WO 2018/195211 discloses compositions, kits, and methods for constructing libraries for simultaneous detection of genomic variants and DNA methylation status on limited DNA inputs, such as circulating polynucleotide fragments in the body of a subject, including circulating tumor DNA.

WO 2011/070441, WO 2017/006317, WO 2019/142193 and WO 2020/188561, assigned to the Applicant of the present invention, disclose methods for detecting methylation changes in DNA samples.

It would be highly beneficial to have methods and systems that can produce various types of genetic and epigenetic data from cell-free DNA obtained from a single sampling from a subject and using sequencing data from a single run.

SUMMARY

The present disclosure provides methods and systems for profiling genetic and epigenetic characteristics of DNA samples, particularly cell-free DNA samples obtained from biological fluids such as plasma and urine. The methods and systems of the present disclosure involve digestion with at least one methylation-sensitive restriction enzyme, preferably a plurality of methylation-sensitive restriction enzymes applied simultaneously, preparation of a sequencing library using library preparation methods that preserve the sequence information at the ends of DNA molecules in the sample, high throughput sequencing and analysis of sequence reads.

The present disclosure provides a more simple and accurate assay compared to hitherto described methods, yielding high-quality sequencing data compared to bisulfite sequencing enabling sensitive detection of cancer-associated changes. Vast amount of information from a single run and based on the same sequencing data can be obtained, including methylation data, mutation data and more, thus avoiding the need of parallel assays to obtain comprehensive genetic and epigenetic information. Importantly, it was surprisingly found that high-quality sequencing data can be obtained even from very low amounts of DNA, without a need for amplification prior to library preparation. As exemplified hereinbelow, the methods disclosed herein are able to detect genetic and epigenetic changes of early-stage cancer, when the amounts of tumor-derived DNA in the plasma are very low, based on the amount of cell-free DNA that can be achieved from a single standard blood test tube.

The methods and systems of the present disclosure do not involve or require bisulfite conversion. The methods and systems of the present disclosure do not require changing the sequence of the DNA and enable co-analysis of, e.g., methylation, mutation, copy number and nucleosome positioning, based on the same sequencing data.

As exemplified hereinbelow, a comparison of sequencing data obtained for cell-free DNA samples subjected to methylation-sensitive enzymatic digestion followed by high-throughput sequencing to sequencing data obtained following bisulfite conversion and high-throughput sequencing showed significantly better sequencing metrics (number of reads, mapping rate, etc.), copy number integrity and nucleosome positioning integrity for enzyme-treated DNA compared to bisulfite-treated DNA. Analysis of pooled cell-free DNA samples compared to individual samples showed that there is loss of information when low amounts of DNA are used. However, while the quality of the sequencing data of the enzyme-treated DNA samples remained high and enabled reliable analysis, bisulfite sequencing showed significantly reduced number of reads and mapping rate, and practically lost all copy number and nucleosome positioning information. In addition, sequencing noise was high for bisulfite-treated DNA samples, such that mutations were indistinguishable from the sequencing noise. With respect to methylation detection, methylation analysis in enzyme-treated DNA samples detected significantly more methylation changes in the plasma compared to bisulfite-treated DNA samples.

As further exemplified hereinbelow, while bisulfite-treated DNA samples resulted in broader CG coverage at the lower end of sequencing depths compared to enzyme-treated samples, a continuous and sharp decrease was seen in the number of CGs that were covered in bisulfite-treated DNA samples as the depth increased. In contrast, enzyme-treated samples showed substantially constant coverage even at depths over 250-300. At high depths, methylation-sensitive digestion provided significantly better CG coverage compared to bisulfite. Methylation-sensitive digestion provided coverage of millions of CGs at very high depths, thus enabling the detection of rare methylation signals, for example, methylated DNA molecules from a tumor in the plasma at an early stage of the tumor, which may be present in the plasma at very low amounts - 1% or even less of the total cell-free DNA. The data showed that at depths required for identification of rare signals, bisulfite does not provide sufficient coverage, and such rare signals are likely to be missed when using bisulfite sequencing on low amounts of DNA.

