

CLAIMS

1. A method for preparing a cell-free nucleic acid sample to enable identification of at least one histone modification in a nucleosome contained therein using a DNA sequencing-based technique, the method comprising:

(a) providing a cell-free nucleic acid sample containing a plurality of nucleosomes each comprising a cfDNA molecule wound around a histone core;

(b) ligating adapters comprising terminal hybridizing regions to the ends of the cfDNA molecules, thereby providing a modified cell-free nucleic acid sample comprising adapter-ligated cfDNA molecules each wound around a histone core;

(c) providing a proximity probe comprising, at a first terminus, a histone modification binding domain that specifically binds to a histone modification of interest; at a second terminus, a nucleic acid binding domain complementary to a terminal hybridizing region; and a non-hybridizing region therebetween comprising a nucleic acid sequence that corresponds to the histone modification of interest and thereby serves as a histone modification barcode, wherein the proximity probe is dimensioned to allow for simultaneous binding of the histone modification binding domain to the histone modification of interest and hybridization of the complementary nucleic acid binding domain with the hybridizing nucleic acid region;

(d) incubating the modified cell-free nucleic acid sample with the proximity probe under conditions effective to facilitate (i) binding of the histone modification binding domain to the histone modification and (ii) hybridization of the complementary nucleic acid binding domain with the hybridizing nucleic acid region to form a dsDNA segment with a 5' terminus originating with the cell-free DNA and a 3' terminus originating with the proximity probe and comprising the histone modification barcode; and

(e) extending the 5' terminus of the dsDNA segment along the non-hybridizing region of the proximity probe and the histone modification barcode by adding a polymerase and a mixture of dNTPs, thereby providing a histone modification-barcoded dsDNA template molecule for amplification and sequencing.

2. The method of claim 1, wherein step (c) comprises providing a plurality of proximity probes each targeting a different histone modification.

3. The method of claim 2, further comprising amplifying the histone modification-barcoded dsDNA template molecules.

4. The method of claim 3, further comprising sequencing the amplified, histone modification-barcoded dsDNA template molecules and determining information about the type and location of histone modifications from the histone modification barcodes observed in the sequence reads generated.

5. A combined workflow process for extracting multiple types of data from a single, cell-free nucleic acid sample, wherein the process comprises:

(a) ligating DNA adapters comprising a first adapter sequence that includes at least one UFI sequence onto the ends of end-blunted cfDNA in the sample to provide adapter-ligated DNA, wherein the at least one UFI sequence includes a source identifier barcode;

(b) synthesizing cDNA from RNA in the sample and covalently attaching a cDNA adapter comprising the source identifier barcode and an RNA indicator barcode to at least one terminus of the cDNA, thereby providing adapter-bound cDNA in a cell-free admixture that also comprises the adapter-ligated DNA;

(c) functionalizing 5-hydroxymethylcytosine (5hmC) residues in the sample with an affinity tag that allows selective removal of 5hmC-containing DNA from the cell-free sample;

(d) removing the 5hmC-containing DNA from the cell-free sample, with untagged DNA and adapter-bound cDNA remaining in the sample;

(e) appending a 5hmC process barcode to the 5hmC-containing DNA; and

(f) amplifying and sequencing the 5hmC-containing DNA, the untagged DNA, and the adapter-bound cDNA.

6. The method of claim 5, wherein the at least one UFI sequence additionally comprises a molecular barcode, a strand identifier barcode, a histone modification barcode, or a combination thereof.

7. The process of claim 5, wherein the 5hmC-containing DNA provided in step (e) is pooled with the untagged DNA and adapter-bound cDNA prior to amplification.

