CLAIMS

What is claimed:

1. A method of generating a predictive cancer model, comprising:
   (a) sequencing nucleic acid compositions of one or more subjects' biological samples thereby generating one or more sequencing reads;
   (b) filtering the one or more sequencing reads with a human genome database thereby producing one or more filtered sequencing reads;
   (c) generating a plurality of k-mers from the one or more filtered sequencing reads; and
   (d) generating a predictive cancer model by training a predictive model with the plurality of k-mers and corresponding clinical classification of the one or more subjects.

2. The method of claim 1, further comprising determining an abundance of the plurality of k-mers and training the predictive model with the abundance of the plurality of k-mers.

3. The method of claim 1, wherein filtering is performed by exact matching between the one or more sequencing reads and the human reference genome database.

4. The method of claim 3, wherein exact matching comprises computationally filtering of the one or more sequencing reads with the software program Kraken or Kraken2.

5. The method of claim 3, wherein exact matching comprises computationally filtering of the one or more sequencing reads with the software program bowtie 2 or any equivalent thereof.

6. The method of claim 1, further comprising performing in-silico decontamination of the one or more filtered sequencing reads thereby producing one or more decontaminated sequencing reads.

~~7. The method of claim 6, further comprising mapping the one or more decontaminated sequencing reads to a build of a human reference genome database to produce a plurality of mutated human sequence alignments.~~

~~8. The method of claim 7, wherein mapping is performed by bowtie 2 sequence alignment tool or any equivalent thereof.~~

~~9. The method of claim 7, wherein mapping comprises end-to-end alignment, local alignment, or any combination thereof.~~

~~10. The method of claim 7, further comprising identifying cancer mutations in the plurality of mutated human sequence alignments by querying a cancer mutation database.~~

~~11. The method of claim 10, further comprising generating a cancer mutation abundance table with the cancer mutations.~~

~~12. The method of claim 1, wherein the plurality of k-mers comprise non-human k-mers, human mutated k-mers, non-classified DNA k-mers, or any combination thereof.~~

~~13. The method of claim 1, wherein the biological samples comprise a tissue sample, a liquid biopsy sample, or any combination thereof.~~

~~14. The method of claim 1, wherein the one or more subjects are human or non-human mammal.~~

~~15. The method of claim 1, wherein the nucleic acid composition comprises DNA, RNA, cell-free DNA, cell-free RNA, exosomal DNA, exosomal RNA, circulating tumor cell DNA, circulating tumor cell RNA, or any combination thereof.~~

~~16. The method of claim 1, wherein the human reference genome database is GRCh38.~~

~~17. The method of claim 2, wherein an output of the predictive cancer model provides a diagnosis of a presence or an absence of cancer, a cancer body site location, cancer somatic mutations, or any combination thereof associated with the presence or the absence of cancer of a subject.~~

~~18. The method of claim 17, wherein the output of the predictive cancer model comprises an analysis of the cancer somatic mutations, the abundance of the plurality of k-mers, or any combination thereof.~~

~~19. The method of claim 1, wherein the trained predictive model is trained with a set of cancer mutation and k-mer abundances that are known to be present or absent with a characteristic abundance in a cancer of interest.~~

~~20. The method of claim 12, wherein the non-human k-mers originate from the following domains of life: bacterial, archaeal, fungal, viral, or any combination thereof domains of life.~~

~~21. The method of claim 1, wherein the predictive cancer model is configured to determine a presence or lack thereof one or more types of cancer of a subject.~~

~~22. The method of claim 21, wherein the one or more types of cancer are at a low-stage.~~

~~23. The method of claim 22, wherein the low-stage comprises stage I, stage II, or any combination thereof stages of cancer.~~

~~24. The method of claim 1, wherein the predictive cancer model is configured to determine a presence or lack thereof one or more subtypes of cancer in a subject.~~

~~25. The method of claim 1, wherein the predictive cancer model is configured to predict a subject's stage of cancer, cancer prognosis, or any combination thereof.~~

~~26. The method of claim 1, wherein the predictive cancer model is configured to predict a therapeutic response of a subject when administered a therapeutic compound to treat cancer.~~

~~27. The method of claim 1, wherein the predictive cancer model is configured to determine an~~

optimal therapy for a subject.

28. The method of claim 1, wherein the predictive cancer model is configured to longitudinally model a course a subject's one or more cancers' response to a therapy, thereby producing a longitudinal model of the course of the subject's one or more cancers' response to the therapy.