The present disclosure further discloses an improved method for determining methylation values for genomic loci of interest. Methylation analysis according to the present disclosure is carried out for restriction loci, namely, restriction sites of the restriction enzyme(s) used in the assay. Methylation analysis as disclosed herein is based on analyzing alignments covering a predefined genomic region of at least 50 bps in length, preferably at least 100 bps in length, that contains a restriction locus of interest, and determining a read count of sequence reads covering the predefined genomic region. Such alignments represent DNA molecules of at least 50 bps in length (preferably at least 100 bps in length), in which the analyzed restriction locus, as well as any additional restriction loci within the DNA molecule, were all methylated in the DNA sample and therefore the DNA molecules remained intact following digestion with the enzymes used in the assay. Analyzing alignments which are at least 50 or at least 100 bps in length and containing a plurality of restriction loci which were all methylated in the DNA sample increases the specificity of the cancer-related hypermethylation signal and enables an improved, more accurate detection of differences between normal and cancerous samples. In addition, the analysis of such alignments is advantageous for evaluating nucleosome positioning in cell-free DNA in addition to methylation because the copy numbers of such alignments reflect nucleosomal boundaries, wherein a high copy number is typical of the middle of the nucleosome, and a low copy number is typical of the boundaries between nucleosomes.

The present disclosure further discloses a method for direct calculation of both methylated and unmethylated levels of DNA based on sequencing data generated following methylation-sensitive/-dependent restriction of a DNA sample. Advantageously, the methods and systems of the present disclosure allow independently determining the methylated and unmethylated levels of DNA in a single assay and based on the same sequencing data, thus providing an improved identification of methylation changes. More particularly, the methods and systems disclosed herein comprise according to some examples digestion of a DNA sample with at least one methylation-sensitive restriction endonuclease, followed by high-throughput sequencing producing a plurality of sequence reads. Sequence reads may be aligned against a reference genome and restriction loci, namely, restriction sites within the genome, are selected and analyzed. The level of methylated DNA at the selected restriction loci is determined based on the read count of each restriction locus, which represents the number of DNA molecules in the

sample in which the restriction locus was methylated and therefore remained intact. The level of

unmethylated DNA at the selected restriction loci is determined by a unique analysis of the ends of sequence reads, by determining the number of reads starting or ending at a nucleotide within each restriction locus. This number of reads represents the number of DNA molecules in the sample in which the restriction locus was unmethylated and therefore cut by the restriction endonuclease. Such direct analysis of unmethylated DNA molecules is advantageous over indirect assessment based on the level of methylated DNA, as carried out by existing methods. The direct determination of unmethylation in addition to methylation using the same sequencing data provides complementary methylation information of genomic regions and thus improved methylation profiling, a more accurate and valid assessment of potential DNA methylation markers, and better detection of methylation differences between samples. It also provides an increased sensitivity of methylation analysis, particularly beneficial for genomic regions with extremely high or extremely low methylation levels.

In one aspect, the present disclosure provides a method for profiling genetic and epigenetic characteristics of a cell-free DNA (cfDNA) sample from a subject, the method comprising:

(a) subjecting the cell-free DNA sample to digestion with at least one methylation-sensitive restriction endonuclease, to obtain restriction endonuclease-treated DNA in which methylated sites are intact and unmethylated sites are cut;

(b) preparing a sequencing library from the restriction endonuclease-treated DNA while preserving the sequence information at the ends of the DNA molecules, wherein preparing the sequencing library comprises ligating sequencing adapters to DNA molecules in the restriction endonuclease-treated DNA, wherein each adapter is capable of ligation to both the digested and undigested DNA molecules;

(c) sequencing the sequencing library by a high-throughput sequencing method to provide sequencing data; and

(d) determining from the sequencing data a methylation value for at least one restriction locus,

wherein an amount of cell-free DNA comprising 3000 haploid equivalents is sufficient for the method, and wherein the cell-free DNA sample is not subjected to amplification prior to library preparation.

