assessed in cancers including lung, colon, and breast. These may suffer from low sensitivities as they may be insufficiently prevalent in the tumors.

[0006] There remains a need for more sensitive and specific screening tools for detecting early-stage or low tumor-burden colorectal cancer tumor signals in relapse and primary screening in at risk populations.

SUMMARY

[0007] The present disclosure provides methods and systems directed to methylation-profiling of genes associated with colorectal cancer detection and disease progression.

[0007a] In an aspect, the present disclosure provides a machine learning classifier capable of distinguishing a population of subjects without a colorectal cancer from subjects with the colorectal cancer, comprising: a) sets of measured values representative of differentially-methylated genomic regions disclosed herein where the measured values are obtained from methylation sequencing data from subjects without the colorectal cancer and subjects having the colorectal cancer; b) wherein the measured values are used to generate a set of features corresponding to properties of the differentially-methylated genomic regions and wherein the features are inputted to a machine learning or statistical model; c) wherein the machine learning or statistical model provides a feature vector useful as a classifier capable of distinguishing a population of subjects without a colorectal cancer from individuals having the colorectal cancer.

[0007b] In an example, the present disclosure provides 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.

[0008] In an example, the present disclosure provides a methylation signature panel characteristic of a 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.

[0008a] Thus, in one aspect, there is provided a methylation signature panel characteristic of a colorectal cancer comprising: one or more genomic regions selected from the group consisting of **Table 12**, wherein the one or more genomic regions are more methylated in a cell-free deoxyribonucleic acid (cfDNA) sample from a subject having the colorectal cancer or colorectal cancer subtypes, and are less methylated in a cfDNA sample from a subject not having the colorectal cancer.

[0009] In some examples, the biological sample is a nucleic acid, DNA, ribonucleic acid (RNA), or cell-free nucleic acid (e.g., cfDNA or cfRNA).

[0010] In some examples, the genomic region is a non-coding region, a coding region, or a non-transcribed or regulator region.

[0011] In some examples, the signature panel comprises increased methylation in two or more genomic regions selected from the group consisting of **Table 11**.

[0012] In some examples, 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.

[0013] In some examples, 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. In some embodiments, the colon cell proliferative disorder comprises the colorectal cancer.

[0014] In some examples, 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.

[0015] In some examples, the signature panel comprises 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**.

[0016] In some embodiments, the signature panel comprises genomic regions methylated in colorectal cancer comprising methylated regions in one or more genomic regions selected from the group consisting of ITGA4, EMBP1, TMEM163, SFMBT2, ELMO, and ZNF543.

[0017] In some embodiments, the regions methylated in colorectal cancer comprise methylated regions in both ITGA4 and EMBP1 genomic regions.

[0018] In some embodiments, the regions methylated in colorectal cancer comprise methylated regions in one or more genomic regions selected from the group consisting of ITGA4, EMBP1, TMEM163, SFMBT2, ELMO, ZNF543, CHST10, CCNA1, BEND4, KRBA1, S1PR1, and PPP1R16B.

[0019] In some examples, 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**.

[0020] In another example, the present disclosure provides a methylation signature panel characteristic of a colon cell proliferative disorder comprising: two or more methylated genomic regions in **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.

[0020a] Thus, in an aspect, there is provided a methylation signature panel characteristic of a colorectal cancer comprising: two or more genomic regions selected from the group consisting of **Table 12**, wherein the two or more genomic regions are more methylated in a cell-free deoxyribonucleic acid (cfDNA) sample from a subject having the colorectal cancer or colorectal cancer subtypes, and are less methylated in a cfDNA sample from a subject not having the colorectal cancer.

[0021] In some examples, the biological sample is a nucleic acid, DNA, ribonucleic acid (RNA), or cell-free nucleic acid (cfDNA or cfRNA).

[0022] In some examples, the genomic region is a non-coding region, a coding region, or a non-transcribed or regulator region.

[0023] In some examples, the signature panel comprises increased methylation in 6 or more, or 12 or more genomic regions in **Tables 1-11**.

[0024] In some examples, 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.

[0025] In some examples, 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. In some embodiments, the colon cell proliferative disorder comprises the colorectal cancer.

[0026] In some examples, 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.

[0027] In some examples, the signature panel comprises 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**.

