Claims

1. A method of analyzing a tumor sample for a somatic mutation, comprising:

(a) acquiring a library comprising a plurality of nucleic acid molecules from a tumor sample;

(b) enriching the library for preselected sequences by contacting the library with a plurality of bait sets to provide selected tumor nucleic acid molecules, thereby producing a library catch;

(c) acquiring a read for a subgenomic interval comprising a somatic mutation from a tumor nucleic acid molecule from said library or library catch by a next generation sequencing method;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for a preselected nucleotide position,

wherein a read from each of X unique subgenomic intervals is aligned with a unique alignment method, wherein unique subgenomic interval means different from the other X-1 subgenomic intervals, and wherein unique alignment method means different from the other X-1 alignment methods, and X is at least 2, 10, 15, 20, 30, 50, 100, 500, or 1,000,

wherein said method of analyzing a tumor sample comprises application of two or more of the following alignment methods:

(i) a first alignment method that is selected responsive to, or is optimized for, a variant which appears at a frequency of 5% or less, in cells in the tumor sample, the nucleic acid molecules of the library, or the nucleic acid molecules of the library catch;

(ii) a second alignment method that is selected responsive to, or is optimized for, a variant which appears at a frequency of greater than 5% and up to 10%, 15% or 20%, in cells in the tumor sample, the nucleic acid molecules of the library, or the nucleic acid molecules of the library catch;

(iii) a third alignment method that is selected responsive to, or is optimized for, a variant which is:

(1) a pharmacogenomic SNP that is associated with the ability of patient to respond to or metabolize a drug;

(2) a genomic SNP that uniquely identifies (fingerprints) a patient; or

(3) a genomic SNP/locus that is used to assess copy number gains/losses of genomic DNA and LOH;

(iv) a fourth alignment method that is selected responsive to, or is optimized for, a structural breakpoint; and

(v) a fifth alignment method that is selected responsive to, or is optimized for, a target for which sparse coverage can improve the ability to detect copy number changes.

2. The method of claim 1, wherein the method of analyzing a tumor sample comprises application of the fourth alignment method.

3. The method of claim 1 or 2, wherein the structural breakpoint is a rearrangement, a translocation, or an indel.

4. The method of any of claims 1-3, wherein the target is an intron target.

5. A method of analyzing a tumor sample for a somatic mutation, comprising:

(a) acquiring a library comprising a plurality of nucleic acid molecules from a tumor sample;

(b) enriching the library for preselected sequences by contacting the library with a plurality of bait sets to provide selected tumor nucleic acid molecules, thereby producing a library catch;

(c) acquiring a read for a subgenomic interval comprising a somatic mutation from a tumor nucleic acid molecule from said library or library catch by a next generation sequencing method;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for the preselected nucleotide position, thereby analyzing said tumor sample,

wherein the alignment method comprises aligning the read with a rearrangement alignment sequence which is preselected to align with a preselected rearrangement,

wherein a read from each of at least X unique subgenomic intervals is aligned with a unique alignment method and X is 2, 10, 15, 20, 30, 50, 100, 500, or 1,000,

wherein unique subgenomic interval means different from the other X-1 subgenomic intervals, and wherein unique alignment method means different from the other X-1 alignment methods, and

wherein the reads from the at least X unique subgenomic intervals are analyzed using at least X unique preselected rearrangement alignment sequences, and wherein unique means different from the other X-1.

6. The method of claim 5, wherein the preselected rearrangement is a preselected indel.

7. The method of claim 5 or 6, wherein:

(i) the preselected rearrangement alignment sequence comprises a sequence or its complement corresponding to said preselected rearrangement;

(ii) the preselected rearrangement alignment sequence comprises a simulated sequence that is other than the sequence of the preselected rearrangement or its complement selected to align with a read of said preselected rearrangement sequence; or

(iii) the rearrangement alignment sequence is not identical to the preselected rearrangement.

8. The method of any of claims 5-7, wherein subgenomic intervals from at least X genes from Table 1 or 1A having the priority 1 annotation are aligned with unique alignment methods, and X is equal to 10, 15, 20, or 30.

9. The method of any of claims 1-8, comprising:

a) applying a first unique alignment method to a first subgenomic interval, a variant of which is associated with a tumor phenotype, wherein the variant is a point mutation from Table 6;

b) applying a second unique alignment method to a second subgenomic interval, a variant of which is associated with a tumor phenotype, wherein the variant is a rearrangement chosen from a deletion, insertion, or translocation on Table 5; and

c) applying a third unique alignment method to a third subgenomic interval comprising a subgenomic interval in which variants are not associated with a tumor phenotype or with a tumor of the type in said sample.

10. The method of any of claims 1-9, wherein a subgenomic interval being analyzed comprises a nucleotide position with a rearrangement, and the method comprises using an alignment method that includes:

selecting a rearrangement reference sequence for alignment with a read, wherein said rearrangement reference sequence is preselected to align with a preselected rearrangement, wherein the reference sequence is not identical to the genomic rearrangement;

comparing a read with said preselected rearrangement reference sequence; and

determining if said read meets a predetermined alignment criterion, wherein the predetermined criterion is an alignment to said preselected rearrangement reference with less than a preselected level of mismatch or gaps;

thereby analyzing a read.

11. The method of any of claims 1-10, wherein a first alignment method differs from a second alignment method by a value for a match reward, mismatch penalty, gap penalty, expected threshold, word size, filter, or mask.

12. A method of analyzing a tumor sample comprising a somatic mutation, comprising:

(a) acquiring a library comprising a plurality of nucleic acid molecules from a tumor sample;

(b) enriching the library for preselected sequences by contacting the library with a plurality of bait sets to provide selected tumor nucleic acid molecules, thereby providing a library catch;

(c) acquiring a read for a subgenomic interval comprising a somatic mutation from a tumor nucleic acid molecule from said library or library catch by a next generation sequencing method;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for the preselected nucleotide position,

wherein a nucleotide value assigned for a nucleotide position in each of X unique subgenomic intervals by a unique calling method,

wherein unique subgenomic interval means different from the other X-1 subgenomic intervals, and wherein unique calling method means different from the other X-1 calling methods, and X is at least 2.

13. The method of claim 12, wherein the calling methods differ by relying on different Bayesian prior values chosen from a first calling method applied to a first nucleotide position that

is a function of a first Bayesian prior and a second calling method applied to a second nucleotide position that is a function of a second Bayesian prior.

14. The method of claim 12 or 13, wherein assigning said nucleotide value is a function of a value which is or represents the prior expectation of observing a read showing a preselected variant at said preselected nucleotide position in a tumor of type.

15. The method of any of claims 12-14, wherein the method comprises assigning a nucleotide value for at least X, where X is 10, 20, 40, 50, 60, 70, 80, 90, or 100, preselected nucleotide positions, wherein each assignment is a function of a unique value which is or represents the prior expectation of observing a read showing a preselected variant at said preselected nucleotide position in a tumor of type, wherein unique value means different from the values for the other X-1 assignments.

16. The method of any of claims 12-15, wherein assigning said nucleotide value is a function of a set of values which represent the probabilities of observing a read showing said preselected variant at said preselected nucleotide position if the variant is present in the sample at a frequency of less than 10% and/or if the variant is absent.

17. The method of any of claims 12-16, wherein the calling method comprises:

acquiring, for a preselected nucleotide position in each of said X unique subgenomic intervals:

(i) a first value which is or represents a prior expectation of observing a read showing a preselected variant at said preselected nucleotide position in a type of tumor; and

(ii) a second set of values which represent the probabilities of observing a read showing said preselected variant at said preselected nucleotide position if the variant is present in the sample at a frequency of at least 1%, at least 5%, at least 10%, or more, or if the variant is absent;

responsive to said values, assigning a nucleotide value from said reads for each of said preselected nucleotide positions by weighing the comparison among the values in the second set using the first value, thereby analyzing said sample.

18. The method of claim 17, wherein the preselected nucleotide position and the type of tumor are according to Tables 6, 9 and 10.

