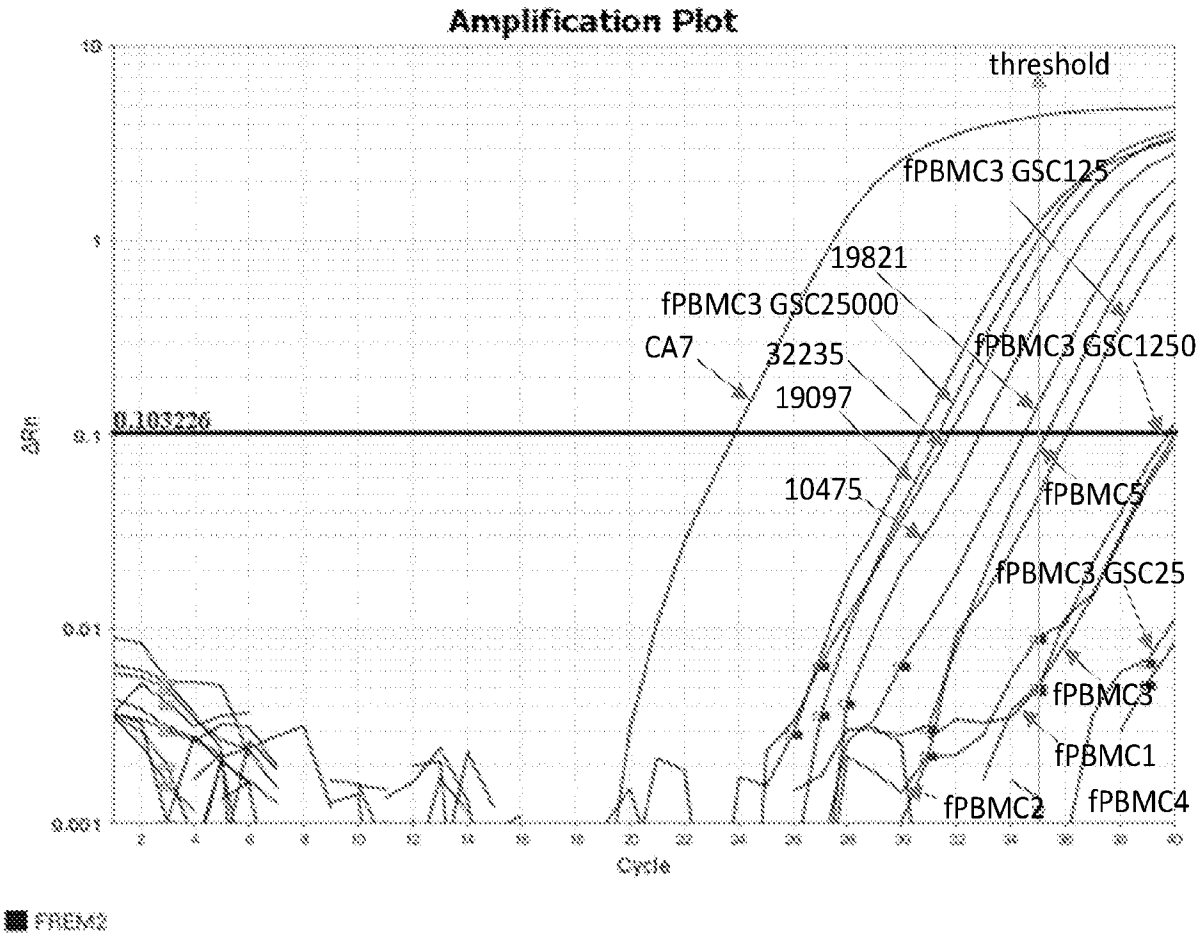


FIG. 10
Gene expression approach: FREM2



