

requires two different sequencing assays and two different pools of cfDNA. This can be cost-prohibitive in terms of plasma/cfDNA input and associated costs. In addition, destruction of DNA by bisulfite can reduce the sensitivity of variant-calling methods that can work on bisulfite-converted DNA sequencing data (relative to enzymatic conversion). Thus, improved methods for analyzing methylation of cfDNA are needed to preserve the integrity of sample nucleic acid and enable improved accuracy of methylation state analysis at the whole genome or targeted level.

SUMMARY

[0008a] In a first aspect, provided herein is a method of methylation sequencing a nucleic acid molecule of a biological sample comprising:

- a) ligating a nucleic acid adapter to the nucleic acid molecule, wherein the nucleic acid molecule comprises unconverted nucleic acids, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5Mc)-containing bases or 5-hydroxymethylcytosine (5hmC)-containing bases;
- b) converting unmethylated cytosines to uracils in the nucleic acid molecule using an enzymatic conversion method, thereby generating converted nucleic acids;
- c) amplifying the converted nucleic acids at least in part by polymerase chain reaction, thereby generating amplified converted nucleic acids;
- d) contacting the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci, thereby generating enriched converted nucleic acids;
- e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x; and
- f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to determine a methylation profile of the nucleic acid molecule of the biological sample.

[0008b] In a second aspect, provided herein is a method for determining a targeted methylation pattern in a nucleic acid molecule of a biological sample from a subject comprising:

- a) ligating a nucleic acid adapter to the nucleic acid molecule, wherein the nucleic acid molecule comprises unconverted nucleic acids, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine,

- thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5Mc)-containing bases or 5-hydroxymethylcytosine (5hmC)-containing bases;
- b) converting unmethylated cytosines to uracils in the nucleic acid molecule using an enzymatic conversion method to thereby produce converted nucleic acids;
 - c) amplifying the converted nucleic acids at least in part by polymerase chain reaction to thereby generate amplified converted nucleic acids;
 - d) contacting the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci;
 - e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x; and
 - f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to determine the targeted methylation pattern of the nucleic acid molecule of the biological sample from the subject.

[0009] Methods and systems provided herein address limitations of bisulfite-based methylation sequencing by improving the quality and accuracy of nucleic acid methylation sequencing and uses thereof for detection of disease. More accurate and complete information regarding methylation state permits higher quality feature generation for use in machine learning models and classifier generation.

[0010] In one example, a method is provided for performing methylation sequencing of a nucleic acid sample comprising:

- a) ligating a nucleic acid adapter comprising a unique molecular identifier to the nucleic acid molecule, wherein the nucleic acid molecule comprises unconverted nucleic acids;
- b) converting unmethylated cytosines to uracils in the nucleic acid molecule using a minimally-destructive conversion method, thereby generating converted nucleic acids;
- c) amplifying the converted nucleic acids by polymerase chain reaction, thereby generating amplified converted nucleic acids;
- d) probing the amplified converted nucleic acids with nucleic acid probes that are complementary to a pre-identified panel of CpG or CH loci to enrich for sequences corresponding to the panel, thereby generating probed converted nucleic acids;
- e) determining the nucleic acid sequence of the probed converted nucleic acids at a depth of >100x; and
- f) comparing the nucleic acid sequence of the probed converted nucleic acids to a reference

nucleic acid sequence of the pre-identified panel of CpG or CH loci to determine the methylation profile of the nucleic acid molecule of the biological sample.

[0011] In one example, the nucleic acid molecule is plasma cfDNA.

[0012] In one example, the minimally-destructive conversion method comprises enzymatic conversion, TAPS, or CAPS.

[0013] In one example, the unique molecular identifier is 4 bp to 6 bp in length and has a 5' thymidine overhang.

acid sequence of the pre-identified panel of CpG or CH loci to determine the methylation profile of the cell-free DNA (cfDNA) sample from the subject.

[0025] In one example, the determining of the nucleic acid sequence of the converted nucleic acids comprises duplex-like error correction.

[0026] In one example, the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine and cytosine bases, and not comprising 5mC-containing or 5hmC-containing bases.

[0027] In one example, the pre-identified panel of CpG or CH loci comprises loci associated with transcription factor start sites.

[0028] In one example, the targeted methylation pattern comprises hemi-methylated CpG loci.