In another aspect, the present disclosure provides a method for detecting cancer-related genetic and epigenetic changes in a cell-free DNA sample (cfDNA) from a subject,

the method comprising: profiling methylation and optionally at least one additional genetic and epigenetic characteristics of the cfDNA sample according to any one of claims 1-13, to obtain a genetic and epigenetic profile of the cfDNA sample; and comparing the genetic and epigenetic profile of the cfDNA sample to one or more reference genetic and epigenetic profile selected from a cancer profile and a non-cancer profile, to detect cancer-associated genetic and epigenetic changes in the cfDNA sample.

According to one example, the present disclosure provides a method for profiling genetic and epigenetic characteristics of a cell-free DNA (cfDNA) sample from a subject, the method comprising:

(a) subjecting the cell-free DNA sample to digestion with at least one methylation-sensitive restriction endonuclease, to obtain restriction endonuclease-treated DNA in which methylated restriction sites are intact and unmethylated restriction sites are cut;

(b) preparing a sequencing library from the restriction endonuclease-treated DNA while preserving the sequence information at the ends of the DNA molecules, wherein preparing the sequencing library comprises ligating sequencing adapters to DNA molecules in the restriction endonuclease-treated DNA, wherein each adapter is capable of ligation to both the digested and undigested DNA molecules;

(c) sequencing the sequencing library by a high-throughput sequencing method to provide sequencing data; and

(d) determining from the sequencing data a methylation value for at least one restriction locus and optionally at least one additional genetic or epigenetic characteristic of the cell-free DNA sample selected from DNA mutation, copy number variation and nucleosome positioning,

wherein an amount of cell-free DNA comprising 3000 haploid equivalents is sufficient for the method, wherein the cell-free DNA sample is not subjected to amplification prior to library preparation, and wherein determining the methylation value

and the at least one additional genetic or epigenetic characteristic of the cell-free DNA sample is carried out based on the same sequencing data.

According to another example, the present disclosure provides a method for processing a cell-free DNA sample to obtain sequencing data for genetic and epigenetic analysis, the method comprising:

(a) subjecting the cell-free DNA sample to digestion with at least one methylation-sensitive restriction endonuclease, to obtain restriction endonuclease-treated DNA in which methylated restriction sites are intact and unmethylated restriction sites are cut;

(b) preparing a sequencing library from the restriction endonuclease-treated DNA while preserving the sequence information at the ends of the DNA molecules, wherein preparing the sequencing library comprises ligating sequencing adapters to DNA molecules in the restriction endonuclease-treated DNA, wherein each adapter is capable of ligation to both the digested and undigested DNA molecules; and

(c) sequencing the sequencing library by a high-throughput sequencing method to obtain sequencing data,

wherein an amount of cell-free DNA comprising 3000 haploid equivalents is sufficient to achieve at least one of: unique mapping rate of at least 85%, a copy number integrity characterized by Pearson correlation of at least 0.65 compared to undigested sample and nucleosome positioning integrity characterized by Pearson correlation of at least 0.55 compared to undigested sample,

and wherein genetic and epigenetic analysis is performed based on the same sequencing data.

In some examples, an amount of cell-free DNA comprising 6,000 haploid equivalents is sufficient for the methods disclosed herein.

In some examples, the cell-free DNA is plasma cell-free DNA, and the amount of the cell-free DNA is an amount obtained from 9-10 ml of blood.

In some examples, the amount of cell-free DNA is between 10-200 ng. In additional examples, the amount of cell-free DNA is between 20-100 ng.

In some examples, the at least one methylation-sensitive restriction endonuclease produces non-blunt ends, and the method further comprises subjecting the restriction endonuclease-treated DNA to end repair prior to the ligation of sequencing adapters, to obtain DNA molecules with blunt ends.

In some examples, the high-throughput sequencing is whole genome high-throughput sequencing.