40 cycle samples

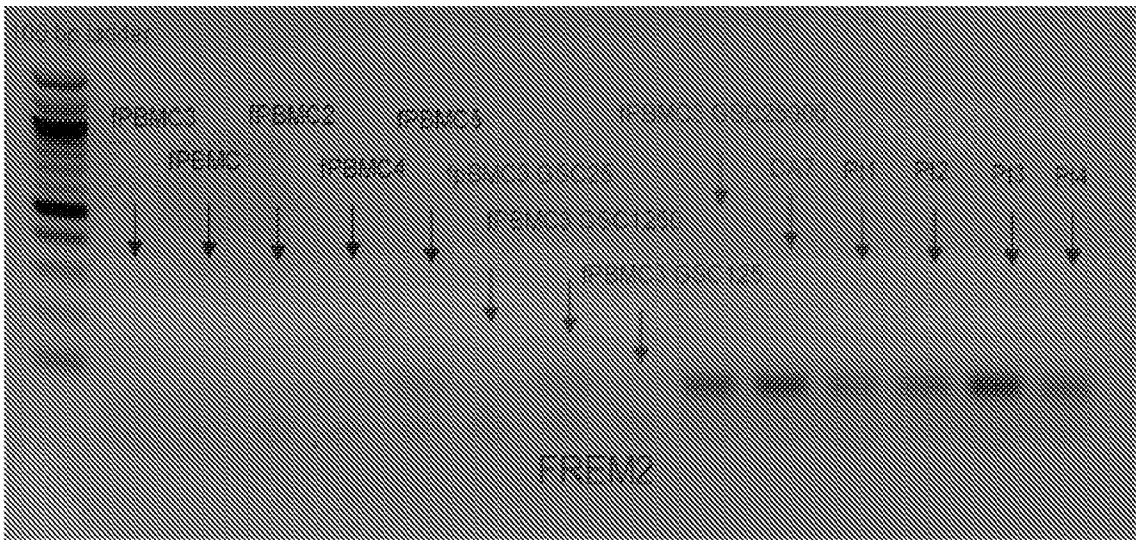
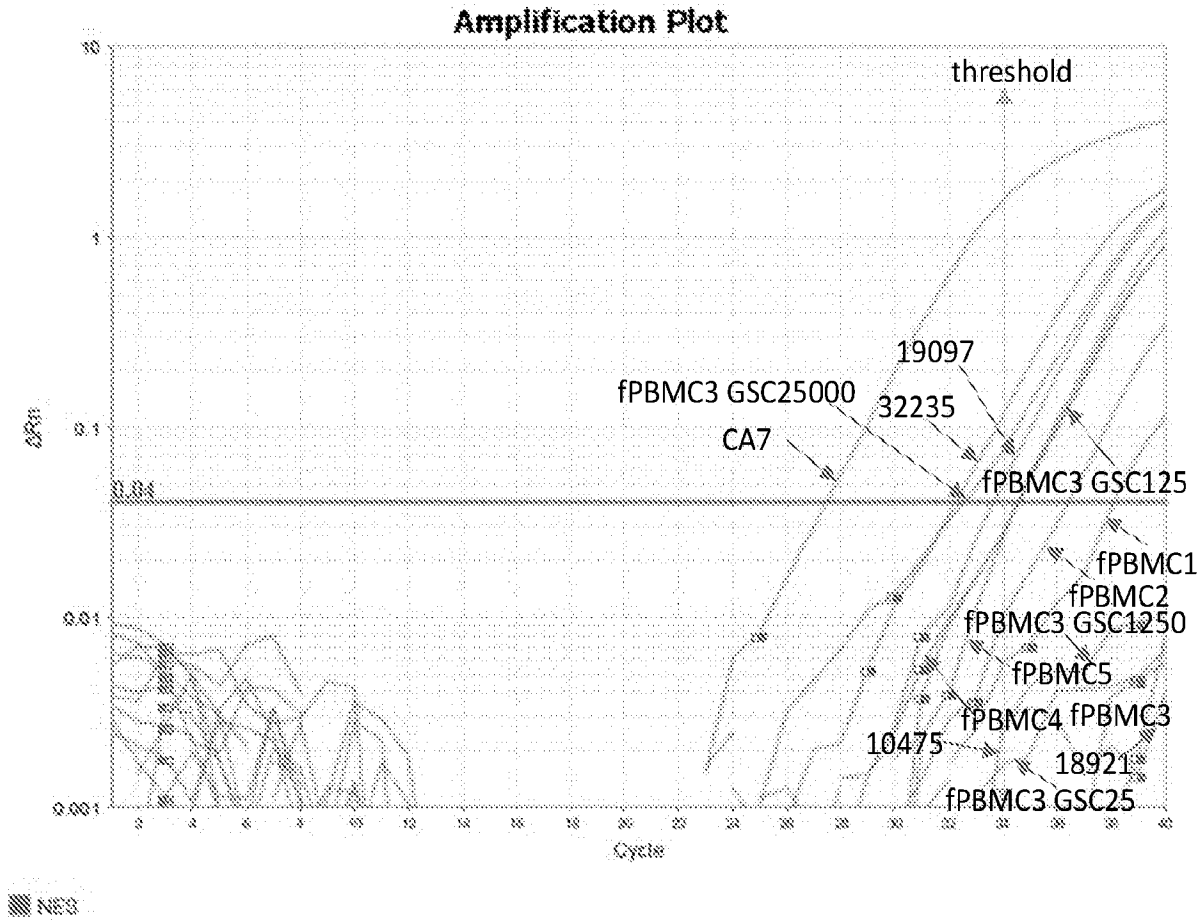


