

SEQUENCING INDEPENDENT COMPOSITIONS FOR DETECTING CARCINOMAS USING METHYLATION MARKERS

FIELD OF INVENTION

5 [001] The present invention provides sequencing independent quantitative methylation profiling compositions, epigenetic biomarker primers, epigenetic biomarker probes, multiplex compositions, kit, and uses thereof.

BACKGROUND OF THE DISCLOSURE

10 [002] Liver cancer is one of the leading causes of cancer mortality globally. Liver cancer comprises mainly of two primary carcinomas namely Hepatocellular Carcinoma (HCC) and Cholangiocarcinoma (CCA). HCC is one of the leading causes of cancer deaths worldwide and is responsible for more than 800,000 deaths annually representing approximately 80%-85% of liver cancer. On the hand, CCA represent 10%-15% of liver cancer. Treatment options for HCC and CCA usually includes but not limited to liver resection, liver transplant, 15 radiofrequency ablation, radioembolization, small molecule sorafenib and trans-arterial chemoembolization (TACE). However, most cases of HCC and CCA are usually diagnosed at late stages for patients suffering from Chronic Liver Disease (CLD) or because of ineffective blood-based diagnostic or screening methods used that either lack specificity or sensitivity or both. Indeed, both the carcinomas are extremely aggressive in nature compared to other cancers. HCC is reported to have recurrence rate as high as 70% at 5-years and a 20 very low 5-year survival [1]. On the other hand, CCA has very poor prognosis and early detection can improve the overall survival. In some cases, HCC and CCA co-occur called as mixed tumor HCC-CCA and the prognosis is extremely poor.

25 [003] Towards early and accurate detection of HCC, CCA and HCC-CCA, various new strategies are being explored. The current blood-based screening of HCC is carried out by detecting the level of a biomarker called as Alpha-fetoprotein (AFP) from blood. As per the current practices, serum level of AFP at cut-off value above 5 ng/ml – 10ng/ml is considered to be elevated and an indication of carcinoma. However, the AFP suffers from a low sensitivity of approximately 60%-70% and specificity of 70%-90%. Another blood-based 30 biomarker protein induced by vitamin K absence or antagonist-II (PIVKA-II) can also be used as for diagnostic purposes which has a 60-80% sensitivity and 80%-90% specificity. Multiple studies combined power of the two biomarkers have reported the sensitivity and specificity having marginal increase.

[004] In the case of CCA, CA 19-9 is the only blood-based biomarker that is available. Reports have indicated a wide range of sensitivity for CA 19-9 ranging from 50% to 90% and in several cases, is subjective. In addition, CA-19-9 is also elevated in other cancer such as but not limited to pancreatic and gallbladder cancer.

5 [005] With advances in the field of genetics, attempts on diagnosis using mutations were carried out. For several cancers such as lung cancer and breast cancer wherein the mutations are highly specific, diagnosis through mutational detection is a promising way forward. However, hepatocellular carcinoma is an extremely complex genetic alteration process, and occurs from multiple genetic alterations. To use mutational marker for genetic screening for
10 detecting HCC would require screening of multiple mutations in multiple genes for achieving sensitivity levels above the current standards. For instance, a recent study reported 65% sensitivity and 10% specificity using 33 study subjects through detecting multiple driver and novel mutations using Next Generation Sequencing (NGS)[2]. In combination with AFP results, the sensitivity slightly improved to 73%. Numerous studies on HCC detection
15 mutation suggests such screening tool would not only be expensive but also will have only marginal improvement in the sensitivity thereby rendering such screening or diagnostic tool not-so-feasible for clinical translation.

[006] The recent developments in the field of epigenetics are now being explored for application in the field of cancer diagnostics. Indeed, it is now recognized that epigenetic
20 alterations in conjunction with genetic aberrations are the cause for cancer. The epigenetic modifications usually occur through DNA and RNA methylation, histone modifications and non-coding RNAs regulating tumorigenesis and cancer progression. For using epigenetics as a methodology for diagnosis of cancer, cell free DNA from blood has been widely used. Specifically for the case of HCC and CCA, use of cell free DNA and various epigenetic
25 markers have been reported.

[007] JP6369857B2 reported the analysis of methylation status in CpG island present in the promoter region of CELF6 gene using conventional PCR methods and methylation specific PCR.

[008] US10513739B2 reported detection of methylation pattern by specifically using Set-1:
30 cg26668608, cg11606215, cg18196829, cg25459300, cg25574765, cg11397370, and cg11825899, and one or more markers selected from the group consisting of Set-2: cg10428836, cg25754195, cg05205842, cg24067911, cg24067911, cg23211949, cg17213048, cg23461741, cg06482904, cg07459019, cg20490031, and cg01643250 wherein the first set of nucleic acid probes hybridizes to Set-1 markers present in DNA and the second

set of nucleic acid probes hybridizes to at-least one of the Set-2 markers. The invention uses the known method of bisulfite padlock probes combined with deep sequencing for assessment of methylation profiles.

[009] EP2821487B1 reported the analysis of methylation status of CpG site in promoter region of at least one gene ACTG1, EPHA4 and TSC22D1 for determining the status of hepatocellular carcinoma.

[010] CN105755150A reported the use of IGFBP7 gene for detecting HBV related hepatocellular carcinoma using conventional PCR methods.

[011] US10184154B2 reported detecting CCA using multiple methylation markers targeting EMX1, PRKCB, CYP26C1, LOC645323, ZNF781, ST8SIA1, chr7.25896389-25896501, VSTM2B.764, KCNA1, BMP3, SALL1, PTGDR, HIST1H1D, KLHDC7B, LBX2, chr5.77268600, chr6.28175437, PNMAL2, SP9, TRIM36, and RYR2 using genomic DNA.

[012] It is noteworthy to mention that several works included but not limited to the studies cited above have specific conditional requirements for methylation profiling using epigenetic markers such as but not limited to targeting the specific gene promoter region, detecting HCC for specific conditions such as HBV related HCC, or have sequencing conditions that are highly complex requiring techniques such as Next Generation Sequencing, pyrosequencing, or microarray sequencing for analysis. In cases wherein pyrosequencing is required, it is known and reported that analysis using cfDNA from blood usually results in relatively low sensitivity deeming the application to be non-suitable for clinical application [3]. Use of sequencing or deep sequencing is extremely restrictive and limits the scope of screening and diagnostic solutions due to several factors such as but not limited to sample requirements, complexity, cost, and time.

[013] As such improved methods for detecting HCC or CCA or combination thereof including combined HCC-CCA is needed.

SUMMARY OF THE DISCLOSURE

[014] The present disclosure, in certain embodiments, provides composition for Sequencing Independent Quantitative Methylation Profiling, herein referred as qSIMP, comprising of nucleotide sequences targeted towards methylation specific markers including sequencing independent multiplex assay compositions comprising of combinations of multiple nucleotide sequences targeted towards various individual methylation specific markers, kits and uses thereof. For instance, disclosed herein is a qSIMP composition to determine methylation status containing multiple gene targeting primers and probes for quantifying methylation

levels and methylation patterns of subjects to detect the presence or absence of liver cancer including HCC or CCA or HCC-CCA or to prognose HCC or CCA or HCC-CCA or combination thereof.

[015] In some embodiments, also described herein are the uses of Sequencing Independent Quantitative Methylation Profiling (qSIMP) in general and qSIMP compositions to determine methylation status for specific aspects such as but not limited to profiling epigenetic methylations, detecting the levels and patterns of methylation. Specifically, the use of Sequencing Independent Quantitative Methylation Profiling (qSIMP) compositions to determine methylation status with PCR and more specifically, the use of Sequencing Independent Quantitative Methylation Profiling (qSIMP) compositions to determine methylation status with digital PCR for detecting the levels of methylation and patterns of methylation is disclosed herein.

[016] In some embodiments, mentioned herein, are the aspects of Sequencing Independent Quantitative Methylation Profiling (qSIMP) composition to determine methylation status of subjects for prognosing carcinomas. Disclosed herein, are the specific aspects of using Sequencing Independent Quantitative Methylation Profiling (qSIMP) compositions to determine methylation status with PCR methods for prognosing HCC or CCA or HCC-CCA.

[017] Additionally, in some embodiments, also disclosed herein are the tools for deriving risk levels based on Sequencing Independent Quantitative Methylation Profiling (qSIMP) of subjects carried out through qSIMP composition and kits, based on the methylation pattern and methylation levels for use in screening, diagnosing, or prognosing of hepatocellular carcinoma (HCC). In certain aspects mentioned herein, are the uses based on methylation pattern and methylation levels measured using qSIMP compositions with digital PCR to determine methylation status for assessing risk level of subjects who have been subjected to medical intervention or treatment regimen such as but not limited to surgical resection, transplant, chemotherapy, radiotherapy, Transarterial Chemoembolization (TACE), Transarterial Radioembolization (TARE), small molecule drugs, large molecule drugs or combination thereof.

[018] In certain embodiments disclosed herein, is the qSIMP compositions to determine the methylation status using PCR techniques for monitoring the recurrence of HCC in subjects who have undergone medical treatment for HCC. Some aspects disclosed herein is the use of qSIMP compositions to determine methylation status for surveillance of subjects who have a risk of developing HCC and specifically for those subjects who have liver disorders or

subjects with one or more diseases that is known to affect liver or subjects with Chronic Liver Diseases (CLD). The embodiment is extendable to CCA and HCC-CCA.

[019] In certain embodiments disclosed herein, is the contacting qSIMP composition comprising of plurality of primers and probes directed towards epigenetic methylation markers with DNA obtained from a subject at stringent conditions or contacting cell free DNA obtained from the blood sample of a subject with multiplex qSIMP composition comprising of plurality of primers and probes, specifically designed for HCC or CCA or HCC-CCA, to determine the methylation status. The multiplex composition comprises of least two sets of primers and probes targeting two or more regions of same gene or at least one region of two or more separate genes each at stringent conditions. Herein, (i) the DNA obtained from subject is first treated wherein the treatment of DNA comprises of but not limited to bisulfite conversion involving the deamination of unmodified cytosines to uracil, leaving the modified bases 5-mC and 5-hmC intact, (ii) contacting the treated DNA with specific compositions comprising of methylation targeted primers and probes, and amplification of the treated DNA under stringent conditions, (iii) analyzing the amplified DNA region or non-amplified DNA region or combination thereof for determining methylation levels or methylation pattern or combination thereof, hereafter referred to as methylation profile, (iv) using the data obtained for methylation profile as an input to a model wherein the model relates to methylation profiles of HCC subjects and non-HCC subjects or HCC subjects, CCA subjects and non-HCC, non-CCA subjects and, (v) assess the risk status based on the methylation status obtained from the model wherein the risk status can be quantitative or qualitative or combination thereof.

[020] Certain aspects disclosed herein is qSIMP compositions comprising of plurality of primers and probes that are contacted with the treated DNA separately or in combinations wherein at least one set of primer and one set of probe targets hypermethylated gene for determining methylation status. For instance, the plurality of primers and probes referred herein that are in contact with treated DNA targets at least one of the genes selected from RUNX, TBXT, HOX, ST8SIA, SFRP, RGS, PAX, VIM, TWIST and NEUROG specifically comprising of RUNX2, TBXT, HOXD9, ST8SIA6, SFRP1, RGS10, PAX9, VIM, TWIST-1 and NEUROG1

[021] In certain aspects, disclosed herein are the details of the uniqueness of the nucleic acid sequences of primers and probes used in qSIMP compositions used for determining methylation. In certain embodiments, disclosed herein are the details of the kit composition comprising of the primers and probes that hybridize to treated DNA. In some aspects, the

sequences are disclosed. In some more aspects, the composition including but not limited to the concentrations, amount and nature of the compositions are disclosed.