In some examples, the high-throughput sequencing is target-specific high-throughput sequencing.

In some examples, determining a methylation value for at least one restriction locus comprises:

(i) selecting at least one restriction locus and determining the number of sequence reads covering a predefined genomic region of at least 50 bps in length that contains said restriction locus; and

(ii) calculating a methylation value for the at least one restriction locus based on the read count determined in step (i) and a reference read count.

In some examples, step (i) comprises determining the number of sequence reads covering a predefined genomic region of at least 100 bps in length that contains said restriction locus.

In some examples, the at least one restriction locus is a plurality of restriction loci.

In some examples, the at least one methylation-sensitive restriction endonuclease is a plurality of methylation-sensitive restriction endonucleases, and the digestion with the plurality of methylation-sensitive restriction endonucleases is a simultaneous digestion.

In some examples, the plurality of methylation-sensitive restriction endonucleases comprises HinP1I. In additional examples, the plurality of methylation-sensitive restriction endonucleases comprises AciI. In additional examples, the digestion is carried out using HinP1I and AciI. In some examples, the digestion is carried out using HinP1I and AciI at a ratio between 1:1 to 5:1 (enzyme units) (Hinp:AciI).

In some examples, the step of subjecting the cell-free DNA sample to digestion with at least one methylation-sensitive restriction endonuclease further comprises determining digestion efficacy, and proceeding to preparing a sequencing library if the digestion efficacy is above a predefined threshold.

According to another example, the present disclosure provides a method for detecting cancer-related genetic and epigenetic changes in a cell-free DNA sample (cfDNA) from a subject, the method comprising: profiling methylation and

optionally at least one additional genetic and epigenetic characteristics of the cfDNA sample as disclosed herein, to obtain a

genetic and epigenetic profile of the cfDNA sample; and comparing the genetic and epigenetic profile of the cfDNA sample to one or more reference genetic and epigenetic profile selected from a cancer profile and a non-cancer profile, to detect cancer-associated genetic and epigenetic changes in the cfDNA sample.

In some examples, the cell-free DNA sample is from a subject suspected of having cancer or at risk of having cancer, and the method further comprises administering to the subject active cancer surveillance and follow-up testing when cancer-associated changes are detected, the cancer surveillance and follow-up testing comprises one or more of blood tests, urine tests, cytology, imaging, endoscopy and biopsy.

According to a further example, the present disclosure provides a method for assessing the presence or absence of cancer in a subject, the method comprising:

(a) subjecting a cell-free DNA (cfDNA) sample of the subject to digestion with at least one methylation-sensitive restriction endonuclease, to obtain restriction endonuclease-treated DNA in which methylated restriction sites are intact and unmethylated restriction sites are cut;

(b) sequencing the restriction endonuclease-treated DNA by a high-throughput sequencing method;

(c) selecting at least one multiomic genomic region, comprising a tumor hypermethylated restriction locus and a tumor mutation locus within 150 bps of each other; and

(d) determining the likelihood that the subject has cancer based on analysis of sequence reads covering the at least one multiomic region.

In some examples, the at least one multiomic region comprises a tumor hypermethylated restriction locus and a tumor mutation locus within 100 bps of each other.

In some examples, analysis of sequence reads covering the at least one multiomic region comprises:

- determining for each multiomic region at least one of:

(i) the number of methylated-mutated sequence reads covering said multiomic region, which include all nucleotides of the restriction locus and present a mutated genotype at the mutation locus;

(ii) the number of methylated-wild type sequence reads covering said multiomic region, which include all nucleotides of the restriction locus and present a wild type genotype at the mutation locus;

(iii) the number of unmethylated-mutated sequence reads covering said multiomic region, which start or end at a nucleotide within the restriction locus and present a mutated genotype at the mutation locus; and

(iv) the number of unmethylated-wild type sequence reads covering said multiomic region, which start or end at a nucleotide within the restriction locus and present a wild type genotype at the mutation locus;

and

- comparing the number of reads determined in (i)-(iv) to reference values of cancer patients and/or healthy individuals in order to assess the likelihood that the subject has cancer.

