

method comprising: (a) providing a set of sequence reads from deoxyribonucleic acid (DNA) extracted from the subject; (b) using said set of sequence reads to generate a coverage pattern for a transcription factor; (c) processing the coverage pattern to provide a signal; and (d) processing the signal with a reference signal, wherein the signal and the reference signal have different frequencies, thereby determining a transcription factor binding profile in the sample.

[0012] In some examples, the DNA is cell-free DNA.

[0013] In some examples, (c) comprises using a low-pass filter. In some examples, (c) comprises using a Savitzky-Golay filter.

[0014] In some examples, the subject is a human.

[0015] In a first aspect, the present disclosure provides a computer-implemented method for detecting a presence or absence of a disease-cancer in a subject, the method comprising: (a) providing a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject; (b) using said set of sequence reads to generate a coverage pattern ~~for at binding sites of a~~ transcription factor; (c) processing the coverage pattern to ~~provide a signal~~ calculate an accessibility score for each of said binding sites of said transcription factor; and (d) ~~processing the signal with a reference signal, wherein the signal and the reference signal have different frequencies, thereby~~ detecting said presence or absence of said disease-cancer in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

[0016] In some examples, the DNA is cell-free DNA.

[0017] In some examples, the disease is cancer.

[0018] In some examples, (b) comprises aligning the set of sequence reads to a reference sequence to provide an aligned sequence pattern, selecting regions of the aligned sequence pattern that correspond to binding sites of the transcription factor, and normalizing the aligned sequence pattern in the regions. In some examples, (d) comprises calculating an accessibility score for each of the binding sites of the transcription factor.

[0019] In some examples, (c) comprises using a low-pass filter. In some examples, (c) comprises using a Savitzky-Golay filter.

[0020] In some examples, the subject is a human.

[0021] In some examples, the transcription factor is a cancer-specific transcription factor. In some examples, the transcription factor is selected from the group consisting of GRH-L2, ASH-2, HOX-B13, EVX2, PU.1, Lyl-1, Spi-B, FOXA1, HNF-1A, HNF-4A, HNF-4G, and DLX-2.

[0022] In some examples, the accessibility scores for at least 2, or at least 5, or at least 10, or at least 15, or at least 20, or at least 25 transcription factor binding sites are determined and inputted

into a machine learning model to train a classifier capable of distinguishing between healthy subjects and cancer patients, between disease progressors and non-progressors, between a

[0050] In some examples, an increased accessibility of transcription factors GLIS1, SOX2, and SOX11 are indicative of small-cell neuroendocrine prostate cancer.

[0051] In a second aspect, the present disclosure provides a system comprising a computing device comprising at least one computer processor, an operating system configured to perform executable instructions, a memory, and a computer program including instructions executable by the computing device to provide a computer application for detecting a presence or absence of a disease-cancer in a subject, the computer application comprising: a sequence module programmed to obtain a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject; a coverage module programmed to use the set of sequence reads to generate a coverage pattern for at binding sites of a transcription factor; a signal module programmed to process the coverage pattern to ~~provide a signal~~ calculate an accessibility score for each of said binding sites of said transcription factor; a detection module programmed to ~~process the signal with a reference signal, wherein the signal and the reference signal have different frequencies, thereby~~ detecting the presence or absence of the disease-cancer in the subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

[0052] In a third aspect, the present disclosure provides a non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for detecting a presence or absence of a disease-cancer in a subject, the method comprising: (a) providing a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject; (b) using said set of sequence reads to generate a coverage pattern for at binding sites of a transcription factor; (c) processing the coverage pattern to ~~provide a signal~~ calculate an accessibility score for each of said binding sites of said transcription factor; and (d) ~~processing the signal with a reference signal, wherein the signal and the reference signal have different frequencies, thereby~~ detecting said presence or absence of said disease-cancer in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

[0053] In a fourth aspect, the present disclosure provides a system for detecting a presence or absence of a disease-cancer in a subject, the system comprising: a database comprising a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject; and one or more computer processors operatively coupled to the database, wherein the one or more computer processors are individually or collectively programmed to: (a) use the set of sequence reads to generate a coverage pattern for at binding sites of a transcription factor; (b) process the coverage pattern to ~~provide a signal~~ calculate an accessibility score for each of said binding sites

of said transcription factor; (c) ~~process the signal with a reference signal, wherein the signal and the reference signal have different frequencies, thereby~~ detecting the presence or absence of the disease-cancer in the subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

[0054] In a fifth aspect, the present disclosure provides a computer-implemented method for monitoring a progression or regression of a disease-cancer in a subject, the method comprising:
(a) providing a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from cell-free DNA extracted from the subject at

a second time that is later than the first time; (b) using the first set of sequence reads to generate a first coverage pattern ~~for at binding sites of~~ a transcription factor and using the second set of sequence reads to generate a second coverage pattern ~~for at binding sites of~~ the transcription factor; (c) processing the first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor and processing the second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and (d) ~~processing the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies~~; (e) ~~processing the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies~~; and (f) ~~based on the processing of the first signal and the second signal with the reference signal~~, monitoring the progression or regression of the ~~disease cancer~~ in the subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0055] In some examples, the accessibility scores for at least 2, or at least 5, or at least 10, or at least 15, or at least 20, or at least 25 transcription factor binding sites are determined and inputted into a machine learning model to train a classifier capable of distinguishing between disease progressors and non-progressors, between a plurality of disease subtypes, between a plurality of disease stages, or any combination thereof.