[022] Certain embodiments disclosed herein are for analyzing the obtained methylations profiles and convert to a meaningful qualitative data or quantitative data or combination thereof.

[023] In some aspects, the features of Sequencing Independent quantitative Methylation Profiling (qSIMP) compositions for use in screening, diagnosis and prognosis of carcinomas are disclosed.

[024] The disclosure herein provides a novel and unique composition for facile and cost-effective methylation profiling. Methylation Profiling using Sequencing Independent Quantitative Methylation Profiling (qSIMP) and allied compositions avoids the need for deep sequencing techniques for methylation profiling. Also, qSIMP compositions to determine methylation status circumvents the constraints posed by Bisulfite Sequencing PCR and Methylation Specific PCR especially for liquid biopsies. To our knowledge, there are no studies reported in the literature that allows for methylation profiling (i) without the need of sequencing techniques for any genes wherein the DNA is highly fragmented (cell free DNA) and (ii) without several constraints posed by Bisulfite Sequencing PCR and Methylation Specific PCR due to primer and probe limitations restricting methylation profiling of several regions in genes, and thereby limiting the use of methylation profiling for several applications such as screening, diagnosis and prognosis .

DEFINITIONS

[025] As used herein, the term sample relates to material or mixture of materials, that is used for investigation. Samples referred herewith are but not limited to tissues and fluids obtained from living organisms. Fluids refers to but not limited to blood, serum, urine, saliva and sweat.

[026] As used herein, the term nucleotide represents moieties containing purine and pyrimidine bases such as adenine (A), thymine (T), guanine (G) and cytosine(C), and other heterocyclic bases that are altered to form modified heterocyclic bases through introduction of variations or functionalization such as but not limited to methylation, alkylation, and acylation. The term nucleotide also includes nucleosides.

[027] As used herein, the term nucleic acid denotes deoxyribonucleic acid (DNA), ribonucleic acid (RNA) or mixture thereof. A nucleic acid may be polynucleotide or oligonucleotide. The nucleic acid may be single stranded or double stranded.

[028] As used herein, the term target DNA or target RNA refers to the region of DNA that is of interest under study and region of RNA that is of interest under study respectively. Target DNA and target polynucleotides or target RNA and target polynucleotides are used interchangeably herewith.

5 [029] As used herein, the term treated DNA refers to the DNA obtained after conversion of DNA using any chemical means or any physical means for determining the methylation status, including the levels and pattern.

[030] As used herein, qSIMP or qSIMP compositions are interchangeably used and is related to the compositions required for methylation profiling.

10 [031] As used herein, fragmented DNA or DNA fragment refers to smaller sizes of DNA in length relative to genomic DNA and is denoted in base pairs (bp). The size of fragmented DNA referred herewith can be greater than 2 bp. Generally, the fragmented DNA described herein refers to the DNA strands obtained from samples usually containing cell free DNA (cfDNA) and circulating tumor DNA (ctDNA). cfDNA, Fragmented DNA and Sheared DNA
15 are used interchangeable herein.

[032] As used herein, cell free DNA refers to all non-encapsulated DNA present in the bodily fluids such as blood stream. A portion of that cell-free DNA that originates from tumor is referred herein as circulating tumor DNA (ctDNA).

20 [033] As used herein, the terms double stranded and duplex as used herein, describes two complementary polynucleotides that are base-paired, i.e., hybridized together.

[034] As used herein, the term plurality refers to two or more.

[035] As used herein, the term amplifying as used herein refers to generating one or more copies of a target nucleic acid, using the target nucleic acid as a template.

25 [036] As used herein, the terms determining, measuring, evaluating, assessing, assaying, and analysing are used interchangeably herein to refer to any form of measurement and include determining if an element is present or not. These terms include both quantitative and/or qualitative determinations. Assessing may be relative or absolute. Assessing the presence of includes determining the amount of something present, as well as determining whether it is present or absent.

30 [037] As used herein, the term library preparation refers to the initial steps of deep sequencing using methods such as but not limited to Next Generation Sequencing or microarrays. Library preparation or Library build allows DNA or cDNA to adhere to the sequencing flow cell and allows the sample to be identified. Two common methods of library preparation are ligation-based library prep and tagmentation-based library prep. However,

other forms of library build following similar principles are to be construed as the term referred herein.

[038] As used herein Sets refers to one forward primer and one reverse primer or multiple forward primers and multiple reverse primers or one probe or multiple probes in combination with primers.

[039] As used herein, CpG cytosine is a term referred to the cytosine present in CpG islands that are methylated and does not undergo deamination following bisulfite or allied treatment. In other words, a CpG cytosine upon bisulfite treatment will remain intact as cytosine because of the presence of a methylation group attached to the cytosine nucleotide.

[040] As used herein, non-CpG cytosine is a term referred to the cytosine present in CpG islands that are unmethylated and undergoes deamination following bisulfite or allied treatment to form Uracil U. In other words, a non-CpG cytosine upon bisulfite treatment will not remain intact as cytosine but convert to Uracil because of the absence of a methylation group attached to the cytosine nucleotide.

[041] As used herein, promoter and promoter core used interchangeably is a minimal stretch of DNA sequence that conations transcription start site and sufficient to directly initiate the transcription. The length of core promoter typically ranges between 60 and 120 base pairs (bp).

[042] As used herein, the term model refers to algorithms that uses simple or complex computations using, but not limited to, statistics, logistic regression, cluster analysis, neural network, nearest neighbour analysis, machine learning and artificial intelligence.

[043] As used herein, the term presence or absence refers to an outcome that is quantitatively determined or qualitatively determined or both.

BRIEF DESCRIPTION OF THE DRAWINGS

[044] Figure 1. Methylation profile obtained by targeting individual genes using qSIMP composition for HCC.

[045] Figure 2. Multiplex methylation profile obtained by targeting multiple genes simultaneously using qSIMP composition for HCC.

[046] Figure 3a. Multiplex methylation profile obtained using qSIMP composition for HCC subject and for Non-HCC/CLD subject.

[047] Figure 3b. Risk status derived using a mathematical model for multiplex methylation profile using qSIMP compositions for HCC subject and for Non-HCC/CLD subject.

- [048] Figure 4a. Multiplex methylation profile using qSIMP compositions for an HCC subject before surgical resection of liver and after surgical resection of liver.
- [049] Figure 4b. Risk status derived using a mathematical model for multiplex methylation profile using qSIMP compositions for an HCC subject before surgical resection of liver and after surgical resection of liver.
- [050] Figure 5. Multiplex methylation profile using qSIMP compositions for HCC subject who had undergone locoregional therapy.
- [051] Figure 6. Multiplex methylation profile using qSIMP compositions for a subject (non-HCC) harboring adenocarcinoma of Hepatic Flexure of Colon.
- [052] Figure 7. Risk status derived using a mathematical model for multiplex methylation profile using qSIMP composition for HCC subjects before and after medical intervention at various time points (14 days, 3 months and 6 months) to check for recurrence.
- [053] Figure 8. Box plot analysis of methylation levels obtained using qSIMP ddPCR for HCC subjects and CLD/Non-HCC subjects show a marked difference in the methylation levels between HCC subjects and CLD/Non-HCC subjects.
- [054] Figure 9. Multiplex methylation profile (top) and respective methylation copies (bottom) obtained using qSIMP compositions for CCA (a) clinical samples derived from CCA subjects, (b) clinical samples derived from HCC subjects and (c) clinical samples derived from (non-CCA, non-HCC) CLD subjects.
- [055] Figure 10. 2D Multiplex methylation profile obtained using qSIMP compositions for HCC and CCA for Non-HCC CCA subject.
- [056] Figure 11. 2D Multiplex methylation profile obtained using qSIMP compositions for HCC and CCA for HCC-CCA mixed tumor subject.

DETAILED DESCRIPTION OF THE DISCLOSURE

- [057] With advancements in the field of epigenetics, the interest to use methylation profiles of genes for cancer diagnostics and screening are now rapidly increasing. Hypermethylation or hypomethylation in several genes have been reported for various cancers and attempts for using the profiles for diagnosis have also been reported. For example, Hao et. al. assessed the utility of DNA methylation patterns for differentiating Tumor-Tissue derived DNA from Normal-Tissue DNA for breast, colon, liver, and lung cancer [4]. It is advantageous to use non-invasive methodologies for screening purposes and to this effect, Tomasz et. al. used Septin9 promoter region of cell free DNA obtained from blood for methylation analysis and detected promoter methylation in 44.3% of the study group with a specificity of 92.3% [5].

[058] Use of methylation profiles have also been reported for prognosis of treatments. One study reported the use of whole genome methylation profile from blood in the differentiating leukemic cancers to improve the prognosis and to predict the survival of patients [6]. Zhang and co-workers reported ctDNA methylation markers for diagnosis and prognosis of HCC using deep sequencing method with bisulfite padlock probes that yielded 83.3% and 90.5% sensitivity and specificity respectively in their validation data set [7]. A bisulfite padlock probe comprises of a common linker sequence to connect two variable capture arms that aides in annealing two adjacent genomic regions of treated DNA.

[059] Methylation profiling is carried out in CpG islands of a gene. CpG islands are clusters of CpG dinucleotides in GC-rich regions and often associated with the 5' end of genes and considered gene markers. Methylation of promoter-associated CGIs plays an important role in gene regulation and carcinogenesis. In general, there are two principles used for methylation profiling namely (i) Bisulfite Sequencing PCR (BSP) and (ii) Methylation specific PCR (MSP). In the case of BSP, primers used for amplification of a target gene sequence is always designed wherein, no CpG Cytosines within their sequence is present. The specific design is mandated to negate any discrimination between methylated or unmethylated DNA of interest. BSP is the foundation for several PCR and sequencing studies reported in the literature for methylation profiling. However, the specific criteria of the mandatory absence of CpG cytosine in the primers limits its application for biopsies such as liquid biopsy wherein the DNA is highly fragmented. The high fragmentation of the DNA results in amplicon lengths of 150 base pairs to 200 base pairs in general and amplification of a GC rich region within the amplicon wherein CpG cytosine is negligible relative to non-CpG Cytosine is extremely challenging.

[060] In the case of MSP, two pairs of primers and probes are needed for profiling one site of one gene wherein one pair of primers are directed towards methylated DNA while the other set is directed towards unmethylated DNA. MSP is a two-tube approach wherein two sets of primers are required for analysis. In general, MSP uses two pairs of primers, with one pair specific for amplification of methylated DNA and the other pair specific for amplification of unmethylated DNA. In addition, one of the conditions also involves the use of same CpG region of DNA for the pair of primers limiting the application for multiple CpG regions of multiple genes. However, the major disadvantage of MSP is that the principle can be used only for qualitative analysis.