FIG. 11

Gene expression approach: NES



40 cycle samples

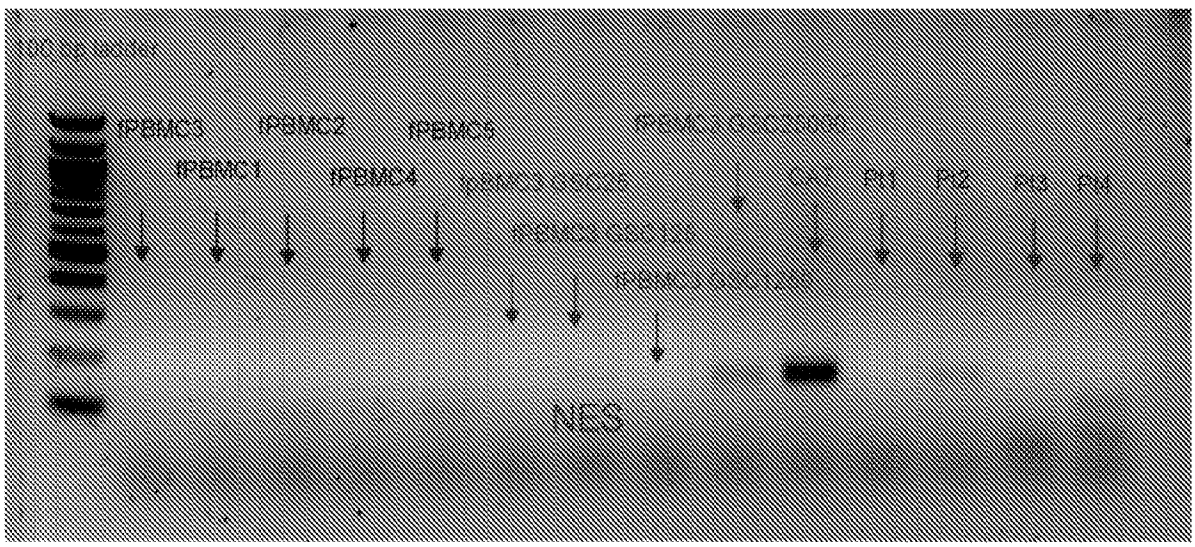
