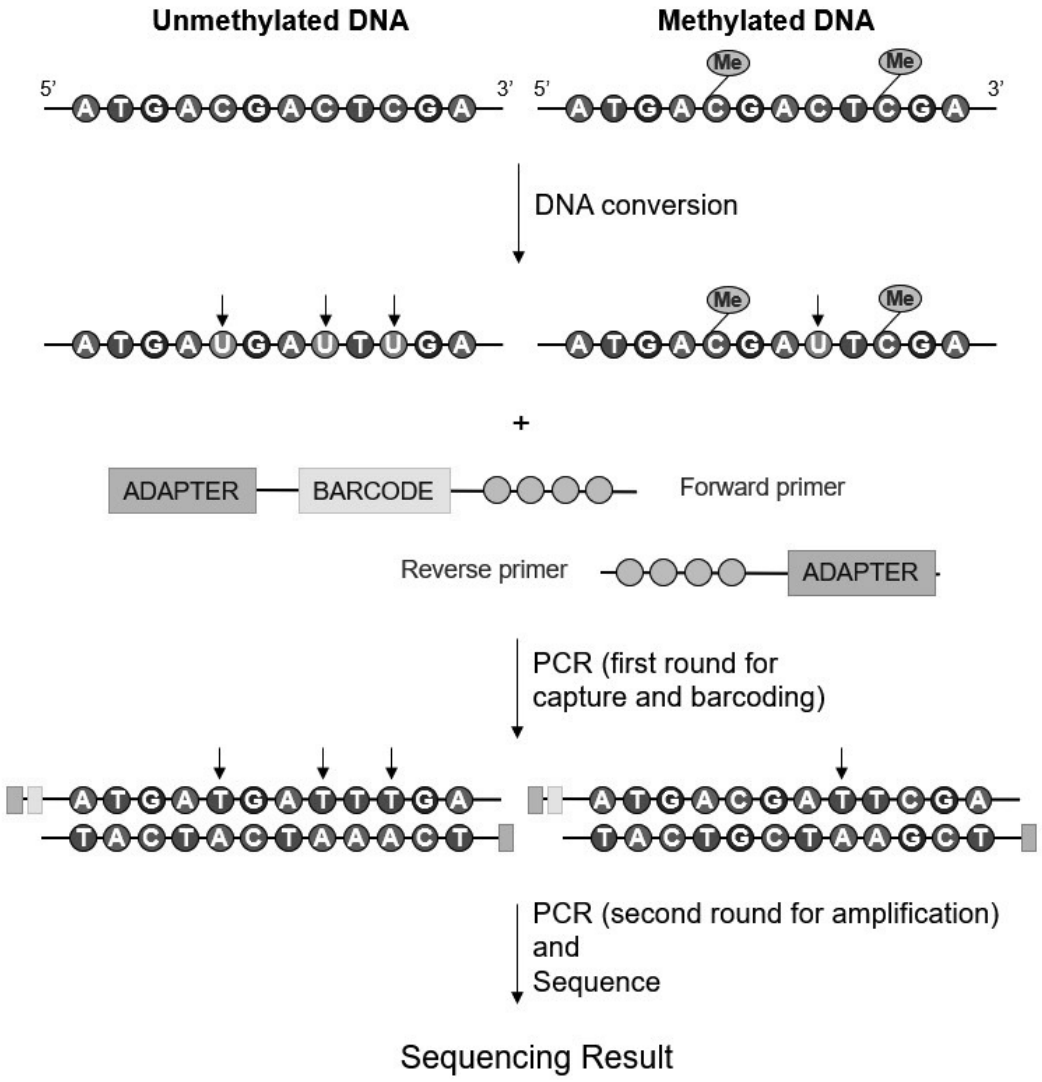

Unmethylated DNA
Methylated DNA
5'
3'
Me
DNA conversion
+
ADAPTER
BARCODE
Forward primer
Reverse primer
ADAPTER
PCR (first round for capture and barcoding)
PCR (second round for amplification) and Sequence
Sequencing Result

Figure 1

CLIP4_methyl_2F *ACACGACGCTCTTCCGATCT*NNNNNNNNNN
GGGTT**Y**GGAGAGGTGTG

CLIP4_methyl_2R *GACGTGTGCTCTTCCGATCT* CCTACCTCCTCC**R**AATCCTC

Indexed Illumina P5 Primers
F-1:
AATGATACGGCGACCACCGAGATCTACACCTAGCGCTACACTCTTTCCCTAC
ACGACGCTCTTCCGATC*T