[0056] In some examples, the accessibility scores for at least 2, or at least 5, or at least 10, or at least 15, or at least 20, or at least 25 transcription factor binding sites are determined and inputted into a machine learning model to train a classifier capable of distinguishing between disease treatment responders and non-responders.

[0057] In some examples, the second coverage pattern indicates phenotypic changes of a tumor during a course of the disease.

[0058] In some examples, the phenotypic change is a change from androgen-dependent to androgen-independent stage of cancer.

[0059] In some examples, the DNA is cell-free DNA.

[0060] In some examples, the disease is cancer.

[0061] In some examples, (b) comprises aligning the first set of sequence reads and the second set of sequence reads to a reference sequence to provide a first aligned sequence pattern and a second aligned sequence pattern, respectively, selecting regions of the first aligned sequence pattern and the second aligned sequence pattern that correspond to binding sites of the transcription factor, and normalizing the first aligned sequence pattern and second aligned sequence pattern in the regions.

[0062] In some examples, (c) comprises using a low-pass filter. In some examples, (c) comprises using a Savitzky-Golay filter.

[0063] In some examples, the subject is a human.

[0064] In some examples, the transcription factor is a cancer-specific transcription factor.

[0065] In some examples, the transcription factor is selected from the group consisting of GRH-L2, ASH-2, HOX-B13, EVX2, PU.1, Lyl-1, Spi-B, and FOXA1.

[0066] In some examples, the transcription factor is selected from the group consisting of HNF-1a, HNF-4a, HNF-4g, EVX-2 and DLX-2.

[0067] In some examples, the method further comprises, based on (f), adjusting a therapeutic regimen for the disease in the subject.

[0068] In a sixth aspect, the present disclosure provides a system comprising a computing device comprising at least one computer processor, an operating system configured to perform executable instructions, a memory, and a computer program including instructions executable by the computing device to provide a computer application for monitoring a progression or regression of a disease-cancer in a subject, the computer application comprising: a sequence module programmed to obtain a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from cell-free DNA extracted from the subject at a second time that is later than the first time; a coverage module programmed to use the first set of sequence reads to generate a first coverage pattern ~~for~~ at binding sites of a transcription factor and use the second set of sequence reads to generate a second coverage pattern ~~for-at binding sites of~~ the transcription factor; a signal module programmed to process the first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and process the second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; ~~and a first processing module programmed to process the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies; a second processing module programmed to process the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies; and a monitoring module programmed to, based on the processing of the first signal and the second signal with the reference signal,~~ monitor the progression or regression of the disease-cancer in the subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0069] In a seventh aspect, the present disclosure provides a non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for monitoring a progression or regression of a disease-cancer in a subject, the method comprising: (a) providing a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of

sequence reads from [cell-free](#) DNA extracted from the subject at a second time that is later than the first time; (b) using the first set of sequence reads to generate a first coverage pattern ~~for~~[at binding sites of](#) a transcription factor and using the second set of sequence reads to generate a second coverage pattern ~~for~~[at binding sites of](#) the transcription factor;

(c) processing the first coverage pattern to calculate a first accessibility score for each of said binding sites of said transcription factor ~~provide a first signal~~ and processing the second coverage pattern to calculate a second accessibility score for each of said binding sites of said transcription factor ~~provide a second signal~~; and (d) ~~processing the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies~~; (e) ~~processing the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies~~; and (f) ~~based on the processing of the first signal and the second signal with the reference signal~~, monitoring the progression or regression of the disease-cancer in the subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0070] In an eighth aspect, the present disclosure provides a system for monitoring a progression or regression of a disease-cancer in a subject, the system comprising: a database comprising a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from cell-free DNA extracted from the subject at a second time that is later than the first time; and one or more computer processors operatively coupled to the database, wherein the one or more computer processors are individually or collectively programmed to: (a) use the first set of sequence reads to generate a first coverage pattern for at binding sites of a transcription factor and use the second set of sequence reads to generate a second coverage pattern for at binding sites of the transcription factor; (b) process the first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and process the second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and (c) ~~process the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies~~; (d) ~~process the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies~~; and (e) ~~based on the processing of the first signal and the second signal with the reference signal~~, monitor the progression or regression of the disease-cancer in the subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0071] In another embodiment, the present disclosure provides a system to determine a transcription factor binding profile in a nucleic acid sample from a subject, the system comprising a processor configured to: (a) analyze a set of sequence reads from deoxyribonucleic acid (DNA) extracted from the subject; (b) using the set of sequence reads to generate a coverage pattern for a transcription factor; (c) processing the coverage pattern to provide a signal; and (d)

processing the signal with a reference signal, wherein the signal and the reference signal have different frequencies, thereby determining a transcription factor binding profile.

