

We claim:

1. A method for diagnosing colon cancer in a sample obtained from a subject, comprising:

5 detecting an expression level of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and
comparing the expression level of the first 634 colon cancer-related nucleic acid molecules, or a decision score derived therefrom, to a control threshold indicative of a diagnosis of colon cancer, wherein the expression level, or a decision score derived
10 therefrom, on the same side of the threshold indicates a diagnosis of colon cancer, thereby diagnosing colon cancer in the sample obtained from the subject.

2. The method of claim 1, wherein the control threshold comprises:

a threshold derived from corresponding transcripts from colon cancer-related nucleic
15 acid molecules listed in Table 6 in a known colon cancer sample (or samples), wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a known colon cancer group indicates a diagnosis of colon cancer.

3. A method for classifying a colon cancer sample, comprising:

20 detecting an expression level of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and
comparing the expression level of the first 634 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of known classification, wherein the expression level, or a decision score derived therefrom, on the
25 same side of the threshold permits classification of the colon cancer sample.

4. The method of claim 3, wherein the control threshold comprises:

a threshold derived from corresponding transcripts from colon cancer-related nucleic
acid molecules listed in Table 6 in a colon cancer sample (or samples) of known
30 classification, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) of known classification permits classification of the colon cancer sample.

5. The method of any one of claims 3 or 4, wherein the colon cancer sample is

classified as stage I, stage II, stage III and stage IV.

6. The method of any one of claims 3-5, further comprising choosing a treatment plan that will be effective for the classified colon cancer.

5

7. The method of claim 6, wherein the treatment is surgical resection, chemotherapy, radiation or any combination thereof.

8. A method for predicting a response to a treatment for colon cancer,
10 comprising:

detecting an expression level of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

comparing the expression level of the first 634 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of a known
15 response to treatment, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates a similar response to treatment, thereby predicting response to treatment.

9. The method of claim 8, wherein the control threshold comprises:
20 a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) having a known response to treatment, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) having a known response to treatment indicates a similar response to treatment, thereby predicting response to
25 treatment.

10. The method of any one of claims 8 or 9, wherein the treatment is surgical resection.

30 11. The method of any one of claims 8-10, wherein the treatment is chemotherapy and/or radiation.

12. A method for predicting long term survival of a subject with colon cancer, comprising:

detecting an expression level of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

comparing the expression level of the the first 634 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of having a history of long term survival, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates long term survival of the subject, thereby predicting long term survival of a subject.

13. The method of claim 12, wherein the control threshold comprises:

a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) obtained from a subject (or subjects) having a history of long term survival, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) obtained from a subject (or subjects) having a history of long term survival indicates long term survival of the subject, thereby predicting long term survival of a subject.

14. The method of claim 13, wherein long term survival comprises at least 5 year survival.

15. A method for predicting of recurrence of colon cancer in a subject, comprising:

detecting an expression level of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

comparing the expression level of the the first 634 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of a history of recurrence, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates a recurrence in the subject.

16. The method of claim 15, wherein the control threshold comprises:

a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) having a history of recurrence, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) of known history of recurrence, indicates a recurrence in the subject.

17. A method of preparing a personalized colon cancer genomics profile for a subject, comprising:

5 detecting an expression level of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and
creating a report summarizing the data obtained by the gene expression analysis.

10 18. The method of any one of claims 1-17, wherein the nucleic acids obtained from the subject comprise RNA and/or cDNA transcribed from RNA extracted from a sample of colorectal tissue obtained from the subject.

19. The method of claim 18, wherein the sample is a biopsy sample.

15 20. The method of any one of claims 18 or 19, wherein the sample is a fixed and/or paraffin embedded sample.

21. The method of any one of claims 1-20, wherein the expression level is normalized against a control gene or genes.

20

22. The method of any one of claims 1-21, wherein the level of expression is determined with PCR and/or or microarray-based methods.

23. The method of any one of claims 1-22, wherein detecting the expression level
25 of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 comprises detecting the expression levels for the transcripts listed in Table 1.