19. The method of any of claims 12-18, comprising one or more or all of:

(i) assigning a nucleotide value for at least 10 preselected nucleotide positions, wherein each assignment is based on a unique first and/or second value, as compared with the values for the other assignments;

(ii) the assignment of method of (i), wherein at least 10 of the assignments are made with first values which are a function of a probability of a preselected variant being present of less than 5, 10, or 20% of the cells in a preselected type of tumor;

(iii) assigning a nucleotide value for at least X preselected nucleotide positions, each of which being associated with a preselected variant having a unique probability of being present in the tumor of preselected type as compared with the probability for the other X-1 assignments, wherein each of said X assignments is based on a unique first and/or second value as compared with the values for the other X-1 assignments, wherein X= 10; or

(iv) assigning a nucleotide value at a first and a second nucleotide position, wherein the likelihood of a first preselected variant at said first nucleotide position being present in a tumor of preselected type is at least 2 times greater than the likelihood of a second preselected variant at said second nucleotide position being present, wherein each assignment is based on a unique first and/or second value as compared with the values for the other assignments.

20. The method of any of claims 12-19, comprising assigning a nucleotide value to a plurality of preselected nucleotide positions, wherein said plurality comprises an assignment for variants falling into at least 3, or all, of the following probability ranges:

less than 0.01 %;

greater than 0.01 and less than or equal to 0.02 %;

greater than 0.02 and less than or equal to 0.03 %;

greater than 0.03 and less than or equal to 0.04 %;

greater than 0.04 and less than or equal to 0.05 %;

greater than 0.05 and less than or equal to 0.1 %;

greater than 0.1 and less than or equal to 0.2 %;

greater than 0.2 and less than or equal to 0.5 %;

greater than 0.5 and less than or equal to 1.0 %;

greater than 1.0 and less than or equal to 2.0 %;

greater than 2.0 and less than or equal to 5.0 %;

greater than 5.0 and less than or equal to 10.0 %;

greater than 10.0 and less than or equal to 20.0 %;

greater than 20.0 and less than or equal to 50.0 %; and

greater than 50 and less than or equal to 100.0 %;

wherein, a probability range is the range of probabilities that a preselected variant at a preselected nucleotide position is present in a tumor of preselected type or the probability that a preselected variant at a preselected nucleotide position is present in the cells in a tumor sample, a library from the tumor sample, or library catch from that library, for a preselected type, and wherein each assignment is based on a unique first and/or second value, wherein unique first or second value means different from the values for the other assignments in a probability range or different from the first and/or second value for one or more or all of the other probability ranges.

21. The method of any of claims 12-20, comprising one or both of:

(i) assigning a nucleotide value for at least 3 preselected nucleotide positions each, independently, having a preselected variant present in less than 5, 4, 3, 2, 1, 0.5, 0.4, 0.3, 0.2, or 0.1 % of the DNA in said sample, wherein each assignment is based on a unique first and/or second value as compared with the values for the other assignments; or

(ii) assigning a nucleotide value at a first and a second nucleotide position, wherein the likelihood of a preselected variant at the first position in the DNA of said sample is at least 2 times greater than a the likelihood of a preselected variant at said second nucleotide position in the DNA of said sample, wherein each assignment is based on a unique first and/or second value as compared with the values for the other assignments.

22. The method of any of claims 12-21, comprising assigning a nucleotide value to at least 3 preselected nucleotide positions having a preselected variant present than X % of the cells in said sample, of the nucleic acid in a library from said sample, or the nucleic acid in a library catch from that library,

wherein X is independently chosen from less than 1.0 %; between and equal to 1.0- 2.0 %; greater than 2.0 % and less than or equal to 3 %; greater than 3.0 % and less than or equal to 4 %; greater than 4.0 % and less than or equal to 5 %; greater than 5.0 % and less than or equal to 10 %;

greater than 10.0 % and less than or equal to 20 %; greater than 20.0 % and less than or equal to 40 %; greater than 40.0 % and less than or equal to 50 %; or greater than 50.0 % and less than or equal to 100 %, of the cells in said sample, of the nucleic acid in a library from said sample, or the nucleic acid in a library catch from that library;

wherein each assignment is based on a unique first and/or second value, wherein unique means different from the values for the other assignments in the range in (i) of less than 1% or different from a first and/or second values for a determination in one or more or all of the other ranges.

23. The method of any of claims 12-22, comprising assigning a nucleotide value at each of X nucleotide positions, each nucleotide position, independently, having a likelihood of a preselected variant being present in the DNA of said sample that is unique as compared with the likelihood for a preselected variant at the other X-1 nucleotide positions, wherein X is equal to or greater than 1, 2, 3, 5, 10, 20, 40, 50, 60, 70, 80, 90, or 100, and wherein each assignment is based on a unique first and/or second value as compared with the values for the other assignments.

24. The method of any of claims 12-23, said method comprising assigning a nucleotide value at at least 3 preselected nucleotide positions, each having, independently, a first value that is a function of a probability that is less than 0.5.

25. The method of any of claims 12-24, said method comprising assigning a nucleotide value at each of at least X nucleotide positions, each independently having a first value that is unique as compared with the other X-1 first values, and wherein each of said X first values is a function of a probability that is less than 0.5, wherein X is equal to or greater than 3.

26. The method of any of claims 12-25, wherein a nucleotide position in at least 20 genes from Table 1 or 1A, or in at least 10 genes from Table 1 or 1A having the priority 1 annotation, is assigned a nucleotide value.

27. The method of any of claims 12-26, wherein a unique first value, second value, or both values are applied to subgenomic intervals in each of at least 10% of said genes analyzed.

28. The method of any of claims 12-27, further comprising 2, 3, 4 or all of the following:

(i) responsive to said first value, second value or both values, assigning a nucleotide value from reads for a first preselected nucleotide position for which a coverage greater than 500X sequencing depth is required to enable a high level of sensitivity for a mutation that appears at a frequency of 5% or less in cells in the sample;

(ii) responsive to said first value, second value or both values, assigning a nucleotide value from reads for a second preselected nucleotide positions for which greater than 200X sequencing depth is required to enable a high level of sensitivity for a mutation that appears at a frequency of greater than 5% and up to 10, 15 or 20% in cells in the sample;

(iii) responsive to said first value, second value or both values, assigning a nucleotide value from reads for a third preselected nucleotide position for which coverage less than that in (i) or (ii) above is required to enable a high level of sensitivity for heterozygous alleles, wherein said third preselected nucleotide position is chosen from (1) a pharmacogenomic SNP associated with the ability of patient to respond to or metabolize a drug; (2) a plurality of genomic SNPs that are used to uniquely identify (fingerprint) a patient; or (3) a genomic SNP or locus that is used to assess copy number gains or losses of genomic DNA and LOH;

(iv) responsive to said first value, second value or both values, assigning a nucleotide value from reads for a fourth preselected nucleotide position chosen from a structural breakpoint, a rearrangement, a translocation or an indel; and

(v) responsive to said first value, second value or both values, assigning a nucleotide value from reads for a fifth preselected nucleotide position for which the coverage is less than in one of (i), (ii) (iii) or (iv) to ensure high detection reliability.

29. The method of any of claims 12-28, further comprising acquiring a threshold value for each of a number of preselected subgenomic intervals, wherein said threshold value is customized based on one or more of the type of cancer in a sample, the gene in which subgenomic interval to be sequenced is located, or the variant to be sequenced.

30. The method of any of claims 12-29, wherein the calling method comprises:

acquiring, for each of said X subgenomic intervals, a threshold value, wherein each of said acquired X threshold values is unique as compared with the other X-1 threshold values, thereby providing X unique threshold values;

for each of said X subgenomic intervals, comparing an observed value which is a function of the number of reads having a preselected nucleotide value at a preselected nucleotide position with its unique threshold value, thereby applying to each of said X subgenomic intervals, its unique threshold value; and

responsive to the result of said comparison, assigning a nucleotide value to a preselected nucleotide position,

wherein X is equal to or greater than 2.

31. The method of claim 30, wherein a unique threshold value is applied to subgenomic intervals in each of at least 10 different genes or in each of at least 10% of said gene analyzed.

32. The method of claim 30 or 31, wherein:

(a) a nucleotide position in at least 10 genes from Table 1 or 1A having the priority 1 annotation is assigned a nucleotide value; and a unique threshold value is applied to a subgenomic interval in each of at least 10% of said genes analyzed;

(b) a nucleotide position in at least 20 genes from Table 1 or 1A is assigned a nucleotide value; and a unique threshold value is applied to a subgenomic interval in each of at least 10% of said genes analyzed.