[0072] In some examples, the DNA is cell-free DNA.

[0073] In some examples, (c) comprises using a low-pass filter. In some examples, (c) comprises using a Savitzky-Golay filter.

[0074] In some examples, the subject is a human.

[0075] In another aspect, the present disclosure provides a system for detecting a presence or absence of a disease in a subject, comprising a processor configured to: (i) use a set of sequence

reads from deoxyribonucleic acid (DNA) extracted from the subject to generate a coverage pattern for a transcription factor; (ii) process the coverage pattern to provide a signal, wherein the signal has a different frequency than a reference signal; and (iii) processing the signal with the reference signal, thereby detecting the presence or absence of the disease in the subject.

[0076] In some examples, the present disclosure provides a system for classifying a tumor by tumor subtype or tumor stage, comprising a processor configured to: (i) use a first set of sequence reads from deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads extracted from DNA from the subject at a second time that is later than the first time to generate a first coverage pattern for a transcription factor and a second coverage pattern for the transcription factor; (ii) process the first coverage pattern to provide a first signal and process the second coverage pattern to provide a second signal, wherein the first signal and the second signal have different frequencies than a reference signal; and (iii) processing the first signal with the reference signal and processing the second signal with the reference signal, to monitor the progression or regression of the disease in the subject.

[0077] In another embodiment, the present disclosure provides a system for monitoring progression or regression of a disease in a subject, comprising a processor configured to: (i) use a first set of sequence reads from deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads extracted from DNA from the subject at a second time that is later than the first time to generate a first coverage pattern for a transcription factor and a second coverage pattern for the transcription factor; (ii) process the first coverage pattern to provide a first signal, and process the second coverage pattern to provide a second signal, wherein the first signal and the second signal have different frequencies than a reference signal; and (iii) processing the first signal with the reference signal and processing the second signal with the reference signal, to monitor the progression or regression of the disease in the subject.

[0078] Another embodiment 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.

[0079] Another embodiment 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.

[0080] In a ninth aspect, the present disclosure provides a computer-implemented method ~~for determining a tumor specific transcription factor binding site (TFBS) pattern, the method~~ comprising: (a) providing a first set of sequence reads from

cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from cell-free DNA extracted from the subject at a second time that is later than the first time; (b) using the first set of sequence reads to generate a first coverage pattern ~~for~~ at binding sites of a transcription factor and using the second set of sequence reads to generate a second coverage pattern ~~for~~ at binding sites of the transcription factor; (c) processing the first coverage pattern to calculate a first accessibility score for each of said binding sites of said transcription factor ~~provide a first signal~~ and processing the second coverage pattern to calculate a second accessibility score for each of said binding sites of said transcription factor ~~provide a second signal~~; and (d) ~~processing the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies;~~ (e) ~~processing the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies;~~ and (f) ~~based on the processing of the first signal and the second signal with the reference signal,~~ determining the tumor-specific TFBS pattern, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0081] In a tenth aspect, the present disclosure provides a system comprising a computing device comprising at least one computer processor, an operating system configured to perform executable instructions, a memory, and a computer program including instructions executable by the computing device to provide a computer application for determining a tumor-specific TFBS pattern, the computer application comprising: a sequence module programmed to obtain a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from cell-free DNA extracted from the subject at a second time that is later than the first time; a coverage module programmed to use the first set of sequence reads to generate a first coverage pattern ~~for~~ at binding sites of a transcription factor and use the second set of sequence reads to generate a second coverage pattern ~~for~~ at binding sites of the transcription factor; a signal module programmed to process the first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and process the second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and a first processing module programmed to process the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies; a second processing module programmed to process the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies; and a detection module programmed to, based on the processing of the first signal and the second signal with the reference signal,

determine the tumor-specific TFBS pattern, [based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score](#).

[0082] In an eleventh aspect, the present disclosure provides a non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for determining a tumor-specific TFBS pattern, the method comprising: (a) providing a first set of sequence reads from [cell-free](#) deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from [cell-free](#) DNA extracted

from the subject at a second time that is later than the first time; (b) using the first set of sequence reads to generate a first coverage pattern ~~for at binding sites of~~ a transcription factor and using the second set of sequence reads to generate a second coverage pattern ~~sequence reads for at binding sites of~~ the transcription factor; (c) processing the first coverage pattern to calculate a first accessibility score for each of said binding sites of said transcription factor ~~provide a first signal~~ and processing the second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and (d) ~~processing the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies~~; (e) ~~processing the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies~~; and (f) ~~based on the processing of the first signal and the second signal with the reference signal~~, determining the tumor-specific TFBS pattern, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0083] In a twelfth aspect, the present disclosure provides a system for determining a tumor-specific TFBS pattern, the system comprising: a database comprising a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from the subject at a first time and a second set of sequence reads from cell-free DNA extracted from the subject at a second time that is later than the first time; and one or more computer processors operatively coupled to the database, wherein the one or more computer processors are individually or collectively programmed to: (a) use the first set of sequence reads to generate a first coverage pattern ~~for at binding sites of~~ a transcription factor and use the second set of sequence reads to generate a second coverage pattern ~~for at binding sites of~~ the transcription factor; (b) process the first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and process the second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and (c) ~~process the first signal with a reference signal, wherein the first signal and the reference signal have different frequencies~~; (d) ~~process the second signal with the reference signal, wherein the second signal and the reference signal have different frequencies~~; and (e) ~~based on the processing of the first signal and the second signal with the reference signal~~, determine the tumor-specific TFBS pattern, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0083a] In a thirteenth aspect, the present disclosure provides a system for detecting a presence or absence of a disease-cancer in a subject, comprising a processor configured to: (i) use sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject to