24. The method of any one of claims 1-23, wherein detecting the expression level
of the first 634 colon cancer-related nucleic acid molecules listed in Table 6 comprises
30 detecting the expression levels for the transcripts listed in Table 2.

25. A nucleic acid probe for detecting a gene expression signature for colon cancer, comprising or consisting substantially of a nucleic acid molecule between 20 and 40 nucleotides in length, capable of specifically hybridizing to one of the nucleic acid sequences

set forth as SEQ ID NOs: 1-634 or its complement.

26. The nucleic acid probe according to claim 25, wherein the probe is labeled.

5 27. The nucleic acid probe according to claim 26, wherein the probe is radiolabeled, fluorescently-labeled, biotin-labeled, enzymatically-labeled, or chemically-labeled.

10 28. A set of probes, for detecting a gene expression signature for colon cancer, comprising 2 or more probes, wherein each probe comprises or consists substantially of a nucleic acid molecule between 20 and 40 nucleotides in length capable of specifically hybridizing to one of the nucleic acid sequences set forth as SEQ ID NOs: 1-634 or its complement.

15 29. The set of probes of claim 28, wherein the probes are labeled.

30. The set of probes of claim 29, wherein the probes are radiolabeled, fluorescently-labeled, biotin-labeled, enzymatically-labeled, or chemically-labeled.

20 31. The set of probes of any one of claims 28-29, wherein the set contains at least one probe complementary to each of the transcripts listed in Table 6.

25 32. The set of probes of claim 31, comprising at least one probe complementary to each transcript in a subset of the transcripts listed in Table 6, wherein the subset comprises at least 1%, 5%, 10%, 25%, 50%, 75% or 95% of the transcripts in Table 6.

33. A device for detecting a gene expression signature for colon cancer, comprising a nucleic acid array comprising the set of probes of any one of claims 28-32.

30 34. A pair of primers for the amplification of a gene expression signature for colon cancer nucleic acid, comprising:

a forward primer 15 to 40 nucleotides in length comprising a nucleic acid sequence that specifically hybridizes to any one of the nucleic acid sequences set forth as SEQ ID NOs: 1-634 or its complement; and

a reverse primer 15 to 40 nucleotides in length comprising a nucleic acid sequence that specifically hybridizes to any one of the nucleic acid sequences set forth as SEQ ID NOs: 1-634 or its complement, wherein the set of primers is capable of directing the amplification of the nucleic acid.

5

35. A set of primer pairs for the amplification of a gene expression signature for colon cancer nucleic acid, comprising at least 2 primer pairs suitable for use in amplification of any one of the nucleic acid sequences set forth as SEQ ID NOs: 1-634.

10

36. The set of primer pairs of claim 35, wherein the set of primer pairs comprises primer pairs suitable for use in amplification of a subset of nucleic acid sequences set forth as SEQ ID NOs: 1-634 wherein the subset comprises at least 1%, 5%, 10%, 25%, 50%, 75% or 95% of the nucleic acid sequences set forth as SEQ ID NOs: 1-636.

15

37. A method for preparing a gene expression profile indicative of colon cancer prognosis, comprising:

detecting the expression level of less than 1000 transcripts in a sample comprising RNA isolated from a colon cancer specimen, wherein the first 634 transcripts listed in Table 6 are detected.

20

38. The method of claim 37, wherein from 400 to 800 transcript expression levels are detected.

25

39. The method of claim 37 or 38, wherein from 500 to 700 transcript expression levels are detected.

40. The method of any one of claim 37-39, wherein the transcripts listed in Table 6 comprise the transcripts listed in Table 1.

30

41. The method of any one of claims 37-40, wherein the colon cancer specimen is a formalin fixed paraffin-embedded tissue sample.

42. The method of any one of claims 37-41, further comprising, scoring the expression level of the transcripts listed in Table 6, or a decision score derived therefrom,

against corresponding levels or scores for high risk and low risk patient populations.

43. The method of claim 42, further comprising, selecting adjuvant chemotherapy where the patient is determined to be in the high risk group.