[0028a] In a third aspect, provided herein is a method for determining a methylation profile of a cell-free deoxyribonucleic acid (cfDNA) sample from a subject comprising:

- a) ligating a nucleic acid adapter to the cfDNA, wherein the cfDNA comprises unconverted nucleic acids, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5mC)-containing or 5-hydroxymethylcytosine (5hmC)-containing bases;
- b) converting unmethylated cytosines to uracils in nucleic acid molecules using an enzymatic conversion method to thereby produce converted nucleic acids;
- c) amplifying the converted nucleic acids at least in part by polymerase chain reaction to thereby generate amplified converted nucleic acids;
- d) contacting the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci, thereby generating enriched converted nucleic acids;
- e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x; and
- f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to determine the methylation profile of the cell-free DNA (cfDNA) sample from the subject.

[0029] In one example, a method is provided for determining a methylation profile of a cell-free DNA (cfDNA) sample from a subject comprising:

- a) ligating a nucleic acid adapter comprising a unique molecular identifier to the cfDNA, wherein the cfDNA comprises unconverted nucleic acids;

- b) enzymatically converting unmethylated cytosines to uracils in nucleic acid molecules to produce converted nucleic acids;
- c) amplifying the converted nucleic acids by polymerase chain reaction;
- d) probing the converted nucleic acids with nucleic acid probes that are complementary to a pre-identified panel of CpG or CH loci to enrich for sequences corresponding to the pre-identified panel of CpG or CH loci;
- e) determining the nucleic acid sequence of the converted nucleic acids at a depth of >100x; and
- f) comparing the nucleic acid sequence of the converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG or CH loci to determine the methylation profile of the cell-free DNA (cfDNA) sample from the subject.

[0030] In one example, the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5mC-containing or 5hmC-containing bases.

[0031] In one example, the unique molecular identifier is 4 bp to 6 bp in length and has a 5' thymidine overhang.

[0032] In one example, the nucleic acid adapter further comprises a unique dual index (UDI) sequence. In one example, the UDI sequence is 4 bp, 5 bp, 6 bp, 7 bp, 8 bp, 9 bp, 10 bp, 11 bp, or 12 bp in length.

[0033] In one example, the method further comprises identifying a tissue-of-origin of the cfDNA sample, identifying a somatic variant in the cfDNA sample, inferring nucleosome

positioning in the cfDNA sample, identifying differentially methylated regions in the cfDNA sample, or identifying a haplotype block in the cfDNA sample.

[0034] Further provided herein is a method of methylation sequencing by duplex sequencing. Duplex sequencing is a tag-based error correction method that can improve sequencing accuracy, for example, methylation sequencing accuracy. In this method, adapters are ligated onto a nucleic acid template and amplified using PCR. In one example, the adapters comprise primer sequences and random 12 bp indices. Deep sequencing provides consensus sequence information from every unique molecular tag. Based on molecular tags and sequencing primers, duplex sequences can be aligned to determine the true sequence of the DNA. Advantages of duplex sequencing include very low error rate and detection and removal of PCR amplification errors. In duplex sequencing, there is also no need for additional library preparation steps after the addition of adapters.

[0035] In a one example, a method of methylation sequencing of a nucleic acid molecule of a biological sample comprising:

- a) preparing a methylation sequencing library from cfDNA fragments of the nucleic acid molecule comprising:
 - i) ligating a duplex adapter to the cfDNA fragments;
 - ii) ligating a duplex unique molecular identifier to the cfDNA fragments; and
 - iii) converting unmethylated cytosines to uracils in the cfDNA fragments using a minimally-destructive conversion method, thereby preparing the methylation sequencing library from the cfDNA of the nucleic acid molecule;
- b) enriching the methylation sequencing library for sequences corresponding to CpG or CH loci, thereby producing an enriched methylation sequencing library;
- c) sequencing the enriched methylation sequencing library at a depth of >100x using single-end or paired-end reads, thereby producing sequenced fragments of single-end or paired-end reads;
- d) for each sequenced fragment of the paired-end reads, correcting a sequencing error that falls within an overlap region of the paired-end reads;
- e) collapsing sequenced fragments into stranded read families to correct errors arising from PCR and sequencing; and
- f) collapsing the stranded read families into duplex read families to identify a methylation discrepancy in an inferred methylation state of symmetric CpG loci in the nucleic acid molecule.