Indexed Illumina P7 Primers
R-1:
CAAGCAGAAGACGGCATACGAGATAACCGCGGGTGACTGGAGTTCAGACGT
GTGCTCTTCCGATC*T

Figure 2

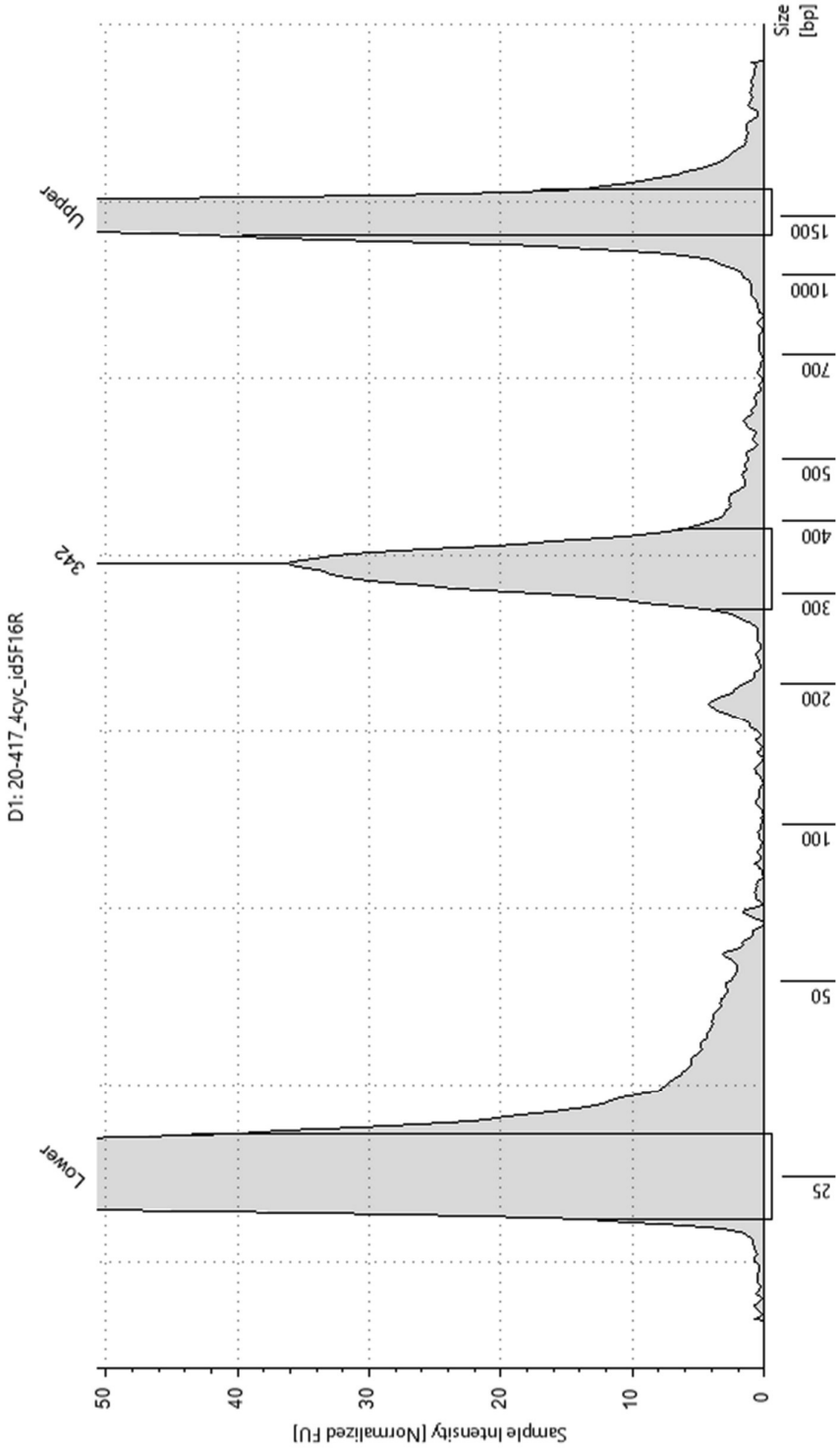

Sample Intensity [Normalized FU]
0
10
20
30
40
50
Size [bp]
25
50
100
200
300
400
500
700
1000
1500
Lower
342
Upper
D1: 20-417_4cyc_id5F16R