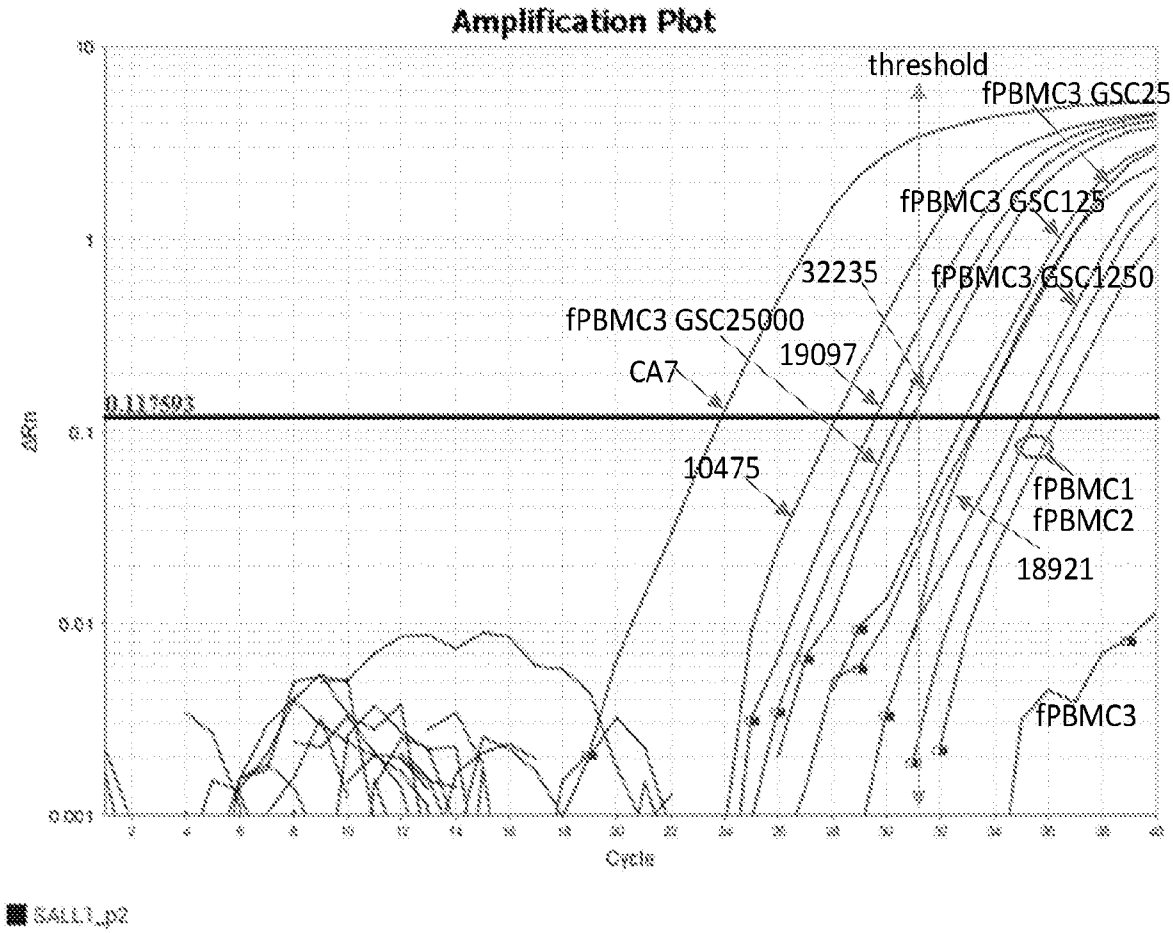
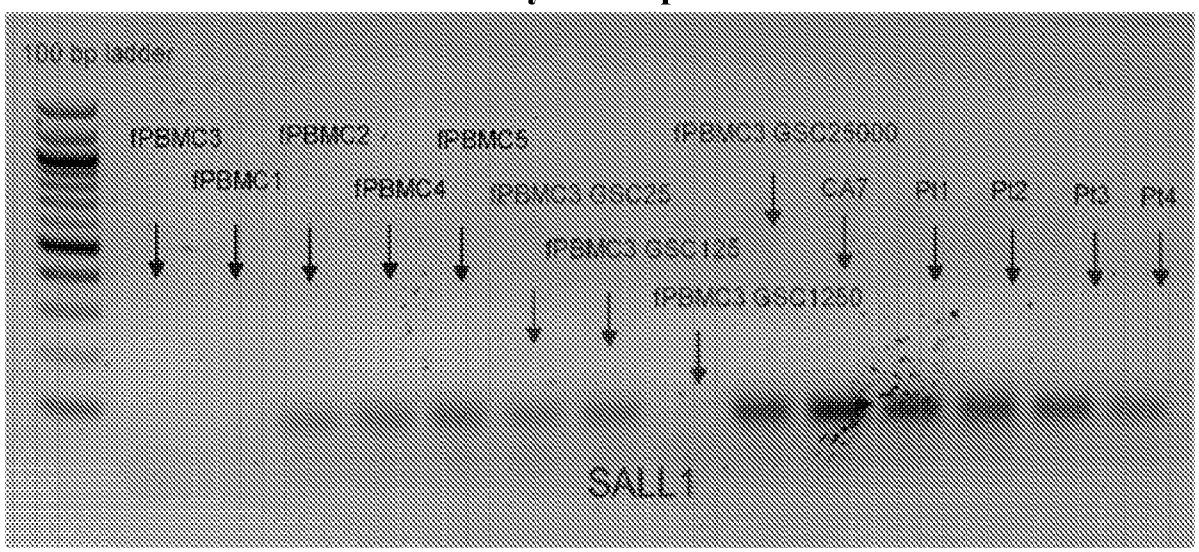