5

44. A method for prognosing colon cancer, comprising:
preparing a gene expression profile for a colon cancer specimen comprising isolated RNA;

and classifying the specimen based the expression levels of the first 634 transcripts
10 listed in Table 6, or a decision score derived therefrom, in a low risk or high risk group.

45. The method of claim 44, wherein the level of less than 800 transcripts are detected in the gene expression profile.

15 46. The method of any one of claims 44-45, wherein the level of less than 700 transcripts are detected in the gene expression profile.

47. The method of any one of claims 44-46, wherein the transcripts listed in Table 6 comprise the transcripts listed in Table 1.

20

48. The method of any one of claims 44-47, wherein the colon cancer specimen is a formalin fixed paraffin-embedded tissue sample.

25 49. The method of any one of claims 44-48, further comprising, selecting adjuvant chemotherapy where the patient is determined to be in the high risk group.

We claim:

1. A method for diagnosing colon cancer in a sample obtained from a subject, comprising:

5 detecting an expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

10 comparing the expression level of the at least 70 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of a diagnosis of colon cancer, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates a diagnosis of colon cancer, thereby diagnosing colon cancer in the sample obtained from the subject.

2. The method of claim 1, wherein the control threshold comprises:

15 a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a known colon cancer sample (or samples), wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a known colon cancer group indicates a diagnosis of colon cancer.

20 3. A method for classifying a colon cancer sample, comprising:

detecting an expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

25 comparing the expression level of the at least 70 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of known classification, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold permits classification of the colon cancer sample.

30 4. The method of claim 3, wherein the control threshold comprises:

a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) of known classification, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) of

known classification permits classification of the colon cancer sample.

5. The method of any one of claims 3 or 4, wherein the colon cancer sample is classified as stage I, stage II, stage III and stage IV.

5

6. The method of any one of claims 3-5, further comprising choosing a treatment plan that will be effective for the classified colon cancer.

10

7. The method of claim 6, wherein the treatment is surgical resection, chemotherapy, radiation or any combination thereof.

8. A method for predicting a response to a treatment for colon cancer, comprising:

15

detecting an expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

20

comparing the expression level of the at least 70 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of a known response to treatment, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates a similar response to treatment, thereby predicting response to treatment.

25

9. The method of claim 8, wherein the control threshold comprises: a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) having a known response to treatment, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) having a known response to treatment indicates a similar response to treatment, thereby predicting response to treatment.

30

10. The method of any one of claims 8 or 9, wherein the treatment is surgical resection.

11. The method of any one of claims 8-10, wherein the treatment is

chemotherapy and/or radiation.

12. A method for predicting long term survival of a subject with colon cancer, comprising:

5 detecting an expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a subject; and

10 comparing the expression level of the at least 70 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of having a history of long term survival, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates long term survival of the subject, thereby predicting long term survival of a subject.

13. The method of claim 12, wherein the control threshold comprises:

15 a threshold derived from corresponding transcripts from colon cancer-related nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) obtained from a subject (or subjects) having a history of long term survival, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold as a colon cancer sample (or samples) obtained from a subject (or subjects)
20 having a history of long term survival indicates long term survival of the subject, thereby predicting long term survival of a subject.

14. The method of claim 13, wherein long term survival comprises at least 5 year survival.

25 15. A method for predicting of recurrence of colon cancer in a subject, comprising:

detecting an expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 in a sample comprising nucleic acids obtained from a
30 subject; and

comparing the expression level of the at least 70 colon cancer-related nucleic acid molecules, or a decision score derived therefrom to a control threshold indicative of a history of recurrence, wherein the expression level, or a decision score derived therefrom, on the same side of the threshold indicates a recurrence in the subject.

16. The method of claim 15, wherein the control threshold comprises:
a threshold derived from corresponding transcripts from colon cancer-related
nucleic acid molecules listed in Table 6 in a colon cancer sample (or samples) having
5 a history of recurrence, wherein the expression level, or a decision score derived
therefrom, on the same side of the threshold as a colon cancer sample (or samples) of
known history of recurrence, indicates a recurrence in the subject.