Figure 3

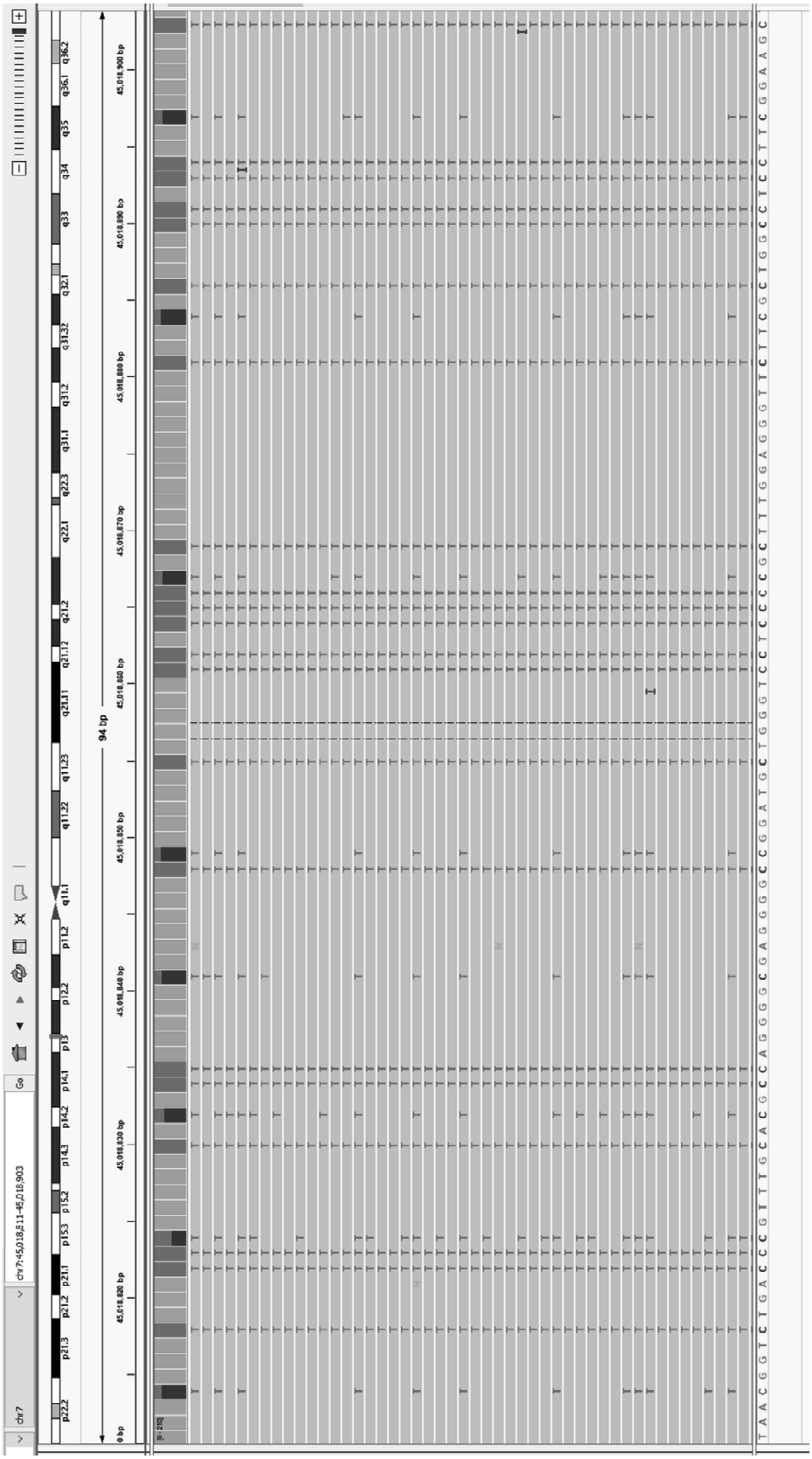


Figure 4 (a)