generate a coverage pattern ~~for at binding sites of~~ a transcription factor; (ii) process said coverage pattern to ~~provide a signal~~ calculate an accessibility score for each of said binding sites of said transcription factor, wherein said signal has a different frequency than a reference signal; and (iii) ~~process said signal with said reference signal, to~~ provide a diagnosis of said ~~disease-cancer~~ in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

[0083b] In a fourteenth aspect, the present disclosure provides a system for monitoring a progression or regression of a ~~disease-cancer~~ in a subject, comprising a processor configured to: (i) use a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads extracted from cell-free DNA from said subject at a second time that is later than said first time to generate a first coverage pattern ~~for at binding sites of~~ a transcription factor and a second coverage pattern ~~for at binding sites of~~ said transcription factor; (ii) process said first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and process said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor, wherein said first signal and said second signal have different frequencies than a reference signal; and (ii) ~~process said first signal with said reference signal and process said second signal with said reference signal, to~~ monitor said progression or regression of said ~~disease-cancer~~ in said subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

[0084] Another aspect 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.

[0085] Another aspect 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.

[0086] 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 examples of the present disclosure are shown and described. As will be realized, the present

CLAIMS

WHAT IS CLAIMED IS:

1. A computer-implemented method for detecting a presence or absence of a disease-cancer in a subject, the method comprising:
 - (a) providing a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject;
 - (b) using said set of sequence reads to generate a coverage pattern for at binding sites of a transcription factor;
 - (c) processing said coverage pattern to calculate an accessibility score for each of said binding sites of said transcription factor~~provide a signal~~; and
 - (d) ~~processing said signal with a reference signal, wherein said signal and said reference signal have different frequencies, thereby~~ detecting said presence or absence of said disease-cancer in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.
- ~~2. The method of claim 1, wherein said DNA is cell-free DNA.~~
- ~~3. The method of claim 1 or 2, wherein said disease is cancer.~~
- ~~42.~~ The method of ~~any one of claims 1-31~~, wherein (b) comprises aligning said set of sequence reads to a reference sequence to provide an aligned sequence pattern, selecting regions of said aligned sequence pattern that correspond to binding sites of said transcription factor, and normalizing said aligned sequence pattern in said regions.
- ~~5. The method of claim 4, wherein (d) comprises calculating an accessibility score for each of said binding sites of said transcription factor.~~
- ~~6. The method of claim 5, wherein said accessibility scores for at least 2, or at least 5, or at least 10, or at least 15, or at least 20, or at least 25 transcription factor binding sites are determined and inputted into a machine learning model to train a classifier capable of distinguishing between healthy subjects and cancer patients, between disease progressors and non-progressors, between a plurality of disease subtypes, between a plurality of disease stages, between disease treatment responders and non-responders, or any combination thereof.~~
- ~~3.~~ The method of any one of claims 1-2, wherein (c) comprises splitting said coverage pattern into a low-frequency signal and a high-frequency signal, and calculating said accessibility score for each of said binding sites of said transcription factor using said high-frequency signal.
- ~~74.~~ The method of any one of claims 1-~~63~~, wherein (c) comprises using a low-pass filter.
- ~~85.~~ The method of claim ~~74~~, wherein (c) comprises using a Savitzky-Golay filter.
- ~~6.~~ The method of claim 5, wherein said Savitzky-Golay filter comprises a third-order Savitzky-Golay filter.

7. The method of claim 5, further comprising obtaining said low-frequency signal using a first Savitzky-Golay filter with a first window size, and obtaining said high-frequency signal using a second Savitzky-Golay filter with a second window size, wherein said first window size is larger than said second window size.

8. The method of claim 3, further comprising performing a locally weighted smoothing on said high-frequency signal.

9. The method of claim 8, wherein said locally weighted smoothing comprises a locally estimated scatterplot smoothing (LOESS).

10. The method of any one of claims 1-9, wherein said accessibility score that deviates from said reference accessibility score by greater than ± 3 mean of the standard deviation, is indicative of said presence of said cancer in said subject.

11. The method of any one of claims 1-10, wherein said subject is a human.

12. The method of any one of claims 1-11, wherein said transcription factor is an epithelial transcription factor.

13. The method of any one of claims 1-12, wherein said transcription factor is a cancer-specific transcription factor.

14. The method of claim 13, wherein said transcription factor is selected from the group consisting of FOXA1, GRHL1, GRHL2, and GRHL3.