17. A method of preparing a personalized colon cancer genomics profile
10 for a subject, comprising:
detecting an expression level of at least 70 colon cancer-related nucleic acid
molecules listed in Table 6 in a sample comprising nucleic acids obtained from a
subject; and
creating a report summarizing the data obtained by the gene expression
15 analysis.

18. The method of any one of claims 1-17, wherein the nucleic acids
obtained from the subject comprise RNA and/or cDNA transcribed from RNA
extracted from a sample of colorectal tissue obtained from the subject.

19. The method of claim 18, wherein the sample is a biopsy sample.

20. The method of any one of claims 18 or 19, wherein the sample is a
fixed and/or paraffin embedded sample.

21. The method of any one of claims 1-20, wherein the expression level is
normalized against a control gene or genes.

22. The method of any one of claims 1-21, wherein the level of expression
30 is determined with PCR and/or or microarray-based methods.

23. The method of any one of claims 1-22, wherein detecting the
expression level of at least 70 colon cancer-related nucleic acid molecules listed in
Table 6 comprises detecting the expression levels for MUM1 and SIGMAR1

transcripts.

24. The method of any one of claims 1-23, wherein detecting the expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 comprises detecting the expression levels for MUM1, SIGMAR1, ARSD, SULT1C2 and PPFIBP1 transcripts.

25. The method of any one of claims 1-24, wherein detecting the expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 comprises detecting the expression levels for ARSD, CXCL9, PCLO, SLC2A3, FCGBP, SLC2A14, SLC2A3, BCL9L and antisense sequences of MUC3A, OLFM4 and RNF39 transcripts.

26. The method of any one of claims 1-25, wherein detecting the expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 comprises detecting the expression levels for the transcripts listed in Table 1.

27. The method of any one of claims 1-26, wherein detecting the expression level of at least 70 colon cancer-related nucleic acid molecules listed in Table 6 comprises detecting the expression levels for the transcripts listed in Table 2.

28. A nucleic acid probe for detecting a gene expression signature for colon cancer, comprising or consisting substantially of a nucleic acid molecule between 20 and 40 nucleotides in length, capable of specifically hybridizing to one of the nucleic acid sequences set forth as SEQ ID NOs: 1-636 or its complement.

29. The nucleic acid probe according to claim 28, wherein the probe is labeled.

30. The nucleic acid probe according to claim 29, wherein the probe is radiolabeled, fluorescently-labeled, biotin-labeled, enzymatically-labeled, or chemically-labeled.

31. A set of probes, for detecting a gene expression signature for colon

cancer, comprising 2 or more probes, wherein each probe comprises or consists substantially of a nucleic acid molecule between 20 and 40 nucleotides in length capable of specifically hybridizing to one of the nucleic acid sequences set forth as SEQ ID NOs: 1-636 or its complement.

5

32. The set of probes of claim 31, wherein the probes are labeled.

33. The set of probes of claim 32, wherein the probes are radiolabeled, fluorescently-labeled, biotin-labeled, enzymatically-labeled, or chemically-labeled.

10

34. The set of probes of any one of claims 31-33, wherein the set contains at least one probe complementary to each of the transcripts listed in Table 6.

15

35. The set of probes of claim 34, comprising at least one probe complementary to each transcript in a subset of the transcripts listed in Table 6, wherein the subset comprises at least 1%, 5%, 10%, 25%, 50%, 75% or 95% of the transcripts in Table 6.

20

36. A device for detecting a gene expression signature for colon cancer, comprising a nucleic acid array comprising the set of probes of any one of claims 31-35.

25

37. A pair of primers for the amplification of a gene expression signature for colon cancer nucleic acid, comprising:

a forward primer 15 to 40 nucleotides in length comprising a nucleic acid sequence that specifically hybridizes to any one of the nucleic acid sequences set forth as SEQ ID NOs: 1-636 or its complement; and

30

a reverse primer 15 to 40 nucleotides in length comprising a nucleic acid sequence that specifically hybridizes to any one of the nucleic acid sequences set forth as SEQ ID NOs: 1-636 or its complement, wherein the set of primers is capable of directing the amplification of the nucleic acid.