[0028] In some embodiments, the signature panel comprises genomic regions methylated in colorectal cancer comprising methylated regions in one or more genomic regions selected from the group consisting of ITGA4, EMBP1, TMEM163, SFMBT2, ELMO, and ZNF543.

[0029] In some embodiments, the regions methylated in colorectal cancer comprise methylated regions in both ITGA4 and EMBP1 genomic regions.

[0030] In some embodiments, the regions methylated in colorectal cancer comprise methylated regions in one or more genomic regions selected from the group consisting of ITGA4, EMBP1,

TMEM163, SFMBT2, ELMO, ZNF543, CHST10, CCNA1, BEND4, KRBA1, S1PR1, and PPP1R16B.

[0031] In some examples, the signature panel comprises methylated 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**.

[0032] In another example, the present disclosure provides a classifier (e.g., 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 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; and 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.

[0032a] Thus, in one aspect, there is provided a machine learning classifier capable of distinguishing a population of subjects without a colorectal cancer from subjects with the colorectal cancer, 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 subjects without the colorectal cancer and subjects having the colorectal cancer; b) wherein the measured values are used to generate a set of features corresponding to properties of the differentially-methylated genomic regions and wherein the features are inputted to a machine learning or statistical model; c) wherein the machine learning or statistical model provides a feature vector useful as a classifier capable of distinguishing a population of subjects without the colorectal cancer from subjects having the colorectal cancer.

[0033] In some embodiments, 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, and methylation levels in one or more genomic regions such as chrM, LINE1, or ALU), 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 lx and 10x mean genomic coverage (for S4 runs), mean CpG coverage (depth) globally, and mean coverage at CpG islands, CGI shelves, and CGI shores.

[0034] In some embodiments, the machine learning model comprising the classifier is 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.

[0035] In some embodiments, the classifier is provided in a system 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.

[0036] In some embodiments, the system comprises a classification circuit that is configured as a machine learning classifier 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 principal component analysis classifier.

[0037] In some embodiments, the computer-readable medium is a non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements any of the methods above or elsewhere herein.

[0038] In some embodiments, the system comprises one or more computer processors and computer memory coupled thereto. The computer memory comprises machine-executable code that, upon execution by the one or more computer processors, implements any of the methods described herein.

[0039] In another example, the present disclosure provides 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.

[0039a] Hence, in one aspect, there is provided a method for determining a methylation profile of a cell-free deoxyribonucleic acid (cfDNA) sample from a subject, 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 acid molecules; b) contacting the plurality of converted nucleic acid molecules 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 **Table 12** to enrich for sequences corresponding to the pre-identified methylation 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 subject.

[0040] In some embodiments, a nucleic acid sequencing library is prepared before the amplification. In some embodiments, the method further comprises amplifying the plurality of

converted nucleic acids. In some embodiments, the amplifying comprises polymerase chain reaction (PCR). In some embodiments, the method further comprises 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. In some embodiments, the reference nucleic acid sequence is at least a portion of a human reference genome. In some embodiments, the human reference genome is hg18.

[0041] In some embodiments, the methylation profile is associated with a colon cell proliferative disorder and provides classification of a subject as having a colon cell proliferative disorder.

[0042] In some embodiments, a nucleic acid adapter comprising a unique molecular identifier is ligated to unconverted nucleic acids in a cfDNA sample before a).

[0043] In some embodiments, the nucleic acid molecules are subjected to cytosine-to-uracil conversion conditions using chemical methods, enzymatic methods, or a combination thereof.

[0044] In some embodiments, the cfDNA in a biological sample is treated with a reagent selected from the group consisting of bisulfite, hydrogen sulfite, disulfite, and combinations thereof.

[0045] In some examples, 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.

[0046] In some examples, the method comprises applying the measured methylation signature panel from the subject against a database of measured methylation signature panels from normal subjects, wherein the database is stored on a computer system; determining that the subject has an increased risk of having a colon cell proliferative disorder by measuring a change of at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, or at least 20% in the methylation status of the methyl signature panel relative to methylation status from normal subjects.

[0047] In some examples, 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.** In some embodiments, 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**. In some embodiments, the methylation profile is indicative of a presence or an absence of a colon cell proliferative disorder in the individual.

[0048] In some examples, 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. In some embodiments, the colon cell proliferative disorder comprises the colorectal cancer.

[0049] In some examples, 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.

