

CLAIMS

WHAT IS CLAIMED IS:

1. A method comprising:
 - (a) obtaining a stabilized biological sample comprising a nucleic acid molecule complexed to at least one nucleic acid binding protein;
 - (b) contacting the stabilized biological sample to a non-specific endonuclease to cleave the nucleic acid molecule into a plurality of segments;
 - (c) attaching a first segment and a second segment of the plurality of segments at a junction; and
 - (d) subjecting the plurality of segments to size selection to obtain a plurality of selected segments.
2. The method of claim 1, wherein the plurality of selected segments is about 145 to about 600 bp.
3. The method of claim 1, wherein the plurality of selected segments is about 100 to about 2500 bp.
4. The method of claim 1, wherein the plurality of selected segments is about 100 to about 600 bp.
5. The method of claim 1, wherein the plurality of selected segments is about 600 to about 2500 bp.
6. The method of claim 1, further comprising, prior to step (d), preparing a sequencing library from the plurality of segments.
7. The method of claim 6, further comprising subjecting the sequencing library to a size selection to obtain a size-selected library.
8. The method of claim 7, wherein the size-selected library is from about 350 bp to about 1000 bp in size.
9. The method of any one of claims 1 to 8, wherein the size selection is conducted with gel electrophoresis, capillary electrophoresis, size selection beads, or a gel filtration column.
10. The method of any one of claims 1 to 9, wherein the method further comprises analyzing the plurality of selected segments to obtain a QC value.
11. The method of claim 10, wherein the QC value is a chromatin digest efficiency (CDE) based on the proportion of segments from 100 bp to 2500 bp in size prior to step (d).
12. The method of claim 11, wherein the method further comprises selecting a sample for further analysis when the CDE value is at least 65%.
13. The method of claim 10, wherein the QC value is a chromatin digest index (CDI) based on the ratio of a number of mononucleosome-sized segments to a number of dinucleosome-sized segments prior to step (d).
14. The method of claim 13, wherein the method further comprises selecting a sample for further analysis when the CDI value is greater than -1.5 and less than 1.

15. The method of claim 1, further comprising, subsequent to the contacting the stabilized biological sample to a non-specific endonuclease, binding the plurality of segments to one or more surfaces.
16. The method of claim 15, wherein the one or more surfaces comprise one or more beads.
17. The method of claim 16, wherein the one or more beads are solid phase reversible immobilization (SPRI) beads.
18. The method of any one of claims 1 to 14, wherein the stabilized biological sample comprises a stabilized cell lysate.
19. The method of any one of claims 1 to 14, wherein the stabilized biological sample comprises a stabilized intact cell.
20. The method of any one of claims 1 to 14, wherein the stabilized biological sample comprises a stabilized intact nucleus.
21. The method of claim 19 or claim 20, wherein step (b) is conducted prior to lysis of the intact cell or the intact nucleus.
22. The method of claim 1, further comprising, prior to step (c), lysing cells and/or nuclei in the stabilized biological sample.
23. The method of any one of claims 1 to 20, wherein the stabilized biological sample comprises fewer than 3,000,000 cells.
24. The method of any one of claims 1 to 23, wherein the stabilized biological sample comprises fewer than 1,000,000 cells.
25. The method of any one of claims 1 to 24, wherein the stabilized biological sample comprises fewer than 100,000 cells.
26. The method of any one of claims 1 to 25, wherein the stabilized biological sample comprises less than 10 µg DNA.
27. The method of any one of claims 1 to 26, wherein the stabilized biological sample comprises less than 1 µg DNA.
28. The method of any one of claims 1 to 27, wherein the non-specific endonuclease is DNase.
29. The method of claim 28, wherein the DNase is DNase I.
30. The method of claim 28, wherein the DNase is DNase II.
31. The method of claim 28, wherein the DNase is micrococcal nuclease.
32. The method of claim 28, wherein the DNase is selected from one or more of DNase I, DNase II, and micrococcal nuclease.
33. The method of any one of claims 1 to 32, wherein the stabilized biological sample has been treated with a crosslinking agent.
34. The method of claim 33, wherein the crosslinking agent is a chemical fixative.
35. The method of claim 34, wherein the chemical fixative comprises formaldehyde.
36. The method of claim 34, wherein the chemical fixative comprises psoralen.