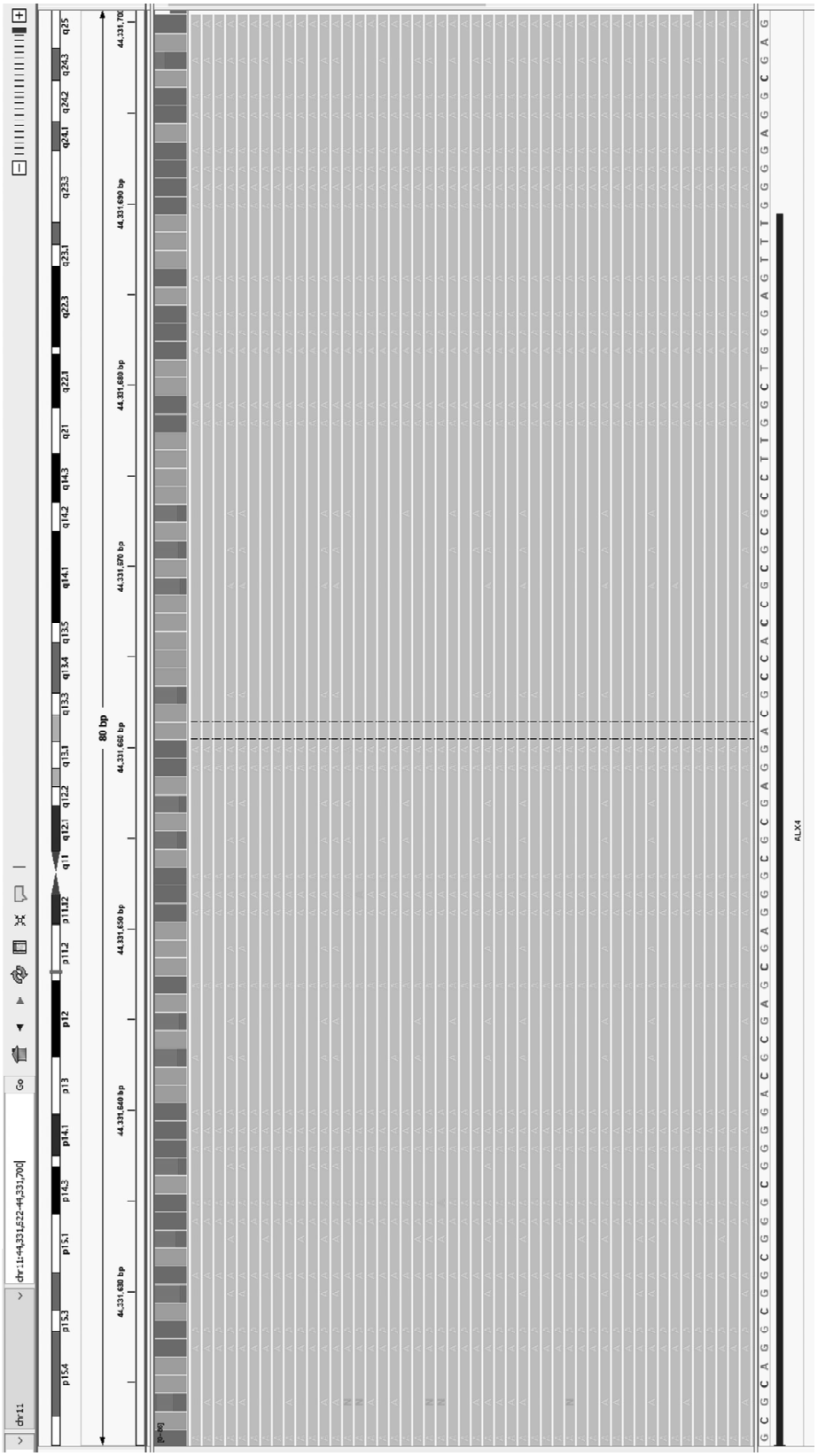


Figure 4 (b)