[0050] In another example, the present disclosure provides 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 plurality of 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 the methylation profile of the individual; and e) applying a trained machine learning model to the methylation profile, wherein the trained machine learning model 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.

[0050a] Thus, in one aspect, there is provided a method for detecting a presence or an absence of a colorectal cancer in a subject, comprising: a) providing conditions capable of converting unmethylated cytosines to uracils in nucleic acid molecules of a cell-free deoxyribonucleic acid (cfDNA) sample obtained or derived from the subject to produce a plurality of converted nucleic acid molecules; b) contacting the plurality of converted nucleic acid molecules with nucleic acid probes complementary to a pre-identified methylation signature panel of at least two genomic regions selected from the group consisting of **Table 12** to enrich for sequences corresponding to the pre-identified methylation 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 subject; 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 subjects without the colorectal cancer and subjects with the colorectal cancer to provide an output value associated with the presence of the colorectal cancer, thereby detecting the presence or the absence of the colorectal cancer in the subject.

[0051] In some embodiments, a nucleic acid sequencing library is prepared before the amplification. In some embodiments, the method further comprises amplifying the plurality of converted nucleic acids. In some embodiments, the amplifying comprises polymerase chain

reaction (PCR). In some embodiments, the method further comprises 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. In some embodiments, the reference nucleic acid sequence is at least a portion of a human reference genome. In some embodiments, the human reference genome is hg18.

[0052] In some examples, 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.

[0053] In some examples, the method comprises applying the measured methylation signature panel from the subject against a database of measured methylation signature panels from normal subjects, wherein the database is stored on a computer system; determining that the subject has an increased risk of having a colon cell proliferative disorder by measuring a change of at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, or at least 20% in the methylation status of the methyl signature panel relative to methylation status from normal subjects.

[0054] In some examples, 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**. In some embodiments, 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**. In some embodiments, the methylation profile is indicative of a presence or an absence of a colon cell proliferative disorder in the individual. In some embodiments, the method further comprises 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.

[0055] In some examples, 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. In some embodiments, the colon cell proliferative disorder comprises the colorectal cancer.

[0056] In some embodiments, 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.

[0057] In some examples, 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.

[0058] In another example, the present disclosure provides a method for monitoring minimal residual disease in a subject previously treated for 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.

[0059] In some examples, the minimal residual disease is selected from the group consisting of response to treatment, tumor load, residual tumor post-surgery, relapse, secondary screen, primary screen, and cancer progression.

[0060] In another example, a method is provided for determining response to treatment.

[0061] In another example, a method is provided for monitoring tumor load.

[0062] In another example, a method is provided for detecting residual tumor post-surgery.

[0063] In another example, a method is provided for detecting relapse.

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

[0065] In another example, a method is provided for use as a primary screen.

[0066] In another example, a method is provided for monitoring cancer progression.

[0067] In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a sensitivity of at least about 80%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a sensitivity of at least about 90%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a sensitivity of at least about 95%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a positive predictive value (PPV) of at least about 70%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a positive predictive value (PPV) of at least about 80%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a positive predictive value (PPV) of at least about 90%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a positive predictive value (PPV) of at least about 95%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a positive predictive value (PPV) of at least about 99%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a negative predictive value (NPV) of at least about 80%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a negative predictive value (NPV) of at least about 90%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a negative predictive value (NPV) of at least about 95%. In some examples, the dataset is indicative of the presence or susceptibility of the colorectal cancer at a negative predictive value (NPV) of at least about 99%. In some examples, the trained algorithm determines the presence or susceptibility of the colorectal cancer of the subject with an Area Under Curve (AUC) of at least about 0.90. In some examples, the trained algorithm determines the presence or susceptibility of the colorectal cancer of the subject with an Area Under Curve (AUC) of at least about 0.95. In some examples, the trained algorithm determines the presence or susceptibility of the colorectal cancer of the subject with an Area Under Curve (AUC) of at least about 0.99.

[0068] In some examples, the method further comprises presenting a report a graphical user interface of an electronic device of a user. In some embodiments, the user is the subject, individual or patient.

[0069] In some examples, the method further comprises determining a likelihood of the determination of a presence or susceptibility of colorectal cancer in the subject, individual, or patient. For example, the likelihood may be a probability value between 0% and 100%.