37. The method of claim 34, wherein the chemical fixative comprises disuccinimidyl glutarate (DSG).
38. The method of claim 34, wherein the chemical fixative comprises ethylene glycol bis(succinimidyl succinate) (EGS).
39. The method of claim 34, wherein the chemical fixative comprises disuccinimidyl glutarate (DSG) and ethylene glycol bis(succinimidyl succinate) (EGS).
40. The method of claim 33, wherein the crosslinking agent is ultraviolet light.
41. The method of any one of claims 1 to 40, wherein the stabilized biological sample is a crosslinked paraffin-embedded tissue sample.
42. The method of any one of claims 1 to 41, further comprising contacting the plurality of selected segments to an antibody.
43. The method of claim 1, further comprising conducting immunoprecipitation on the plurality of segments.
44. The method of claim 43, wherein the immunoprecipitation is conducted subsequent to the attaching.
45. The method of any one of claims 1 to 42, wherein attaching comprises filling in sticky ends using biotin tagged nucleotides.
46. The method of any one of claims 1 to 42, wherein attaching comprises filling in sticky ends using untagged nucleotides.
47. The method of any one of claims 1 to 42, wherein attaching comprises ligating blunt ends.
48. The method of any one of claims 1 to 42, wherein attaching comprises adding overhangs.
49. The method of claim 48, wherein the adding overhangs comprises adenylation.
50. The method of any one of claims 1 to 45, wherein attaching comprises contacting at least the first segment and the second segment to at least one bridge oligonucleotide.
51. The method of claim 50, wherein the bridge oligonucleotide is at least 10 bp in length.
52. The method of claim 50, wherein the bridge oligonucleotide is at least 12 bp in length.
53. The method of claim 50, wherein the bridge oligonucleotide is 12 bp in length.
54. The method of claim 50, wherein the bridge oligonucleotide comprises a barcode sequence.
55. The method of claim 50, wherein the bridge oligonucleotide comprises an affinity tag.
56. The method of claim 55, wherein the affinity tag is biotin.
57. The method of claim 50, wherein attaching comprises contacting at least the first segment and the second segment to multiple bridge oligonucleotides in series.
58. The method of claim 55, wherein the attaching results in samples, cells, nuclei, chromosomes, or nucleic acid molecules of the stabilized biological sample receiving a unique sequence of bridge oligonucleotides.
59. The method of any one of claims 50 to 56, wherein the at least one bridge oligonucleotide is coupled to an immunoglobulin binding protein or a fragment thereof.

60. The method of any one of claims 50 to 57, wherein the at least one bridge oligonucleotide is coupled or fused to two or more immunoglobulin binding proteins or fragments thereof.

61. The method of claim 57 or claim 58, wherein the immunoglobulin binding protein is selected from a Protein A, a Protein G, a Protein A/G, and a Protein L.

62. The method of any one of claims 1 to 45, wherein attaching comprises contacting at least the first segment and the second segment to a barcode.

63. The method of any one of claims 1 to 60, wherein the method does not comprise a shearing step.

64. The method of claim 1, further comprising:

(e) obtaining at least some sequence on each side of the junction to generate a first read pair.

65. The method of claim 62, further comprising:

(f) mapping the first read pair to a set of contigs; and

(g) determining a path through the set of contigs that represents an order and/or orientation to a genome.

66. The method of claim 62, further comprising:

(f) mapping the first read pair to a set of contigs; and

(g) determining, from the set of contigs, a presence of a structural variant or loss of heterozygosity in the stabilized biological sample.

67. The method of claim 62, further comprising:

(f) mapping the first read pair to a set of contigs; and

(g) assigning a variant in the set of contigs to a phase.

68. The method of claim 65, wherein the variant is a human leukocyte antigen (HLA) variant.

69. The method of claim 65, wherein the variant is a killer-cell immunoglobulin-like receptor (KIR) variant.

70. The method of claim 62, further comprising:

(f) mapping the first read pair to a set of contigs;

(g) determining, from the set of contigs, a presence of a variant in the set of contigs; and

(h) conducting a step selected from one or more of: (1) identifying a disease stage, a prognosis, or a course of treatment for the stabilized biological sample; (2) selecting a drug based on the presence of the variant; or (3) identifying a drug efficacy for the stabilized biological sample.

71. The method of any one of claims 1 to 68, wherein the DNase is coupled or fused to an immunoglobulin binding protein or a fragment thereof.

72. The method of any one of claims 1 to 69, wherein the DNase is coupled to two or more immunoglobulin binding proteins or fragments thereof.

73. The method of claim 69 or claim 70, wherein the immunoglobulin binding protein is selected from a Protein A, a Protein G, a Protein A/G, and a Protein L.