15. The method of claim 13, wherein said transcription factor is selected from the group consisting of GRHL2, PU.1, LYL1, SPIB, AR, HOXB13, and NKX3-1.

16. The method of claim 13, wherein said transcription factor is selected from the group consisting of GRHL2, ASH2, HOXB13, EVX2, PU.1, Lyl-1, Spi-B, FOXA1, HNF-1A, HNF-4A, HNF-4G, and DLX-2.

17. The method of claim 13, wherein said transcription factor is selected from the group consisting of GRHL2, FOXA1, and ZNF121; and wherein said cancer comprises breast cancer.

18. The method of claim 13, wherein said transcription factor is selected from the group consisting of EVX2, DLX2, HNF1A, HNF4A, GRHL2, and HNF4G; and wherein said cancer comprises colon cancer.

19. The method of claim 13, wherein said transcription factor is selected from the group consisting of LYL1 and PU.1, and wherein said cancer comprises colon cancer.

20. The method of claim 13, wherein said transcription factor is selected from the group consisting of GRHL2, FOXA1, HOXB13, AR, and NKX3-1; and wherein said cancer comprises prostate adenocarcinoma.

21. The method of claim 13, wherein said transcription factor is selected from the group consisting of REST, GRHL2, FOXA1, HOXB13, AR, and NKX3-1; and wherein said cancer comprises prostate small-cell neuroendocrine prostate cancer.

22. The method of claim 13, wherein said transcription factor is selected from the group consisting of GLIS1, SOX2, and SOX11; and wherein said cancer comprises prostate cancer.

23. The method of any one of claims 1-~~22~~, wherein (d) comprises identifying ~~a sign of higher-an increased~~ accessibility score of a binding site of said transcription factor as compared to said reference accessibility score.

24. The method of claim 1, further comprising detecting said presence or absence of said ~~disease-cancer~~ in said subject with an accuracy of at least about 70%.

25. The method of claim 1, further comprising detecting said presence or absence of said ~~disease-cancer~~ in said subject with an accuracy of at least about 80%.

26. The method of claim 1, further comprising detecting said presence or absence of said ~~disease-cancer~~ in said subject with an accuracy of at least about 90%.

27. The method of claim 1, further comprising detecting said presence of said ~~disease-cancer~~ in said subject with a sensitivity of at least about 70%.

28. The method of claim 1, further comprising detecting said presence of said ~~disease-cancer~~ in said subject with a sensitivity of at least about 80%.

29. The method of claim 1, further comprising detecting said presence of said ~~disease-cancer~~ in said subject with a sensitivity of at least about 90%.

30. The method of claim 1, further comprising detecting said absence of said ~~disease-cancer~~ in said subject with a specificity of at least about 70%.

31. The method of claim 1, further comprising detecting said absence of said ~~disease-cancer~~ in said subject with a specificity of at least about 80%.

32. The method of claim 1, further comprising detecting said absence of said ~~disease-cancer~~ in said subject with a specificity of at least about 90%.

33. The method of claim 1, further comprising detecting said presence of said ~~disease-cancer~~ in said subject with a positive predictive value (PPV) of at least about 70%.

34. The method of claim 1, further comprising detecting said presence of said ~~disease-cancer~~ in said subject with a positive predictive value (PPV) of at least about 80%.

35. The method of claim 1, further comprising detecting said presence of said ~~disease-cancer~~ in said subject with a positive predictive value (PPV) of at least about 90%.

36. The method of claim 1, further comprising detecting said absence of said ~~disease-cancer~~ in said subject with a negative predictive value (NPV) of at least about 70%.

~~2737.~~ The method of claim 1, further comprising detecting said absence of said [disease-cancer](#) in said subject with a negative predictive value (NPV) of at least about 80%.

~~2838.~~ The method of claim 1, further comprising detecting said absence of said [disease-cancer](#) in said subject with a negative predictive value (NPV) of at least about 90%.

~~2939.~~ The method of claim 1, further comprising detecting said presence or absence of said [disease-cancer](#) in said subject with an Area Under the Receiver Operator Characteristic (AUROC) of at least about 0.70.

~~3040.~~ The method of claim 1, further comprising detecting said presence or absence of said [disease-cancer](#) in said subject with an Area Under the Receiver Operator Characteristic (AUROC) of at least about 0.80.

~~3141.~~ The method of claim 1, further comprising detecting said presence or absence of said [disease-cancer](#) in said subject with an Area Under the Receiver Operator Characteristic (AUROC) of at least about 0.90.