According to a further example, the present disclosure provides a method for characterizing a cell-free DNA (cfDNA) sample of a subject suspected of having cancer or at risk of having cancer, the method comprising:

(a) subjecting the cell-free DNA sample to digestion with at least one methylation-sensitive restriction endonuclease, to obtain restriction endonuclease-treated DNA in which methylated sites are intact and unmethylated sites are cut;

(b) sequencing the restriction endonuclease-treated DNA by a high-throughput sequencing method;

(c) selecting at least one multiomic genomic region, comprising a tumor hypermethylated restriction locus and a tumor mutation locus within 150 bps of each other; and

(d) determining for each multiomic region at least one of:

(i) the number of methylated-mutated sequence reads covering said multiomic region, which include all nucleotides of the restriction locus and present a mutated genotype at the mutation locus;

(ii) the number of methylated-wild type sequence reads covering said multiomic region, which include all nucleotides of the restriction locus and present a wild type genotype at the mutation locus;

(iii) the number of unmethylated-mutated sequence reads covering said multiomic region, which start or end at a nucleotide within the restriction locus and present a mutated genotype at the mutation locus; and

(iv) the number of unmethylated-wild type sequence reads covering said multiomic region, which start or end at a nucleotide within the restriction locus and present a wild type genotype at the mutation locus,

thereby characterizing the cell-free DNA sample.

According to another example, the present disclosure provides a method for profiling methylation of a DNA sample from a subject, the method comprising:

(a) subjecting the DNA sample to digestion with at least one methylation-sensitive restriction endonuclease, to obtain restriction endonuclease-treated DNA in which methylated sites are intact and unmethylated sites are cut;

(b) preparing a sequencing library from the restriction endonuclease-treated DNA, wherein preparing the sequencing library comprises ligating sequencing adapters to DNA fragments in the restriction endonuclease-treated, wherein each adapter is capable of ligation to both the digested and undigested DNA molecules;

(c) sequencing the sequencing library by a high-throughput sequencing method to obtain sequence reads;

(d) selecting at least one restriction locus and determining the number of sequence reads covering a predefined genomic region of at least 50 bps in length that contains said restriction locus; and

(e) calculating a methylation value for the at least one restriction locus based on the read count determined in step (d) and a reference read count,

thereby profiling methylation of the cell-free DNA sample.

In some examples, the predefined region covering the restriction locus starts at least 25 bps upstream of the cut site within the restriction locus and ends at least 25 bps downstream of the cut site within the restriction locus.

In some examples, step (d) comprises determining the number of sequence reads covering a predefined genomic region of at least 100 bps in length that contains said restriction locus. In some examples, the predefined region covering the restriction locus starts at least 50 bps upstream of the cut site within the restriction locus and ends at least 50 bps downstream of the cut site within the restriction locus.

In some examples, the at least one restriction locus is located within a CG-island.

In some examples, the reference read count is a read count determined for the predefined genomic region of at least 50 bps in length that contains the restriction locus in an undigested control DNA sample, optionally corrected for sequencing depth differences.

In some examples, the reference read count is a read count determined using a reference region of at least 50 bps in length containing a reference locus that is not cut by the restriction endonuclease.

In some examples, the reference read count is an average read count determined using a plurality of reference regions of at least 50 bps in length containing reference loci that are not cut by the restriction endonuclease.

In some examples, calculating a methylation value comprises normalizing the read count determined in step (d) against a median read count of the DNA sample, to obtain a normalized read count, and calculating a ratio of the normalized read count to a normalized reference read count.

According to another example, the present disclosure provides a method for genetic and epigenetic profiling of a DNA sample, the method comprising determining a methylation value for at least one restriction locus as disclosed herein, and further determining from the sequencing data at least one additional genetic or epigenetic characteristic of the DNA sample selected from DNA mutation, copy number variation and nucleosome positioning.