8. A combined workflow process for extracting multiple types of data from a single, cell-free nucleic acid sample, wherein the process comprises:

(a) ligating DNA adapters comprising a first adapter sequence that includes at least one UFI sequence comprising a source identifier barcode onto the ends of end-blunted cfDNA in the sample to provide adapter-ligated DNA;

(b) synthesizing cDNA from RNA in the sample and covalently attaching a cDNA adapter comprising the source identifier barcode and an RNA indicator barcode to at least one terminus of the cDNA, thereby providing adapter-bound cDNA;

(c) functionalizing 5hmC residues in the sample with a first affinity tag that allows selective removal of 5hmC-containing DNA from the cell-free sample;

(d) removing the tagged 5hmC-containing DNA, with untagged 5-methylcytosine (5mC) containing DNA, untagged unmodified DNA, and adapter-bound cDNA remaining in the sample;

(e) appending a 5hmC process barcode to the tagged 5hmC-containing DNA; and

(f) converting methylcytosine residues in the remaining sample to oxidized methylcytosine residues;

(g) functionalizing the oxidized methylcytosine residues with a second affinity tag that allows selective removal of the functionalized species from the sample;

(h) removing the tagged 5mC-containing DNA, with untagged DNA and adapter-bound cDNA remaining in the sample;

(i) appending a 5mC process barcode to the tagged 5mC-containing DNA; and

(j) amplifying and sequencing the tagged 5hmC-containing DNA, the tagged 5mC-containing DNA, the untagged DNA, and the adapter-bound cDNA.

9. The process of claim 8, wherein the at least one molecular barcode further comprises a fragment identifier sequence, a strand identifier sequence, or both a fragment identifier sequence and a strand identifier sequence.

10. The process of claim 8, wherein the tagged 5hmC-containing DNA and the tagged 5mC-containing DNA are pooled with the untagged DNA and the adapter-bound cDNA prior to sequencing.

11. A combined workflow process for extracting at least two types of data from a single, cell-free nucleic acid sample, wherein the process comprises:

(a) ligating DNA adapters comprising a first adapter sequence that includes at least one molecular barcode comprising a source identifier barcode onto the ends of end-blunted DNA in the sample to provide adapter-ligated DNA;

(b) synthesizing cDNA from RNA in the sample and covalently attaching, to at least one terminus of the cDNA, a cDNA adapter comprising a 5hmC residue, the source identifier barcode, and an RNA indicator barcode, thereby providing barcoded, adapter-bound cDNA;

(c) functionalizing 5hmC residues in the sample with an affinity tag that allows selective removal of 5hmC-containing species from the cell-free sample;

(d) removing the 5hmC-containing DNA and the barcoded, adapter-bound cDNA from the cell-free sample; and

(e) amplifying and sequencing a pooled admixture of the 5hmC-containing DNA and the barcoded-adapter-bound cDNA to provide data on DNA hydroxymethylation and on cfRNA in the same sample.

12. A combined workflow process for extracting multiple types of data from a single, cell-free nucleic acid sample, comprising:

(a) ligating an adapter comprising a hybridizing nucleic acid region to each terminus of nucleosome-associated DNA, thereby providing a modified cell-free nucleic acid sample comprising nucleosomes associated with adapter-ligated DNA;

(b) providing a proximity probe comprising a histone modification binding domain at a first terminus, a nucleic acid binding domain complementary to the hybridizing nucleic acid region at an opposing second terminus, and a non-hybridizing region therebetween comprising a nucleic acid sequence selected to correspond to a specific histone modification and thereby serve as a histone modification barcode, wherein the proximity probe is dimensioned to allow for simultaneous binding of the histone modification binding domain to the histone modification and the hybridization of the complementary nucleic acid binding domain with the hybridizing nucleic acid region;

(c) incubating the modified cell-free nucleic acid sample with the proximity probe under conditions effective to facilitate (i) binding of the histone modification binding domain to the

histone modification and (ii) hybridization of the complementary nucleic acid binding domain with the hybridizing nucleic acid region to form a dsDNA segment with a 5' terminus originating with the cell-free DNA and a 3' terminus originating with the proximity probe; and

(d) extending the 5' terminus of the segment along the non-hybridizing region of the proximity probe and the histone modification barcode by adding a polymerase and a mixture of dNTPs, thereby providing a histone modification-barcoded dsDNA template molecule for further processing and sequencing;

(e) purifying nucleic acids in the sample to provide a composition comprising histone modification-barcoded dsDNA and DNA;

(f) synthesizing a first strand of cDNA from RNA in the sample;

(g) synthesizing a second strand of cDNA complementary to the first strand to provide a cDNA duplex; and

(h) covalently attaching to at least one terminus of the cDNA duplex, in the absence of a ligase, a cDNA adapter comprising a sequence that includes a source identifier barcode and an RNA indicator barcode, thereby providing a nucleic acid composition comprising adapter-bound cDNA and the histone modification-barcoded dsDNA template molecule.

