

CLAIMS

WHAT IS CLAIMED IS:

1. A methylation signature panel characteristic of colon cell proliferative disorder comprising:

one or more methylated genomic regions selected from the group consisting of **Table 11**, wherein the one or more regions are more methylated in a biological sample from an individual having a colon cell proliferative disorder or colon cell proliferative disorder subtypes, and are less methylated in normal tissues and normal blood cells in an individual not having a colon cell proliferative disorder.
2. The methylation signature panel of claim 1, wherein the biological sample is a nucleic acid, DNA, RNA, or cell-free nucleic acid (cfDNA or cfRNA).
3. The methylation signature panel of claim 1, wherein the signature panel comprises increased methylation in two or more genomic regions selected from the group consisting of **Table 11**.
4. The methylation signature panel of claim 1, wherein the colon cell proliferative disorder is selected from the group consisting of adenoma (adenomatous polyps), sessile serrated adenoma (SSA), advanced adenoma, colorectal dysplasia, colorectal adenoma, colorectal cancer, colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.
5. The methylation signature panel of claim 1, wherein the colon cell proliferative disorder is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, and stage 4 colorectal cancer.
6. The methylation signature panel of claim 1, wherein the signature panel includes two or more methylated genomic regions in **Tables 1-11**, three or more methylated genomic regions in **Tables 1-11**, four or more methylated genomic regions in **Tables 1-11**, five or more methylated genomic regions in **Tables 1-11**, six or more methylated genomic regions in **Tables 1-11**, seven or more methylated genomic regions in **Tables 1-11**, eight or more methylated genomic regions in **Tables 1-11**, nine or more methylated genomic regions in **Tables 1-11**, ten or more methylated genomic regions in **Tables 1-11**, eleven or more methylated genomic regions in **Tables 1-11**, twelve or more methylated genomic regions in **Tables 1-11**, or thirteen or more methylated genomic regions in **Tables 1-11**.

7. The methylation signature panel of claim 1, wherein the signature panel includes genomic regions methylated in colorectal cancer comprising methylated regions in one or more genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, and FLI1.

8. The methylation signature panel of claim 1, wherein the regions methylated in colorectal cancer comprise methylated regions selected from the group consisting of IKZF1, KCNQ5, and ELMO1 genomic regions.

9. The methylation signature panel of claim 1, wherein the regions methylated in colorectal cancer comprise methylated regions in one or more genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, FLI1, CLIP4, ELOVL5, FAM72B, and ST3GAL1.

10. The methylation signature panel of claim 1, wherein the signature panel comprises methylated genomic regions selected from the group consisting of **Table 1, Table 2, Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, Table 9, Table 10, and Table 11.**

11. A methylation signature panel characteristic of a colon cell proliferative disorder comprising:

two or more methylated genomic regions selected from the group consisting of **Tables 1-11**, wherein the two or more regions are more methylated in a biological sample from an individual having a colon cell proliferative disorder or colon cell proliferative disorder subtypes, and are less methylated in normal tissues and normal blood cells in an individual not having a colon cell proliferative disorder.

12. The methylation signature panel of claim 11, wherein the biological sample is a nucleic acid, DNA, RNA, or cell-free nucleic acid.

13. The methylation signature panel of claim 11, wherein the signature panel comprises increased methylation in 6 or more genomic regions selected from the group consisting of **Tables 1-11.**

14. The methylation signature panel of claim 11, wherein the colon cell proliferative disorder is selected from the group consisting of adenoma (adenomatous polyps), sessile serrated adenoma (SSA), advanced adenoma, colorectal dysplasia, colorectal adenoma, colorectal cancer, colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

15. The methylation signature panel of claim 11, wherein the colon cell proliferative disorder is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, and stage 4 colorectal cancer.

16. The methylation signature panel of claim 11, wherein the signature panel includes three or more methylated genomic regions in **Tables 1-11**, four or more methylated genomic regions in **Tables 1-11**, five or more methylated genomic regions in **Tables 1-11**, six or more methylated genomic regions in **Tables 1-11**, seven or more methylated genomic regions in **Tables 1-11**, eight or more methylated genomic regions in **Tables 1-11**, nine or more methylated genomic regions in **Tables 1-11**, ten or more methylated genomic regions in **Tables 1-11**, eleven or more methylated genomic regions in **Tables 1-11**, twelve or more methylated genomic regions in **Tables 1-11**, or thirteen or more methylated genomic regions in **Tables 1-11**.