29. The method of claim 28, wherein the predictive cancer model is configured to determine an adjustment to the course of therapy of a subject's one or more cancers based at least in part on the longitudinal model.

30. The method of claim 1, wherein the predictive cancer model is configured to determine the presence or lack thereof: acute myeloid leukemia, adrenocortical carcinoma, bladder urothelial carcinoma, brain lower grade glioma, breast invasive carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, esophageal carcinoma, glioblastoma multiforme, head and neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectum adenocarcinoma, sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumors, thymoma, thyroid carcinoma, uterine carcinosarcoma, uterine corpus endometrial carcinoma, uveal melanoma, or any combination thereof cancer of a subject.

31. The method of claim 6, wherein the in-silico decontamination identifies and removes non-human contaminant features, while retaining other non-human signal features.

32. The method of claim 13, wherein the liquid biopsy comprises: plasma, serum, whole blood, urine, cerebral spinal fluid, saliva, sweat, tears, exhaled breath condensate, or any combination thereof.

33. The method of claim 10, wherein the cancer mutation database is derived from the Catalogue of Somatic Mutations in Cancer (COSMIC), the Cancer Genome Project (CGP), The Cancer Genome Atlas (TGCA), the International Cancer Genome Consortium (ICGC) or any combination thereof.

~~34. The method of claim 2, wherein determining the abundance of the plurality of k-mers is performed by Jellyfish, UCLUST, GenomeTools (Tallymer), KMC2, Gerbil, DSK or any combination thereof.~~

~~35. The method of claim 1, wherein the clinical classification of the one or more subjects comprises healthy, cancerous, non-cancerous disease, or any combination thereof classification.~~

~~36. The method of claim 1, wherein the one or more filtered sequencing reads comprise non-exact matches to a reference human genome, non-human sequencing reads, non-matched non-human sequencing reads, or any combination thereof.~~

~~37. The method of claim 36, wherein the non-matched non-human sequencing reads comprise sequencing reads that do not match to a non-human reference genome database.~~

~~38.~~1. A method of diagnosing cancer of a subject, comprising:

(a) ~~determining~~ receiving a plurality of somatic mutation~~s~~ and non-human nucleic acid molecule k-mer sequences of a sample obtained from a subject~~'s sample~~;

(b) comparing the plurality of somatic mutation~~s~~ and ~~the plurality of~~ non-human nucleic acid molecule k-mer sequences ~~of~~ the subject with a plurality of somatic mutation~~s~~ and non-human nucleic acid molecule k-mer sequences for a given cancer; and

(c) diagnosing cancer of the subject by providing a probability of the ~~presence or lack thereof~~ cancer based at least in part on the comparison of the subject's plurality of somatic mutation~~s~~ and non-human nucleic acid molecule k-mer sequences and the plurality of somatic mutation~~s~~ and non-human nucleic acid molecule k-mer sequences for the given cancer.

~~39.~~2. The method of claim ~~38~~1, ~~wherein~~ further comprising comparing a count of the plurality of somatic mutation nucleic acid molecule k-mer sequence of the subject to a count of the plurality of somatic mutation nucleic acid molecule k-mer sequences of the given cancer. ~~determining the plurality of somatic mutations further comprises counting somatic mutations of the subject's sample.~~

~~40.~~3. The method of claim ~~38~~1, ~~wherein~~ further comprising comparing a count of ~~determining~~ the plurality non-human nucleic acid molecule k-mer sequences ~~comprises counting the non-~~

~~human k-mer sequences of the subject's sample~~of the subject to a count of the plurality of non-human nucleic acid molecule k-mer sequences of the given cancer .

~~41.~~4. The method of claim ~~38~~1, wherein diagnosing the cancer of the subject further comprises determining a category or location of the subject's cancer.

~~42.~~5. The method of claim ~~38~~1, wherein diagnosing the cancer of the subject further comprises determining one or more types of the subject's cancer.

~~43.~~6. The method of claim ~~38~~1, wherein diagnosing the cancer of the subject further comprises determining one or more subtypes of the subject's cancer.

~~44.~~7. The method of claim ~~38~~1, wherein diagnosing the cancer of the subject further comprises determining the stage of the subject's cancer, cancer prognosis, or any combination thereof.

~~45.~~8. The method of claim ~~38~~1, wherein diagnosing the cancer of the subject further comprises determining a type of cancer at a low-stage.

~~46.~~9. The method of claim ~~45~~8, wherein the type of cancer at the low-stage comprises stage I~~,~~ or stage II cancers.