(c) a nucleotide position for at least 20 variants or codons from Table 1, 1A, 2, or 3, is assigned a nucleotide value; and a unique threshold value is applied to a subgenomic interval in each of at least 10% of said genes analyzed; or

(d) a nucleotide position for at least 10 variants from the bottom half or bottom third of Table 9 is assigned a nucleotide value; and a unique threshold value is applied to a subgenomic interval in each of at least 10% of said genes analyzed.

33. A method of analyzing a tumor sample, comprising:

(a) acquiring a library comprising a plurality of tumor nucleic acid molecules from a tumor sample;

(b) contacting the library with a bait set to provide selected tumor nucleic acid molecule, thereby providing a library catch;

(c) acquiring a read for a subgenomic interval comprising a somatic mutation from a tumor nucleic acid molecule from said library or library catch by a next generation sequencing method;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for the preselected nucleotide position,

wherein the method comprises contacting the library with at least two of the following bait sets:

(i) a first bait set that selects a high-level target chosen from one or more tumor nucleic acid molecules that comprise a subgenomic interval comprising a somatic mutation that appears at a frequency of about 5% or less of the cells from the sample harboring the alteration in their genome;

(ii) a second bait set that selects a mid-level target chosen from one or more tumor nucleic acid molecules that comprise a subgenomic interval comprising a somatic mutation that appears at a frequency of about 10% or higher of the cells from the sample harboring the alteration in their genome;

(iii) a third bait set that selects a low-level target chosen from one or more pharmacogenomic (PGx) nucleic acid molecules that comprise a subgenomic interval chosen from one or more of:

a) a PGx single nucleotide polymorphism (SNP) that distinguishes the ability of a patient to metabolize different drugs,

b) a genomic SNP that uniquely identifies (fingerprints) a patient, or

c) a genomic SNP/locus that is used to assess copy number gains or losses of genomic DNA and loss-of-heterozygosity (LOH);

(iv) a fourth bait set that selects a nucleic acid molecule that comprises an intron sequence that detects a structural breakpoint; or

(v) a fifth bait set that selects a one-copy deletion of several terminal exons,

wherein each bait set of said plurality has a unique preselected efficiency for selection for its target as compared with the other bait sets in the plurality.

34. The method of claim 33, wherein the efficiency of selection is modified by one or more of:

(i) differential representation of different bait sets;

(ii) differential overlap of bait subsets;

(iii) differential bait parameters comprising sequence modifications or shorter length to reduce capture efficiency and lower the relative target coverage depths;

(iv) mixing of different bait sets at different molar ratios to enhance or reduce relative target coverage depths; or

(v) using different types of oligonucleotide baits.

35. The method of claim 34, wherein different types of oligonucleotide baits are chosen from one or more of:

(a) one or more chemically (non-enzymatically) individually synthesized baits,

(b) one or more baits synthesized in an array,

(c) one or more *in vitro* transcribed baits;

(d) any combination of (a), (b) [[and/]]or (c),

(e) one or more naturally or non-naturally occurring DNA oligonucleotide,

(f) one or more naturally or non-naturally occurring RNA oligonucleotide,

(g) a combination of (e) and (f), or

(h) a combination of any of the above.

36. The method of claim 35, wherein the different types of oligonucleotide baits is mixed at a ratio chosen from 1:1, 1:2, 1:3, 1:4, 1:5, 1:10, 1:20, 1:50; 1:100, or 1:1000, or wherein the ratio of chemically-synthesized bait to array-generated bait is chosen from 1:5, 1:10, or 1:20.

37. The method of claim 35, wherein the non-naturally occurring DNA or RNA oligonucleotides are chosen from one or more of: a locked nucleic acid (LNA); a peptide nucleic acid (PNA); a DNA or RNA oligonucleotide modified to capture low GC regions; a bicyclic nucleic acid (BNA); a crosslinked oligonucleotide; a modified 5-methyl deoxycytidine; or 2,6-diaminopurine.

38. The method of claim 34, wherein one or more of the following bait parameters is modified:

(i) increasing/decreasing bait representation or overlap to enhance or reduce coverage of tumor nucleic acid molecules, which are under or over-covered relative to other targets in the same category;

(ii) for low coverage, hard to capture target sequences comprising high GC content sequences, targeting with the bait sets covering adjacent sequences;

(iii) modifying a bait sequence to reduce secondary structure of the bait and enhance its efficiency for selection;

(iv) modifying a bait length to equalize melting hybridization kinetics of different baits within the same category;

(v) modifying baits of different orientation for the same target region, forward and reverse strand, to have different binding efficiencies;

(vi) modifying the amount of a binding entity present on each bait that affects its binding efficiency.

(vii) modifying the type of nucleotide used for different baits to affect binding affinity to the target, and enhance or reduce the relative target coverage; or

(viii) using modified oligonucleotide baits having more stable base pairing to equalize melting hybridization kinetics between areas of low or normal GC content relative to high GC content.

39. The method of any of claims 1-38, wherein the bait set is chosen from one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve or all A-M of the following:

A. a bait set that selects an exon sequence that includes a single nucleotide alteration associated with a cancerous phenotype;

B. a bait set that selects an in-frame deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more codons from a reference nucleotide sequence;

C. a bait set that selects an intragenic deletion;

D. a bait set that selects an intragenic insertion;

E. a bait set that selects a deletion of a full gene;

F. a bait set that selects an inversion;

G. a bait set that selects an interchromosal translocation;

H. A bait set that selects a tandem duplication;

I. a bait set that selects a nucleotide sequence of interest flanked by adjacent non-repetitive sequences;

J. a bait set that selects one or more subgenomic intervals corresponding to a fusion sequence;

K. a bait set that selects a subgenomic interval adjacent to a nucleotide sequence that includes an undesirable feature chosen from a nucleotide sequence of high GC content, or a nucleotide sequence including one or more repeated elements and/or inverted repeats;

L. a bait set that selects a genomic rearrangement that includes an intron sequence from; or

M. a bait set that selects a subgenomic interval that includes an exon adjacent to a cancer associated gene fusion.

40. The method of any of claims 1-39, wherein the subgenomic interval comprises or consists of one or more of: a single nucleotide position; an intragenic region or an intergenic region; an exon or an intron, or a fragment thereof, typically an exon sequence or a fragment thereof; a coding region or a non-coding region, or a fragment thereof; a promoter; an enhancer; a 5' untranslated region (5' UTR); a 3' untranslated region (3' UTR); a cDNA or a fragment thereof; an SNP; a somatic mutation, a germ line mutation or both; an alteration comprising a point or a single mutation; a deletion mutation; an in-frame deletion, an intragenic deletion, a full gene deletion; an insertion mutation; an intragenic insertion; an inversion mutation; an intra-chromosomal inversion; a linking mutation; a linked insertion mutation; an inverted duplication mutation; a tandem duplication; an intrachromosomal tandem duplication; a translocation; a chromosomal translocation, a non-reciprocal translocation; a rearrangement; a genomic rearrangement; a rearrangement of one or more introns, or a fragment thereof; a rearranged intron including a 5'- or a 3'- UTR, or a combination thereof.

41. The method of any of claims 1-40, wherein the tumor nucleic acid molecule includes one or more alterations which are associated with cancer, or predictive of responsiveness to anti-cancer therapeutics.

42. The method of claim 41, wherein the alteration comprises an altered nucleotide sequence, a chromosomal translocation, an intra-chromosomal inversion, a change in copy number, a change in expression level, a change in protein level, a change in protein activity, or a change in methylation status, in a cancer tissue or cancer cell, or encodes an altered amino acid sequence, as compared to a normal, healthy tissue or cell.

43. The method of claim 41 or 42, wherein the alteration is associated, or not associated, with one or more of cancer risk, cancer progression, cancer treatment or resistance to cancer treatment; a genetic risk factor for cancer; a positive treatment response predictor; a negative treatment response predictor; a positive prognostic factor; a negative prognostic factor, or a diagnostic factor.

