a benzene, and a toluene. considering the kit components listed herein and substantially equivalent variants thereof, alternative kits are contemplated wherein at least one commercially available kit component is excluded, being replaced by instructions for successful use of the remaining components in combination with reagents independently obtained.

[0010a] Provided herein is a method, comprising: obtaining a preserved sample from a subject, the sample comprising protein-DNA complexes; treating the preserved sample with a proteinase at a temperature not greater than 40 °C to isolate nucleic acids such that protein-DNA complexes are not destroyed, such that a first segment and a second segment are held together independent of a phosphodiester backbone, and such that the first segment and the second segment each have at least one exposed nucleic acid end; and deriving genomic structural information by analyzing nucleic acids of the protein-DNA complexes in the sample, wherein the preserved sample is a formalin fixed paraffin-embedded (FFPE) sample.

INCORPORATION BY REFERENCE

[0011] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference in its entirety as well as any references cited therein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1A depicts an exemplary schematic of a formalin fixed, paraffin embedded (FFPE) tissue sample.

[0013] FIG. 1B depicts an exemplary schematic of a protocol for chromatin-based next generation sequencing (NGS) library preparation.

[0014] FIG. 2A and **FIG. 2B** depict exemplary simple kernels that can be used for finding reciprocal translocations.

[0015] FIG. **3** depicts an image with a signal of a reciprocal translocation between ETV6 and NTRK3.

[0016] FIG. 4A, **FIG. 4B**, and **FIG. 4C** depict image analysis-based results at the same pair of chromosomes compared in three different samples.

[0017] FIG. 5A, **FIG 5B**, and **FIG. 5C** depict median normalized read density (over 10 samples) for chromosome 1 versus chromosome 7 (**FIG. 5A**), chromosome 2 versus chromosome 5 (**FIG. 5B**), and chromosome 1 versus chromosome 1 (**FIG. 5C**).

[0018] FIG. 6A and **FIG. 6B** depict various bin handling approaches. **FIG. 6A** shows equal bin sizes and **FIG. 6B** shows bin interpolation.

[0019] FIG. **7** depicts analysis by a genome-wide scanning analysis pipeline.

[0020] FIG. 8A and **FIG. 8B** depict read pair distance frequency data derived from FFPE-based 'Chicago' read pair libraries (**FIG. 8A**) and classic 'Chicago' based read pair libraries (**FIG. 8B**).

[0021] FIG. 9A and **FIG. 9B** illustrate the mapped locations on the GRCh38 reference sequence of read pairs are plotted in the vicinity of structural differences between GM12878 and the

CLAIMS

WHAT IS CLAIMED IS:

1. A method, comprising

obtaining a preserved sample from a subject, the sample comprising protein-DNA complexes;

treating the preserved sample with a proteinase at a temperature not greater than 40 °C to isolate nucleic acids such that protein-DNA complexes are not destroyed, such that a first segment and a second segment are held together independent of a phosphodiester backbone, and such that the first segment and the second segment each have at least one exposed nucleic acid end; and

deriving genomic structural information by analyzing nucleic acids of the protein-DNA complexes in the sample,

wherein the preserved sample is a formalin fixed paraffin-embedded (FFPE) sample.

2. The method of claim 1, wherein the preserved sample is crosslinked.

3. The method of claim 2, wherein the preserved sample is crosslinked using at least one of a formaldehyde, a formalin, UV light, mitomycin C, nitrogen mustard, melphalan, 1,3-butadiene diepoxide, cis diaminedichloroplatinum(II) and cyclophosphamide.

4. The method of claim 1, wherein the preserved sample maintains positional information as to nucleic acids within it.

5. The method of claim 1, wherein the genomic structural information is indicative of at least one of an inversion, an insertion, a deletion, and a translocation relative to a reference genome.

6. The method of claim 5, wherein the reference genome is a wild type genome of a species common to the subject.

7. The method of claim 5, wherein the reference genome is obtained from a reference tissue of the subject.

8. The method of claim 1, comprising deriving information indicative of phase status for a first segment and a second segment of the nucleic acids.