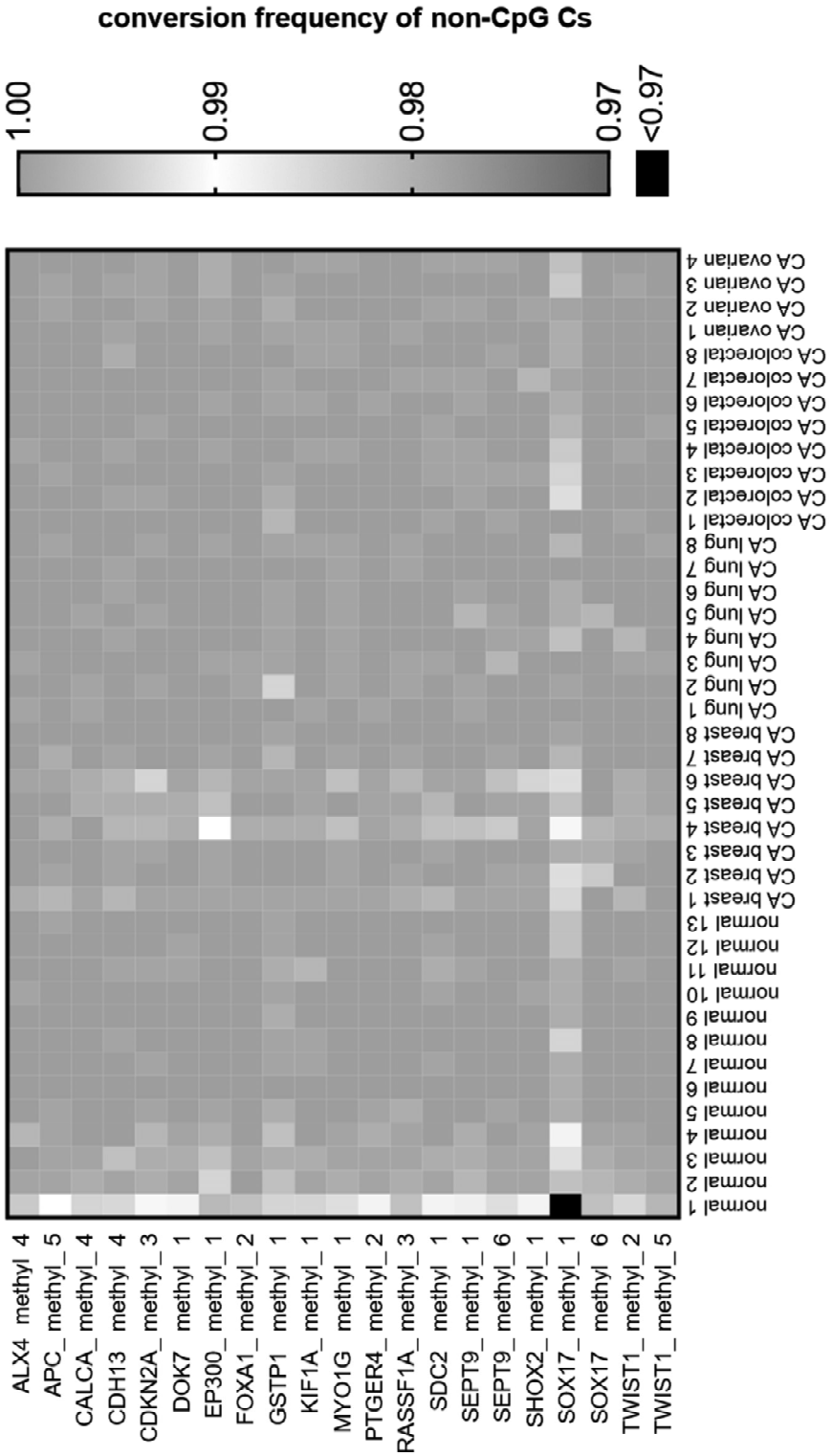

conversion frequency of non-CpG Cs
1.00
0.99
0.98
0.97
<0.97
normal 1
normal 2
normal 3
normal 4
normal 5
normal 6
normal 7
normal 8
normal 9
normal 10
normal 11
normal 12
normal 13
CA breast 1
CA breast 2
CA breast 3
CA breast 4
CA breast 5
CA breast 6
CA breast 7
CA breast 8
CA lung 1
CA lung 2
CA lung 3
CA lung 4
CA lung 5
CA lung 6
CA lung 7
CA lung 8
CA colorectal 1
CA colorectal 2
CA colorectal 3
CA colorectal 4
CA colorectal 5
CA colorectal 6
CA colorectal 7
CA colorectal 8
CA ovarian 1
CA ovarian 2
CA ovarian 3
CA ovarian 4
ALX4 methyl 4
APC_ methyl_5
CALCA_ methyl_4
CDH13 methyl 4
CDKN2A_ methyl_3
DOK7 methyl 1
EP300_ methyl_1
FOXA1_ methyl_2
GSTP1 methyl 1
KIF1A_ methyl_1
MYO1G methyl 1
PTGER4_ methyl_2
RASSF1A_ methyl_3
SDC2 methyl 1
SEPT9_ methyl_1
SEPT9_ methyl_6
SHOX2_ methyl_1
SOX17_ methyl_1
SOX17 methyl 6
TWIST1_ methyl_2
TWIST1_ methyl_5