44. The method of any of claims 1-43, wherein the tumor sample comprises one or more of: premalignant or malignant cells; cells from a solid tumor, a soft tissue tumor or a metastatic lesion; tissue or cells from a surgical margin; a histologically normal tissue; one or more circulating tumor cells (CTC); a normal adjacent tissue (NAT); a blood sample from the same subject having or at risk of having the tumor; an FFPE sample.

45. The method of any of claims 1-44, wherein the methods comprises sequencing of a cDNA derived from an RNA acquired from a tumor sample.

46. The method of claim 45, further comprising performing a DNA sequencing step and an RNA sequencing step.

47. The method of claim 46, further comprising a step chosen from one or more of: detecting a change in the level of a gene or gene product comprising an alteration, enriching a sample for a target RNA, or depleting the sample of certain high abundance RNAs.

48. The method of any of claims 1-47, further comprising one or more of:

(i) fingerprinting the tumor sample;

(ii) quantifying the abundance of a gene or gene product;

(iii) quantifying the relative abundance of a transcript in the tumor sample;

(iv) identifying the tumor sample as belonging to a particular subject;

(v) identifying a genetic trait in the tumor sample comprising one or more of a subject's genetic make-up, ethnicity, race, or familial traits;

(vi) determining the ploidy in the tumor sample;

(vii) determining a loss of heterozygosity in the tumor sample;

(viii) determining the presence or absence of a gene duplication event in the tumor sample;

(ix) determining the presence or absence of a gene amplification event in the tumor sample; or

(x) determining the level of tumor/normal cellular admixture in the tumor sample.

49. The method of any of claims 1-48, wherein the method comprises sequencing a subgenomic interval chosen from at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 or all of the following:

A) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty or more subgenomic intervals from a mutated or wild-type gene or gene product chosen from at least five or more of: ABL1, AKT1, AKT2, AKT3, ALK, APC, AR, BRAF, CCND1, CDK4, CDKN2A, CEBPA, CTNNB1, EGFR, ERBB2, ESR1, FGFR1, FGFR2, FGFR3, FLT3, HRAS, JAK2, KIT, KRAS, MAP2K1, MAP2K2, MET, MLL, MYC, NF1, NOTCH1, NPM1, NRAS, NTRK3, PDGFRA, PIK3CA, PIK3CG, PIK3R1, PTCH1, PTCH2, PTEN, RB1, RET, SMO, STK11, SUFU, or TP53;

B) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty, thirty-five, forty, forty-five, fifty, fifty-five, sixty, sixty-five, seventy, seventy-five, eighty, eighty-five, ninety, ninety-five, one hundred, one hundred and five, one hundred and ten, one hundred and fifteen, one hundred and twenty or more of subgenomic intervals from a mutated or wild type gene or gene product chosen from at least five or more of: ABL2, ARAF, ARFRP1, ARID1A, ATM, ATR, AURKA, AURKB, BAP1, BCL2, BCL2A1, BCL2L1, BCL2L2, BCL6, BRCA1, BRCA2, CBL, CARD11, CBL, CCND2, CCND3, CCNE1, CD79A, CD79B, CDH1, CDH2, CDH20, CDH5, CDK6, CDK8, CDKN2B, CDKN2C, CHEK1, CHEK2, CRKL, CRLF2, DNMT3A, DOT1L, EPHA3, EPHA5, EPHA6, EPHA7, EPHB1, EPHB4, EPHB6, ERBB3, ERBB4, ERG, ETV1, ETV4, ETV5, ETV6, EWSR1, EZH2, FANCA, FBXW7, FGFR4, FLT1, FLT4, FOXP4, GATA1, GNA11, GNAQ, GNAS, GPR124, GUCY1A2, HOXA3, HSP90AA1, IDH1, IDH2, IGF1R, IGF2R, IKBKE, IKZF1, INHBA, IRS2, JAK1, JAK3, JUN, KDM6A, KDR, LRP1B, LRP6, LTK, MAP2K4, MCL1, MDM2, MDM4, MEN1, MITF, MLH1, MPL, MRE11A, MSH2, MSH6, MTOR, MUTYH, MYCL1, MYCN, NF2, NKX2-1, NTRK1, NTRK2, PAK3, PAX5, PDGFRB, PKHD1, PLCG1, PRKDC, PTPN11, PTPRD, RAF1, RARA, RICTOR, RPTOR, RUNX1, SMAD2, SMAD3, SMAD4, SMARCA4, SMARCB1, SOX10, SOX2, SRC, TBX22, TET2, TGFBR2, TMPRSS2, TNFAIP3, TNK, TNKS2, TOP1, TSC1, TSC2, USP9X, VHL, or WT1;

C) at least five, six, seven, eight, nine, ten, fifteen, twenty, or more subgenomic intervals from a gene or gene product according to Table 1, 1A, 2, 3 or 4;

D) at least five, six, seven, eight, nine, ten, fifteen, twenty, or more subgenomic intervals from a gene or gene product chosen from one or more of: ABL1, AKT1, ALK, AR, BRAF,

BRCA1, BRCA2, CEBPA, EGFR, ERBB2, FLT3, JAK2, KIT, KRAS, MET, NPM1, PDGFRA, PIK3CA, RARA, AKT2, AKT3, MAP2K4, NOTCH1, and TP53;

E) at least five, six, seven, eight, nine, ten, or more subgenomic intervals including a mutated or a wild type codon chosen from one or more of: codon 315 of the ABL1 gene; codon 1114, 1338, 1450 or 1556 of APC; codon 600 of BRAF; codon 32, 33, 34, 37, 41 or 45 of CTNNB1; codon 719, 746-750, 768, 790, 858 or 861 of EGFR; codon 835 of FLT3; codon 12, 13, or 61 of HRAS; codon 617 of JAK2; codon 816 of KIT; codon 12, 13, or 61 of KRAS; codon 88, 542, 545, 546, 1047, or 1049 of PIK3CA; codon 130, 173, 233, or 267 of PTEN; codon 918 of RET; codon 175, 245, 248, 273, or 306 of TP53;

F) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty, or more of subgenomic intervals from a mutated or wild type gene or gene product chosen from: ABCB1, BCC2, ABCC4, ABCG2, C1orf144, CYP1B1, CYP2C19, CYP2C8, CYP2D6, CYP3A4, CYP3A5, DPYD, ERCC2, ESR2, FCGR3A, GSTP1, ITPA, LRP2, MAN1B1, MTHFR, NQO1, NRP2, SLC19A1, SLC22A2, SLCO1B3, SOD2, SULT1A1, TPMT, TYMS, UGT1A1, or UMPS;

G) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty, or more of subgenomic intervals from a mutated or wild type PGx gene or gene product associated with one or more of: (i) better survival of a cancer patient treated with a drug; (ii) paclitaxel metabolism; (iii) toxicity to a drug; or (iv) a side effect to a drug;

H) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 110 or more genes or gene products according to Table 3;

J) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 110 or more genes or gene products according to Table 3 in a solid tumor sample from the cancer types specified therein;

K) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more genes or gene products according to Table 4;

L) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more genes or gene products according to Table 4 in a heme tumor sample from the cancer types specified therein;

M) at least five genes or gene products selected from Table 1, 1A-4, wherein an allelic variation is associated with a preselected type of tumor and wherein said allelic variation is present in less than 5% of the cells in said tumor type;

N) at least five genes or gene products selected from Table 1, 1A-4, which are embedded in a GC-rich region; or

O) at least five genes or gene products indicative of a genetic factor for developing cancer chosen from one or more of BRCA1, BRCA2, EGFR, HRAS, KIT, MPL, ALK, PTEN, RET, APC, CDKN2A, MLH1, MSH2, MSH6, NF1, NF2, RB1, TP53, VHL or WT1.

50. The method of any of claims 1-49, wherein the method further comprises providing a report in electronic, web-based, or paper form, to the patient or to another person or entity, a caregiver, a physician, an oncologist, a hospital, clinic, third-party payor, insurance company or government office.