[0036] In one example, the minimally-destructive conversion comprises enzymatic conversion, TAPS, or CAPS.

[0036a] In a fourth aspect, provided herein is a method of producing a classifier comprising:

- a) ligating a nucleic acid adapter to a nucleic acid molecule of a biological sample obtained from healthy subjects and biological samples from subjects having a cancer, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5mC)-containing or 5-hydroxymethylcytosine (5hmC)-containing bases;
- b) converting unmethylated cytosines to uracils in the nucleic acid molecule using an enzymatic conversion method, thereby generating converted nucleic acids;
- c) amplifying the converted nucleic acids at least in part by polymerase chain reaction, thereby generating amplified converted nucleic acids;
- d) probing the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci, thereby generating enriched converted nucleic acids;
- e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x;
- f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to obtain sets of measured values of machine learning input features that are representative of methylation profiles from the healthy subjects and from the subjects having the cancer; and
- g) training, using the machine learning input features, a machine learning model to produce the classifier that distinguishes between the healthy subjects and the subjects having the cancer.

[0037] In one example, a method is provided for producing a classifier comprising:

- a) ligating a nucleic acid adapter comprising a unique molecular identifier to a nucleic acid molecule of a biological sample obtained from healthy subjects and biological samples from subjects having a cancer;
- b) converting unmethylated cytosines to uracils in the nucleic acid molecule using a minimally-destructive conversion method, thereby generating converted nucleic acids;
- c) amplifying the converted nucleic acids with polymerase chain reaction, thereby generating amplified converted nucleic acids;
- d) probing the amplified converted nucleic acids with nucleic acid probes that are complementary to a pre-identified panel of CpG or CH loci to enrich for sequences

- e) corresponding to the panel, thereby generating probed converted nucleic acids;
- f) determining the nucleic acid sequence of the probed converted nucleic acids at a depth of >100x;
- g) comparing the nucleic acid sequence of the probed converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG or CH loci to obtain sets of measured values of input features that are representative of methylation profiles from the healthy subjects and from the subjects having the cancer; and
- h) training a machine learning model to produce the classifier that distinguishes between the healthy subjects and the subjects having the cancer.

[0038] In one example, the pre-identified panel of CpG or CH loci comprises loci associated with transcription start sites.

[0039] In one example, the method further comprises determining hemi-methylated CpG or CH loci.

[0040] In one example, the method further comprises identifying a tissue-of-origin for the nucleic acid molecule.

[0041] In one example, the method further comprises identifying genomic position and fragment length for the nucleic acid molecule.

[0042] In one example, the unique molecular identifier is 4 bp to 6 bp in length and has a 5' thymidine overhang.

[0043] In one example, the nucleic acid adapter further comprises a unique dual index (UDI) sequence. In one example, the UDI sequence is 4 bp, 5 bp, 6 bp, 7 bp, 8 bp, 9 bp, 10 bp, 11 bp, or 12 bp in length.

[0044] In one example, the amplifying of the converted nucleic acids comprises using primers that contain a unique dual index (UDI) sequence. In one example, the UDI sequence is 4 bp, 5 bp, 6 bp, 7 bp, 8 bp, 9 bp, 10 bp, 11 bp, or 12 bp in length.

[0045] In one example, the input features are selected from base wise methylation % for CpG, base wise methylation % for CHG, base wise methylation % for CHH, the count or rate of observing fragments with different counts or rates of methylated CpGs in a region, conversion efficiency (e.g., 100-Mean methylation % for CHH), hypomethylated blocks, methylation levels for CPG, methylation levels for CHH, methylation levels for CHG, fragment length, fragment midpoint, methylation levels for chrM, methylation levels for LINE1, methylation levels for ALU, dinucleotide coverage (e.g., normalized coverage of dinucleotide), evenness of coverage (e.g., unique CpG sites at 1x and 10x mean genomic coverage (e.g., for S4 runs)), mean CpG coverage (e.g., depth) globally, and mean coverage at CpG islands, CGI shelves, and CGI shores.

[0046] In one example, a classifier is provided that distinguishes a population of healthy individuals from individuals with a cancer comprising: sets of measured values representative of methylation profiles from methylation sequencing data from healthy subjects and subjects having the cancer,

wherein the measured values are used to generate a set of features corresponding to properties of the methylation profiles, wherein the set of features are inputted to a machine learning or statistical model, wherein the model provides a feature vector useful as a classifier that distinguishes the population of healthy individuals from individuals having the cancer.