17. The methylation signature panel of claim 11, wherein the signature panel includes genomic regions methylated in colorectal cancer comprising methylated regions in one or more genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, and FLI1.

18. The methylation signature panel of claim 11, wherein the regions methylated in colorectal cancer comprise methylated regions selected from the group consisting of IKZF1, KCNQ5, and ELMO1 genomic regions.

19. The methylation signature panel of claim 11, wherein the regions methylated in colorectal cancer comprise methylated regions in one or more genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, FLI1, CLIP4, ELOVL5, FAM72B, and ST3GAL1.

20. The methylation signature panel of claim 11, wherein the signature panel comprises methylated genomic regions selected from the group consisting of **Table 1, Table 2, Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, Table 9, Table 10, and Table 11**.

21. A machine learning classifier capable of distinguishing a population of healthy individuals from individuals with colon cell proliferative disorder, comprising:

a) sets of measured values representative of differentially-methylated genomic regions of claim 1 where the measured values are obtained from methylation sequencing data from healthy subjects and subjects having a colon cell proliferative disorder;

b) wherein the measured values are used to generate a set of features corresponding to properties of the differentially-methylated genomic regions and where the features are inputted to a machine learning or statistical model;

c) wherein the model provides a feature vector useful as a classifier capable of distinguishing a population of healthy individuals from individuals having a colon cell proliferative disorder.

22. The classifier of claim 21, wherein the sets of measured values describe characteristics of the methylated regions selected from the group consisting of: base wise methylation percent for CpG, CHG, CHH, the count or rate of observing fragments with different counts or rates of methylated CpGs in a region, conversion efficiency (100-Mean methylation percent for CHH), hypomethylated blocks, methylation levels (global mean methylation for CPG, CHH, CHG, fragment length, fragment midpoint, number of methylated CpGs per fragment, fraction of CpG methylation to total CpG per fragment, fraction of CpG methylation to total CpG per region, fraction of CpG methylation to total CpG in panel, dinucleotide coverage (normalized coverage of dinucleotide), evenness of coverage (unique CpG sites at 1x and 10x mean genomic coverage (for S4 runs), mean CpG coverage (depth) globally and mean coverage at CpG islands, CGI shelves, and CGI shores.

23. A system comprising a machine learning model classifier for detecting a colon cell proliferative disorder, comprising:

a) a computer-readable medium comprising a classifier operable to classify subjects as having the colon cell proliferative disorder or not having the colon cell proliferative disorder based on a methylation signature panel; and

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

24. The system of claim 23, comprising the classifier of claim 21 loaded into a memory of a computer system, the machine learning model trained using training vectors obtained from training biological samples, a first subset of the training biological samples identified as having a colon cell proliferative disorder and a second subset of the training biological samples identified as not having a colon cell proliferative disorder.

25. A method for determining a methylation profile of a cell-free deoxyribonucleic acid (cfDNA) sample from an individual, comprising:

a) providing conditions capable of converting unmethylated cytosines to uracils in nucleic acid molecules of the cfDNA sample to produce a plurality of converted nucleic acids;

b) contacting the plurality of converted nucleic acids with nucleic acid probes complementary to a pre-identified methylation signature panel of at least two differentially methylated regions selected from the group consisting of **Tables 1-11** to enrich for sequences corresponding to the signature panel;

c) determining nucleic acid sequences of the plurality of converted nucleic acid molecules; and

d) aligning the nucleic acid sequences of the plurality of converted nucleic acid molecules to a reference nucleic acid sequence, thereby determining the methylation profile of the individual.

26. The method of claim 25, further comprising amplifying the plurality of converted nucleic acids.

27. The method of claim 26, wherein the amplifying comprises polymerase chain reaction (PCR).

28. The method of claim 25, further comprising determining the nucleic acid sequences of the converted nucleic acid molecules at a depth of greater than 1000x, greater than 2000x, greater than 3000x, greater than 4000x, or greater than 5000x.

29. The method of claim 25, wherein the reference nucleic acid sequence is at least a portion of a human reference genome.

30. The method of claim 29, wherein the human reference genome is hg18.

31. The method of claim 25, wherein the pre-identified methylation signature panel includes three or more methylated genomic regions in **Tables 1-11**, four or more methylated genomic regions in **Tables 1-11**, five or more methylated genomic regions in **Tables 1-11**, six or more methylated genomic regions in **Tables 1-11**, seven or more methylated genomic regions in **Tables 1-11**, eight or more methylated genomic regions in **Tables 1-11**, nine or more methylated genomic regions in **Tables 1-11**, ten or more methylated genomic regions in **Tables 1-11**, eleven or more methylated genomic regions in **Tables 1-11**, twelve or more methylated genomic regions in **Tables 1-11**, or thirteen or more methylated genomic regions in **Tables 1-11**.