51. The method of claim 50, wherein said report comprises one or more of:

(a) output from the method, comprising the identification of nucleotide values, the indication of presence or absence of an alteration, mutation, or wildtype sequence for subgenomic intervals associated with a tumor of the type of the sample;

(b) information on the role of a sequence, an alteration, mutation, or wildtype sequence, in a disease, wherein said information comprises information on prognosis, resistance, or potential or suggested therapeutic options;

(c) information on the likely effectiveness of a therapeutic option, the acceptability of a therapeutic option, or the advisability of applying the therapeutic option to a patient having a sequence, alteration or mutation identified in the test, and in embodiments, identified in the report;

(d) information, or a recommendation on, the administration of a drug, the administration at a preselected dosage or in a preselected treatment regimen, in combination with other drugs, to the patient;

(e) wherein not all mutations identified in the method are identified in the report, the report can be limited to mutations in genes having a preselected level of correlation with the occurrence, prognosis, stage, or susceptibility of the cancer to treatment, with a preselected therapeutic option; or

(f) is provided to the patient or to another person or entity within 7, 14, or 21 days from receipt of the sample by the entity practicing the method.

52. The method of any of claims 1-51, wherein said contacting step (b) is performed in solution.

53. The method of any of claims 1-52, wherein the method comprises 1, 2, 3, 4 or 5 of:

a) sequencing a first subgenomic interval to provide for about 500X or higher sequencing depth, thereby detecting a mutation present in no more than 5% of the cells from the sample;

b) sequencing a second subgenomic interval to provide for about 200X-500X sequencing depth, thereby detecting a mutation present in no more than 10% of the cells from the sample;

c) sequencing a third subgenomic interval to provide for about 200X-500X sequencing depth, thereby detecting a subgenomic interval chosen from one or more of: a) a pharmacogenomic (PGx) single nucleotide polymorphism (SNP) that distinguishes the ability of a patient to metabolize different drugs, or b) a plurality of genomic SNPs that uniquely identifies a patient;

d) sequencing a fourth subgenomic interval to provide for about 5X-50X sequencing depth, thereby detecting a structural breakpoint; or

e) sequencing a fifth subgenomic interval to provide for about 100X-300X sequencing depth, thereby detecting a copy number change.

54. The method of claim 53, wherein the detection of the copy number change comprises detection of a genomic SNP or locus that is used to assess copy number gains or losses of genomic DNA or loss-of-heterozygosity (LOH).

What is claimed is:

1. A method of analyzing a tumor sample, comprising:

(a) acquiring a library comprising a plurality of tumor members from a tumor sample;

(b) contacting the library with a plurality of bait sets to provide selected members, thereby providing a library catch;

(c) acquiring a read for a subgenomic interval from a tumor member from said library or library catch;

(d) aligning said read; and

(e) assigning a nucleotide value from said read for a preselected nucleotide position for a preselected nucleotide position in each of a plurality of subgenomic intervals,

thereby analyzing said tumor sample,

wherein

each of X nucleotide positions is analyzed under a unique set of conditions for one or a combination of steps (b), (c), (d), or (e) and wherein X is at least 10, 20, 30, 40, 50, 100, 200, 300 or 500, wherein unique means different from the other X-1 sets of conditions.

2. The method of claim 1, wherein:

a first nucleotide position is analyzed with a first set of bait conditions, a first alignment method, and a first mutation calling method;

a second nucleotide position is analyzed with said first set of bait conditions, a second alignment method, and said first mutation calling method; and

a third nucleotide position is analyzed with said first set of bait conditions, said first alignment method, and a second mutation calling method,

to provide three nucleotide positions each analyzed under unique, as compared to the other two, conditions.

3. The method of claim 1, wherein the conditions comprise one or more of those wherein:

a first bait set is used for the first subgenomic interval and a second bait set is used for the second subgenomic interval;

a first alignment method is applied to a read for the first subgenomic interval and a second alignment method is applied to a read for second subgenomic interval; or

a first mutation calling method is applied to a nucleotide position of the first subgenomic interval and a second mutation calling method is applied to a nucleotide position of the second subgenomic interval.

4. The method of claim 1, further comprising one or more of:

nucleotide positions in at least two subgenomic intervals are analyzed with different bait sets,

nucleotide positions in at least two subgenomic intervals are analyzed with different alignment methods; or

nucleotide positions in at least two subgenomic intervals are analyzed with different mutation calling methods.

5. The method of claim 1, wherein:

for each of X nucleotide positions, responsive to a characteristic of a preselected alteration, that occurs at the nucleotide position, the nucleotide position is analyzed under a unique set of conditions (wherein unique means different from the other X-1 sets of conditions and wherein X is at least 10, 20, 30, 40, 50, 100, 200, 300 or 500).

6. The method of claim 5, wherein said characteristic is selected from one or more of:

(i) the type of alteration present in the subgenomic interval;

(ii) sequence in or near said the nucleotide position of the nucleotide position being evaluated comprises a sequence which can affect the expected propensity for misalignment for the subgenomic interval;

(iii) a prior expectation of observing a read showing the alteration in a tumor of preslected type;

(iv) the probability of observing a read showing the alteration due to base-calling error alone;

(v) a preselected depth of sequencing desired for detecting the alteration; or

(vi) the gene, or type of gene, in which the alteration is located, wherein the gene or type of gene can be chosen from one or more of: (a) an oncogene or tumor suppressor, or (b) a gene or type of gene characterized by a preselected alteration.

7. The method of claim 1, wherein said method is performed on a preserved tumor sample under conditions that allow for 95, 98, or 99% sensitivity or specificity for nucleotide positions in at least 10, 20, 50 or 100 subgenomic intervals.

8. A method of analyzing a tumor sample, comprising:

(a) acquiring a library comprising a plurality members from a tumor sample;

(b) enriching the library for preselected sequences by contacting the library with a plurality of bait sets to provide selected members, thereby producing a libray catch;

(c) acquiring a read for a subgenomic interval from a tumor member from said library or library catch by a next generation sequencing method;

(d) aligning said read by an alignment method, and

(e) assigning a nucleotide value from said read for the preselected nucleotide position, thereby analyzing said tumor sample,

wherein

a read from each of X unique subgenomic intervals is aligned with a unique alignment method and X is at least 2,

wherein unique subgenomic interval means different from the other X-1 subgenomic intervals, and wherein unique alignment method means different from the other X-1 alignment methods.

9. The method of claim 8, wherein X is at least 10, 15, 20, 30, 50, 100, 500, or 1,000.

10. The method of claim 8, wherein subgenomic intervals from at least X genes from Table 1 or 1A having the priority 1 annotation, are aligned with unique alignment methods, and X is equal to 10, 15, 20, or 30.

11. The method of claim 8, comprising:

a) applying a first unique alignment method to a first genomic interval, a variant of which is associated with a tumor phenotype, wherein the variant is a point mutation from Table 6;

b) applying a second unique alignment method to a second genomic interval, a variant of which is associated with a tumor phenotype, wherein the variant is a rearrangement chosen from a deletion, insertion, or translocation on Table 5; and

c) applying a third unique alignment method to a third genomic interval comprising a genomic interval in which variants are not associated with a tumor phenotype or with a tumor of the type in said sample.

12. The method of claim 8, wherein a subgenomic interval being analyzed comprises a nucleotide position with a rearrangement, and the method comprises using an alignment method that includes:

selecting a rearrangement reference sequence for alignment with a read, wherein said rearrangement reference sequence is preselected to align with a preselected rearrangement, wherein the reference sequence is not identical to the genomic rearrangement;

comparing a read with said preselected rearrangement reference sequence; and

determining if said read meets a predetermined alignment criterion, wherein the predetermined criterion is an alignment to said preselected rearrangement reference with less than a preselected level of mismatch or gaps;

thereby analyzing a read.

13. The method of claim 8, wherein a first alignment method differs from a second alignment method by a value for a match reward, mismatch penalty, gap penalty, expected threshold, word size, filter, or mask.

14. A method of analyzing a tumor sample, comprising:

(a) acquiring a library comprising a plurality members from a tumor sample;

(b) enriching the library for preselected sequences by contacting the library with a plurality of bait sets to provide selected members, thereby providing a library catch;

(c) acquiring a read for a subgenomic interval from a tumor member from said library or library catch;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for the preselected nucleotide position, thereby analyzing said tumor sample.

wherein a nucleotide value assigned for a nucleotide position in each of X unique subgenomic intervals is assigned by a unique calling method and X is at least 2,

wherein unique subgenomic interval means different from the other X-1 subgenoimc intervals, and wherein unique calling method means different from the other X-1 calling methods.