~~47.~~10. The method of claim ~~38~~1, wherein diagnosing the cancer of the subject further comprises determining ~~the~~ a mutation status of the subject's cancer.

~~48.~~11. The method of claim ~~38~~1, ~~wherein~~ further comprising ~~diagnosing the cancer of the subject further comprises~~ determining the subject's response to therapy to treat the subject's cancer.

~~49.~~12. The method of claim ~~38~~1, wherein the given cancer or the subject's cancer comprises: acute myeloid leukemia, adrenocortical carcinoma, bladder urothelial carcinoma, brain lower grade glioma, breast invasive carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, esophageal carcinoma, glioblastoma multiforme, head and neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectum

adenocarcinoma, sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumors, thymoma, thyroid carcinoma, uterine carcinosarcoma, uterine corpus endometrial carcinoma, uvea melanoma, or any combination thereof.

~~50.~~13. The method of claim ~~38~~1, wherein the subject is a non-human mammal.

~~51.~~14. The method of claim ~~38~~1, wherein the subject is a human.

~~52.~~15. The method of claim ~~38~~1, where the subject is mammal.

~~53.~~16. The method of claim ~~38~~1, wherein the plurality of non-human k-mer sequences originate from the following non-mammalian domains of life: viral, bacterial, archaeal, fungal, or any combination thereof.

17. The method of claim 1, wherein the comparing is conducted with a predictive model.

18. The method of claim 17, wherein the predictive model comprises an ensemble of predictive models.

19. The method of claim 1, wherein the sample of the subject comprises a tissue sample, a liquid biopsy sample, or combination thereof.

20. The method of claim 19, wherein the liquid biopsy comprises: plasma, serum, whole blood, urine, cerebral spinal fluid, saliva, sweat, tears, exhaled breath condensate, or any combination thereof.

~~54. A method of generating a predictive cancer model, comprising:~~

~~(a) providing one or more nucleic acid sequencing reads of one or more subjects' biological samples;~~

~~(b) filtering the one or more nucleic acid sequencing reads with a human genome database thereby producing one or more filtered sequencing reads;~~

~~(c) generating a plurality of k-mers from the one or more filtered sequencing reads; and~~

~~(d) generating a predictive cancer model by training a predictive model with the plurality of k-mers and corresponding clinical classification of the one or more subjects.~~

55. The method of claim 54, further comprising determining an abundance of the plurality of k-mers and training the predictive model with the abundance of the plurality of k-mers.

56. The method of claim 54, wherein filtering is performed by exact matching between the one or more nucleic acid sequencing reads and the human reference genome database.

57. The method of claim 56, wherein exact matching comprises computationally filtering of the one or more nucleic acid sequencing reads with the software program Kraken or Kraken2.

58. The method of claim 56, wherein exact matching comprises computationally filtering of the one or more nucleic acid sequencing reads with the software program bowtie 2 or any equivalent thereof.

59. The method of claim 54, further comprising performing in-silico decontamination of the one or more filtered sequencing reads thereby producing one or more decontaminated sequencing reads.

60. The method of claim 59, further comprising mapping the one or more decontaminated sequencing reads to a build of a human reference genome database to produce a plurality of mutated human sequence alignments.

61. The method of claim 60, wherein mapping is performed by bowtie 2 sequence alignment tool or any equivalent thereof.

62. The method of claim 60, wherein mapping comprises end-to-end alignment, local alignment, or any combination thereof.

63. The method of claim 60, further comprising identifying cancer mutations in the plurality of mutated human sequence alignments by querying a cancer mutation database.

64. The method of claim 63, further comprising generating a cancer mutation abundance table with the cancer mutations.

65. The method of claim 54, wherein the plurality of k-mers may comprise non-human k-mers, human mutated k-mers, non-classified DNA k-mers, or any combination thereof.

66. The method of claim 54, wherein the one or more biological samples comprises a tissue sample, a liquid biopsy sample, or any combination thereof.