~~32.—The method of claim 1, further comprising applying a trained classifier to said signal to detect said presence or absence of said disease in said subject.~~

~~33.—The method of claim 5, further comprising applying a trained classifier to said accessibility scores of said binding sites of said transcription factor to detect said presence or absence of said disease in said subject.~~

~~34.—The method of claim 32 or 33, wherein said trained classifier comprises a trained machine learning classifier.~~

~~35.—The method of claim 34, wherein said trained machine learning classifier comprises a supervised machine learning algorithm.~~

~~36.—The method of claim 35, wherein said supervised machine learning algorithm comprises one or more of: a regression, a support vector machine, a tree-based method, a neural network, and a random forest.~~

~~3742.~~ A system comprising a computing device comprising at least one computer processor, an operating system configured to perform executable instructions, a memory, and a computer program including instructions executable by the computing device to provide a computer application for detecting a presence or absence of a [disease-cancer](#) in a subject, the computer application comprising:

a sequence module programmed to obtain a set of sequence reads from [cell-free](#) deoxyribonucleic acid (DNA) extracted from said subject;

a coverage module programmed to use said set of sequence reads to generate a coverage pattern ~~for~~ [at binding sites of](#) a transcription factor;

a signal module programmed to process said coverage pattern to calculate an accessibility score for each of said binding sites of said transcription factor~~provide a signal~~; and

a detection module programmed to ~~process said signal with a reference signal, wherein said signal and said reference signal have different frequencies, thereby~~ detecting said presence or absence of said disease-cancer in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

3843. A non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for detecting a presence or absence of a disease-cancer in a subject, the method comprising:

- (a) providing a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject;
- (b) using said set of sequence reads to generate a coverage pattern at binding sites of ~~for~~ a transcription factor;
- (c) processing said coverage pattern to calculate an accessibility score for each of said binding sites of said transcription factor~~provide a signal~~; and
- (d) ~~processing said signal with a reference signal, wherein said signal and said reference signal have different frequencies, thereby~~ detecting said presence or absence of said disease-cancer in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

3944. A system for detecting a presence or absence of a disease-cancer in a subject, the system comprising:

a database comprising a set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject; and

one or more computer processors operatively coupled to the database, wherein the one or more computer processors are individually or collectively programmed to:

- (a) use said set of sequence reads to generate a coverage pattern ~~for~~ at binding sites of a transcription factor;
- (b) process said coverage pattern to calculate an accessibility score for each of said binding sites of said transcription factor~~provide a signal~~; and
- (c) ~~process said signal with a reference signal, wherein said signal and said reference signal have different frequencies, thereby~~ detecting said presence or absence of said disease-cancer in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

4045. A computer-implemented method for monitoring a progression or regression of a disease cancer in a subject, the method comprising:

- (a) providing a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads from cell-free DNA extracted from said subject at a second time that is later than said first time;
- (b) using said first set of sequence reads to generate a first coverage pattern ~~for at~~ binding sites of a transcription factor and using said second set of sequence reads to generate a second coverage pattern ~~for at binding sites of~~ said transcription factor;
- (c) processing said first coverage pattern to calculate a first accessibility score for each of said binding sites of said transcription factor ~~provide a first signal~~, and processing said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and
- ~~(d) processing said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~
- ~~(e) processing said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and~~
- ~~(f)(d)~~ based on said processing of said first signal and said second signal with said reference signal, monitoring said progression or regression of said disease cancer in said subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

~~41. The method of claim 40, wherein said DNA is cell-free DNA.~~

~~42. The method of claim 40 or 41, wherein said disease is cancer.~~

4346. The method of ~~any one of claims 40-42~~45, wherein (b) comprises aligning said first set of sequence reads and said second set of sequence reads to a reference sequence to provide a first aligned sequence pattern and a second aligned sequence pattern, respectively, selecting regions of said first aligned sequence pattern and said second aligned sequence pattern that correspond to binding sites of said transcription factor, and normalizing said first aligned sequence pattern and second aligned sequence pattern in said regions.

47. The method of any one of claims 45-46, wherein (c) comprises splitting said coverage pattern into a low-frequency signal and a high-frequency signal, and calculating said accessibility score for each of said binding sites of said transcription factor using said high-frequency signal.

4448. The method of any one of claims 4045-4347, wherein (c) comprises using a low-pass filter.

4549. The method of claim 4148, wherein (c) comprises using a Savitzky-Golay filter.

50. The method of claim 49, wherein said Savitzky-Golay filter comprises a third-order Savitzky-Golay filter.

51. The method of claim 49, further comprising obtaining said low-frequency signal using a first Savitzky-Golay filter with a first window size, and obtaining said high-frequency signal using a second Savitzky-Golay filter with a second window size, wherein said first window size is larger than said second window size.

52. The method of claim 47, further comprising performing a locally weighted smoothing on said high-frequency signal.

53. The method of claim 52, wherein said locally weighted smoothing comprises a locally estimated scatterplot smoothing (LOESS).

54. The method of any one of claims 45-53, wherein said accessibility score that deviates from said reference accessibility score by greater than ± 3 mean of the standard deviation, is indicative of said presence of said cancer in said subject.

4655. The method of any one of claims 4045-4554, wherein said subject is a human.

4756. The method of any one of claims 4045-4655, wherein said transcription factor is a cancer-specific transcription factor.

57. The method of claim 56, wherein said transcription factor is selected from the group consisting of FOXA1, GRHL1, GRHL2, and GRHL3.