[0047] In one example, a method is provided for detecting cancer in a population of subjects comprising:

- a) assaying by using targeted minimally-destructive conversion methyl sequencing nucleic acids of a biological sample from a subject to obtain a methylation profile of the nucleic acids;
- b) classifying the biological sample by inputting the methylation profile to a trained algorithm that classifies samples from healthy subjects and subjects having the cancer; and
- c) outputting a report on a computer screen that identifies the biological sample as negative for the cancer if the trained algorithm classifies the biological sample as negative for the cancer at a specified confidence level.

[0048] In one example, the cancer is colorectal cancer.

[0049] In one example, the present disclosure provides a system for classifying individuals based on methylation state comprising:

- a) a computer readable medium product comprising a classifier, wherein the classifier comprises: sets of measured values representative of methylation profiles from methylation sequencing data from healthy subjects and subjects having a cancer, wherein

the measured values are used to generate a set of features corresponding to properties of the methylation profiles from healthy subjects and subjects having the cancer, wherein the features are inputted to a machine learning or statistical model, wherein the model provides a feature vector useful as a classifier that distinguishes a population of healthy individuals from individuals having the cancer; and

- b) one or more processors for executing instructions stored on the computer readable medium product.

[0050] In one example, the system comprises a classification circuit that is configured as a machine learning classifier selected from a linear discriminant analysis (LDA) classifier, a quadratic discriminant analysis (QDA) classifier, a support vector machine (SVM) classifier, a random forest (RF) classifier, a linear kernel support vector machine classifier, a first order polynomial kernel support vector machine classifier, a second order polynomial kernel support vector machine classifier, a ridge regression classifier, an elastic net algorithm classifier, a sequential minimal optimization algorithm classifier, a naive Bayes algorithm classifier, and a non-negative matrix factorization (NMF) predictor algorithm classifier.

[0051] In one example, the system comprises means for performing any of the above methods.

[0052] In one example, the system comprises one or more processors configured to perform any of the above methods.

[0053] In one example, the system comprises modules that respectively perform the steps of any of the above methods.

[0054] In another example, the present disclosure provides a method for monitoring minimal residual disease status in a subject previously treated for a disease comprising: determining a methylation profile as described herein as a baseline methylation state and repeating an analysis to determine the methylation profile at one or more pre-determined time points, wherein a change from baseline indicates a change in the minimal residual disease status at baseline in the subject.

[0055] In one example, the present disclosure provides a method for monitoring minimal residual disease status in a subject previously treated for a disease comprising:

- a) determining a baseline methylation profile of a biological sample obtained from the subject at a baseline methylation state;
- b) determining a test methylation profile of a biological sample obtained from the subject at one or more pre-determined time points following the baseline methylation state; and
- c) determining a change in the test methylation profile as compared to the baseline methylation profile, wherein the change indicates a change in the minimal residual disease status of the subject.

[0056] In some examples, the disease is a cancer. In some examples, the disease is a colorectal cancer.

[0057] In another example, the present disclosure provides a method for monitoring minimal residual disease status in a subject previously treated for a colorectal cancer comprising: detecting a methylated fragment in a biological sample from the subject, wherein the methylated fragment in the biological sample indicates a change in the minimal residual disease status at baseline for the colorectal cancer in the subject.

[0058] In some examples, the minimal residual disease status is selected from response to treatment, tumor load, residual tumor post-surgery, relapse, secondary screen, primary screen, and cancer progression.

[0059] In another example, a method is provided for determining response to treatment for a subject.

[0060] In another example, a method is provided for monitoring tumor load in a subject.

[0061] In another example, a method is provided for detecting residual tumor post-surgery in a subject.

[0062] In another example, a method is provided for detecting relapse in a subject.

[0063] In another example, a method is provided for use as a secondary screen of a subject.

[0064] In another example, a method is provided for use as a primary screen of a subject.

[0065] In another example, a method is provided for monitoring cancer progression in a subject.

[0066] In one example, the present disclosure provides a kit for detecting a tumor comprising reagents for carrying out the aforementioned methods, and instructions for detecting tumor signals, for example, methylation signatures. Reagents may include, for example, primer sets, PCR reaction components, sequencing reagents, minimally-destructive conversion reagents, and library preparation reagents.