Figure 5

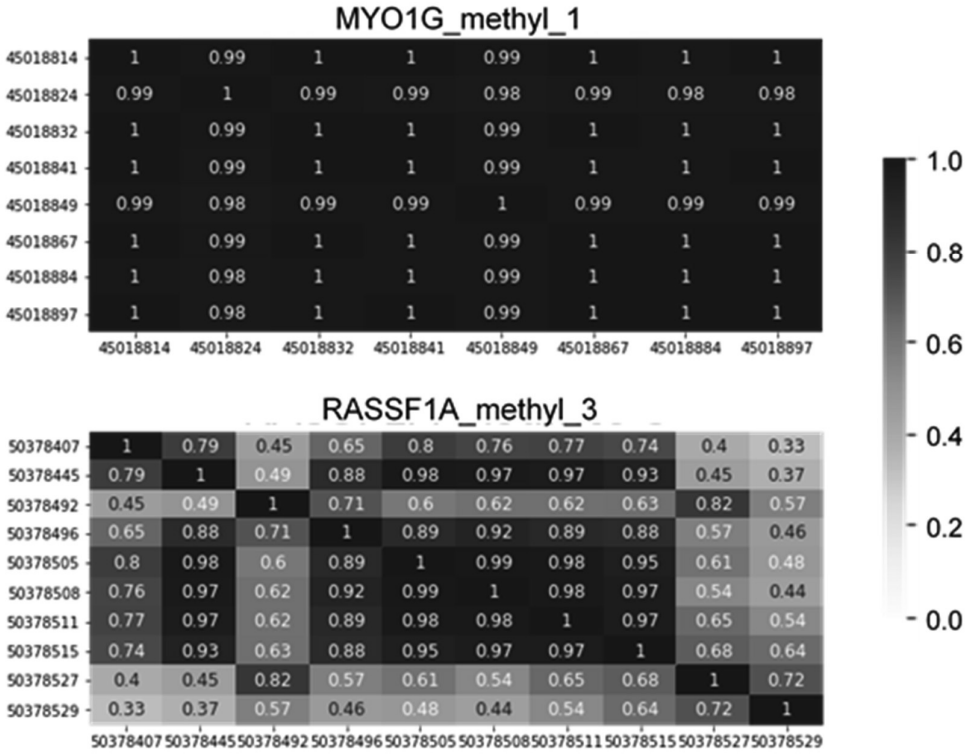

MYO1G_methyl_1
45018814 1 0.99 1 1 0.99 1 1 1
45018824 0.99 1 0.99 0.99 0.98 0.99 0.98 0.98
45018832 1 0.99 1 1 0.99 1 1 1
45018841 1 0.99 1 1 0.99 1 1 1
45018849 0.99 0.98 0.99 0.99 1 0.99 0.99 0.99
45018867 1 0.99 1 1 0.99 1 1 1
45018884 1 0.98 1 1 0.99 1 1 1
45018897 1 0.98 1 1 0.99 1 1 1
45018814 45018824 45018832 45018841 45018849 45018867 45018884 45018897
RASSF1A_methyl_3
50378407 1 0.79 0.45 0.65 0.8 0.76 0.77 0.74 0.4 0.33
50378445 0.79 1 0.49 0.88 0.98 0.97 0.97 0.93 0.45 0.37
50378492 0.45 0.49 1 0.71 0.6 0.62 0.62 0.63 0.82 0.57
50378496 0.65 0.88 0.71 1 0.89 0.92 0.89 0.88 0.57 0.46
50378505 0.8 0.98 0.6 0.89 1 0.99 0.98 0.95 0.61 0.48
50378508 0.76 0.97 0.62 0.92 0.99 1 0.98 0.97 0.54 0.44
50378511 0.77 0.97 0.62 0.89 0.98 0.98 1 0.97 0.65 0.54
50378515 0.74 0.93 0.63 0.88 0.95 0.97 0.97 1 0.68 0.64
50378527 0.4 0.45 0.82 0.57 0.61 0.54 0.65 0.68 1 0.72
50378529 0.33 0.37 0.57 0.46 0.48 0.44 0.54 0.64 0.72 1
50378407 50378445 50378492 50378496 50378505 50378508 50378511 50378515 50378527 50378529
1.0
0.8
0.6
0.4
0.2
0.0