[061] Considering the cell free DNA obtained from bodily fluids is highly fragmented, random, and usually contains 150-200 base pairs (B.P), it is challenging to design primers

and probes satisfying any of the above criteria for optimal efficiency that includes high sensitivity and high specificity that is required for clinical diagnostic applications. In addition, since it is mandatory to target the CpG sites of the genes of interest wherein the methylation profiling is to be carried out and use of MSP or BSP for amplification and identification of aberrant DNA methylation is extremely restrictive especially for cell free DNA. To increase sensitivity and specificity, molecular inversion padlock probes have been reported that is used in conjunction with deep sequencing for CpG methylation profiling using genomic DNA [8]. The principle of padlock probes is that the designed probes hybridize to the target DNA and thereafter, extension and ligation occurs to form a circularized probe. The DNA is then digested, the probes are amplified, and the extensive library building is carried out prior to deep sequencing [9]. Indeed, preparation of library for sequencing is expensive and has been a bottleneck for rapid scale-up. In addition, methylation profiling is usually carried out using deep sequencing techniques such as next generation sequencing increasing the cost of analysis drastically. Also, the ubiquitous technology used for methylation profiling includes Illumina's Methylation microarray platform wherein the whole genome DNA methylation profiling can be carried out. However, it is known that there are several probes whose signal may be affected due to cross hybridization or due to genetic variation thereby decreasing the sensitivity or specificity of the compositions for the purpose of diagnosis or prognosis. The key requirements for a diagnostic or a screening tool, specifically using methylation profiling include high sensitivity, high specificity, low cost, facile workflows, and ease of scale-up for effective clinical translation.

[062] In some embodiments, disclosed herein are Sequencing Independent Quantitative Methylation Profiling (qSIMP) to assess methylation profiles of genes for use in screening, diagnosis, or prognosis of cancer. For instance, the disclosure provides the compositions and uses thereof using qSIMP compositions to determine methylation status for screening, diagnosing or prognosing HCC using HCC specific methylation markers or using qSIMP compositions to determine methylation status for screening, diagnosing or prognosing CCA or using qSIMP compositions to determine methylation status for screening, diagnosing or prognosing HCC-CCA. The disclosure provides unique compositions for amplification of specific regions of nucleotide sequences present in DNA and readout to obtain methylation profiles, and specifically targeting CpG regions that may also be present in upstream or downstream of promoter core of DNA. In some embodiments, the compositions of qSIMP disclosed herein, target specific genes and may be used to determine the presence or absence

of HCC or CCA or HCC-CCA through the obtained methylation status and comprises of at least one of the gene selected from RUNX, TBXT, HOX, ST8SIA, SFRP, RGS, PAX, VIM, TWIST and NEUROG or combination thereof. In some more instance, the compositions of qSIMP disclosed herein, target specific genes to evaluate methylation status that may be used to determine presence or absence of HCC or CCA or HCC-CCA and comprises of at least one of the genes selected from RUNX2, TBXT, HOXD9, ST8SIA6, SFRP1, RGS10, PAX9, VIM, TWIST-1 and NEUROG1 or at least the combination thereof.

[063] In some embodiments, disclosure herein provides unique forward primers and reverse primers generally referred to as the set of primers for each treated DNA region used for amplification comprised within the compositions disclosed herein. In some embodiments, disclosure herein provides the composition containing unique probes that hybridizes to the amplifiable region enabling absorbance-based, fluorescence-based, or luminescence-based signal readout or a combination thereof. In some embodiments, the disclosure provides a set of primers in combination with at least one probe wherein the nucleotide sequence of untreated DNA comprises of at least 1 CpG Cytosine and at least 1 non CpG Cytosine wherein the set of primers undergoes annealing and amplification of treated DNA and the probe contains at least 1 CpG Cytosine and at least 1 non CpG Cytosine wherein the probe anneals and amplifies to the amplifiable treated DNA for producing signal.

[064] In some embodiments, the disclosure herein provides the composition of a mixture comprising all the primers and probes to carry out methylation profiling of multiple genes of the sample DNA simultaneously also referred as qSIMP multiplex methylation assays.

[065] In some embodiments, disclosed herein is the use of qSIMP compositions in methylation profiling that can be carried out using PCR such as but not limited to RT-PCR, digital PCR including but not limited to droplet digital PCR and without the need of deep sequencing techniques or whole genome sequencing techniques using techniques such as but not limited to Next Generation Sequencing or microarrays. However, amplicon-based methylation profiling using the said compositions with techniques such as NGS or microarrays can be carried out. Importantly, the invention provided herein avoids the use of library preparation step for methylation profiling. The invention provides a novel methylation marker-based composition for determining methylation status without the need of deep sequencing and its associated sample preparation steps to facilitate easy and cost-effective methylation profiling that may be used for clinical applications such as screening, diagnosis and prognosis of subjects suspected of the cancer or suspected of developing the cancer.

- [066] In some instance, disclosed herein are the qSIMP compositions performed using droplet digital PCR for determining methylation status. In some more embodiments, contacting the qSIMP compositions with treated DNA and thereupon subjecting the product to stringent PCR cycles may result in amplification of a single copy of a single gene of treated DNA in individual partitions that may produce signal or amplification of multiple copies of a single gene of treated DNA in individual partitions that may produce signal or amplification of single copy of multiple genes in individual partitions that may produce signal or amplification of multiple copies of multiple genes in individual partitions that may produce signal and wherein the signal is used for determining the methylation status.
- [067] In some embodiments, disclosed herein are the compositions of qSIMP used in conjunction with specific PCR techniques wherein absolute copies of methylated DNA is determined without determining the copies of unmethylated DNA for detecting the presence or absence of carcinomas such as HCC or CCA or HCC-CCA using methylation markers.
- [068] In some embodiments, disclosed herein are the kits and compositions for methylation profiling for use in screening, diagnosing or prognosing the presence or absence of HCC or CCA or HCC-CCA.
- [069] In some embodiments, disclosed herein are the use of qSIMP compositions for determining the methylation status to assess risk level of subjects that may suggest the presence or absence or prognose HCC or CCA or HCC-CCA using multivariate analysis wherein the multivariate analysis uses parameters such as but not limited to copies of methylated genes, amount of DNA and amount of plasma derived from blood and optionally routine blood-derived parameters including but not limited to AFP, CA19-9, and PIVKA.
- [070] It is noteworthy to mention that prior to the detailing of the present invention, it is to be understood that this invention is not limited to embodiments described herein and can be extrapolated to other similar mechanisms and methods known to personnel skilled in the art to which this invention belongs.
- [071] Also, the terminology and definitions used herein are for describing particular embodiments generally, and therefore must not be construed to be limiting. In addition, unless defined otherwise, all scientific terms used herein have the same meaning as commonly understood by one of skilled in the art to which this invention belongs.
- [072] Those skilled in the art of this invention will be aware that the present disclosure is subject to variations and modifications other than those specifically described. It is to be understood that the present disclosure includes all such variations and modifications.

[073] Unless the context requires otherwise, throughout the specification which follow, the word comprise and variations thereof such as comprises and comprising are to be construed in an open and inclusive.

[074] Reference throughout this disclosure to one embodiment or an embodiment means that a particular feature structure or characteristic described in connection with the embodiment is included in at least one embodiment. Thus, the appearance of the phrases in one embodiment or in an embodiment in various places throughout this disclosure are not necessarily all referring to the same embodiment.

[075] As used in the description herein, and throughout the disclosure, the meaning of a, an and the includes plural references unless the context clearly dictates otherwise. Also, as used in the description herein, the meaning of in includes in and on unless the context clearly dictates otherwise.

Screening, Diagnosis and Prognosis of liver cancer using methylation markers.

[076] In some embodiments, disclosed herein is the use of qSIMP composition to determine methylation status for screening a subject for HCC or CCA or HCC-CCA comprising (a) isolating DNA from biological specimen and specifically isolating cell free DNA (cfDNA) from blood or plasma, (b) treating the DNA through a method called as bisulfite conversion involving the deamination of unmodified cytosines to uracil, leaving the modified bases 5-mC (5-methylcytosine) and 5-hmC (5-hydroxymethylcytosine) intact to obtain treated DNA, (c) contacting qSIMP compositions comprising of one or more set of primers and one or more set of probes targeting a specific region of gene, as listed in Table 1, with the treated DNA, (d) amplification of the region present in the treated DNA under stringent conditions wherein the set of primers and probes hybridize, obtaining the methylation profile of the amplified product as a fluorescence, absorbance or luminescence readout, and (f) analyzing the methylation profile using suitable models to evaluate the risk of developing HCC or CCA or HCC-CCA or to detect the presence or absence of HCC or CCA or HCC-CCA.

[077] In some embodiments, one gene or plurality of genes are targeted for methylation profiling. In some cases, the methylation profiling is carried out by targeting at least RUNX2, TBXT, HOXD9, ST8SIA6, SFRP1, RGS10, PAX9, VIM, TWIST-1 and NEUROG1 using a specific mixture of primers and probes designed for hybridizing to the CpG islands of the genes.

[078] In some embodiments, the methylation profiling is carried out by contacting individual sets of primers and probes to treated DNA for hybridizing with individual genes and thereafter combining the output for obtaining overall methylation profile for diagnosing the

presence or absence of HCC or CCA or HCC-CCA or prognosing HCC or CCA or HCC-CCA. In further embodiments, the methylation profiling is carried out by contacting the mixture of multiple primer sets and multiple probes to treated DNA targeting various regions of various genes to obtain multiplexed methylation profile for diagnosing the presence or
5 absence of HCC or CCA or HCC-CCA.

[079] In some embodiments, the biological sample used from the subject is obtained from cells, blood, plasma, serum, bodily fluids, tissues, and the like. In some instance, the DNA is extracted from the biological sample through means known to a person skilled in the art. Specifically, the cell free DNA, that may or may not contain aberrations related to the subject
10 related to this disclosure, is extracted from plasma derived from blood of a subject suspected of having HCC or CCA or HCC-CCA or suspected of developing HCC or CCA or HCC-CCA through means known to a person skilled in the art. In some embodiments, the concentration and the amount of extracted cell free DNA is quantified through absorbance or florescence or luminescence or combination thereof.

15 [080] In some cases, the biological sample comprises of tissue biopsy sample. In some other cases, the biological sample comprises of circulating tumor cells.

[081] In some embodiments, the cell free DNA is treated with deamination agents, also known as bisulfite conversion, to obtain treated cell free DNA generally referred to as treated DNA. In some embodiments, the treated DNA is purified to remove agents used for obtaining
20 treated DNA using methods known to a person skilled in the art.

[082] In some embodiments, the treated DNA is contacted with sets of primers and probes under stringent conditions for hybridization, amplification, and signal readout. In some embodiments, the treated DNA is contacted with sets of primers and probes under stringent conditions and droplet digital PCR method is used for hybridization, amplification, and signal
25 readout.

[083] In some embodiments, the methylation profile is generated using single model or multiple models to determine the status of carcinoma. Specifically, the results obtained from the PCR method is input to the model to generate methylation profile to diagnose or prognose HCC or CCA or HCC-CCA. In some embodiments, the model generates an output based on
30 the methylation profile suggesting extent of the cancer or stage of the cancer.

[084] In some embodiments, the methylation profile is used to differentiate a subject harboring carcinoma originating due to HCC or CCA or HCC-CCA from a subject harboring carcinoma that is not originating from hepatocellular carcinoma HCC or CCA or HCC-CCA. In some embodiments, the methylation profile is used to differentiate a subject harboring

other liver-based cancer such as but not limited to CCA from a subject harboring HCC or HCC-CCA.

[085] In some embodiments, the methylation profile is used as a treatment response marker. The treatment includes but not limited to surgical resection of liver, liver transplant, trans
5 arterial chemoembolization (TACE), trans arterial radioembolization (TARE), chemotherapy or radiotherapy.