BRIEF DESCRIPTION OF DRAWING

[0067] **FIG. 1** provides a flow diagram showing conventional methyl bisulfite conversion and degradation compared to the modified methods that preserve fragment length information as described herein.

[0068] **FIG. 2** provides a schematic showing staggered adapters useful in the methods described herein.

[0069] **FIG. 3** provides a schematic showing duplex-like error correction methods described

CLAIMS

WHAT IS CLAIMED IS:

1. A method of methylation sequencing a nucleic acid molecule of a biological sample comprising:
 - a) ligating a nucleic acid adapter to the nucleic acid molecule, wherein the nucleic acid molecule comprises unconverted nucleic acids, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5Mc)-containing bases or 5-hydroxymethylcytosine (5hmC)-containing bases;
 - b) converting unmethylated cytosines to uracils in the nucleic acid molecule using an enzymatic conversion method, thereby generating converted nucleic acids;
 - c) amplifying the converted nucleic acids at least in part by polymerase chain reaction, thereby generating amplified converted nucleic acids;
 - d) contacting the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci, thereby generating enriched converted nucleic acids;
 - e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x; and
 - f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to determine a methylation profile of the nucleic acid molecule of the biological sample.
2. The method of claim 1, wherein the nucleic acid molecule is plasma cell-free deoxyribonucleic acid (cfDNA).
3. The method of claim 1, wherein the enzymatic conversion method comprises TET-assisted pyridine borane sequencing (TAPS) or chemical-assisted pyridine borane sequencing (CAPS).
4. The method of claim 1, wherein the nucleic acid adapter further comprises a unique molecular identifier.
5. The method of claim 4, wherein the unique molecular identifier is 4 nucleotides to 6 nucleotides in length and comprises ' 5' thymidine overhang.
6. The method of claim 5, wherein the nucleic acid adapter further comprises a unique dual index (UDI) sequence.

7. The method of claim 6, wherein the UDI sequence is 4 nucleotides, 5 nucleotides, 6 nucleotides, 7 nucleotides, 8 nucleotides, 9 nucleotides, 10 nucleotides, 11 nucleotides, or 12 nucleotides in length.
8. The method of claim 1, wherein the amplifying of the converted nucleic acids comprises using primers that contain a unique dual index (UDI) sequence.
9. The method of claim 1, wherein the nucleic acid probes comprise unmethylated nucleic acid probes.
10. The method of claim 1, wherein the nucleic acid probes hybridize to target regions of interest that are consistent with unmethylated cytosines at CpG sites in the reference nucleic acid sequence.
11. The method of claim 1, wherein the nucleic acid probes comprise the target regions of interest that are consistent with methylated cytosines at CpG sites in the reference nucleic acid sequence.
12. The method of claim 1, further comprising chemically or enzymatically altering the nucleic acid probes, wherein the chemically or enzymatically altered nucleic acid probes are methylated or unmethylated nucleic acid probes.
13. The method of claim 1, wherein one or more cytosines in CG loci of the enriched nucleic acids are converted to thymines, and wherein one or more cytosines in CH loci of the enriched nucleic acids are converted to thymines.
14. The method of claim 1, wherein the converting of the unmethylated cytosines to the uracils further comprises performing sequential Ten-eleven translocation/apolipoprotein B MRNA-editing enzyme (TET/APOBEC) enzymatic conversion.
15. The method of claim 1, wherein the converting of the unmethylated cytosines to uracils comprises performing TET-assisted pyridine borane sequencing (TAPS).
16. A method for determining a targeted methylation pattern in a nucleic acid molecule of a biological sample from a subject comprising:
 - a) ligating a nucleic acid adapter to the nucleic acid molecule, wherein the nucleic acid molecule comprises unconverted nucleic acids, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5Mc)-containing bases or 5-hydroxymethylcytosine (5hmC)-containing bases;
 - b) converting unmethylated cytosines to uracils in the nucleic acid molecule using an enzymatic conversion method to thereby produce converted nucleic acids;
 - c) amplifying the converted nucleic acids at least in part by polymerase chain reaction to thereby generate amplified converted nucleic acids;