Figure 6

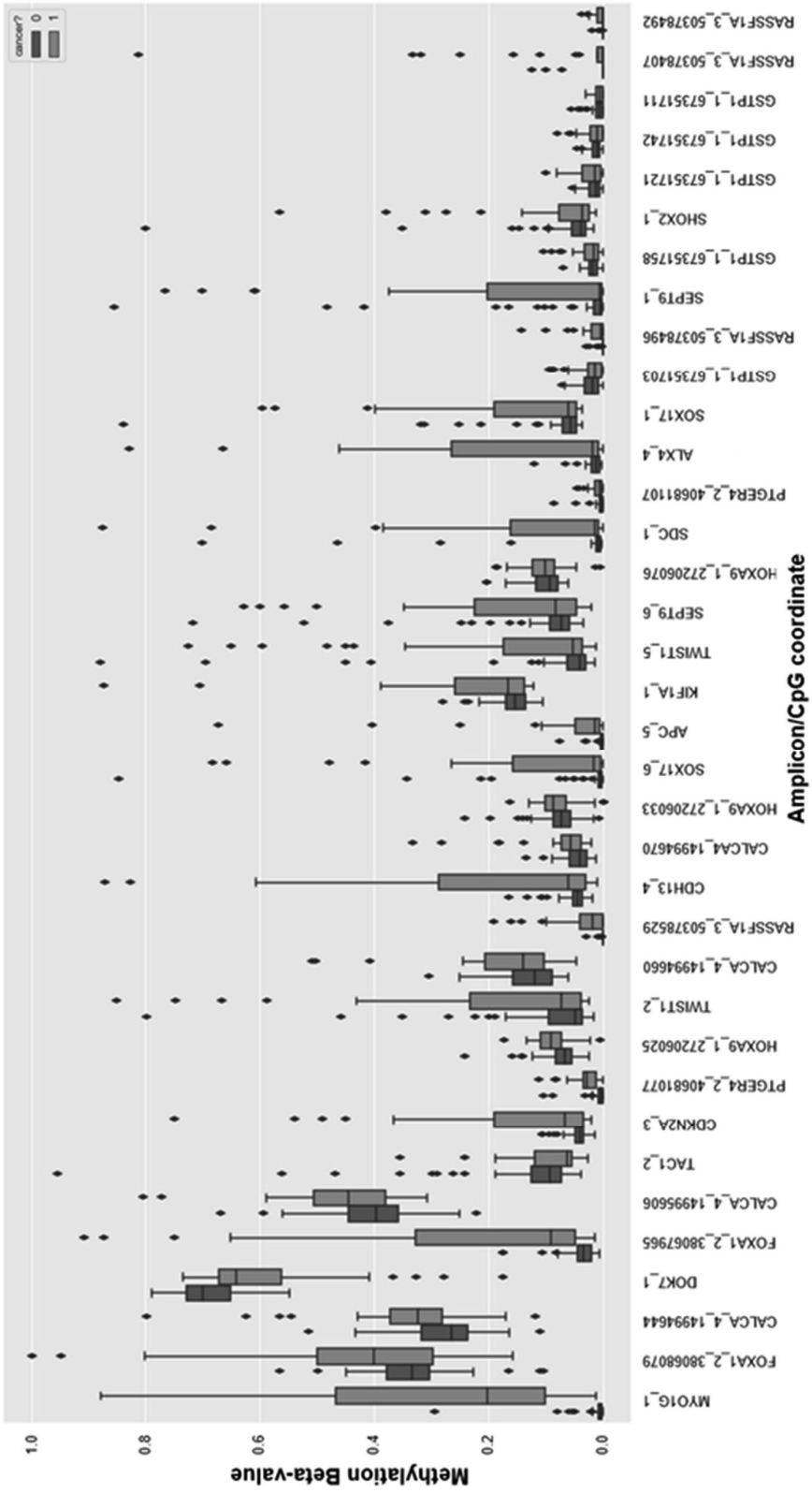

cancer?
0
1
Methylation Beta-value
1.0
0.8
0.6
0.4
0.2
0.0
Amplicon/CpG coordinate
MYO1G_1
FOXA1_2_38068079
CALCA_4_14994644
DOK7_1
FOXA1_2_38067965
CALCA_4_14995606
TAC1_2
CDKN2A_3
PTGER4_2_40681077
HOXA9_1_27206025
TWIST1_2
CALCA_4_14994660
RASSF1A_3_50378529
CDH13_4
CALCA4_14994670
HOXA9_1_27206033
SOX17_6
APC_5
KIF1A_1
TWIST1_5
SEPT9_6
HOXA9_1_27206076
SDC_1
PTGER4_2_40681107
ALX4_4
SOX17_1
GSTP1_1_67351703
RASSF1A_3_50378496
SEPT9_1
GSTP1_1_67351758
SHOX2_1
GSTP1_1_67351721
GSTP1_1_67351742
GSTP1_1_67351711
RASSF1A_3_50378407
RASSF1A_3_50378492