67. The method of claim 54, wherein the one or more subjects are human or non-human mammal.

~~68. The method of claim 54, wherein the one or more nucleic acid sequencing reads comprise DNA, RNA, cell-free DNA, cell-free RNA, exosomal DNA, exosomal RNA, circulating tumor cell DNA, circulating tumor cell RNA, or any combination thereof.~~

~~69. The method of claim 54, wherein the human reference genome database is GRCh38.~~

~~70. The method of claim 54, wherein an output of the predictive cancer model provides a diagnosis of a presence or an absence of cancer, a cancer body site location, cancer somatic mutations, or any combination thereof associated with the presence or the absence of cancer of a subject.~~

~~71. The method of claim 70, wherein the output of the predictive cancer model comprises an analysis of the cancer somatic mutations, the abundance of the plurality of k-mers, or any combination thereof.~~

~~72. The method of claim 54, wherein the trained predictive model is trained with a set of cancer mutation and k-mer abundances that are known to be present or absent with a characteristic abundance in a cancer of interest.~~

~~73. The method of claim 65, wherein the non-human k-mers originate from the following domains of life: bacterial, archaeal, fungal, viral, or any combination thereof domains of life.~~

~~74. The method of claim 54, wherein the predictive cancer model is configured to determine the presence or lack thereof one or more types of cancer of a subject.~~

~~75. The method of claim 74, wherein the one or more types of cancer are at a low-stage.~~

~~76. The method of claim 75, wherein the low-stage comprises stage I, stage II, or any combination thereof stages of cancer.~~

~~77. The method of claim 54, wherein the predictive cancer model is configured to determine the presence or lack thereof one or more subtypes of cancer of a subject.~~

~~78. The method of claim 54, wherein the predictive cancer model is configured to predict a~~

subject's stage of cancer, cancer prognosis, or any combination thereof.

79. The method of claim 54, wherein the predictive cancer model is configured to predict a therapeutic response of a subject when administered a therapeutic compound to treat cancer.

80. The method of claim 54, wherein the predictive cancer model is configured to determine an optimal therapy for a subject.

81. The method of claim 54, wherein the predictive cancer model is configured to longitudinally model a course of a subject's one or more cancers' response to a therapy, thereby producing a longitudinal model of the course of a subject's one or more cancers' response to the therapy.

82. The method of claim 81, wherein the predictive cancer model is configured to determine an adjustment to the course of therapy of a subject's one or more cancers based at least in part on the longitudinal model.

83. The method of claim 54, wherein the predictive cancer model is configured to determine the presence or lack thereof: acute myeloid leukemia, adrenocortical carcinoma, bladder urothelial carcinoma, brain lower grade glioma, breast invasive carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, esophageal carcinoma, glioblastoma multiforme, head and neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, lymphoid neoplasm diffuse large B cell lymphoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectum adenocarcinoma, sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumors, thymoma, thyroid carcinoma, uterine carcinosarcoma, uterine corpus endometrial carcinoma, uveal melanoma, or any combination thereof cancer of a subject.

84. The method of claim 59, wherein the in-silico decontamination identifies and removes non-human contaminant features, while retaining other non-human signal features.

85. The method of claim 66, wherein the liquid biopsy comprises: plasma, serum, whole blood, urine, cerebral spinal fluid, saliva, sweat, tears, exhaled breath condensate, or any combination

thereof.

86. The method of claim 63, wherein the cancer mutation database is derived from the Catalogue of Somatic Mutations in Cancer (COSMIC), the Cancer Genome Project (CGP), The Cancer Genome Atlas (TGCA), the International Cancer Genome Consortium (ICGC) or any combination thereof.

87. The method of claim 55, wherein determining the abundance of the plurality of k-mers is performed by Jellyfish, UCLUST, GenomeTools (Tallymer), KMC2, Gerbil, DSK, or any combination thereof.

88. The method of claim 54, wherein the clinical classification of the one or more subjects comprises healthy, cancerous, non-cancerous disease, or any combination thereof.

89. The method of claim 54, wherein the one or more filtered sequencing reads comprise non-human sequencing reads, non-matched non-human sequencing reads, or any combination thereof.

90. The method of claim 89, wherein the non-matched non-human sequencing reads comprise sequencing reads that do not match to a non-human reference genome database.

91. A method of diagnosing cancer of a subject using a trained predictive model, comprising:

(a) receiving a plurality of somatic mutations and non-human k-mer sequences of a first one or more subjects' nucleic acid samples;

(b) providing as an input to a trained predictive model the first one or more subjects' plurality of somatic mutations and non-human k-mer sequences, wherein the trained predictive model is trained with a second one or more subjects' plurality of somatic mutation sequences, non-human k-mer sequences, and corresponding clinical classifications of the second one or more subjects, and wherein the first one or more subjects and the second one or more subjects are different subjects; and

(c) diagnosing cancer of the first one or more subjects based at least in part on an output of the trained predictive model.

92. The method of claim 91, wherein receiving the plurality of somatic mutations further

comprises counting somatic mutations of the first one or more subjects' nucleic acid samples.