In some examples, the DNA is cell-free DNA extracted from a biological fluid sample. In additional examples, the DNA is DNA extracted from a tumor sample.

According to a further example, the present disclosure provides a method for identifying genomic regions differentially methylated between a first and second source of DNA, the method comprising:

profiling methylation of at least one DNA sample from the first source as disclosed herein, to obtain a first DNA methylation profile;

profiling methylation of at least one DNA sample from the second source as disclosed herein, to obtain a second DNA methylation profile; and

comparing the first and second DNA methylation profiles to identify genomic regions differentially methylated between the first and second sources of DNA.

In some examples, the first source of DNA is a cancer DNA and the second source of DNA is a non-cancer DNA. In some examples, the first source of DNA is plasma cell-free DNA of a cancer patient and the second source of DNA is plasma cell-free DNA of one or more healthy individuals. In additional examples, the first and second sources of DNA are different stages of a cancer.

According to a further example, the present disclosure provides a method for profiling methylation of a DNA sample from a subject, the method comprising:

(a) providing a DNA sample from a subject;

(b) subjecting the DNA sample to digestion with at least one methylation-sensitive restriction endonuclease, thereby obtaining restriction endonuclease-treated DNA comprising restriction endonuclease-generated DNA fragments;

(c) performing high-throughput sequencing of the endonuclease-treated DNA to obtain sequence reads;

(d) determining from the sequence reads a read count of at least one restriction locus, the read count representing the number of DNA molecules in the DNA sample in which said at least one restriction locus was methylated and therefore remained intact;

(e) determining from the sequence reads a read count of sequence reads or ending at a nucleotide within the at least one restriction locus, the read count representing the number of DNA molecules in the DNA sample in which said at least one restriction locus was unmethylated and therefore cut by the restriction endonuclease;

(f) calculating a level of methylated DNA at the at least one restriction locus based on the read count of the at least one restriction locus determined in step (d) and a level of unmethylated DNA at the at least one restriction locus based on the read count of sequence reads starting or ending at a nucleotide within the at least one restriction locus determined in step (e),

thereby profiling methylation of the DNA sample.

In some examples, steps (c)-(e) comprise:

preparing a sequencing library from the restriction endonuclease-treated DNA using sequencing adapters ligated to a plurality of the restriction endonuclease-generated DNA fragments, and subjecting the sequencing library to high-throughput sequencing to obtain sequence reads;

mapping a plurality of the sequence reads against a reference genome to produce mapped sequence reads, and selecting at least one restriction locus within the reference genome;

determining from the mapped sequence reads a read count of the at least one restriction locus, the read count representing the number of DNA molecules in the DNA sample in which said at least one restriction locus was methylated and therefore remained intact; and

determining from the mapped sequence reads a read count of sequence reads starting or ending at a nucleotide within the at least one restriction locus, the read count representing

the number of DNA molecules in the DNA sample in which said at least one restriction locus was unmethylated and therefore cut by the restriction endonuclease.

In some examples, the high-throughput sequencing is whole genome high-throughput sequencing. In other examples, the high-throughput sequencing is target-specific high-throughput sequencing.

In some examples, the reference genome is the complete human genome.

In some examples, the DNA is cell-free DNA extracted from a biological fluid sample. In some examples, the biological fluid sample is plasma, serum or urine. Each possibility of the biological sample is a separate example of the present disclosure.

In some examples, the DNA is DNA extracted from a tumor sample.

In some examples, calculating a level of methylated DNA at the at least one restriction locus comprises calculating a ratio of the read count of the at least one restriction locus determined in step (d) to an expected read count of the at least one restriction locus.

In some examples, calculating a level of unmethylated DNA at the at least one restriction locus comprises calculating a difference between the read count of sequence reads starting or ending at a nucleotide within the at least one restriction locus determined in step (e) and an expected read count of sequences starting or ending at a nucleotide within the at least one restriction locus, and subsequently dividing the difference by an expected read count of the at least one restriction locus.