[0070] In some examples, the trained algorithm (e.g., machine learning model or classifier) comprises a supervised machine learning algorithm. In some embodiments, the supervised

machine learning algorithm comprises a deep learning algorithm, a support vector machine (SVM), a neural network, or a Random Forest.

[0071] In some examples, the method further comprises providing said subject with a therapeutic intervention based at least in part on the methylation profile or analysis, such as a therapeutic intervention to treat a patient with colorectal cancer (e.g., chemotherapy, radiotherapy, immunotherapy, or surgery).

[0072] In some examples, the method further comprises monitoring the presence or susceptibility of the colorectal cancer, wherein said monitoring comprises assessing the presence or susceptibility of the colorectal cancer of said subject at a plurality of time points, wherein the assessing is based at least on the presence or susceptibility of the colorectal cancer determined each of the plurality of time points.

[0073] In some examples, a difference in the assessment of the presence or susceptibility of the colorectal cancer of the subject among the plurality of time points is indicative of one or more clinical indications selected from the group consisting of: (i) a diagnosis of the presence or susceptibility of the colorectal cancer of the subject, (ii) a prognosis of the presence or susceptibility of the colorectal cancer of the subject, and (iii) an efficacy or non-efficacy of a course of treatment for treating the presence or susceptibility of the colorectal cancer of the subject.

[0074] In some examples, the method further comprises stratifying the colorectal cancer of the subject by using the trained algorithm to determine a sub-type of the colorectal cancer of the subject from among a plurality of distinct subtypes or stages of colorectal cancer.

[0075] Another example of the present disclosure provides a non-transitory computer readable medium comprising machine executable code that, upon execution by one or more computer processors, implements any of the methods above or elsewhere herein.

[0076] Another example of the present disclosure provides a system comprising one or more computer processors and computer memory coupled thereto. The computer memory comprises machine executable code that, upon execution by the one or more computer processors, implements any of the methods above or elsewhere herein.

[0077] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure.

CLAIMS

WHAT IS CLAIMED IS:

1. A methylation signature panel characteristic of a colorectal cancer comprising:

one or more genomic regions selected from the group consisting of **Table 12**, wherein the one or more genomic regions are more methylated in a cell-free deoxyribonucleic acid (cfDNA) sample from a subject having the colorectal cancer or colorectal cancer subtypes, and are less methylated in a cfDNA sample from a subject not having the colorectal cancer.

2. The methylation signature panel of claim 1, wherein the methylation signature panel comprises increased methylation in two or more genomic regions selected from the group consisting of **Table 12**.

3. The methylation signature panel of claim 1, wherein the colorectal cancer is selected from the group consisting of colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

4. The methylation signature panel of claim 1, wherein the colorectal cancer is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, and stage 4 colorectal cancer.

5. The methylation signature panel of claim 1, wherein the methylation signature panel comprises two or more genomic regions in **Table 12**.

6. The methylation signature panel of claim 1, wherein the methylation signature panel comprises genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, and FLI1.

7. The methylation signature panel of claim 1, wherein the methylation signature panel comprises genomic regions selected from the group consisting of IKZF1, KCNQ5, and ELMO1.

8. The methylation signature panel of claim 1, wherein the methylation signature panel comprises genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, FLI1, CLIP4, ELOVL5, FAM72B, and ST3GAL1.

9. A methylation signature panel characteristic of a colorectal cancer comprising:

two or more genomic regions selected from the group consisting of **Table 12**, wherein the two or more genomic regions are more methylated in a cell-free deoxyribonucleic acid

(cfDNA) sample from a subject having the colorectal cancer or colorectal cancer subtypes, and are less methylated in a cfDNA sample from a subject not having the colorectal cancer.

10. The methylation signature panel of claim 9, wherein the methylation signature panel comprises increased methylation in 6 or more genomic regions selected from the group consisting of genomic regions of **Table 12**.

11. The methylation signature panel of claim 9, wherein the colorectal cancer is selected from the group consisting of colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

12. The methylation signature panel of claim 9, wherein the colorectal cancer is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, and stage 4 colorectal cancer.

13. The methylation signature panel of claim 9, wherein the methylation signature panel comprises three or more genomic regions in **Table 12**.

14. The methylation signature panel of claim 9, wherein the methylation signature panel comprises genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, and FLI1.

15. The methylation signature panel of claim 9, wherein the methylation signature panel comprises genomic regions selected from the group consisting of IKZF1, KCNQ5, and ELMO1.