93. The method of claim 91, wherein receiving the plurality of non-human k-mer sequences comprises counting the non-human k-mer sequences of the first one or more subjects' nucleic acid samples.

94. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining a category or location of the first one or more subjects' cancers.

95. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining one or more types of first one or more subjects' cancers.

96. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining one or more subtypes of the first one or more subjects' cancers.

97. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining the first one or more subjects' stage of cancer, cancer prognosis, or any combination thereof.

98. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining a type of cancer at a low-stage.

99. The method of claim 98, wherein the type of cancer at the low-stage comprises stage I, or stage II cancers.

100. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining the mutation status of the first one or more subjects' cancers.

101. The method of claim 91, wherein diagnosing the cancer of the first one or more subjects further comprises determining the first one or more subjects' response to therapy to treat the first one or more subjects' cancers.

102. The method of claim 91, wherein the cancer comprises: acute myeloid leukemia, adrenocortical carcinoma, bladder urothelial carcinoma, brain lower grade glioma, breast invasive carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, esophageal carcinoma, glioblastoma multiforme, head and neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell

carcinoma, kidney renal papillary cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectum adenocarcinoma, sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumors, thymoma, thyroid carcinoma, uterine carcinosarcoma, uterine corpus endometrial carcinoma, uveal melanoma, or any combination thereof.

103. The method of claim 91, wherein the first one or more subjects and the second one or more subjects are non-human mammal.

104. The method of claim 91, wherein the first one or more subjects and the second one or more subjects are human.

105. The method of claim 91, wherein the first one or more subject and the second one or more subjects are mammal.

106. The method of claim 91, wherein the plurality of non-human k-mer sequences originate from the following non-mammalian domains of life: viral, bacterial, archaeal, fungal, or any combination thereof.

107. A computer-implemented method for utilizing a trained predictive model to determine the presence or lack thereof cancer of one or more subjects, the method comprising:

(a) receiving a plurality of somatic mutations and non-human k-mer sequences of a first one or more subjects' nucleic acid samples;

(b) providing as an input to a trained predictive model the first one or more subjects' plurality of somatic mutations and non-human k-mer sequences, wherein the trained predictive model is trained with a second one or more subjects' plurality of somatic mutation sequences, non-human k-mer sequences, and corresponding clinical classifications of the second one or more subjects, and wherein the first one or more subjects and the second one or more subjects are different subjects; and

(c) determining the presence or lack thereof cancer of the first one or more subjects based at

least in part on an output of the trained predictive model.

108. The computer-implemented method of claim 107, wherein receiving the plurality of somatic mutations further comprises counting somatic mutations of the first one or more subjects' nucleic acid samples.

109. The computer-implemented method of claim 107, wherein receiving the plurality of non-human k-mer sequences comprises counting the non-human k-mer sequences of the first one or more subjects' nucleic acid samples.

110. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining a category or location of the first one or more subjects' cancers.

111. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining one or more types of the first one or more subjects' cancer.

112. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining one or more subtypes of the first one or more subjects' cancers.

113. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining the stage of the cancer, cancer prognosis, or any combination thereof.

114. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining a type of cancer at a low-stage.

115. The computer-implemented method of claim 114, wherein the type of cancer at the low-stage comprises stage I, or stage II cancers.

116. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining the mutation status of the first one or more subjects' cancers.

~~117. The computer-implemented method of claim 107, wherein determining the presence or lack thereof cancer of the first one or more subjects further comprises determining the first one or more subjects' response to a therapy to treat the first one or more subjects' cancers.~~

~~118. The computer-implemented method of claim 107, wherein the cancer comprises: acute myeloid leukemia, adrenocortical carcinoma, bladder urothelial carcinoma, brain lower grade glioma, breast invasive carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, esophageal carcinoma, glioblastoma multiforme, head and neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectum adenocarcinoma, sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumors, thymoma, thyroid carcinoma, uterine carcinosarcoma, uterine corpus endometrial carcinoma, uveal melanoma, or any combination thereof.~~

~~119. The computer-implemented method of claim 107, wherein the first one or more subjects and the second one or more subjects are non-human mammal.~~

~~120. The computer-implemented method of claim 107, wherein the first one or more subjects and the second one or more subjects are human.~~

~~121. The computer-implemented method of claim 107, wherein the first one or more subject and the second one or more subjects are mammal.~~

~~122. The computer-implemented method of claim 107, wherein the plurality of non-human k-mer sequences originate from the following non-mammalian domains of life: viral, bacterial, archaeal, fungal, or any combination thereof.~~