58. The method of claim 56, wherein said transcription factor is selected from the group consisting of GRHL2, PU.1, LYL1, SPIB, AR, HOXB13, and NKX3-1.

4859. The method of claim 4756, wherein said transcription factor is selected from the group consisting of GRH-L2, ASH-2, HOX-B13, EVX2, PU.1, Lyl-1, Spi-B, FOXA1, HNF-1a, HNF-4a, HNF-4g, and DLX-2.

60. The method of claim 56, wherein said transcription factor is selected from the group consisting of GRHL2, FOXA1, and ZNF121; and wherein said cancer comprises breast cancer.

61. The method of claim 56, wherein said transcription factor is selected from the group consisting of EVX2, DLX2, HNF1A, HNF4A, GRHL2, and HNF4G; and wherein said cancer comprises colon cancer.

62. The method of claim 56, wherein said transcription factor is selected from the group consisting of LYL1 and PU.1, and wherein said cancer comprises colon cancer.

63. The method of claim 56, wherein said transcription factor is selected from the group consisting of GRHL2, FOXA1, HOXB13, AR, and NKX3-1; and wherein said cancer comprises prostate adenocarcinoma.

64. The method of claim 56, wherein said transcription factor is selected from the group consisting of REST, GRHL2, FOXA1, HOXB13, AR, and NKX3-1; and wherein said cancer comprises prostate small-cell neuroendocrine prostate cancer.

65. The method of claim 56, wherein said transcription factor is selected from the group consisting of GLIS1, SOX2, and SOX11; and wherein said cancer comprises prostate cancer.

~~49~~66. The method of any one of claims ~~40~~45-48~~65~~, further comprising, based on (~~fd~~), adjusting a therapeutic regimen for said ~~disease-cancer~~ in said subject.

~~50~~67. A system comprising a computing device comprising at least one computer processor, an operating system configured to perform executable instructions, a memory, and a computer program including instructions executable by the computing device to provide a computer application for monitoring a progression or regression of a ~~disease-cancer~~ in a subject, the computer application comprising:

a sequence module programmed to obtain a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads from cell-free DNA extracted from said subject at a second time that is later than said first time;

a coverage module programmed to use said first set of sequence reads to generate a first coverage pattern ~~for at binding sites of~~ a transcription factor and use said second set of sequence reads to generate a second coverage pattern ~~for at binding sites of~~ said transcription factor;

a signal module programmed to process said first coverage pattern to ~~provide a first signal~~calculate a first accessibility score for each of said binding sites of said transcription factor, and process said second coverage pattern to ~~provide a second signal~~calculate a second accessibility score for each of said binding sites of said transcription factor; and

~~— a first processing module programmed to process said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~

~~— a second processing module programmed to process said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and~~

a monitoring module programmed to, ~~based on said processing of said first signal and said second signal with said reference signal,~~ monitor said progression or regression of said ~~disease-cancer~~ in said subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

~~51~~68. A non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for monitoring a progression or regression of a ~~disease-cancer~~ in a subject, the method comprising:

- (a) providing a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads from cell-free DNA extracted from said subject at a second time that is later than said first time;
- (b) using said first set of sequence reads to generate a first coverage pattern ~~for-at~~ binding sites of a transcription factor and using said second set of sequence reads to generate a second coverage pattern ~~for-at~~ binding sites of said transcription factor;
- (c) processing said first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and processing said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and
- ~~(d)~~ processing said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;
- ~~(e)~~ processing said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and
- ~~(f)~~(d) based on said processing of said first signal and said second signal with said reference signal, monitoring said progression or regression of said ~~disease-cancer~~ in said subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

~~52~~69. A system for monitoring a progression or regression of a ~~disease-cancer~~ in a subject, the system comprising:

a database comprising a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads from cell-free DNA extracted from said subject at a second time that is later than said first time; and

one or more computer processors operatively coupled to the database, wherein the one or more computer processors are individually or collectively programmed to:

- (a) use said first set of sequence reads to generate a first coverage pattern ~~for-at~~ binding sites of a transcription factor and use said second set of sequence reads to generate a second coverage pattern ~~for-at~~ binding sites of said transcription factor;
- (b) process said first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and

processing said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and

~~(e) processing said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~

~~(d) processing said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and~~

~~(e) based on said processing of said first signal and said second signal with said reference signal, (c) monitor said progression or regression of said disease-cancer in said subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.~~

5370. A computer-implemented method ~~for determining a tumor-specific transcription factor binding site (TFBS) pattern, the method~~ comprising:

(a) providing a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads extracted from cell-free DNA from said subject at a second time that is later than said first time;

(b) using said first set of sequence reads to generate a first coverage pattern ~~for at~~ binding sites of a transcription factor and using said second set of sequence reads to generate a second coverage pattern ~~for at binding sites of~~ said transcription factor;

(c) processing said first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and processing said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and

~~(d) — processing said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~

~~(e) — processing said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and~~

~~(f)(d) — based on said processing of said first signal and said second signal with said reference signal, determining said tumor-specific TFBS pattern, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.~~