15. The method of claim 14, wherein the calling methods can differ by relying on different Bayesian prior values chosen from a first calling method applied to a first nucleotide position is a function of a first Bayesian prior and a second calling method applied to a second nucleotide position is a function of a second Bayesian prior.

16. The method of claim 14, wherein assigning said nucleotide value is a function of a value which is or represents the prior (e.g., literature) expectation of observing a read showing a preselected variant, e.g., a mutation, at said preselected nucleotide position in a tumor of type.

17. The method of claim 14, wherein the method comprises assigning a nucleotide value for at least X, where X is 10, 20, 40, 50, 60, 70, 80, 90, or 100, preselected nucleotide positions, wherein each assignment is a function of a unique (as opposed to the value for the other X-1 assignements) value which is or represents the prior expectation of observing a read showing a preselected variant at said preselected nucleotide position in a tumor of type.

18. The method of claim 14, wherein assigning said nucleotide value is a function of a set of values which represent the probabilities of observing a read showing said preselected variant at said preselected nucleotide position if the variant is present in the sample at a frequency of less than 10% and/or if the variant is absent.

19. A method of analyzing the sequence of a tumor sample comprising:

(a) acquiring a plurality of duplicates of a subgenomic interval;

(b) acquiring a read for each of said plurality of duplicates to provide a plurality of duplicate reads;

(c) comparing the nucleotide values at a first nucleotide position in each of said plurality of duplicate reads;

(d) comparing the nucleotide values at second nucleotide position in each of said plurality of duplicate reads,

wherein, for one of the nucleotide positions, each of said plurality of reads does not have the same nucleotide value, and, for the other of said nucleotide positions, each of said plurality of reads has the same nucleotide value;

(e) assigning a first classifier, said first classifier having a quality score or a duplicate-adjusted nucleotide value, to the nucleotide value at the position not having the same nucleotide value in all of said plurality of reads,

(f) assigning a second classifier, said second classified having a quality score or a duplicate-adjusted nucleotide value, to the nucleotide value at the position having the same nucleotide value for each plurality of the reads,

wherein said first classifier indicates a first level of quality or confidence that the nucleotide value to which it is assigned is correct and said second classifier indicates a second level of quality or confidence that the nucleotide value to which it is assigned is correct and said first level is equal to or lower than a preselected criterion.

20. A method of analyzing a tumor sample, comprising:

acquiring a library comprising a plurality of tumor members from a tumor sample; selected members, thereby providing a library catch;

acquiring a read for a subgenomic interval from a tumor member from said library or library catch by a next generation sequencing method;

aligning said read; and

assigning a nucleotide value from said read for a preselected nucleotide position in each of a plurality of subgenomic intervals,

thereby analyzing said tumor sample,

wherein

wherein the method comprises one or 2, 3, 4 or 5 of:

a) sequencing a first subgenomic interval to provide for about 500X or higher sequencing depth, thereby detecting a mutation present in no more than 5% of the cells from the sample;

b) sequencing a second subgenomic interval to provide for about 200X-about 500X sequencing depth, thereby detecting a mutation present in no more than 10% of the cells from the sample;

c) sequencing a third subgenomic interval to provide for about 10-100X sequencing depth, wherein the subgenomic interval is chosen from one or more of: a) a pharmacogenomic (PGx) single nucleotide polymorphism (SNP) that distinguishes the ability of a patient to metabolize different drugs, or b) a genomic SNPs that uniquely identifies a patient;

d) sequencing a fourth subgenomic interval to provide for about 5-50X sequencing depth to detect a structural breakpoint; or

e) sequencing a fifth subgenomic interval to provide for about 100-300X sequencing depth, to detect copy number changes, *e.g.*, to detect a genomic SNPs/loci that is used to assess copy number gains/losses of genomic DNA or loss-of-heterozygosity (LOH).

21. A method of analyzing a tumor sample, comprising:

(a) acquiring a library comprising a plurality of tumor members from a tumor sample;

(b) contacting the library with a bait set to provide selected members, thereby providing a library catch;

(c) acquiring a read for a subgenomic interval from a tumor member from said library or library catch by a next generation sequencing method;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for the preselected nucleotide position, thereby analyzing said tumor sample,

wherein the method comprises contacting the library with at least two, three, four, or five, of bait sets, wherein each bait set of said plurality has a unique (as opposed to the other bait sets in the plurality), preselected efficiency for selection for its target.

22. The method of claim 21, wherein the efficiency of selection of a first bait set in the plurality differs from the efficiency of a second bait set in the plurality by at least 2 fold.

23. In an embodiment, the first and second bait sets provide for a depth of sequencing that differs by at least 2 fold.

24. The method of claim 21, wherein the method comprises contacting one or a plurality of the following bait sets with the library:

a) a first bait set that selects a high-level target chosen from one or more tumor members that comprise a subgenomic interval having an alteration that appears at a frequency of about 5% or less of the cells from the sample harbor the alteration in their genome;

b) a second bait set that selects a mid-level target chosen from one or more tumor members that comprise a subgenomic interval having an alteration that appears at a higher frequency of about 10% of the cells from the sample harbor the alteration in their genome;

c) a third bait set that selects a low-level target chosen from one or more PGx members that comprise a subgenomic interval chosen from one or more of: a) a pharmacogenomic (PGx) single nucleotide polymorphism (SNP) that distinguishes the ability of a patient to metabolize different drugs, b) a genomic SNPs that uniquely identifies a patient, c) a genomic SNPs/loci that is used to assess copy number gains/losses of genomic DNA and loss-of-heterozygosity (LOH);

d) a fourth bait set that selects a member that includes an intron sequence that detects a structural breakpoint; or

e) a fifth bait set that selects a one-copy deletion of several terminal exons.

25. The method of claim 21, wherein the value for efficiency of selection is modified by one or more of:

(i) differential representation of different bait sets;

(ii) differential overlap of bait subsets;

(iii) differential bait parameters;

(iv) mixing of different bait sets at different molar ratios to enhance or reduce relative target coverage depths;

(v) using different types of oligonucleotide baits.

26. The method of claim 25, wherein different types of oligonucleotide baits can be chosen from one or more of:

(a) one or more chemically (non-enzymatically) individually synthesized baits,

(b) one or more baits synthesized in an array,

(c) one or more *in vitro* transcribed baits;

(d) any combination of (a), (b) and/or (c),

(e) one or more naturally or non-naturally occurring DNA oligonucleotide,

(f) one or more naturally or non-naturally occurring RNA oligonucleotide,

(g) a combination of (e) and (f), or

(h) a combination of any of the above.

27. The method of claim 26, wherein the different types of oligonucleotide baits is mixed at a ratio chosen from 1:1, 1:2, 1:3, 1:4, 1:5, 1:10, 1:20, 1:50; 1:100, or 1:1000.

28. The method of claim 27, wherein the ratio of chemically-synthesized bait to array-generated bait is chosen from 1:5, 1:10, or 1:20.

29. The method of claim 26, wherein the non-naturally occurring DNA or RNA oligonucleotides are chosen from one or more of: a locked nucleic acid (LNA); a peptide nucleic acid (PNA); a DNA or RNA oligonucleotide modified to capture low GC regions; a bicyclic nucleic acid (BNA); a crosslinked oligonucleotide; a modified 5-methyl deoxycytidine; or 2,6-diaminopurine.

30. The method of claim 21, wherein one or more of the following bait parameters is modified:

(i) Increasing/decreasing bait representation or overlap can be used to enhance/reduce coverage of tumor members, which are under/over-covered relative to other targets in the same category;

(ii) For low coverage, hard to capture target sequences (including high GC content sequences), targeting with the bait sets covering adjacent sequences;

(iii) Modifying a bait sequence can be made to reduce secondary structure of the bait and enhance its efficiency of selection;

(iv) Modifying a bait length can be used to equalize melting hybridization kinetics of different baits within the same category;

(v) Modifying baits of different orientation for the same target region (forward and reverse strand) may have different binding efficiencies;

(vi) Modifying the amount of a binding entity present on each bait may affect its binding efficiency. Increasing/decreasing the tag level of baits targeting a specific target may be used to enhance/reduce the relative target coverage;

(vii) Modifying the type of nucleotide used for different baits can be altered to affect binding affinity to the target, and enhance/reduce the relative target coverage; or

(viii) Using modified oligonucleotide baits having more stable base pairing, can be used to equalize melting hybridization kinetics between areas of low or normal GC content relative to high GC content.