FIG. 12

Gene expression approach: SALL1



40 cycle samples



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US20/17062

A. CLASSIFICATION OF SUBJECT MATTER

IPC - G01N 33/50; C12Q 1/6886; C12N 15/113; G16H 50/20 (2020.01)

CPC - G01N 33/5091, 33/57488; C12Q 1/6886; C12N 15/113; G16H 50/20

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)
See Search History document

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	(LI, M et al.) Discovery and Identification of Serum Potential Biomarkers for Colorectal Cancer Using TMT Quantitative Proteomics. International Journal of Cancer and Oncology. 26 April 2016, Vol. 4, No. 2; pages 213- 219; abstract; page 213, 2nd column, 1st paragraph; page 214, 1st column, 4th paragraph; page 216, 1st column, 1st paragraph and 2nd column, 1st paragraph; DOI: 10.15436/2377-0902.17.1341	6-7, 15-16, 18/15-16, 19/18/15-16 ---------- 8, 17, 18/17, 19/18/17
Y	WO 2018/211409 A1 (UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INC.) 22 November 2018; paragraphs [0016], [0017], [0054], [0056], [0058], [0076]; Figs. 13, 14A	2-3, 4/2-3, 8, 17, 18/17, 19/18/17
Y	US 2010/0099085 A1 (BABAIAN, R et al.) 22 April 2010; abstract; paragraphs [0010], [0015], [0016], [0018], [0023], [0041], [0047]; Fig. 3	1-3, 4/1-3, 9-11, 12/9-11, 13/12/9-11,
Y	(RUGGERO, K et al.) Epigenetic Regulation in Prostate Cancer Progression. Current Molecular Biology Reports. Epub 18 April 2018, Vol. 4, No. 2; pages 101-115; Table 1; DOI: 10.1007/s40610-018-0095-9	1-3, 9-11,
Y	(EHRLICH, M) DNA methylation in cancer: too much, but also too little. Oncogene. 12 August 2002, Vol. 21, No. 35; pages 5400-5413; page 5406, 1st column, 1st paragraph; DOI: 10.1038/sj.onc.1205651	1-3, 4/1-3, 9-11, 12/9-11, 13/12/9-11

☒ 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
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 March 2020 (31.03.2020)	15 MAY 2020
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US20/17062

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	(SPEKTOR, R et al.) Methyl-ATAC-seq measures DNA methylation at accessible chromatin. bioRxiv. 17 October 2018; pages 1-21; page 3, 1st paragraph; page 5, 2nd and 3rd paragraphs; DOI: 10.1101/445486	4/1-3, 12/9-11, 13/12/9-11
A	(GEVAERT, O et al.) Identifying Master Regulators Of Cancer And Their Downstream Targets By Integrating Genomic And Epigenomic Features. Pacific Symposium Biocomputing. 2013; pages 1-13	1-3, 4/1-3, 6-11, 12/9-11, 13/12/9-11, 15-17, 18/15-17, 19/18/15-17
A	(DE SOUZA, CF et al.) A Distinct DNA Methylation Shift in a Subset of Glioma CpG Island Methylator Phenotypes during Tumor Recurrence. Cell Reports.10 April 2018, Vol. 23, No. 2; pages 637-651; DOI: 10.1016/j.celrep.2018.03.107	1-3, 4/1-3, 6-11, 12/9-11, 13/12/9-11, 15-17, 18/15-17, 19/18/15-17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US20/17062

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.: 5, 14, 20-22
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:

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.