13. The process of claim 12, further comprising:

(i) amplifying and sequencing the histone modification-barcoded dsDNA template molecule and the adapter-bound cDNA.

14. The process of claim 13, wherein following step (d), adapters comprising an additional barcode are ligated to the histone modification-barcoded dsDNA.

15. The process of claim 14, wherein the additional barcode comprises a source identifier sequence.

16. The process of claim 15, wherein the additional barcode further comprises a fragment identifier sequence, a strand identifier sequence, or both a fragment identifier sequence and a strand identifier sequence.

17. The process of claim 12, further comprising:

(i) functionalizing 5hmC residues in the nucleic acid composition with a first affinity tag that allows selective removal of 5hmC-containing species;

(j) removing the tagged 5hmC-containing DNA from the composition, with untagged DNA and adapter-bound cDNA remaining; and

(k) appending a 5hmC process barcode to the tagged 5hmC-containing DNA.

18. The process of claim 17, further including:

(l) amplifying and sequencing the 5hmC-containing DNA, the untagged DNA, and the adapter-bound cDNA.

19. The process of claim 18, wherein the barcoded 5hmC-containing DNA provided in step (k) is pooled with the untagged DNA and adapter-bound cDNA prior to sequencing.

20. The process of claim 17, further comprising:

(l) converting methylcytosine residues in the remaining sample to oxidized methylcytosine residues;

(m) functionalizing the oxidized methylcytosine residues with a second affinity tag that allows selective removal of the functionalized species from the sample;

(n) removing the tagged 5mC-containing DNA, with untagged DNA and adapter-bound cDNA remaining; and

(o) appending a 5mC process barcode to the tagged 5mC-containing DNA.

21. The process of claim 20, further comprising:

(p) amplifying and sequencing the tagged 5hmC-containing DNA, the tagged 5mC-containing DNA, the untagged DNA, and the adapter-bound cDNA.

22. The process of claim 21, wherein the tagged 5hmC-containing DNA and the tagged 5mC-containing DNA are pooled with the untagged DNA and the adapter-bound cDNA prior to sequencing.

23. The process of claim 12, wherein the cell-free nucleic acid sample is obtained from a blood sample.

24. The process of claim 23, further comprising carrying out a plasma protein analysis on a fraction of the blood sample.

25. The process of claim 24, wherein plasma proteins are analyzed by generating a dsDNA template molecule comprising a protein identifier barcode corresponding to a protein analyte.

26. The process of claim 25, further comprising amplifying and sequencing the histone modification-barcoded dsDNA template molecule, the adapter-bound cDNA, and the protein-barcoded dsDNA template molecule.

27. The process of claim 26, wherein the histone modification-barcoded dsDNA template molecule, the adapter-bound cDNA, and the protein-barcoded dsDNA template molecule are pooled prior to sequencing.

28. The process of claim 17, wherein the cell-free nucleic acid sample is obtained from a blood sample.

29. The process of claim 28, further comprising carrying out a plasma protein analysis on a fraction of the blood sample.

30. The process of claim 29, wherein plasma proteins are analyzed by generating a dsDNA template molecule that comprises a protein identifier barcode corresponding to a protein analyte.

31. The process of claim 30, further comprising amplifying and sequencing the 5hmC-containing DNA, the untagged DNA, and the adapter-bound cDNA.

32. The process of claim 31, wherein the 5hmC-containing DNA, the untagged DNA, and the adapter-bound cDNA are pooled prior to sequencing.

33. The process of claim 20, wherein the cell-free nucleic acid sample is obtained from a blood sample.

34. The process of claim 33, further comprising carrying out a plasma protein analysis on the blood sample.

35. The process of claim 34, wherein plasma proteins are analyzed by generating a dsDNA template molecule that comprises a protein identifier barcode corresponding to a protein analyte.

36. The process of claim 35, further comprising amplifying and sequencing the 5hmC-containing DNA, the 5mC-containing DNA, the untagged DNA, and the adapter-bound cDNA.

37. The process of claim 36, wherein the 5hmC-containing DNA, the 5mC-containing DNA, the untagged DNA, and the adapter-bound cDNA are pooled prior to sequencing.