CLAIMS

WHAT IS CLAIMED IS:

1. A method of analyzing a nucleic acid population comprising at least two forms of nucleic acid selected from double-stranded DNA, single-stranded DNA and single-stranded RNA, the method, wherein each of the at least two forms comprises a plurality of molecules, comprising:

(a) linking at least one of the forms of nucleic acid with at least one tag nucleic acid to distinguish the forms from one another,

(b) amplifying the forms of nucleic acid at least one of which is linked to at least one nucleic acid tag, wherein the nucleic acids and linked nucleic acid tag, are amplified, to produce amplified nucleic acids, of which those amplified from the at least one form are tagged;

(c) assaying sequence data of the amplified nucleic acids at least some of which are tagged; wherein the assaying obtains sequence information sufficient to decode the tag nucleic acid molecules of the amplified nucleic acids to reveal the forms of nucleic acids in the population providing an original template for the amplified nucleic acids linked to the tag nucleic acid molecules for which sequence data has been assayed.

2. The method of claim 1, further comprising the step of decoding the tag nucleic acid molecules of the amplified nucleic acids to reveal the forms of nucleic acids in the population providing an original template for the amplified nucleic acids linked to the tag nucleic acid molecules for which sequence data has been assayed.

3. The method of claim 1 or 2, further comprising enriching for at least one of the forms relative to one or more of the other forms.

4. The method of claim 1 or 2, wherein at least 70% of the molecules of each form of nucleic acid in the population are amplified in step (b).

5. The method of claim 1 or 2, wherein at least three forms of nucleic acid are present in the population and at least two of the forms are linked to different tag nucleic acid forms distinguishing each of the three forms from one another.

6. The method of claim 5 wherein each of the at least three forms of nucleic acid in the population is linked to a different tag.

7. The method of claim 1 or 2, wherein each molecule of the same form is linked to a tag comprising the same identifying tag.

8. The method of claim 1 or 2, wherein molecules of the same form are linked to different types of tags.

9. The method of claim 1 or 2, wherein step (a) comprising: subjecting the population to reverse transcription with a tagged primer, wherein the tagged primer is incorporated into cDNA generated from RNA in the population.

10. The method of claim 9, wherein the reverse transcription is sequence-specific.

11. The method of claim 9, wherein the reverse transcription is random.

12. The method of claim 9, further comprising degrading RNA duplexed to the cDNA.

13. The method of claim 5 further comprising separating single-stranded DNA from double-stranded DNA and ligating nucleic acid tags to the double-stranded DNA.

14. The method of claim 13, wherein the single-stranded DNA is separated by hybridization to one or more capture probes.

15. The method of claim 5, further comprising circularizing single-stranded DNA with a circligase and ligating nucleic acid tags to the double-stranded DNA.

16. The method of claim 1, comprising, before assaying, pooling tagged nucleic acids comprising different forms of nucleic acid.

17. The method of claims 1-16 wherein the nucleic acid population is from a bodily fluid sample.

18. The method of claim 17, wherein the bodily fluid sample is blood, serum, or plasma.

19. The method of claim 1 or 2, wherein the nucleic acid population is a cell free nucleic acid population.

20. The method of claim 18, wherein the bodily fluid sample is from a subject suspected of having a cancer.

21. The method of claims 1-20, wherein the sequence data indicates presence of a somatic or germline variant.

22. The method of claims 1-21, wherein the sequence data indicates presence of a copy number variation.

23. The method of claims 1-22 , wherein the sequence data indicates presence of a single nucleotide variation (SNV), indel or gene fusion.

24. A method of analyzing a nucleic acid population comprising nucleic acids with different extents of modification, comprising:

contacting the nucleic acid population with an agent that preferentially binds to nucleic acids bearing the modification,

separating a first pool of nucleic acids bound to the agent from a second pool of nucleic acids unbound to the agent, wherein the first pool of nucleic acids is overrepresented for the modification, and the nucleic acids in the second pool are underrepresented for the modification;

linking the nucleic acids in the first pool and/or second pool to one or more nucleic acid tags that distinguish the nucleic acids in the first pool and the second pool to produce a population of tagged nucleic acids;

amplifying the labelled nucleic acids, wherein the nucleic acids and the linked tags are amplified;

assaying sequence data of the amplified nucleic acids and linked tags; wherein the assaying obtains sequence data for decoding the tags to reveal whether the nucleic acids for

which sequence data has been assayed were amplified from templates in the first or the second pool.

25. The method of Claim 24, comprising the step of decoding the tags to reveal whether the nucleic acids for which sequence data has been assayed were amplified from templates in the first or the second pool.

26. The method of claim 25 or 26, wherein the modification is binding of nucleic acids to a protein.

27. The method of claim 25 or 26, wherein the protein is a histone or transcription factor.

28. The method of claim 25 or 26, wherein the modification is a post-replication modification to a nucleotide.

29. The method of claim 27, wherein the post-replication modification is 5-methyl-cytosine, and the extent of binding of the agent to nucleic acids increases with the extent of 5-methyl-cytosines in the nucleic acid.

30. The method of claim 27, wherein the post-replication modification is 5-hydroxymethyl-cytosine, and the extent of binding of the agent to nucleic acid increases with the extent of 5-hydroxymethyl-cytosine in the nucleic acid.

31. The method of claim 27, wherein the post-replication modification is 5-formyl-cytosine or 5-carboxyl-cytosine and the extent of binding of the agent increases with the extent of 5-formyl-cytosine or 5-carboxyl-cytosine in the nucleic acid.

32. The method of claim 25 or 26, further comprising washing nucleic acids bound to the agent and collecting the wash as a third pool including nucleic acids with the post replication modification at an intermediate extent relative to the first and second pools.

33. The method of claim 25 or 26, comprising, before assaying, pooling tagged nucleic acids from the first and second pools.

34. The method of claim 25 or 26, wherein the agent is 5-methyl- binding domain magnetic beads.

35. The method of claims 24-34, wherein the nucleic acid population is from a bodily fluid sample.

36. The method of claim 35, wherein the bodily fluid sample is blood, serum, or plasma.

37. The method of claim 25 or 26, wherein the nucleic acid population is a cell free nucleic acid population.

38. The method of claim 35, wherein the bodily fluid sample is from a subject suspected of having a cancer.

39. The method of claims 25-38, wherein the sequence data indicates presence of a somatic or germline variant.

40. The method of claims 25-39, wherein the sequence data indicates presence of a copy number variation.

41. The method of any 25-39, wherein the sequence data indicates presence of a single nucleotide variation (SNV), indel or gene fusion.

iv) nucleic acids, wherein the nucleic acids and the linked tags are amplified; and assaying the molecular tagged partitions with the NGS instrument;

v) a software module for generating sequence data for decoding the tag; and

vi) a software module for analyzing the sequence data to decode the tags to reveal whether the nucleic acids for which sequence data has been assayed were amplified from templates in the first or the second pool.