9. The method of claim 1, comprising tagging the exposed nucleic acid end of the sample so as to convey physical linkage information.

10. The method of claim 9, wherein the tagging comprises ligating oligonucleotides to a DNA protein complex released from the preserved sample such that the oligonucleotides convey information indicative of a common complex.

11. The method of claim 10, wherein the oligonucleotides comprise base sequence specific to a complex.

12. The method of claim 10, wherein the oligonucleotides comprise base sequence unique to a complex.

13. The method of claim 9, wherein the tagging comprises ligating a first nucleic acid segment of the protein-DNA complex to a second segment of the protein-DNA complex to form a paired end molecule.

14. The method of claim 13, comprising sequencing a portion of the first nucleic acid segment and a portion of the second nucleic acid segment.

15. The method of claim 14, comprising assigning contigs having unique sequence common to the portion of the first nucleic acid segment and contigs having unique sequence

common to the portion of the second nucleic acid segment to a common scaffold in a nucleic acid assembly.

16. The method of claim 13, comprising contacting the paired end nucleic acid molecule to a set of nucleic acid probes.

17. The method of claim 16, wherein the set of nucleic acid probes are fluorescent probes.

18. The method of claim 16, wherein the set of nucleic acid probes anneal to a first locus and a second locus implicated in a genome structural rearrangement.

19. The method of claim 18, wherein the first locus and the second locus are not adjacent in a genome unaffected by the genome structural rearrangement.

20. The method of claim 18, wherein the first locus and the second locus are adjacent in a genome unaffected by the genome structural rearrangement.

21. The method of any one of claims 16-20, comprising sequencing nucleic acids of the sample when contacting the set of nucleic acid probes indicates a rearrangement.

22. The method of claim 13, comprising contacting the paired end nucleic acid molecule to a set of nucleic acid primers.

23. The method of claim 22, wherein the set of nucleic acid primers anneal to a first locus and a second locus implicated in a genome structural rearrangement.

24. The method of claim 23, wherein the set of nucleic acid primers yield an amplicon in a nucleic acid amplification reaction when the first locus and the second locus form a ligated paired end molecule.

25. The method of claim 23, wherein the set of nucleic acid primers do not yield an amplicon in a nucleic acid amplification reaction when the first locus and the second locus do not form a ligated paired end molecule.

26. The method of claim 18, wherein the first locus and the second locus are not adjacent in a genome unaffected by the genome structural rearrangement.

27. The method of claim 18, wherein the first locus and the second locus are adjacent in a genome unaffected by the genome structural rearrangement.

28. The method of any one of claims 22-27, comprising sequencing nucleic acids of the sample when an amplicon is generated from the set of nucleic acid primers contacted to the paired end nucleic acid molecule.

29. The method of claim 1, wherein the preserved sample is treated by contacting the preserved tissue sample to xylene.

30. The method of claim 1, wherein the preserved sample is treated by contacting the preserved tissue sample to ethanol.

31. The method of claim 1, wherein the preserved sample is treated by protecting the sample from boiling conditions.

32. The method of claim 1, wherein the preserved sample is treated by contacting the preserved tissue sample to at least one of an anthranilate and a phosphanilate.

33. The method of claim 1, wherein the protein-DNA complexes comprise chromatin.

34. The method of claim 1, wherein the preserved sample preserves positional information reflective of its configuration in a tissue.

35. The method of claim 1, wherein the preserved sample is not homogenized prior to isolating nucleic acids.

36. The method of claim 1, wherein the preserved sample is stored for at least one week prior to isolating nucleic acids.

37. The method of claim 1, wherein the preserved sample is stored for at least 6 months prior to isolating nucleic acids.

38. The method of claim 1, wherein the preserved sample is transported from a collection point prior to isolating nucleic acids.

39. The method of claim 1, wherein the preserved sample is collected in a sterile environment.

40. The method of claim 1, wherein the preserved sample is positioned in a nonsterile environment prior to isolating nucleic acids.

41. The method of claim 1, wherein the proteinase comprises proteinase K.

42. The method of claim 1, wherein the preserved sample is treated with the proteinase for not greater than 1 hour.