Gene panels and detection

[086] In some embodiments, disclosed herein is the use of qSIMP composition to determine methylation status for screening a subject for HCC or CCA or HCC-CCA comprising (a)
10 isolating DNA from biological specimen and specifically isolating cell free DNA (cfDNA) from blood or plasma, (b) treating the DNA through a method called as bisulfite conversion involving the deamination of unmodified cytosines to uracil, leaving the modified bases 5-mC (5-methylcytosine) and 5-hmC (5-hydroxymethylcytosine) intact to obtain treated DNA, (c) contacting the treated DNA with a qSIMP composition comprising of at least one or more
15 set of HCC or CCA specific primers and one or more sets of HCC or CCA specific probes to obtain amplicons for methylation profiling using HCC or CCA methylation markers. The sequence of resultant amplicons generated using qSIMP compositions are listed in Table 2, (d) amplification of the region present in the treated DNA under stringent conditions wherein the set of primers and probes hybridize, (e) obtaining the methylation profile of the amplified
20 product as a fluorescence, absorbance or luminescence readout, and (f) analyzing the methylation profile, specifically hypermethylated gene profile, using suitable models to evaluate the risk of developing HCC or CCA or HCC-CCA or to detect the presence or absence of HCC or CCA or HCC-CCA or combination thereof.

[087] In some embodiments, qSIMP compositions can be generated for targeting other
25 relevant genes as a biomarker for methylation profiling in addition to the gene biomarkers and sequences as listed in table 1 and table 2.

[088] In some embodiments, disclosed herein is the uniqueness of the compositions comprising a set of primers and probes that are contacted with treated DNA to obtain methylation profile is carried out without the need of using a set of unmethylated specific
30 primers and probes that are contacted with treated DNA.

[089] In some embodiments, multiple primers, multiple probes, and PCR reagents are combined in appropriate proportions to obtain a qSIMP composition to obtain methylation profiles by performing PCR reactions.

[090] In some embodiments, the detection after amplification is carried out by PCR methods including but not limited to reverse transcription PCR, digital PCR, droplet digital (ddPCR) and ligation mediated PCR. PCR reagents are usually available from commercial entities. Machines used for specific PCR are also commercially available.

5 [091] In certain embodiments, the extent of methylation is determined using copies of treated DNA that are methylated. The obtained methylation values measured for specific biomarkers are combined with other relevant parameters such as but not limited to amount of cell free DNA, amount of blood, amount of plasma, and characteristics of plasma to obtain methylation pattern.

10 [092] In certain embodiments, the methylation values combined with methylation profile and is analyzed for determining the presence or absence of HCC or CCA or HCC-CCA or to prognose HCC or CCA or HCC-CCA. In some embodiments, the subject's parameters or methylation profile or combination thereof is used for building models using various mathematic models such as but not limited to regression models, K-nearest neighbor
15 classifiers, machine learning, neural networks, and artificial intelligence to determine the presence or absence of HCC or CCA or HCC-CCA or to prognose HCC or CCA or HCC-CCA.

Composition and Kits

[093] In some instance, provided herein are the compositions and kits for determining
20 methylation profile by targeting the CpG regions of any of the multiple genes selected from RUNX2, TBXT, HOXD9, ST8SIA6, SFRP1, RGS10, PAX9, VIM, TWIST-1 and NEUROG1. For instance, the primers and probes of the qSIMP composition used for determining methylation status to detect the presence or absence of HCC or CCA is listed in SEQ ID NOs: 001-133. In some embodiments, the kit contains qSIMP reagents comprising of
25 plurality of primers in combination with plurality of probes directed towards multiple regions of a single gene as mentioned or multiple regions of multiple genes as mentioned to detect multiplex methylation profile and status of samples. In some embodiments, at least one set of primers in combination with at least one probe when in contact with aberrantly methylated treated DNA region results in amplification of the DNA and allows for signal readout using
30 PCR reagents and methods.

[094] In some embodiments, the concentration of individual primers ranges from 5nM to 1500nM in the plurality of the mixture present in the kit and the concentration of probes ranges from 5nM to 1500nM in the plurality of the mixture present in the kit. In some embodiments, the concentration of individual set of primers and probes targeting one specific

region of the gene may be different from the concentrations of other primers and probes present in the mixture. In some embodiments, the kit uses equimolar concentration of all primers and equimolar concentrations of all probes.

[095] In some embodiments, the kits also include buffering agents, stabilizing agents, enzymes, reaction mixture generally referred to as PCR supermix and other components required for stability and efficiency of the kit including but limited to thermal stability for preserving or extending the shelf-life of the product.

[096] In some embodiments, the kit also includes unmethylated DNA control to serve as negative control, methylated DNA control to serve as positive control and negative template controls.

EXAMPLES

[097] The examples provided in the disclosure is for illustrative purposes only and does not limit the scope of the disclosure.

Example 1. qSIMP based methylation levels of individual biomarker genes.

[098] 250ng of DNA was taken and bisulfite conversion was carried out. The treated DNA was eluted in 50 μ l of elution buffer to obtain treated DNA concentration of approximately 5ng/ μ l. 0.5 μ l (2.5 ng) of the treated DNA was contacted with 1000nM of gene specific unique set of primers and 250nM of unique probes 1000nM of gene specific unique set of primers selected from SEQ ID 1-35 and 250nM of unique probes selected from SEQ ID 90-106 individually and along with other requisite droplet digital PCR reagents including supermix (no dUTP). Droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Methylation levels of individual genes in four separate reactions was carried out using droplet digital PCR and the methylation levels and pattern of individual genes are shown in figure 1 for representation. Each individual dots above the horizontal threshold line represent hypermethylated copies of the genes.

Example 2. Amplification of methylated DNA.

[099] The optimized PCR cycle and the reagents pertaining to amplification of methylated

DNA for one standard and one sample is as follows:

Reagent	Volume (μl)	Final Concentration
Standard	1	-
RNAse free water	9	-
qSIMP Composition	1.2	1X
Droplet PCR Supermix, 200 reactions	10	1X
Total Volume	21.2	

Reagent	Volume (μl)	Final Concentration
Sample	10	-
RNAse free water	0	-
qSIMP Composition	1.2	1X
Droplet PCR Supermix, 200 reactions	10	1X
Total Volume	21.2	

Steps	Temperature (°C)	Cycle	Time
Initial denaturation/Enzyme activation	95	1	10 min
Denaturation	94	40	30 s
Annealing/Extension	64	40	1 min
Enzyme deactivation	98	1	10 min
Hold	4	-	∞

5 **Example 3. Multiplexed methylation profile for plurality of biomarker genes.**

[100] 250ng of DNA was taken and bisulfite conversion was carried out. The treated DNA was eluted in 50 μl of elution buffer to obtain treated DNA concentration of approximately 5ng/ μl. 0.5 μl (2.5 ng) of the treated DNA was contacted with 1000nm of primers selected from SEQ ID 1-35 and 250nm of probes selected from SEQ ID 90-106 and along with other

10 requisite droplet digital PCR reagents including supermix (no dUTP). Droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Four gene multiplexing analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of all genes as shown in

figure 2 for representation. Each individual dots above the horizontal threshold line represent hypermethylated copies of the genes.

Example 4. Multiplexed methylation profile using plurality of biomarker genes for HCC subject and non-HCC subject.

- 5 [101] 20 ml of blood was collected from subjects and about 10 ml of plasma was isolated. cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition comprising of 1000nm
- 10 of primers selected from SEQ ID 1-35 and 250nm of probes selected from SEQ ID 90-106 was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Four gene multiplexing analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of HCC subject and non-HCC/CLD subject as shown in
- 15 representative figure 3a. Each individual dots above the horizontal threshold line represent hypermethylated copies of the genes. Based on the methylation copies obtained from qSIMP ddPCR, a mathematical model was used to assess the risk status of subjects as shown in figure 3b.

Example 5. Multiplexed methylation profile using plurality of biomarker genes for HCC subject before and after surgical resection.

- 20 [102] 20 ml of blood was collected from subjects and about 10 ml of plasma was isolated before surgical resection of liver and 14 Days after surgical resection of liver. Cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA
- 25 Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition comprising of 1000nm of primers selected from SEQ ID 1-35 and 250nm of probes selected from SEQ ID 90-106 was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle.
- 30 Four gene multiplexing analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of an HCC subject before and after surgical resection of liver as shown in representative figure 4a. Each individual dots above the horizontal threshold line

represent hypermethylated copies of the genes. Based on the methylation copies obtained from qSIMP ddPCR, a mathematical model was used to assess the risk status of the subject before and after surgical resection as shown in figure 4b.

Example 6. Multiplexed methylation profile using plurality of biomarker genes for HCC subject after TACE.

[103] 20 ml of blood was collected from subjects prior to surgical resection but who had undergone Transarterial chemoembolization (TACE) previously. About 10 ml of plasma was isolated from blood and cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition comprising of 1000nm of primers selected from SEQ ID 1-35 and 250nm of probes selected from SEQ ID 90-106 was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Four gene multiplexing analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of HCC subject after Transarterial chemoembolization (TACE). Explant analysis of the subject through histopathaology showed 100% non-viable tumor. The methylation profile for the subject pre-surgery/post-TACE is shown in figure 5 for representation. Each individual dots above the horizontal threshold line represent hypermethylated copies of the genes. Based on the methylation copies obtained from qSIMP ddPCR, a mathematical model was used to assess the risk status of the subject (Validation Set) Pre-resection/Post-TACE and was found to be Low Risk.

Example 7. Multiplexed methylation profile using plurality of biomarker genes for HCC for subjects harboring other carcinomas.

[104] 20 ml of blood was collected from a subject harboring adenocarcinoma of hepatic flexure of colon and about 10 ml of plasma was isolated. cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition comprising of 1000nm of primers selected from SEQ ID 1-35 and 250nm of probes selected from SEQ ID 90-106 was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and

the mixture was then subjected to an optimized PCR thermal cycle. Four gene multiplexing analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of the subject. The methylation profile for the subject is shown in figure 6 for representation. Each individual dots above the horizontal threshold line represent hypermethylated copies of the genes. Based on the methylation copies obtained from qSIMP ddPCR, a mathematical model was used to assess the risk status of the subject harboring adenocarcinoma of hepatic flexure of colon (Validation Set) and was found to be Low Risk.

Example 8. Multiplexed methylation profile as a surveillance tool for monitoring HCC recurrence using plurality of biomarker genes.

[105] 20 ml of blood was collected from subjects and about 10 ml of plasma was isolated before surgical resection of liver or liver transplant, 14 days post-surgical resection of liver or liver transplant, 3 months post-surgical resection or liver transplant and 6 months post-surgical resection or liver transplant. cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition comprising of 1000nm of primers selected from SEQ ID 1-35 and 250nm of probes selected from SEQ ID 90-106 was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Four gene multiplexing analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of HCC subjects before and after surgical resection of liver at various time points for surveillance and the risk status derived using a mathematical model is shown in representative figure 7.

Example 9. Multiplexed methylation profile-based Box-Plot analysis.

[106] Methylation levels and patterns obtained using qSIMP ddPCR 4 gene multiplex HCC assay is shown as box plot for HCC subject and Non-HCC/CLD subjects (Figure 8). Student t-test for 3 methylation patterns as shown in Figure 8 was carried out and the result was found to be extremely significant with $p < 0.0001$.

Example 10. Multiplex methylation profile using qSIMP composition for CCA using samples from CCA subjects, HCC subjects and CLD subjects.

