CLAIMS

WHAT IS CLAIMED IS:

1. A method of analyzing a cell free 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 second pool to one or more nucleic acid tags that distinguish the nucleic acids in the first pool and the second pool and that distinguish, alone or in combination with information from the nucleic acid to which they are attached, different nucleic acids within the pools from one another 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.

2. The method of claim 1, wherein the modification is binding of nucleic acids to a protein.

3. The method of claim 2, wherein the protein is a histone or transcription factor.

4. The method of claim 1, wherein the modification is a post-replication modification to a nucleotide.

5. The method of claim 4, 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.

6. The method of claim 4, 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.

7. The method of claim 4, 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.

8. The method of claim 1, 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.

9. The method of claim 1 or 2, wherein the amplified labelled nucleic acids are enriched for target sequences prior to the assaying sequence data step.

10. The method of claim 9, comprising use of one or a plurality of bait sets, wherein the bait sets comprise target-specific oligonucleotide baits, and wherein the bait sets are combined with the amplified labeled nucleic acids under conditions that allow hybridization of the target sequences with the baits.

11. The method of claim 1, comprising, before assaying, pooling tagged nucleic acids from the first and second pools.

12. The method of claim 1, wherein the agent is 5-methyl- binding domain magnetic beads.

13. The method of any one of claims 1-12, wherein the nucleic acid population is from a bodily fluid sample.

14. The method of claim 13, wherein the bodily fluid sample is blood, serum, or plasma.

15. The method of claim 13, wherein the bodily fluid sample is from a subject suspected of having a cancer.

16. The method of any one of claims 1-15, wherein the sequence data indicates presence of a somatic or germline variant.

17. The method of any one of claims 1-16, wherein the sequence data indicates presence of a copy number variation.

18. The method of any one of claims 1-16, wherein the sequence data indicates presence of a single nucleotide variation (SNV), indel or gene fusion.

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

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

iii) 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.