16. The methylation signature panel of claim 9, wherein the methylation signature panel comprises genomic regions selected from the group consisting of IKZF1, KCNQ5, ELMO1, CHST2, PRKCB, FLI1, CLIP4, ELOVL5, FAM72B, and ST3GAL1.

17. A machine learning classifier capable of distinguishing a population of subjects without a colorectal cancer from subjects with the colorectal cancer, 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 subjects without the colorectal cancer and subjects having the colorectal cancer;

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

c) wherein the machine learning or statistical model provides a feature vector useful as a classifier capable of distinguishing a population of subjects without the colorectal cancer from subjects having the colorectal cancer.

18. The machine learning classifier of claim 17, wherein the sets of measured values describe characteristics of the differentially-methylated genomic 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 lx and 10x mean genomic coverage (for S4 runs), mean CpG coverage (depth) globally and mean coverage at CpG islands, CGI shelves, and CGI shores.

19. A method for determining a methylation profile of a cell-free deoxyribonucleic acid (cfDNA) sample from a subject, 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 acid molecules;

b) contacting the plurality of converted nucleic acid molecules 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 **Table 12** to enrich for sequences corresponding to the pre-identified methylation 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 subject.

20. The method of claim 19, further comprising amplifying the plurality of converted nucleic acid molecules.

21. The method of claim 20, wherein the amplifying comprises polymerase chain reaction (PCR).

22. The method of claim 19, 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.

23. The method of claim 19, wherein the reference nucleic acid sequence is at least a portion of a human reference genome.

24. The method of claim 23, wherein the human reference genome is hg18.

25. The method of claim 19, wherein the pre-identified methylation signature panel comprises three or more methylated genomic regions in **Table 12**.

26. The method of claim 25, wherein the pre-identified methylation signature panel comprises one or more methylated genomic regions in **Table 12**, two or more methylated genomic regions in **Table 12**, or three methylated genomic regions in **Table 12**.

27. The method of claim 19, wherein the methylation profile is indicative of a presence or an absence of the colorectal cancer in the subject.

28. The method of claim 27, wherein the colorectal cancer is selected from the group consisting of colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

29. The method of claim 27, wherein the colorectal cancer is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, or stage 4 colorectal cancer.

30. A method for detecting a presence or an absence of a colorectal cancer in a subject, comprising:

a) providing conditions capable of converting unmethylated cytosines to uracils in nucleic acid molecules of a cell-free deoxyribonucleic acid (cfDNA) sample obtained or derived from the subject to produce a plurality of converted nucleic acid molecules;

b) contacting the plurality of converted nucleic acid molecules with nucleic acid probes complementary to a pre-identified methylation signature panel of at least two genomic regions selected from the group consisting of **Table 12** to enrich for sequences corresponding to the pre-identified methylation 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 subject; 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 subjects without the colorectal cancer and subjects with the colorectal cancer to provide an output value associated with the presence of the colorectal cancer, thereby detecting the presence or the absence of the colorectal cancer in the subject.

31. The method of claim 30, further comprising amplifying the plurality of converted nucleic acid molecules.

32. The method of claim 31, wherein the amplifying comprises polymerase chain reaction (PCR).

33. The method of claim 30, 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.

34. The method of claim 30, wherein the reference nucleic acid sequence is at least a portion of a human reference genome.

35. The method of claim 34, wherein the human reference genome is hg18.

36. The method of claim 30, wherein the pre-identified methylation signature panel comprises three or more methylated genomic regions in **Table 12**.

37. The method of claim 36, wherein the pre-identified methylation signature panel comprises one or more methylated genomic regions in **Table 12**, two or more methylated genomic regions in **Table 12**, or three methylated genomic regions in **Table 12**.

38. The method of claim 30, wherein the colorectal cancer is selected from the group consisting of colon cancer, rectal cancer, colorectal carcinoma, colorectal adenocarcinoma, carcinoid tumors, gastrointestinal carcinoid tumors, gastrointestinal stromal tumors (GISTs), lymphomas, and sarcomas.

39. The method of claim 30, wherein the colorectal cancer is selected from the group consisting of stage 1 colorectal cancer, stage 2 colorectal cancer, stage 3 colorectal cancer, and stage 4 colorectal cancer.

40. The method of claim 30, 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.