[107] 20 ml of blood was collected from subjects and about 10 ml of plasma was isolated. cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing

purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition comprising of primers selected from SEQ ID 31-89 and probes selected from SEQ ID 91-133 was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Analysis using droplet digital PCR was carried out to obtain methylation levels and patterns of CCA subjects, HCC subjects and non-HCC/non-CCA/CLD subjects as shown in representative figure 9. Each individual dots above the horizontal threshold line represent hypermethylated copies of the genes. Based on the methylation copies obtained, a mathematical model was used to assess the risk status of subjects.

Example 11. Multiplex methylation profile using qSIMP composition for HCC and CCA using sample from CCA subject.

[108] 20 ml of blood was collected from the subject and about 10 ml of plasma was isolated. Cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition for HCC and CCA was contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Analysis using droplet digital PCR was carried out to obtain methylation levels and 2D methylation patterns of CCA subject as shown in representative figure 10. Each individual dots in the respective quadrant represent hypermethylated copies of the genes except for the blank copy quadrant. Based on the methylation copies obtained, a mathematical model was used to assess the risk status of subjects.

Example 12. Multiplex methylation profile using qSIMP composition for HCC and CCA using sample from HCC-CCA subject.

[109] 20 ml of blood was collected from the subject and about 10 ml of plasma was isolated. cell free DNA was extracted using Qiagen nucleic acid extraction kit. The eluant containing purified cell free DNA was subjected deamination process for bisulfite conversion using Zymo EZ DNA Methylation-Gold kit. 10 µl of the DNA containing theoretical DNA amount more than 10 ng was taken for ddPCR reaction. qSIMP composition for HCC and CCA was

- contacted with the DNA and droplets were generated according to the manufacturer's protocol (Bio-Rad) and the mixture was then subjected to an optimized PCR thermal cycle. Analysis using droplet digital PCR was carried out to obtain methylation levels and 2D methylation patterns of HCC-CCA subject as shown in representative figure 11. Each individual dots in the respective quadrant represent hypermethylated copies of the genes except for the blank copy quadrant. Based on the methylation copies obtained, a mathematical model was used to assess the risk status of subjects.

Table 1

Ref Gene	Flanked Amplicon Sequence	SEQ ID NO	Base pairs	Non-CpG C	CpG C
RGS10	TCGGGAGTTTGAAGTTGGGCGGGGCGGGTCGGGG CGGTATTTTAGGCGAGGGTTAGGGATTTTCGGGAC GCGTTAATCGTTTTTCGCGGGGTTATTAAGTT	134	99	12	8
	GGGGGAGTACGGTTGGATTTCGGGTATTCGATTGG GTTTTTGGGTTTGGGGCGTCGGGGGTTTTTCGTTTTT GTCGCGAGGAGTTAAA	135	86	14	8
	GTCGGTTGAGTAGGAAGCGGTCGTCGTTAGGTAA GCGGTTTCGGGGGTTTCGCGGGTTTCGGGGTCGGGG CGTCGTTTCGTTTCGTTTCGGTTTTTTGTTTT	136	98	18	9
	AAGGCGTAGCGTTTACGAAGTCGAGGAGTAGTTA GGTTTTAGTAGGCGGGGATTTTGGTGACGAGGAG GTAAAGGTTATAGGTTG	137	85	19	6
	GATTGGAGTTTCGGTCGATCGTATTATTTTTTTTCG TTGGGTATCGTATTTTCGCGGTTTTAGTCGCGACG GGAGCGTAGAGTTTTTTTCGGGAG	138	94	19	9
	TTAGATTGGAGTTTCGGTCGATCGTATTATTTTTTT TCGTTGGGTATCGTATTTTCGCGGTTTTAGTCGCG ACGGGAGCGTAGAGTTTTTTTCGGG	139	95	17	9
	TTGTAGAATTTTTTAGATTTTTTCGAGTTTTGGCG TTCGCGCGGTACGGTCGGGGAGAGGGCGTTATTC GCGGCGGTTTCGGTTAGTTTAGCGTAC	140	97	20	8
	CGGGGTTGGATTGTTGAATTTATATCGTTTTATTAT CGGATTTTAATTGGGAGGGGGCGTTAGCGTATTTA GGACGCGGTGTTTTAAGGGATTTT	141	95	23	6
RUNX 2	CGGGGTTGGATTGTTGAATTTATATCGTTTTATTAT CGGATTTTAATTGGGAGGGGGCGTTAGCGTATTTA GGACGCGGTGTTTTAAGGGATT	142	93	22	6
	GGGAAAGAGGATGGAGGTGATGGGGTTGAGGAG GTGAGAAATTAAGTTTGATGAGGTCGATTGTCGG GCGGAGTTTGATGGGTTAATAGAGTTTGCGCA	143	99	9	4
	TTTAGGTTTCGGAGTAGGAAGGTTTCGAGGGCGTTT TTATTTTATTCGTTTATTTTTTTCGTTTTTTCGTTAGG TTTTTATTGGTTGGCGCG	144	90	33	8
VIM	TTTAGGTTTCGGAGTAGGAAGGTTTCGAGGGCGTTT TTATTTTATTCGTTTATTTTTTTCGTTTTTTCGTTAGG TTTTTATTGGTTGGCGCG				

	GGTAAGTCGCGTTTTGGGGGATTTTTACGAGGAGG AGATGCGGGAGTTGCGTCGGTAGGTGGATTAGTT AATTAACGATAAAGTTCGCGTCGAGGTGG	145	97	12	10
	CGCGAGAGTTGTTACGTTTTTGGGGATGTGGTCGG GGGGAGGTTTGTAGGGAGATAGCGGAGAGCGGG GTTGTGGTTGTGGTGCGTAGTTTC	146	94	14	7
	CGAAGTTTCGTTGAAATTTGTCGAGGGTAGTAGGT TTGAAAGTTGTAGGCGTTAGTTGCGCGGAGGTGG CGTAGTTGTTTTGGAGGCGTAGA	147	92	14	8
	TGTCGAGGGTAGTAGGTTTGAAAGTTGTAGGCGTT AGTTGCGCGGAGGTGGCGTAGTTGTTTTGGAGGC GTAGAGCGAATAC	148	82	10	7
ST8SI A6	TCGTAGGGAGGCCGATTTTTTCGATCGGTGGACGGC GGTTGGGGAGTAGAGGGATCGGTTAGGGTTGCGG GGGTTTTAGCGCGTATCGT	149	87	14	10
	GCGGTTGTATCGGTTTGGATTTCGTTGGGAGTAGGT AGGTTGGGACGGCGCGAGTGAGTGTAAGAGTGTG CGTGTGTTTGAAGAGG	150	85	7	5
	TCGGTTTGGATTTCGTTGGGAGTAGGTAGGTTGGGA CGGCGCGAGTGAGTGTAAGAGTGTGCGTGTGTTT GAAGAGGGCGCG	151	81	5	7
TBXT	TTGTTTTTGGTCGTCGTTTTTTCGAGAAAAGGGGT TTTTTGATCGAGATTTGCGACGGTTTTTCGGGTTT CGGGTTTCGGTATAGATTCGGGAGGAGG	152	98	27	10
	GGTCGCGTACGTTTTGAAGTGTCGGGTAGTTTTTT ATTGGTCGAGAGCGGTTATATTAGATTTAGTCGGG CGGGTTTGGGAGGTCGGGGGTTAATAATGGGTTTT CG	153	107	26	10
	AGGAGGGATTTTCGTTTCGGGATTCGAGGGATCGA TTCGTTTCGCGGTATGCGGAGGGTTGTTAGGAGTT AG	154	71	15	8
	TTTTTTTTAGATGGTGAGAGTCGCGGGGATATTCG ACGTCGGGGTAGGTTGATTTACGATTTTGGGTGTG CGTAACGTCGTTTGGGGTTTCGT	155	93	20	10
	CGCGGGAAGAGTTTGTAGTATCGAGTGGAATTATT TGTTGAGCGTCGTGGAGAATGAGTTGTAGGCGGG TAGCGAGAAGGGCGATTTTAT	156	90	18	8
	AATTATCGTTGGAAGTACGTGAACGGGGAATGGG TGTCGGGGGGTAAGTCGGAGTCGTAGGCGTTTAG TTGCGTTTATATTTATTTTCGATTCGTTTAATTTTCGG GGTTT	157	109	28	11
	TTAGGTAGGTGTGAGGAGTTCGTTAGGTTTGGGAT TAGCGAGGAGGGGGTCGCGTTGCGTTTTTTGTACG GGGGTTGTGG	158	80	15	6
HOXD 9	AGGTGTAGTTTGCGAATTAGTCGGTGGTTTCGGGC GTCGGCGGGGAGTTGTTTCGGCGGCGGATAGTGTA ATGTTGGGTGGGAGTGCGGGACGTTTTAAATGTT	159	104	14	11
	TGTTTTTAGTGTATTGGACGCGTTGTTTTTTTTTTT GAAGGTTGGGTTTCGCGTGGGCGGTTCGCGGGTGGT	160	86	22	8

	GGTTTTTTCGGTTTTT				
	GCGGGATTTTCGGTGGAGATTTTTAGTGAGGTTTAG AGGAGTTTAGGGTTTCGGGCGGGTTGGGGTTTGT TTTAGTGATTGGACGCGT	161	89	16	7
	GGTTTATGTGTTTCGGTTCGGTTCGGTTCGATTGGT TTGGTTCGTTTAGCGGTTGTTAGGATTATAGACGC GATCGTATTTAGAGGAGGCG	162	91	28	10
	GTAGAGGTTCGTTTATCGTCGTCGTTGTCGGGGGTT GGGAGGGTATCGCGGATTACGAGGTAGAGAATAC GGTCGATTTCGGGGTAAAATTATACGAGGC	163	108	19	12
SFRP1	TAAGGCGTAGAGTTGTCGTTTTTCGGGTTAGTTTCG TCGTGATTTTGGGTTCGCGAGGGGCGTTGAGCGATA TTTGGGATAGATTAGACGCGT	164	87	23	11
	TTCGATGTTTATGTCGGTTTTCGTTTTGTTTTTCG CGACGTCGGGGTTGTTTTTCGTCGTTTTTCGCGCG CGTTTTGTCGTAAATTTTAGGGATTTTCGGGG	165	104	33	15
	GTACGAGTCGCGTACGGTTTCGTAGAGTTAGCGAT ACGGGTAGATGGGTTCGGTTTAGGTAGACGGGCGC GAAGAGCGAGTAGAGGAAGATTGGGTGTC	166	99	16	13
TWIST 1	TTAGGTTTTTTGGAAACGGTGTCGGTGTTGTAGAG TTCGCGAGGTGTTTGGGAGTTGGGCGAGAGTTGTA GATTTGGAGGTT	167	82	16	5
	TTCGTTTACGATTGGGTGCGTTGGCGTTTTTCGTA CGTTGGTTATGATTCGTTGCGTTTGTAGTTTTTCGT AAGATTGCGGATTTTCGTCG	168	91	23	12
	GTTAACGTGCGGGAGCGTTAGCGTATTTAGTCGTT GAACGAGGCGTTCGTCGCGTTGCGGAAGATTATTT TTACGTTGTTTTTCGGATAAGTTGAGTAAGA	169	96	23	13
PAX9	GTTTTTCGTTTTTCGGAGTCGGTACGGGGGTTGGTGG GTGGCGGGATGAATTCGGGGAAGGCGATAGGTTG GAAACGTAGTTTTTCGA	170	86	13	8
	ATTTTATTTTTTCGTAGGGTTTCGGTTCGTGTCGTTGT CGGATCGTAGAACGTCGGGACGTATTTTCGAGTTC GGTTCGATTTAATTCGT	171	89	25	13
	CGTTGTTTTTTCGCGTTTTTTCGCGGGTCGTTTTAG CGTCGGGAAGGGGGCGTCGTAGGTGATTGTTTCGC GGTATTGGGACGGTTGATAGATT	172	93	19	13
	TCGGTATTTTCGTTTGGGAGATTCGGGATCGTTTG TTGGCGGACGGCGTGTCGATAAGTATAATGTGTT TTTCGTGA	173	78	16	9
	GAGGTCGTTGTTTAAACGTTATTCGGTTTCGTATCG TGGAATTGGTTTAAATGGGTATTCGATCGTGTGAT ATTAGTCGTTAGTTACGGGTTTCGTACGGTTGC	174	103	27	12
NEUR OG1	GATTAAGGGTTGTTCGGAGGCGGTGTTTCGGGAGC GTTTTTGTGTCGTCGTAGTGCGTTTTTGTGTTTCG GTTTTT	175	76	25	8
	TTAAAATCGAGACGTTGCGTTTCGTTTATAATTAT ATTTGGGTTTTGGTCGAGATATTGCGTTTGCCGGA TTAAGGGTTGTTTCGG	176	85	21	8