- d) contacting the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci;
 - e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x; and
 - f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to determine the targeted methylation pattern of the nucleic acid molecule of the biological sample from the subject.
17. The method of claim 16, wherein the determining of the nucleic acid sequence of the enriched converted nucleic acids comprises duplex-like error correction.
18. The method of claim 16, wherein the nucleic acid adapter further comprises a unique molecular identifier.
19. The method of claim 16, wherein the pre-identified panel of CpG, CHG, or CHH loci comprises loci associated with transcription factor start sites.
20. The method of claim 16, wherein the targeted methylation pattern comprises hemi-methylated CpG loci.
21. A method for determining a methylation profile of a cell-free deoxyribonucleic acid (cfDNA) sample from a subject comprising:
- a) ligating a nucleic acid adapter to the cfDNA, wherein the cfDNA comprises unconverted nucleic acids, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5mC)-containing or 5-hydroxymethylcytosine (5hmC)-containing bases;
 - b) converting unmethylated cytosines to uracils in nucleic acid molecules using an enzymatic conversion method to thereby produce converted nucleic acids;
 - c) amplifying the converted nucleic acids at least in part by polymerase chain reaction to thereby generate amplified converted nucleic acids;
 - d) contacting the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci, thereby generating enriched converted nucleic acids;
 - e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x; and

- f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to determine the methylation profile of the cell-free DNA (cfDNA) sample from the subject.
22. The method of claim 21, wherein the nucleic acid adapter further comprises a unique molecular identifier.
23. The method of claim 21, further comprising identifying a tissue-of-origin of the cfDNA sample, identifying a somatic variant in the cfDNA sample, inferring nucleosome positioning in the cfDNA sample, identifying differentially methylated regions in the cfDNA sample, or identifying a haplotype block in the cfDNA sample.
24. A method of producing a classifier comprising:
- a) ligating a nucleic acid adapter to a nucleic acid molecule of a biological sample obtained from healthy subjects and biological samples from subjects having a cancer, wherein the nucleic acid adapter is a conversion-tolerant adapter comprising guanine, thymine, adenine, and cytosine bases, and not comprising 5-methylcytosine (5mC)-containing or 5-hydroxymethylcytosine (5hmC)-containing bases;
 - b) converting unmethylated cytosines to uracils in the nucleic acid molecule using an enzymatic conversion method, thereby generating converted nucleic acids;
 - c) amplifying the converted nucleic acids at least in part by polymerase chain reaction, thereby generating amplified converted nucleic acids;
 - d) probing the amplified converted nucleic acids with nucleic acid probes that are at least partially complementary to a pre-identified panel of CpG, CHG, or CHH loci to enrich for sequences corresponding to the pre-identified panel of CpG, CHG, or CHH loci, thereby generating enriched converted nucleic acids;
 - e) determining a nucleic acid sequence of the enriched converted nucleic acids at a depth of greater than 100x;
 - f) comparing the nucleic acid sequence of the enriched converted nucleic acids to a reference nucleic acid sequence of the pre-identified panel of CpG, CHG, or CHH loci to obtain sets of measured values of machine learning input features that are representative of methylation profiles from the healthy subjects and from the subjects having the cancer; and
 - g) training, using the machine learning input features, a machine learning model to produce the classifier that distinguishes between the healthy subjects and the subjects having the cancer.

25. The method of claim 24, wherein the nucleic acid adapter further comprises a unique molecular identifier.
26. The method of claim 24, wherein the pre-identified panel of CpG, CHG, or CHH loci comprises loci associated with transcription start sites.
27. The method of claim 24, further comprising determining hemi-methylated CpG, CHG, or CHH loci.
28. The method of claim 24, further comprising identifying a tissue-of-origin for the nucleic acid molecule.
29. The method of claim 24, further comprising identifying genomic position and fragment length for the nucleic acid molecule.
30. The method of claim 24, wherein the input features are selected from base wise methylation % for CpG, base wise methylation % for CHG, base wise methylation % for CHH, the count or rate of observing fragments with different counts or rates of methylated CpGs in a region, conversion efficiency, hypomethylated blocks, methylation levels for CPG, methylation levels for CHH, methylation levels for CHG, fragment length, fragment midpoint, methylation levels for chrM, methylation levels for LINE1, methylation levels for ALU, dinucleotide coverage, evenness of coverage, mean CpG coverage globally, and mean coverage at CpG islands, CGI shelves, and CGI shores.