Figure 7

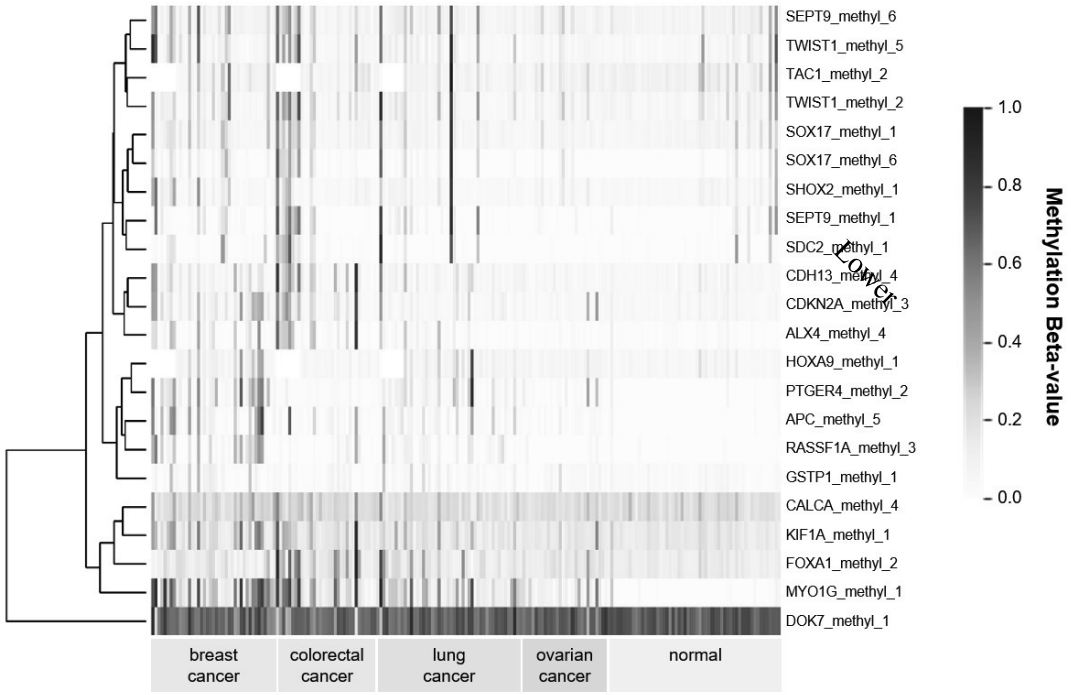
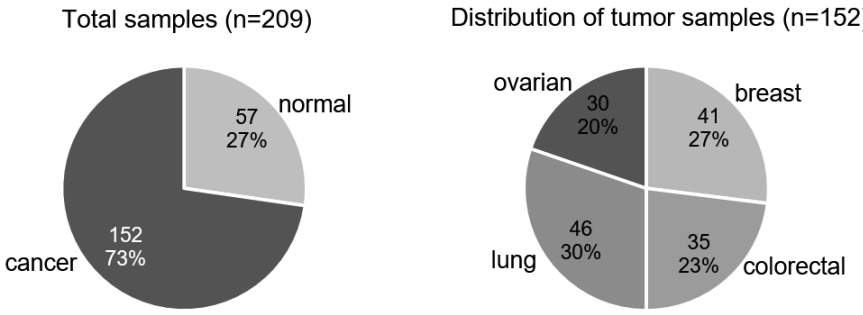

SEPT9_methyl_6
TWIST1_methyl_5
TAC1_methyl_2
TWIST1_methyl_2
SOX17_methyl_1
SOX17_methyl_6
SHOX2_methyl_1
SEPT9_methyl_1
SDC2_methyl_1
CDH13_methyl_4
CDKN2A_methyl_3
ALX4_methyl_4
HOXA9_methyl_1
PTGER4_methyl_2
APC_methyl_5
RASSF1A_methyl_3
GSTP1_methyl_1
CALCA_methyl_4
KIF1A_methyl_1
FOXA1_methyl_2
MYO1G_methyl_1
DOK7_methyl_1
breast cancer
colorectal cancer
lung cancer
ovarian cancer
normal
1.0
0.8
0.6
0.4
0.2
0.0
Methylation Beta-value

Figure 8



Samples used as training set	Logistic regression model for cancer prediction	Random forest model for cancer type identification
All samples (n=209)	Sensitivity: 94.1% Specificity: 87.7%	Sensitivity: 58% Specificity: 90%
Normal and colorectal cancer only (n=92)	Sensitivity: 93.9% Specificity: 96.5%	Sensitivity: 68% Specificity: 88%

Figure 9