AGGTTTTTAAGGCGGGTTGGTATCGTTGCGTTGTGT AGGATCGACGGATAGATAGAAAGGCGTTTAGAGC GTTGTAGTTCGGATTGAGGGTA	177	91	16	8
GTAGGTAGGGGACGTATTGCGGCGGTAGGAGGCG TTTTCGGGTATCGTTTTCGGGTAGTTTTTGATTTCGT TAGGCG	178	76	20	9

Table 2

Ref Gene	Methylation Specific Primer/Probe Sequence	Base Pairs	Non-CpG C	CpG C
RGS10 FP	TTGTAGAATTTTTTAGATTTTTTCGAGT	29	10	1
	GATTGGAGTTTCGGTCGATCGTAT	24	8	3
	AAGGCGTAGCGTTTACGAAGTC	22	4	4
	TCGGTTGAGTAGGAAGCGGT	21	4	2
	GGGGGAGTACGGTTGGATTCT	20	5	3
	TCGGGAGTTTGAAGTTGGGC	20	3	2
	TTAGATTGGAGTTTCGGTCGATC	23	6	3
RGS10 RP	CTCCCGAAAAAACTCTACGCTCC	23	4	2
	CAACCTATAACCTTTACCTCCTCGTC	26	5	1
	AAAACAAAAAACCGAACGAACGAA	24	11	3
	TTTAACTCCTCGCGACAAAAACG	23	5	3
	AACTTAATAACCCCGCGAAAAACG	23	5	3
	CCCGAAAAAACTCTACGCTCC	21	4	2
	GTACGCTAAACTAACCGAACCG	23	5	4
RGS10 Probes	TTCGTTGGGTATCGTATTTTCGCGG	25	7	4
	CGTCGTTAGGTAAGCGGTTTCG	22	3	4
	GTTTGGGGCGTCGGGGGTT	19	4	2
	TCGGGGCGGTATTTTAGGCGA	21	4	3
	TTCGTTGGGTATCGTATTTTCGCGG	25	7	4
	TAGTTAGGTTTTAGTAGGCGGGGATTTTG	29	10	1
	GAGAGGGCGTTTATTCGCGG	20	5	3
RUNX2 FP	CGGGGTGGATTGTGAATTTATATCGT	28	8	2
	GGGAAAGAGGATGGAGGTGATG	22	0	0
RUNX2 RP	AATCCCTTAAACACCGCGTCCTA	23	7	2
	AAAATCCCTTAAACACCGCGTC	23	8	2
	TCGCCAACTCTATTAACCCATCAAAC	27	5	1
RUNX2 Probes	TGGGAGGGGGCGTTAGCGTATTTA	24	7	2
	ATTAAGTTTGATGAGGTCGATTGTGCGGGCGGA	32	4	3
	AGTTTGATGAGGTTCGATTGTGCGGC	25	3	3
VIM FP	CGCGAGAGTTGTTACGTTT	19	6	3
	TTTAGGTTTCGGAGTAGGAAGGTTT	25	8	2
	GGTAAGTCGCGTTTGGGGGATTTTTAC	27	6	3
	CGAAGTTTCGTTGAAATTTGTCGAG	25	6	3
	TGTCGAGGGTAGTAGGTTTGAAAGT	25	5	1
VIM RP	CGCGCCAACCAATAAAAACCTAACG	25	5	3
	CCACCTCGACGCGAACTTTATCGTTA	26	4	4
	GAAACTACGCCACCACAAC	19	5	2
	TCTACGCCTCCAAAACAACCTACG	23	5	2

	GTATTCGCTCTACGCCTCCAAA	22	2	3
VIM Probes	GAGGGCGTTTTTATTTTATTCGTT	24	13	2
	AGGAGATGCGGGAGTTGCGTCGG	23	2	3
	TGTTAGGGAGATAGCGGAGAGCGG	24	3	2
	AAAGTTGTAGGCGTTAGTTGCGCG	24	3	3
	TGTAGGCGTTAGTTGCGCGGAGGTGGCGT	29	3	3
ST8SIA6 FP	GCGGTGTATCGGTTTGGATT	21	5	2
	TCGGTTTGGATTTCGTTGGGAGTAG	24	3	2
	TCGTAGGGAGGCGATTTTTCGATC	23	6	4
ST8SIA6 RP	CCTCTTCAAACACACGCACACT	22	0	0
	CGCGCCCTCTTCAAACACA	19	0	2
	ACGATACGCGCTAAAACCCCCG	22	4	4
ST8SIA6 Probes	GTAGGTTGGGACGGCGCGA	19	2	3
	CGGTTGGGGAGTAGAGGGATCGGTTAG	27	4	2
TBXT FP	TTGTTTTTGGTCGTCGTTTTTCGA	25	9	3
	GGTCGCGTACGTTTTGAAGTGTC	23	4	4
	AGGAGGGATTTTCGTTCCGGGATT	23	6	2
	TTTTTTTTAGATGGTGAGAGTCGCGG	26	5	1
	CGCGGGAAAGAGTTTGTAGTATCGA	25	4	3
	AATTATCGTTGGAAGTACGTGAACGG	26	4	3
	TTAGGTAGGTGTGAGGAGTTCGTT	24	6	1
TBXT RP	CCTCCTCCCGAATCTATACCG	21	4	2
	CGAAAACCCATTATTAACCCCCGA	24	7	2
	CTAACTCCTAACAACCCCTCCGCA	23	5	1
	ACGAAACCCCAAACGACGTTACG	25	5	4
	ATAAAATCGCCCTTCTCGCTACCCG	25	6	3
	AAACCCCGAAATTAAACGAATCGAA	25	10	3
	CCACAACCCCGTACAAAAAACG	23	4	2
TBXT Probes	TCGAGATTTCGACGGTTTTCGGGTTTC	28	8	5
	TAGATTTAGTCGGGCGGGTTTGGA	25	6	2
	AGGGATCGATTTCGTTTCGCGGTA	23	4	4
	ACGTCGGGGTAGGTTGATTTACGATT	26	6	3
	TTGTTGAGCGTCGTGGAGAATGAGTTGT	28	5	3
	TCGGGGGGTAAGTCGGAGTCGTA	26	5	3
	TTTGGGATTAGCGAGGAGGGGGTC	24	4	2
HOXD9 FP	TGTTTTTATGTATTGGACGCGTTGT	26	6	2
	GCGGGATTTTCGGTGGAGATT	20	3	2
	GGTTTATGTGTTTCGGTTCGGTTCGG	25	7	3
	GTAGAGGTCGTTTATCGTCGTCGT	24	6	4
	AGGTTGTAGTTTGCGAATTAGTCGG	25	5	2
HOXD9 RP	CGCTACAAACTAAAACTCCGACGAA	26	7	3
	ACGCGTCCAATACACTAAAAACAAACC	27	5	2
	CGCCTCCTCTAAATACGATCGCGTC	25	4	4
	GCCTCGTATAATTTTACCCCCAAATCGAC	25	5	3
	AACATTTTAAACGTCCCGCACTCC	25	4	2
HOXD9 Probes	TTTTGAAGGTTGGGTTTCGCGTGGGC	25	5	3
	TTTAGAGGAGTTTAGGGTTTCGGGCG	23	6	2
	TCGATTTGGTTTGGTTTCGTTTAGCGGTTGT	30	12	3
	TTGGGAGGGTATCGCGGATTACGAGGT	27	6	2
	AGTTGTTTCGGCGGCGGATAGTGTAAAT	24	3	3

SFRP1 FP	TAAGGCGTAGAGTTGTCGTTTTTCG	24	7	3
	TTCGATGTTTATGTCGGTTTTGCGT	25	9	3
	GTACGAGTCGCGTACGGTTTC	21	4	5
SFRP1 RP	ACGCGTCTAATCTATCCCAAATATCGC	27	6	2
	CCCCGAAAATCCCTAAAAATTTACGAC	27	8	2
	GACACCCAAATCTTCTCTACTCG	24	4	2
SFRP1 Probes	AGTTTCGTCGTGTATTTGGGTCGCGA	26	7	4
	TTGTTTTTCGTCGTTTTTTCGCGCGCGTTT	29	12	6
	TACGGGTAGATGGGTCGGTTTAGGTAGAC	29	5	3
TWIST1 FP	TTAGGTTTTTTTGAAACGGTGTCTCGG	25	7	2
	TTCGTTTAGCGATTGGGTGCGT	22	4	3
	GTTAACGTGCGGGAGCGTTA	20	4	3
TWIST1 RP	AACCTCCAAATCTACAACCTCTCGCC	25	4	1
	CGACGAAAATCCGCAATCTTACGAA	23	6	4
	TCTTACTCAACTTATCCGAAAACAACGTA	29	8	2
TWIST1 Probes	TTGTAGAGTTTCGCGAGGTGTTTGGGAG	27	5	2
	TCGTACGTTGGTTATGATTTCGTTGCGTT	28	8	4
	ATTTAGTCGTTGAACGAGGCGTTTCGT	26	6	4
PAX9 FP	GTTTTTCGTTTTTCGAGTCGGTACG	24	5	4
	ATTTTATTTTTTCGTAGGGTTTCGGTCGT	28	10	3
	CGTTGTTTTTTTCGCGTTTTTTTCGC	24	6	5
	TCGGTATTTTCGTTTGGGAGATTTCG	25	6	3
	GAGGTCGTTGTTTAACGTTATTCGGT	26	9	3
PAX9 RP	TCGAAAAACTACGTTTCCAACCTATCG	27	6	3
	ACGAATTAAATCGAACCCGAACCTCGA	26	10	4
	AATCTATCAACCGTCCCAATACCGC	25	4	2
	TCACGAAAAACACATTATACTTATCGCACA	30	6	2
	GCAACCGTACGAAACCCGTAA	21	4	4
PAX9 Probes	TTGGTGGGTGGCGGGATGAATTC	23	2	2
	TCGTTGTCGGATCGTAGAACGTCG	24	4	5
	GGTCGTTTTAGCGTCGGGAAGG	22	4	3
	ATCGTTTGTTGGCGGACGGC	20	4	4
	TCGTGGAATTGGTTTAATTGGGTATTTCGATCGT	33	8	3
NEUROG 1 FP	GATTAAGGGTTGTTTCGGAGGCG	26	4	2
	TTAAAATCGAGACGTTGCGTTTTTCGT	25	5	4
	GTTTTAAGGCGGGTTGGTATCGT	23	4	2
	GTAGGTAGGGGACGTATTGCG	21	4	2
NEUROG 1 RP	AAAAACCGAACAACAAAAAACGCA	25	12	4
	CCGAACAACCCTTAATCCGCC	21	4	2
	TACCCTCAATCCGAACATAACGC	24	6	2
	GCCTAACGAATCAAAAACCTACCCGAA	26	8	3
NEUROG 1 Probes	GTTCGGGAGCGTTTTTTGTCGTCGT	25	8	4
	TTGGGTTTTGGTCGAGATATTGCGTT	26	8	2
	TTGTGTAGGATCGACGGATAGATAGAAAGGC	31	4	3
	CGGTAGGAGGCGTTTTTCGGGTA	22	5	3