31. The method of claim 21, wherein the bait set is chosen from one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve or all A-M of the following:

A. A bait set that selects an exon sequence that includes a single nucleotide alteration associated with a cancerous phenotype;

B. A bait set that selects an in-frame deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more codons from a reference nucleotide sequence;

C. A bait set that selects an intragenic deletion;

D. A bait set that selects an intragenic insertion;

E. A bait set that selects a deletion of a full gene;

F. A bait set that selects an inversion;

G. A bait set that selects an interchromosal translocation;

H. A bait set that selects a tandem duplication;

I. A bait set that selects a nucleotide sequence of interest flanked by adjacent non-repetitive sequences;

J. A bait set that selects one or more subgenomic intervals corresponding to a fusion sequence;

K. A bait set that selects a subgenomic interval adjacent to a nucleotide sequence that includes an undesirable feature chosen from a nucleotide sequence of high GC content, or a nucleotide sequence including one or more repeated elements and/or inverted repeats;

L. A bait set that selects a genomic rearrangement that includes an intron sequence from a 5' or 3'-UTR; or

M. A bait set that selects a subgenomic interval that includes an exon adjacent to a cancer associated gene fusion.

32. The method of claim 31, wherein the bait set includes a plurality of oligonucleotides between about 100 nucleotides and 300 nucleotides in length.

33. The method of any of claims 1-32, wherein the subgenomic interval comprises or consists of one or more of: a single nucleotide position; an intragenic region or an intergenic region; an exon or an intron, or a fragment thereof, typically an exon sequence or a fragment thereof; a coding region or a non-coding region, *e.g.*, a promoter, an enhancer, a 5' untranslated region (5' UTR), or a 3' untranslated region (3' UTR), or a fragment thereof; a cDNA or a fragment thereof; an SNP; a somatic mutation, a germ line mutation or both; an alteration, *e.g.*, a point or a single mutation; a deletion mutation; an in-frame deletion, an intragenic deletion, a full gene deletion; an insertion mutation; an intragenic insertion; an inversion mutation; an intra-chromosomal inversion; a linking mutation; a linked insertion mutation; an inverted duplication mutation; a tandem duplication; an intrachromosomal tandem duplication; a translocation; a chromosomal translocation, a non-reciprocal translocation; a rearrangement; a genomic rearrangement; a rearrangement of one or more introns, or a fragment thereof; a rearranged intron including a 5'- or a 3'- UTR), or a combination thereof.

34. The method of any of claims 1-32, wherein the tumor member includes one or more alterations which are associated with cancer, or predictive of responsiveness to anti-cancer therapeutics.

35. The method of claim 34, wherein the alteration comprises an altered nucleotide sequence, an altered amino acid sequence, a chromosomal translocation, an intra-chromosomal

inversion, a change in copy number, a change in expression level, a change in protein level, a change in protein activity, or a change in methylation status, in a cancer tissue or cancer cell, as compared to a normal, healthy tissue or cell.

36. The method of claim 34, wherein the alteration is associated (or not associated) with one or more of cancer risk, cancer progression, cancer treatment or resistance to cancer treatment; a genetic risk factor for cancer; a positive treatment response predictor; a negative treatment response predictor; a positive prognostic factor; a negative prognostic factor, or a diagnostic factor.

37. The method of any of claims 1-32, wherein the tumor sample comprises one or more premalignant or malignant cells; cells from a solid tumor, a soft tissue tumor or a metastatic lesion; tissue or cells from a surgical margin; a histologically normal tissue; one or more circulating tumor cells (CTC); a normal adjacent tissue (NAT); a blood sample from the same subject having or at risk of having the tumor.

38. The method of claim 37, wherein the tumor sample is an FFPE-sample.

39. The method of any of claims 1-32, wherein the methods comprises sequencing of a cDNA derived from an RNA acquired from a tumor sample.

40. The method of claim 39, further comprising performing a DNA sequencing step and an RNA sequencing step.

41. The method of claim 40, further comprising a step chosen from one or more of: detecting a change in the level of a gene or gene product comprising the alteration, enriching a sample for a target RNA, or depleting the sample of certain high abundance RNAs.

42. The method of any of claims 20-32, further comprising one or more of:

(i) fingerprinting the nucleic acid sample;

(ii) quantifying the abundance of a gene or gene product;

(iii) quantifying the relative abundance of a transcript in the sample;

(iv) identifying the nucleic acid sample as belonging to a particular subject;

(v) identifying a genetic trait in the nucleic acid sample comprising one or more of a subject's genetic make-up, ethnicity, race, or familial traits;

(vi) determining the ploidy in the nucleic acid sample;

(vii) determining a loss of heterozygosity in the nucleic acid sample;

(viii) determining the presence or absence of a gene duplication event in the nucleic acid sample;

(ix) determining the presence or absence of a gene amplification event in the nucleic acid sample; or

(x) determining the level of tumor/normal cellular admixture in the nucleic acid sample.

43. The method of any of claims 1-32, wherein the method comprises sequencing a subgenomic interval from at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty or more genes or gene products from the sample, wherein the genes or gene products are chosen from: ABL1, AKT1, AKT2, AKT3, ALK, APC, AR, BRAF, CCND1, CDK4, CDKN2A, CEBPA, CTNNB1, EGFR, ERBB2, ESR1, FGFR1, FGFR2, FGFR3, FLT3, HRAS, JAK2, KIT, KRAS, MAP2K1, MAP2K2, MET, MLL, MYC, NF1, NOTCH1, NPM1, NRAS, NTRK3, PDGFRA, PIK3CA, PIK3CG, PIK3R1, PTCH1, PTCH2, PTEN, RB1, RET, SMO, STK11, SUFU, or TP53.

44. The method of any of claims 1-32, wherein the method comprises sequencing a subgenomic interval chosen from at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 or all of the following:

A) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty or more subgenomic intervals from a mutated or wild-type gene or gene product chosen from at least five or more of: ABL1, AKT1, AKT2, AKT3, ALK, APC, AR, BRAF, CCND1, CDK4, CDKN2A, CEBPA, CTNNB1, EGFR, ERBB2, ESR1, FGFR1, FGFR2, FGFR3, FLT3, HRAS, JAK2, KIT, KRAS, MAP2K1, MAP2K2, MET, MLL, MYC, NF1, NOTCH1, NPM1, NRAS, NTRK3,

PDGFRA, PIK3CA, PIK3CG, PIK3R1, PTCH1, PTCH2, PTEN, RB1, RET, SMO, STK11, SUFU, or TP53;

B) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty, thirty-five, forty, forty-five, fifty, fifty-five, sixty, sixty-five, seventy, seventy-five, eighty, eighty-five, ninety, ninety-five, one hundred, one hundred and five, one hundred and ten, one hundred and fifteen, one hundred and twenty or more of subgenomic intervals from a mutated or wild type gene or gene product chosen from at least five or more of: ABL2, ARAF, ARFRP1, ARID1A, ATM, ATR, AURKA, AURKB, BAP1, BCL2, BCL2A1, BCL2L1, BCL2L2, BCL6, BRCA1, BRCA2, CBL, CARD11, CBL, CCND2, CCND3, CCNE1, CD79A, CD79B, CDH1, CDH2, CDH20, CDH5, CDK6, CDK8, CDKN2B, CDKN2C, CHEK1, CHEK2, CRKL, CRLF2, DNMT3A, DOT1L, EPHA3, EPHA5, EPHA6, EPHA7, EPHB1, EPHB4, EPHB6, ERBB3, ERBB4, ERG, ETV1, ETV4, ETV5, ETV6, EWSR1, EZH2, FANCA, FBXW7, FGFR4, FLT1, FLT4, FOXP4, GATA1, GNA11, GNAQ, GNAS, GPR124, GUCY1A2, HOXA3, HSP90AA1, IDH1, IDH2, IGF1R, IGF2R, IKBKE, IKZF1, INHBA, IRS2, JAK1, JAK3, JUN, KDM6A, KDR, LRP1B, LRP6, LTK, MAP2K4, MCL1, MDM2, MDM4, MEN1, MITF, MLH1, MPL, MRE11A, MSH2, MSH6, MTOR, MUTYH, MYCL1, MYCN, NF2, NKX2-1, NTRK1, NTRK2, PAK3, PAX5, PDGFRB, PKHD1, PLCG1, PRKDC, PTPN11, PTPRD, RAF1, RARA, RICTOR, RPTOR, RUNX1, SMAD2, SMAD3, SMAD4, SMARCA4, SMARCB1, SOX10, SOX2, SRC, TBX22, TET2, TGFBR2, TMPRSS2, TNFAIP3, TNK, TNKS2, TOP1, TSC1, TSC2, USP9X, VHL, or WT1;