38. A set of primer pairs for the amplification of a gene expression signature for colon cancer nucleic acid, comprising at least 2 primer pairs suitable for

use in amplification of any one of the nucleic acid sequences set forth as SEQ ID NOs: 1-636.

39. The set of primer pairs of claim 38, wherein the set of primer pairs comprises primer pairs suitable for use in amplification of a subset of nucleic acid sequences set forth as SEQ ID NOs: 1-636 wherein the subset comprises at least 1%, 5%, 10%, 25%, 50%, 75% or 95% of the nucleic acid sequences set forth as SEQ ID NOs: 1-636.

40. A method for preparing a gene expression profile indicative of colon cancer prognosis, comprising:

detecting the expression level of less than 1000 transcripts in a sample comprising RNA isolated from a colon cancer specimen, wherein at least 50 transcripts listed in Table 6 are detected.

41. The method of claim 40, wherein from 400 to 800 transcript expression levels are detected.

42. The method of claim 40 or 41, wherein from 500 to 700 transcript expression levels are detected.

43. The method of any one of claims 40-42, wherein at least 100 transcripts from Table 6 are detected.

44. The method of any one of claims 40-43, wherein at least 200 transcripts from Table 6 are detected.

45. The method of any one of claims 40-44, wherein at least 300 transcripts from Table 6 are detected.

46. The method of any one of claims 40-45, wherein at least 400 transcripts from Table 6 are detected.

47. The method of any one of claims 40-46, wherein at least 500

transcripts from Table 6 are detected.

48. The method of any one of claims 40-47, wherein at least 600 transcripts from Table 6 are detected.

5

49. The method of any one of claims 40-48, wherein all transcripts listed in Table 6 are detected.

50. The method of any one of claim 40-49, wherein the transcripts listed in Table 6 comprise the transcripts listed in Table 1.

10

51. The method of any one of claims 40-50, wherein the colon cancer specimen is a formalin fixed paraffin-embedded tissue sample.

15

52. The method of any one of claims 40-51, further comprising, scoring the expression level of the transcripts listed in Table 6, or a decision score derived therefrom, against corresponding levels or scores for high risk and low risk patient populations.

20

53. The method of claim 52, further comprising, selecting adjuvant chemotherapy where the patient is determined to be in the high risk group.

25

54. A method for prognosing colon cancer, comprising:
preparing a gene expression profile for a colon cancer specimen comprising isolated RNA;
and classifying the specimen based the expression levels of at least 50 transcripts listed in Table 6, or a decision score derived therefrom, in a low risk or high risk group.

30

55. The method of claim 54, wherein the level of less than 1000 transcripts are detected in the gene expression profile.

56. The method of claim 54 or 55, wherein the level of less than 800 transcripts are detected in the gene expression profile.

57. The method of any one of claims 54-56, wherein the level of less than 700 transcripts are detected in the gene expression profile.

5 58. The method of any one of claims 54-57, wherein the specimen is classified based on the expression level of at least 100 transcripts from Table 6.

10 59. The method of any one of claims 54-58, wherein the specimen is classified based on the expression level of at least 200 transcripts from Table 6.

60. The method of any one of claims 54-59, wherein the specimen is classified based on the expression level of at least 300 transcripts from Table 6.

15 61. The method of any one of claims 54-60, wherein the specimen is classified based on the expression level of at least 400 transcripts from Table 6.

62. The method of any one of claims 54-61, wherein the specimen is classified based on the expression level of at least 500 transcripts from Table 6.

20 63. The method of any one of claims 54-62, wherein the specimen is classified based on the expression level of at least 600 transcripts from Table 6.

25 64. The method of any one of claims 54-63, wherein the expression level of all transcripts listed in Table 6 are used to classify the specimen.

65. The method of any one of claims 54-64, wherein the transcripts listed in Table 6 comprise the transcripts listed in Table 1.

30 66. The method of any one of claims 54-65, wherein the colon cancer specimen is a formalin fixed paraffin-embedded tissue sample.

67. The method of any one of claims 54-66, further comprising, selecting adjuvant chemotherapy where the patient is determined to be in the high risk group.