SEQUENCES

- SEQ ID NO: 1
5'-GATTGGAGTTTCGGTCGATCGTAT-3'
- SEQ ID NO: 2
5 5'-AAGGCGTAGCGTTTACGAAGTC-3'
- SEQ ID NO: 3
5'-TCGGTTGAGTAGGAAGCGGT-3'
- SEQ ID NO: 4
5'-GGGGGAGTACGGTTGGATTC-3'
- 10 SEQ ID NO: 5
5'-TCGGGAGTTTGAAGTTGGGC-3'
- SEQ ID NO: 6
5'-TTAGATTGGAGTTTCGGTCGATC-3'
- SEQ ID NO: 7
15 5'-TTGTAGAATTTTTTTAGATTTTTTCGAGT-3'
- SEQ ID NO: 8
5'-CTCCCGAAAAAACTCTACGCTCC-3'
- SEQ ID NO: 9
5'-CAACCTATAACCTTTACCTCCTCGTC-3'
- 20 SEQ ID NO: 10
5'-AAAACAAAAAACCGAACGAACGAA-3'
- SEQ ID NO: 11
5'-TTTAACTCCTCGCGACAAAAACG-3'
- SEQ ID NO: 12
25 5'-AACTTAATAACCCCGCGAAAACG -3'
- SEQ ID NO: 13
5'-CCCGAAAAAACTCTACGCTCC-3'
- SEQ ID NO: 14
5'-GTACGCTAAACTAACCGAACCG-3'
- 30 SEQ ID NO: 15
5'-CGGGGTTGGATTGTTGAATTTATATCGT-3'
- SEQ ID NO: 16
5'-GGGAAAGAGGATGGAGGTGATG-3'
- SEQ ID NO: 17

- 5'-AAAATCCCTTAAAACACCGCGTC-3'
SEQ ID NO: 18
5'-AATCCCTTAAAACACCGCGTCCTA-3'
SEQ ID NO: 19
- 5 5'-TCGCCAAACTCTATTAACCCATCAAAC-3'
SEQ ID NO: 20
5'-TTTAGGTTTCGGAGTAGGAAGGTTC-3'
SEQ ID NO: 21
5'-GGTAAGTCGCGTTTGGGGGATTTTAC-3'
- 10 SEQ ID NO: 22
5'-CGCGAGAGTTGTTACGTTT-3'
SEQ ID NO: 23
5'-CGAAGTTTCGTTGAAATTTGTCGAG-3'
SEQ ID NO: 24
- 15 5'-TGTCGAGGGTAGTAGGTTTGAAAGT-3'
SEQ ID NO: 25
5'-CGCGCCAACCAATAAAAACCTAACG-3'
SEQ ID NO: 26
5'-CCACCTCGACGCGAACTTTATCGTTA-3'
- 20 SEQ ID NO: 27
5'-GAAACTACGCCACCACAAC-3'
SEQ ID NO: 28
5'-TCTACGCCTCCAAAACAACACTACG-3'
SEQ ID NO: 29
- 25 5'-GTATTCGCTCTACGCCTCCAAA-3'
SEQ ID NO: 30
5'-GCGGTTGTATCGGTTTGGATT-3'
SEQ ID NO: 31
5'-TCGGTTTGGATTTCGTTGGGAGTAG-3'
- 30 SEQ ID NO: 32
5'-TCGTAGGGAGGCGATTTTTCGATC-3'
SEQ ID NO: 33
5'-CCTCTTCAAACACACGCACACT-3'
SEQ ID NO: 34

- 5'-CGCGCCCTCTTCAAACACA-3'
SEQ ID NO: 35
- 5'-ACGATACGCGCTAAAACCCCG-3'
SEQ ID NO: 36
- 5 5'-TTGTTTTTGGTCGTCGTTTTTTCGA-3'
SEQ ID NO: 37
- 5'-GGTCGCGTACGTTTTGAAGTGTC-3'
SEQ ID NO: 38
- 5'-AGGAGGGATTTTCGTTCCGGGATT-3'
10 SEQ ID NO: 39
- 5'-TTTTTTTTAGATGGTGAGAGTCGCGG -3'
SEQ ID NO: 40
- 5'-CGCGGGAAAGAGTTTGTAGTATCGA-3'
SEQ ID NO: 41
- 15 5'-AATTATCGTTGGAAGTACGTGAACGG -3'
SEQ ID NO: 42
- 5'-TTAGGTAGGTGTGAGGAGTTCGTT-3'
SEQ ID NO: 43
- 5'-CCTCCTCCCGAATCTATACCG -3'
20 SEQ ID NO: 44
- 5'-CGAAAACCCATTATTAACCCCGA-3'
SEQ ID NO: 45
- 5'-CTAACTCCTAACAACCCTCCGCA-3'
SEQ ID NO: 46
- 25 5'-ACGAAACCCCAAACGACGTTACG-3'
SEQ ID NO: 47
- 5'-ATAAAATCGCCCTTCTCGCTACCCG -3'
SEQ ID NO: 48
- 5'-AAACCCCGAAATTAAACGAATCGAA -3'
30 SEQ ID NO: 49
- 5'-CCACAACCCCGTACAAAAACG -3'
SEQ ID NO: 50
- 5'-TGTTTTTAGTGTATTGGACGCGTTGT-3'
SEQ ID NO: 51

- 5'-GCGGGATTTTCGGTGGAGATT-3'
SEQ ID NO: 52
5'-GGTTTATGTGTTTCGGTTCGGTTCGG-3'
SEQ ID NO: 53
- 5 5'-GTAGAGGTCGTTTATCGTCGTCGT-3'
SEQ ID NO: 54
5'-AGGTTGTAGTTTGC GAATTAGTCGG-3'
SEQ ID NO: 55
5'-CGCTACAACTAAAACTCCGACGAA-3'
10 SEQ ID NO: 56
5'-ACGCGTCCAATACACTAAAAACAAACC-3'
SEQ ID NO: 57
5'-CGCCTCCTCTAAATACGATCGCGTC-3'
SEQ ID NO: 58
- 15 5'-GCCTCGTATAATTTTACCCCCAAATCGAC-3'
SEQ ID NO: 59
5'-AACATTTTAAAACGTCCCGCACTCC-3'
SEQ ID NO: 60
5'-TAAGGCGTAGAGTTGTCTGTTTTTCG-3'
- 20 SEQ ID NO: 61
5'-TTCGATGTTTATGTCTGGTTTTTGCGT-3'
SEQ ID NO: 62
5'-GTACGAGTCGCGTACGGTTTC-3'
SEQ ID NO: 63
- 25 5'-ACGCGTCTAATCTATCCCAAATATCGC-3'
SEQ ID NO: 64
5'-CCCCGAAAATCCCTAAAAATTTACGAC-3'
SEQ ID NO: 65
5'-GACACCCAAATCTTCCTCTACTCG-3'
- 30 SEQ ID NO: 66
5'-TTAGGTTTTTTTGGAACGGTGTCGG -3'
SEQ ID NO: 67
5'-TTCGTTTAGCGATTGGGTGCGT -3'
SEQ ID NO: 68

- 5'-GTTAACGTGCGGGAGCGTTA-3'
SEQ ID NO: 69
5'-AACCTCCAAATCTACAACTCTCGCC-3'
SEQ ID NO: 70
- 5 5'-CGACGAAAATCCGCAATCTTACGAA-3'
SEQ ID NO: 71
5'-TCTTACTCAACTTATCCGAAAACAACGTA-3'
SEQ ID NO: 72
5'-GTTTTTCGTTTTTCGGAGTCGGTACG-3'
- 10 SEQ ID NO: 73
5'-ATTTTATTTTTTCGTAGGGTTTCGGTCGT-3'
SEQ ID NO: 74
5'-CGTTGTTTTTTCGCGTTTTTTCGC-3'
SEQ ID NO: 75
- 15 5'-TCGGTATTTTCGTTTGGGAGATTCG-3'
SEQ ID NO: 76
5'-GAGGTCGTTGTTTAACGTTATTCGGT-3'
SEQ ID NO: 77
5'-TCGAAAAACTACGTTTCCAACCTATCG-3'
- 20 SEQ ID NO: 78
5'-ACGAATTAAATCGAACCCGAACTCGA -3'
SEQ ID NO: 79
5'-AATCTATCAACCGTCCCAATACCGC-3'
SEQ ID NO: 80
- 25 5'-TCACGAAAAACACATTATACTTATCGCACA -3'
SEQ ID NO: 81
5'-GCAACCGTACGAAACCCGTAA -3'
SEQ ID NO: 82
5'-GATTAAGGGTTGTTTCGGAGGCG-3'
- 30 SEQ ID NO: 83
5'-TTAAAATCGAGACGTTGCGTTTCGT-3'
SEQ ID NO: 84
5'-GTTTTAAGGCGGGTTGGTATCGT-3'

- SEQ ID NO: 85
5'-GTAGGTAGGGGACGTATTGCG-3'
- SEQ ID NO: 86
5'-AAAAACCGAACAAACAAAAACGCA-3'
- 5 SEQ ID NO: 87
5'-CCGAACAACCCCTTAATCCGCC-3'
- SEQ ID NO: 88
5'-TACCCTCAATCCGAAC TACAACGC-3'
- SEQ ID NO: 89
- 10 5'-GCCTAACGAATCAAAA ACTACCCGAA-3
SEQ ID NO: 90
5'-TTCGTTGGGTATCGTATTTTCGCGG-3'
- SEQ ID NO: 91
5'-TAGTTAGGTTTTAGTAGGCGGGGATTTTG-3'
- 15 SEQ ID NO: 92
5'-CGTCGTTAGGTAAGCGGTTTCG-3'
- SEQ ID NO: 93
5'-GTTTGGGGCGTCGGGGGTT-3'
- SEQ ID NO: 94
- 20 5'-TCGGGGCGGTATTTTAGGCGA-3'
SEQ ID NO: 95
5'-TTCGTTGGGTATCGTATTTTCGCGG-3'
- SEQ ID NO: 96
5'-GAGAGGGCGTTTATTCGCGG-3'
- 25 SEQ ID NO: 97
5'-TGGGAGGGGGCGTTAGCGTATTTA-3'
- SEQ ID NO: 98
5'-ATTAAGTTTGATGAGGTCGATTGTCGGGCGGA -3'
- SEQ ID NO: 99
- 30 5'-AGTTTGATGAGGTCGATTGTCGGGC-3'
SEQ ID NO: 100
5'-GAGGGCGTTTTTATTTATTCGTT-3'
- SEQ ID NO: 101
5'-AGGAGATGCGGGAGTTGCGTCGG-3'