C) at least five, six, seven, eight, nine, ten, fifteen, twenty, or more subgenomic intervals from a gene or gene product according to Table 1, 1A, 2, 3 or 4;

D) at least five, six, seven, eight, nine, ten, fifteen, twenty, or more subgenomic intervals from a gene or gene product chosen from one or more of: ABL1, AKT1, ALK, AR, BRAF, BRCA1, BRCA2, CEBPA, EGFR, ERBB2, FLT3, JAK2, KIT, KRAS, MET, NPM1, PDGFRA, PIK3CA, RARA, AKT2, AKT3, MAP2K4, NOTCH1, and TP53;

E) at least five, six, seven, eight, nine, ten, or more subgenomic intervals including a mutated or a wild type codon chosen from one or more of: codon 315 of the ABL1 gene; codon 1114, 1338, 1450 or 1556 of APC; codon 600 of BRAF; codon 32, 33, 34, 37, 41 or 45 of CTNNB1; codon 719, 746-750, 768, 790, 858 or 861 of EGFR; codon 835 of FLT3; codon 12, 13, or 61 of HRAS; codon 617 of JAK2; codon 816 of KIT; codon 12, 13, or 61 of KRAS; codon

88, 542, 545, 546, 1047, or 1049 of PIK3CA; codon 130, 173, 233, or 267 of PTEN; codon 918 of RET; codon 175, 245, 248, 273, or 306 of TP53;

F) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty, or more of subgenomic intervals from a mutated or wild type gene or gene product chosen from: ABCB1, BCC2, ABCC4, ABCG2, C1orf144, CYP1B1, CYP2C19, CYP2C8, CYP2D6, CYP3A4, CYP3A5, DPYD, ERCC2, ESR2, FCGR3A, GSTP1, ITPA, LRP2, MAN1B1, MTHFR, NQO1, NRP2, SLC19A1, SLC22A2, SLCO1B3, SOD2, SULT1A1, TPMT, TYMS, UGT1A1, or UMPS;

G) at least five, six, seven, eight, nine, ten, fifteen, twenty, twenty-five, thirty, or more of subgenomic intervals from a mutated or wild type PGx gene or gene product associated with one or more of: (i) better survival of a cancer patient treated with a drug; (ii) paclitaxel metabolism; (iii) toxicity to a drug; or (iv) a side effect to a drug;

H) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 110 or more genes or gene products according to Table 3;

J) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 110 or more genes or gene products according to Table 3 in a solid tumor sample from the cancer types specified therein;

K) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more genes or gene products according to Table 4;

L) a translocation alteration of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more genes or gene products according to Table 4 in a heme tumor sample from the cancer types specified therein;

M) at least five genes or gene products selected from Table 1, 1A-4, wherein an allelic variation is associated with a preselected type of tumor and wherein said allelic variation is present in less than 5% of the cells in said tumor type;

N) at least five genes or gene products selected from Table 1, 1A-4, which are embedded in a GC-rich region; or

O) at least five genes or gene products indicative of a genetic factor for developing cancer chosen from one or more of BRCA1, BRCA2, EGFR, HRAS, KIT, MPL, ALK, PTEN, RET, APC, CDKN2A, MLH1, MSH2, MSH6, NF1, NF2, RB1, TP53, VHL or WT1.

45. A plurality of bait sets chosen from from one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, or all A-M of the following:

A. A bait set that selects an exon sequence that includes a single nucleotide alteration associated with a cancerous phenotype;

B. A bait set that selects an in-frame deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more codons from a reference nucleotide sequence;

C. A bait set that selects an intragenic deletion;

D. A bait set that selects an intragenic insertion;

E. A bait set that selects a deletion of a full gene;

F. A bait set that selects an inversion;

G. A bait set that selects an interchromosal translocation;

H. A bait set that selects a tandem duplication;

I. A bait set that selects a nucleotide sequence of interest flanked by adjacent non-repetitive sequences;

J. A bait set that selects one or more subgenomic intervals corresponding to a fusion sequence;

K. A bait set that selects a subgenomic interval adjacent to a nucleotide sequence that includes an undesirable feature chosen from a nucleotide sequence of high GC content, or a nucleotide sequence including one or more repeated elements and/or inverted repeats;

L. A bait set that selects a genomic rearrangement that includes an intron sequence from a 5' or 3'-UTR; or

M. A bait set that selects a subgenomic interval that includes an exon adjacent to a cancer associated gene fusion.

46. A method of making the bait sets of claim 45, comprising:

selecting one or more tumor bait oligonucleotide sequences corresponding to the subgenomic intervals;

obtaining a pool of tumor bait oligonucleotide sequences.

47. A method for determining the presence or absence of an alteration associated positively or negatively, with a cancerous phenotype, comprising:

(a) acquiring a library comprising a a plurality of tumor members from a tumor sample;

(b) enriching the library for preselected sequences by contacting the library with a plurality of bait sets to provide selected members;

(c) acquiring a read for a subgenomic interval from a tumor member from said library by a next generation sequencing method;

(d) aligning said read by an alignment method; and

(e) assigning a nucleotide value from said read for the preselected nucleotide position, thereby analyzing said tumor sample,

wherein the method comprises sequencing a subgenomic interval from at least twenty, twenty-five, thirty or more genes or gene products from the sample, wherein the genes or gene products are chosen from: ABL1, AKT1, AKT2, AKT3, ALK, APC, AR, BRAF, CCND1, CDK4, CDKN2A, CEBPA, CTNNB1, EGFR, ERBB2, ESR1, FGFR1, FGFR2, FGFR3, FLT3, HRAS, JAK2, KIT, KRAS, MAP2K1, MAP2K2, MET, MLL, MYC, NF1, NOTCH1, NPM1, NRAS, NTRK3, PDGFRA, PIK3CA, PIK3CG, PIK3R1, PTCH1, PTCH2, PTEN, RB1, RET, SMO, STK11, SUFU, or TP53.

48. The method of any of claims 1-32, wherein the method further comprises providing a report in electronic, web-based, or paper form, to the patient or to another person or entity, a caregiver, a physician, an oncologist, a hospital, clinic, third-party payor, insurance company or government office.

49. The method of claim 48, wherein said report comprises output from the method, comprising the identification of nucleotide values, the indication of presence or absence of an alteration, mutation, or wildtype sequence for sugenomic intervals associated with a tumor of the type of the sample.

50. The method of claim 48, wherein said report comprises information on the role of a sequence, an alteration, mutation, or wildtype sequence, in a disease, wherein said information comprises information on prognosis, resistance, or potential or suggested therapeutic options.

51. The method of claim 48, wherein said report comprises information on the likely effectiveness of a therapeutic option, the acceptability of a therapeutic option, or the advisability of applying the therapeutic option to a patient having a sequence, alteration or mutation identified in the test, and in embodiments, identified in the report.

52. The method of claim 48, wherein said report can include information, or a recommendation on, the administration of a drug, the administration at a preselected dosage or in a preselected treatment regimen, in combination with other drugs, to the patient.

53. The method of claim 48, wherein not all mutations identified in the method are identified in the report, the report can be limited to mutations in genes having a preselected level of correlation with the occurrence, prognosis, stage, or suseptability of the cancer to treatment, with a preselected therapeutic option.

54. The method of claim 48, wherein said report is provided to the patient or to another person or entity within 7, 14, or 21 from receipt of the sample by the entity practicing the method.