5471. A system comprising a computing device comprising at least one computer processor, an operating system configured to perform executable instructions, a memory, and a computer

program including instructions executable by the computing device to provide a computer application for determining a tumor-specific transcription factor binding site (TFBS) pattern, the computer application comprising:

a sequence module programmed to obtain a first set of sequence reads from [cell-free](#) deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads extracted from [cell-free](#) DNA from said subject at a second time that is later than said first time;

a coverage module programmed to use said first set of sequence reads to generate a first coverage pattern ~~for~~ [at binding sites of](#) a transcription factor and use said second set of sequence reads to generate a second coverage pattern ~~for~~ [at binding sites of](#) said transcription factor;

a signal module programmed to process said first coverage pattern to ~~provide a first signal~~ [calculate a first accessibility score for each of said binding sites of said transcription factor](#), and processing said second coverage pattern to ~~provide a second signal~~ [calculate a second accessibility score for each of said binding sites of said transcription factor](#); and

~~— a first processing module programmed to process said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~

~~— a second processing module programmed to process said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies;~~
and

a detection module programmed to, ~~based on said processing of said first signal and said second signal with said reference signal~~, determine said tumor-specific TFBS pattern, [based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score](#).

5572. A non-transitory computer-readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for determining a tumor-specific transcription factor binding site (TFBS) pattern, the method comprising:

- (a) providing a first set of sequence reads from [cell-free](#) deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads extracted from [cell-free](#) DNA from said subject at a second time that is later than said first time;
- (b) using said first set of sequence reads to generate a first coverage pattern ~~for~~ [binding sites of](#) a transcription factor and using said second set of sequence reads to generate a second coverage pattern ~~for~~ [at binding sites of](#) said transcription factor;

- (c) processing said first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and processing said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor; and
- ~~(d) — processing said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~
- ~~(e) — processing said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and~~
- ~~(f)(d)~~ based on said processing of said first signal and said second signal with said reference signal, determining said tumor-specific TFBS pattern, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

5673. A system for determining a tumor-specific transcription factor binding site (TFBS) pattern, the system comprising:

a database comprising a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads extracted from cell-free DNA from said subject at a second time that is later than said first time; and

one or more computer processors operatively coupled to the database, wherein the one or more computer processors are individually or collectively programmed to:

- (a) use said first set of sequence reads to generate a first coverage pattern ~~for at~~ binding sites of a transcription factor and use said second set of sequence reads to generate a second coverage pattern ~~for at~~ binding sites of said transcription factor;
- (b) process said first coverage pattern to ~~provide a first signal~~ calculate a first accessibility score for each of said binding sites of said transcription factor, and process said second coverage pattern to calculate a second accessibility score for each of said binding sites of said transcription factor ~~provide a second signal~~; and
- ~~(c) process said first signal with a reference signal, wherein said first signal and said reference signal have different frequencies;~~
- ~~(d) process said second signal with said reference signal, wherein said second signal and said reference signal have different frequencies; and~~
- ~~(e)~~ based on said processing of said first signal and said second signal with said reference signal, determine said tumor-specific TFBS pattern, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

5774. A system for detecting a presence or absence of a ~~disease-cancer~~ in a subject, comprising a processor configured to:

- (i) use sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject to generate a coverage pattern ~~for at binding sites of~~ a transcription factor;
- (ii) process said coverage pattern to ~~provide a signal~~ calculate an accessibility score for each of said binding sites of said transcription factor, wherein said signal has a different frequency than a reference signal; and
- (iii) ~~process said signal with said reference signal, to~~ provide a diagnosis of said ~~disease-cancer~~ in said subject, based at least in part on a deviation of said accessibility score from a reference accessibility score.

5875. The system of claim 5774, wherein (ii) comprises identifying a signal of higher an increased accessibility score of a binding site of said transcription factor as compared to said reference accessibility score.

5976. A system for monitoring a progression or regression of a ~~disease-cancer~~ in a subject, comprising a processor configured to:

- (i) use a first set of sequence reads from cell-free deoxyribonucleic acid (DNA) extracted from said subject at a first time and a second set of sequence reads extracted from cell-free DNA from said subject at a second time that is later than said first time to generate a first coverage pattern ~~for at binding sites of~~ a transcription factor and a second coverage pattern ~~for at binding sites of~~ said transcription factor;
- (ii) process said first coverage pattern to calculate a first accessibility score for each of said binding sites of said transcription factor ~~provide a first signal~~, and process said second coverage pattern to ~~provide a second signal~~ calculate a second accessibility score for each of said binding sites of said transcription factor, wherein said first signal and said second signal have different frequencies than a reference signal; and
- (iii) ~~process said first signal with said reference signal and process said second signal with said reference signal, to~~ monitor said progression or regression of said ~~disease-cancer~~ in said subject, based at least in part on a deviation of said first and second accessibility scores from a reference accessibility score.

6077. The system of claim 5976, wherein said transcription factor is selected from the group consisting of GRH-L2, ASH-2, HOX-B13, EVX2, PU.1, Lyl-1, Spi-B, FOXA1, HNF-1a, HNF-4a, HNF-4g, and DLX-2.