32. The method of claim 31, wherein the pre-identified methylation signature panel includes one or more methylated genomic regions in **Table 11**, two or more methylated genomic regions in **Table 11**, or three methylated genomic regions in **Table 11**.

33. The method of claim 25, wherein the methylation profile is indicative of a presence or an absence of a colon cell proliferative disorder in the individual.

34. The method of claim 33, wherein the colon cell proliferative disorder is selected from the group consisting of adenoma (adenomatous polyps), sessile serrated adenoma (SSA), advanced adenoma, colorectal dysplasia, colorectal adenoma, colorectal cancer, colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

35. The method of claim 33, wherein the colon cell proliferative disorder is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, or stage 4 colorectal cancer.

36. A method for detecting a presence or an absence of a colon cell proliferative disorder in a subject, comprising:

a) providing conditions capable of converting unmethylated cytosines to uracils in nucleic acid molecules of a biological sample obtained or derived from the subject to produce a plurality of converted nucleic acids;

b) contacting the plurality of converted nucleic acids with nucleic acid probes complementary to a pre-identified methylation signature panel of at least two differentially methylated regions selected from the group consisting of **Tables 1-11** to enrich for sequences corresponding to the signature panel;

c) determining nucleic acid sequences of the converted nucleic acid molecules;

d) aligning the nucleic acid sequences of the plurality of converted nucleic acid molecules to a reference nucleic acid sequence, thereby determining a methylation profile of the individual; and

e) applying a trained machine learning classifier to the methylation profile, wherein the trained machine learning classifier is trained to be capable of distinguishing between healthy individuals and individuals with a colon cell proliferative disorder to provide an output value associated with presence of a colon cell proliferative disorder, thereby detecting the presence or the absence of the colon cell proliferative disorder in the subject.

37. The method of claim 36, wherein the biological sample obtained from the subject is selected from the group consisting of cell-free DNA, cell-free RNA, body fluids, stool, colonic

effluent, urine, blood plasma, blood serum, whole blood, isolated blood cells, cells isolated from the blood, and combinations thereof.

38. The method of claim 36, further comprising amplifying the plurality of converted nucleic acids.

39. The method of claim 38, wherein the amplifying comprises polymerase chain reaction (PCR).

40. The method of claim 36, further comprising determining the nucleic acid sequences of the converted nucleic acid molecules at a depth of greater than 1000x, greater than 2000x, greater than 3000x, greater than 4000x, or greater than 5000x.

41. The method of claim 36, wherein the reference nucleic acid sequence is at least a portion of a human reference genome.

42. The method of claim 41, wherein the human reference genome is hg18.

43. The method of claim 36, wherein the pre-identified methylation signature panel includes three or more methylated genomic regions in **Tables 1-11**, four or more methylated genomic regions in **Tables 1-11**, five or more methylated genomic regions in **Tables 1-11**, six or more methylated genomic regions in **Tables 1-11**, seven or more methylated genomic regions in **Tables 1-11**, eight or more methylated genomic regions in **Tables 1-11**, nine or more methylated genomic regions in **Tables 1-11**, ten or more methylated genomic regions in **Tables 1-11**, eleven or more methylated genomic regions in **Tables 1-11**, twelve or more methylated genomic regions in **Tables 1-11**, or thirteen or more methylated genomic regions in **Tables 1-11**.

44. The method of claim 43, wherein the pre-identified methylation signature panel includes one or more methylated genomic regions in **Table 11**, two or more methylated genomic regions in **Table 11**, or three methylated genomic regions in **Table 11**.

45. The method of claim 36, further comprising administering a treatment to the individual for the colon cell proliferative disorder based on detecting the presence of the colon cell proliferative disorder in the individual.

46. The method of claim 36, wherein the colon cell proliferative disorder is selected from the group consisting of adenoma (adenomatous polyps), sessile serrated adenoma (SSA), advanced adenoma, colorectal dysplasia, colorectal adenoma, colorectal cancer, colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors,

gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

47. The method of claim 36, wherein the colon cell proliferative disorder comprises the colorectal cancer.

48. The method of claim 36, wherein the colon cell proliferative disorder is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, and stage 4 colorectal cancer.

49. The method of claim 36, wherein the trained machine learning classifier is selected from the group consisting of a deep learning classifier, a neural network classifier, 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 or 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 principal component analysis classifier.