- SEQ ID NO: 102
5'-TGTTAGGGAGATAGCGGAGAGCGG-3'
- SEQ ID NO: 103
5'-AAAGTTGTAGGCGTTAGTTGCGCG-3'
- 5 SEQ ID NO: 104
5'-TGTAGGCGTTAGTTGCGCGGAGGTGGCGT-3'
- SEQ ID NO: 105
5'-GTAGGTTGGGACGGCGCGA-3'
- SEQ ID NO: 106
10 5'-CGGTTGGGGAGTAGAGGGATCGGTTAG-3'
- SEQ ID NO: 107
5'-TCGAGATTTGCGACGGTTTTTCGGGTTTC-3'
- SEQ ID NO: 108
5'-TAGATTTAGTCGGGCGGGTTTGGGA-3'
- 15 SEQ ID NO: 109
5'-AGGGATCGATTTCGTTTCGCGGTA-3'
- SEQ ID NO: 110
5'-ACGTCGGGGTAGGTTGATTTACGATT-3'
- SEQ ID NO: 111
20 5'-TTGTTGAGCGTCGTGGAGAATGAGTTGT-3'
- SEQ ID NO: 112
5'-TCGGGGGGTAAGTCGGAGTCGTA-3'
- SEQ ID NO: 113
5'-TTTGGGATTAGCGAGGAGGGGGTC-3'
- 25 SEQ ID NO: 114
5'-TTTTGAAGGTTGGGTTTCGCGTGGGC-3'
- SEQ ID NO: 115
5'-TTTAGAGGAGTTTAGGGTTTCGGGCG-3'
- SEQ ID NO: 116
30 5'-TCGATTTGGTTTGGTTTCGTTTAGCGGTTGT-3'
- SEQ ID NO: 117
5'-TTGGGAGGGTATCGCGGATTACGAGGT-3'
- SEQ ID NO: 118
5'-AGTTGTTTCGGCGGCGGATAGTGTAAT-3'

- SEQ ID NO: 119
5'-AGTTTCGTCGTGTATTTGGGTCGCGA-3'
SEQ ID NO: 120
5'-TTGTTTTTCGTCGTTTTTTCGCGCGCGTTT-3'
- 5 SEQ ID NO: 121
5'-TACGGGTAGATGGGTCGGTTTAGGTAGAC-3'
SEQ ID NO: 122
5'-TTGTAGAGTTCGCGAGGTGTTTGGGAG -3'
SEQ ID NO: 123
- 10 5'-TCGTACGTTGGTTATGATTCGTTGCGTT-3'
SEQ ID NO: 124
5'-ATTTAGTCGTTGAACGAGGCGTTCGT-3'
SEQ ID NO: 125
5'-TTGGTGGGTGGCGGGATGAATTC-3'
- 15 SEQ ID NO: 126
5'-TCGTTGTCGGATCGTAGAACGTCG-3'
SEQ ID NO: 127
5'-GGTCGTTTTAGCGTCGGGAAGG-3'
SEQ ID NO: 128
- 20 5'-ATCGTTTGTTGGCGGACGGC-3'
SEQ ID NO: 129
5'-TCGTGGAATTGGTTTAATTGGGTATTCGATCGT-3'
SEQ ID NO: 130
5'-GTTTCGGGAGCGTTTTTTGTCGTCGT-3'
- 25 SEQ ID NO: 131
5'-TTGGGTTTTGGTCGAGATATTGCGTT-3'
SEQ ID NO: 132
5'-TTGTGTAGGATCGACGGATAGATAGAAAGGC-3'
SEQ ID NO: 133
- 30 5'-CGGTAGGAGGCGTTTTTCGGGTA-3'

References

1. Wang, J., et al., *Circulating tumour DNA methylation in hepatocellular carcinoma diagnosis using digital droplet PCR*. J Int Med Res, 2021. **49**(3): p. 300060521992962.
- 5 2. Xiong, Y., et al., *Detection of a novel panel of somatic mutations in plasma cell-free DNA and its diagnostic value in hepatocellular carcinoma*. Cancer Manag Res, 2019. **11**: p. 5745-5756.
3. Estival, A., et al., *Pyrosequencing versus methylation-specific PCR for assessment of MGMT methylation in tumor and blood samples of glioblastoma patients*. Sci Rep, 10 2019. **9**(1): p. 11125.
4. Hao, X., et al., *DNA methylation markers for diagnosis and prognosis of common cancers*. Proc Natl Acad Sci U S A, 2017. **114**(28): p. 7414-7419.
5. Powrozek, T., et al., *Septin 9 promoter region methylation in free circulating DNA-potential role in noninvasive diagnosis of lung cancer: preliminary report*. Med 15 Oncol, 2014. **31**(4): p. 917.
6. Jiang, H., et al., *DNA methylation markers in the diagnosis and prognosis of common leukemias*. Signal Transduct Target Ther, 2020. **5**(1): p. 3.
7. Xu, R.H., et al., *Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma*. Nat Mater, 2017. **16**(11): p. 1155-1161.
- 20 8. Li, J.B., et al., *Multiplex padlock targeted sequencing reveals human hypermutable CpG variations*. Genome Res, 2009. **19**(9): p. 1606-15.
9. Diep, D., et al., *Library-free methylation sequencing with bisulfite padlock probes*. Nat Methods, 2012. **9**(3): p. 270-2.

We Claim:

1. A multiplex composition for Sequencing Independent Quantitative Methylation Profiling (qSIMP) to determine methylation status of a gene in a sample using PCR techniques comprising of
 - 5 (i) Two or more of methylation specific primers selected from the group comprising of SEQ ID NOs: 1-89, or a complement thereof, and
 - (ii) One or more of probes that bind to the amplicons generated by said methylation specific primers, selected from the group comprising of SEQ ID NOs: 90-133, or a complement thereof,
- 10 wherein, the methylation specific primer hybridizes to a target CpG-containing nucleic acid sequence of the gene that has undergone deamination reaction, wherein, the methylation of said target sequence is indicative of the presence or absence of a disorder indicative through signal produced as a result of amplification.
2. The composition according to claim 1, wherein the primer is at least one selected
 - 15 from:
 - i) primers having base sequences, SEQ ID NOs: 1-14, or a complement thereof,
 - ii) primers having base sequences, SEQ ID NOs: 15-19, or a complement thereof,
 - iii) primers having base sequences, SEQ ID NOs: 20-29, or a complement thereof,
 - iv) primers having base sequences, SEQ ID NOs: 30-35, or a complement thereof,
 - 20 v) primers having base sequences, SEQ ID NOs: 36-49, or a complement thereof,
 - vi) primers having base sequences, SEQ ID NOs: 50-59, or a complement thereof,
 - vii) primers having base sequences, SEQ ID NOs: 60-65, or a complement thereof,
 - viii) primers having base sequences, SEQ ID NOs: 66-71, or a complement thereof,
 - ix) primers having base sequences, SEQ ID NOs: 72-81, or a complement thereof
 - 25 x) primers having base sequences, SEQ ID NOs: 82-89, or a complement thereof
3. The composition according to claim 1, wherein the probe is at least one selected from:
 - i) probe having base sequences, SEQ ID NOs: 90-96, or a complement thereof,
 - ii) probe having base sequences, SEQ ID NOs: 97-99, or a complement thereof,
 - iii) probe having base sequences, SEQ ID NOs: 100-104, or a complement thereof,
 - 30 iv) probe having base sequences, SEQ ID NOs: 105-106, or a complement thereof,
 - v) probe having base sequences, SEQ ID NOs: 107-113, or a complement thereof,
 - vi) probe having base sequences, SEQ ID NOs: 114-118, or a complement thereof,
 - vii) probe having base sequences, SEQ ID NOs: 119-121, or a complement thereof,

- viii) probe having base sequences, SEQ ID NOs: 122-124, or a complement thereof,
ix) probe having base sequences, SEQ ID NOs: 125-129, or a complement thereof
x) probe having base sequences, SEQ ID NOs: 130-133, or a complement thereof
4. The multiplex composition of claim 1, wherein the methylation specific primers
5 hybridize to a target CpG-site of nucleic acid sequence of a gene that has undergone
deamination selected from the group comprising of RUNX, TBXT, HOX, ST8SIA,
SFRP, RGS, PAX, VIM, TWIST and NEUROG.
5. The multiplex composition of claim 1, wherein the methylation specific primer
10 hybridizes to a target CpG-site of nucleic acid sequence of a gene that has undergone
deamination selected from the group consisting of RUNX2, TBXT, HOXD9,
ST8SIA6, SFRP1, RGS10, PAX9, VIM, TWIST-1 and NEUROG1 and combinations
thereof.
6. The multiplex composition of claim 1, wherein the probe comprises a detectable label
15 selected from florescence, absorbance or luminescence markers and optionally
contains one or more quenchers.
7. A method for (qSIMP) to determine methylation status comprising
- 20 a. isolating DNA from biological specimen and specifically isolating cell free
DNA (cfDNA) from blood or plasma or serum,
 - b. treating the DNA through a bisulfite conversion method involving the
deamination of unmodified cytosines to uracil, leaving the modified bases 5-
mC (5-methylcytosine) and 5-hmC (5-hydroxymethylcytosine) intact to obtain
treated DNA,
 - 25 c. contacting the multiplex composition comprising methylation specific primers
and probes as claimed in claim 1, with the treated DNA from previous step,
wherein the primers and probes hybridize to the CpG islands of the treated
DNA,
 - d. amplifying the CpG island present in the treated DNA under stringent
conditions wherein the set of primers and probes hybridize to obtain an
amplification product using PCR techniques,
 - 30 e. obtaining the methylation profile of the amplified product as a florescence,
absorbance or luminescence readout, and
 - f. analyzing the methylation profile using a mathematical model.

8. The method of claim 5, wherein the mathematical model is derived from statistical analysis, logistic regression analysis, principal component analysis, nearest neighbor analysis, k-means cluster analysis, neural network model or combination thereof.
9. The method of claim 5, wherein the sample is selected from cells, blood, plasma,
5 serum, bodily fluids, and tissues.
10. The method of claim 5, wherein the presence or absence of methylated CpG island-containing DNA is indicative of a disorder.
11. The composition as claimed in claim 2 wherein the disorder is hepatocellular carcinoma (HCC) or Cholangiocarcinoma (CCA) or combination thereof.
- 10 12. The method of claim 9, wherein the disorder is hepatocellular carcinoma (HCC) or Cholangiocarcinoma (CCA) or combination thereof.
13. A kit for qSIMP to determine methylation status of a gene in a sample comprising
- a. the multiplex composition comprising one or more methylation specific primers and one or more probes as claimed in claim 1;
 - 15 b. buffering agents,
 - c. stabilizing agents,
 - d. polymerase enzymes,
 - e. PCR supermix,
 - f. unmethylated DNA control to serve as negative control,
 - 20 g. methylated DNA control to serve as positive control, and
 - h. negative template controls.
14. The kit as claimed in claim 13, wherein the concentration of individual primers ranges from 5nM to 1500nM in the plurality of the mixture present in the kit and the concentration of probes ranges from 5nM to 1500nM in the plurality of the mixture
25 present in the kit.