In some examples, calculating a level of methylated DNA at the at least one restriction locus comprises:

determining a total fragment number by summing the read count of the at least one restriction locus determined in step (d) and the read count of sequence reads starting or ending at a nucleotide within the at least one restriction locus determined in step (e), and subsequently subtracting from the sum an expected read count of sequences starting or ending at a nucleotide within the restriction locus; and dividing the read count of the at least one restriction locus determined in step (d) by the total fragment number.

In some examples, calculating a level of unmethylated DNA at the at least one restriction locus comprises:

determining a total fragment number as described herein;

calculating a difference between the read count of sequence reads starting or ending at a nucleotide within the at least one restriction locus determined in step (e) and an expected read count of sequences starting or ending at a nucleotide within the restriction locus; and

dividing the difference by the total fragment number.

In some examples, the expected read counts are read counts determined using a reference locus of the same length as the at least one restriction locus, that is not cut by the restriction endonuclease.

In some examples, the expected read counts are average read counts determined using a plurality of reference loci of the same length as the at least one restriction locus, that are not cut by the restriction endonuclease.

In some examples, the expected read counts are read counts determined for the at least one restriction locus in an undigested control DNA sample, optionally corrected for sequencing depth differences.

In some examples, the at least one restriction locus is a plurality of restriction loci.

In some examples, the at least one methylation-sensitive restriction endonuclease is a plurality of methylation-sensitive restriction endonucleases.

In some examples, the method for profiling methylation further comprises identifying the presence or absence of a disease in the subject based on the methylation profile of the DNA sample, by comparing the methylation profile of the DNA sample to one or more reference methylation profile.

In some examples, the method further comprises preparing a report in paper or electronic form based on the methylation profile and communicating the report to the subject and/or to a healthcare provider of the subject.

According to another example, the present disclosure provides a method for detecting methylation changes in a DNA sample, the method comprising: profiling methylation of the DNA sample as disclosed herein, to obtain a methylation profile of the DNA sample; and comparing the methylation profile of the DNA sample to one or more reference methylation profile to detect methylation changes in the DNA sample.

In some examples, the one or more reference methylation profile comprises a healthy DNA methylation profile. In additional examples, the one or more reference methylation profile comprises a disease DNA methylation profile. In

some examples, the DNA sample is from a subject suspected of having the disease and/or a subject at risk of

developing the disease, and detecting methylation changes comprises determining whether the DNA sample is a healthy or disease DNA sample. In some examples, the disease is a cancer.

According to a further example, the present disclosure provides a method for identifying genomic regions differentially methylated between a first and second source of DNA, the method comprising:

profiling methylation of at least one DNA sample from the first source according to the method disclosed herein, to obtain a first DNA methylation profile;

profiling methylation of at least one DNA sample from the second source according to the method disclosed herein, to obtain a second DNA methylation profile; and

comparing the first and second DNA methylation profiles to identify genomic regions differentially methylated between the first and second sources of DNA.

In some examples, the first source of DNA is a disease DNA and the second source of DNA is a non-disease DNA. In additional examples, the first and second sources of DNA are different stages of a disease. In some examples, the disease is a cancer.

According to a further example, the present disclosure provides a method for profiling genetic and epigenetic characteristics of a DNA sample, the method comprising:

profiling methylation of the DNA sample as disclosed herein; and

determining at least one additional genetic or epigenetic characteristic of the DNA sample, wherein the at least one additional genetic or epigenetic characteristic is selected from DNA mutation, copy number variation and nucleosome positioning,

wherein profiling the methylation and determining the at least one additional genetic or epigenetic characteristic are carried out using the same sequencing data,

thereby profiling genetic and epigenetic characteristics of the DNA sample.

These and further examples and features of the present disclosure will become apparent from the detailed description, examples and claims which follow.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1A. Copy number data of pooled plasma cell-free DNA samples subjected to methylation-sensitive digestion, bisulfite conversion or no treatment prior to sequencing. Data are presented as the number of hits (count) per genomic position.