

Supplementary Information for Estécio et al., “Genome Architecture Marked by Retrotransposons Modulates Predisposition to DNA Methylation in Cancer”

Supplemental Figure Legends

Figure S1. Graphic representation of the presence of **(A)** SINE and **(B)** LINE retrotransposons in each 1-kb bin from transcription start site (TSS) of methylation-prone and methylation-resistant genes.

Figure S2. Validation of the predictive method in a 28 cancer cell lines and 32 primary cancer tissues. Methylation status of more than 6,600 gene promoters was determined by MCAM (methylated CpG island amplification microarray). The measured DNA methylation was significantly higher in predicted methylation-prone genes than in predicted methylation-resistant and methylation-intermediate genes in all cell lines except UC15. Methylation is presented as mean $\pm$ SEM of all genes of a predicted methylation group. MB435, breast cancer; H322, lung cancer; HEL, JURKAT, K562, KG1, KG1a, MOLT4, RAJI, TF1, BJAB, CEM and JTAG, leukemia; RKO, SW48 and KM12, colon cancer; all remaining, bladder cancer. Samples 68-1 to 68-4 were collected from acute myelogenous leukemia patients; all remaining primary samples were collected from bladder cancer patients.

Figure S3. Distribution of SINE, LINE and LTR repeats around the TSS of frequently hypermethylated and unmethylated genes identified by MCAM analysis of 28 cancer cell lines and 32 primary cancer tissues. The genome-wide distribution of these elements in CpG island promoter genes is shown in gray.

Figure S4. Frequency of predicted methylation-prone, methylation-resistant and methylation-intermediate genes in the X-chromosome. The observed frequencies are similar to those observed in autosomal chromosomes and are not significantly different between randomly inactivated and genes that escape inactivation.

Figure S5. Abundance of SINE and LINE retrotransposons in the 20-kb sequence window centered in the TSS of 291 methylation-prone and 740 methylation-resistant non-promoter CpG island (exonic, intronic and intergenic CpG island found to be methylation-prone or methylation-resistant in our MCAM database).

Figure S6. Frequency of binding of H3K27me3 (repressive histone mark) in human embryonic stem cells to predicted 2583 methylation-resistant, 3655 methylation-intermediate and 1690 methylation-prone genes for which data were available. Note that genes predicted methylation-prone (thus depleted for SINE and LINE retrotransposons) are preferential targets of H3K27me3. **(B)** Abundance of SINE and LINE retrotransposon in a 20-kb sequence window centered in the gene TSS of CpG island promoters bound or not bound by the PcG protein SUZ12 in ES cells.

Figure S7. (A) Measured promoter CpG island methylation by MCAM of predicted methylation-prone (PatMan+) and methylation-resistant (PatMan-) genes according the PatMan classifier. **(B)** Performance of CTCF binding and retrotransposon marking in prediction promoter CpG methylation status in cancer. CTCF+, bound genes; CTCF-, unbound genes. A gene was considered bound by CTCF if ChIP-Seq data from public databases show binding of this protein in the 2-kb region centered in the gene TSS.

Figure S1

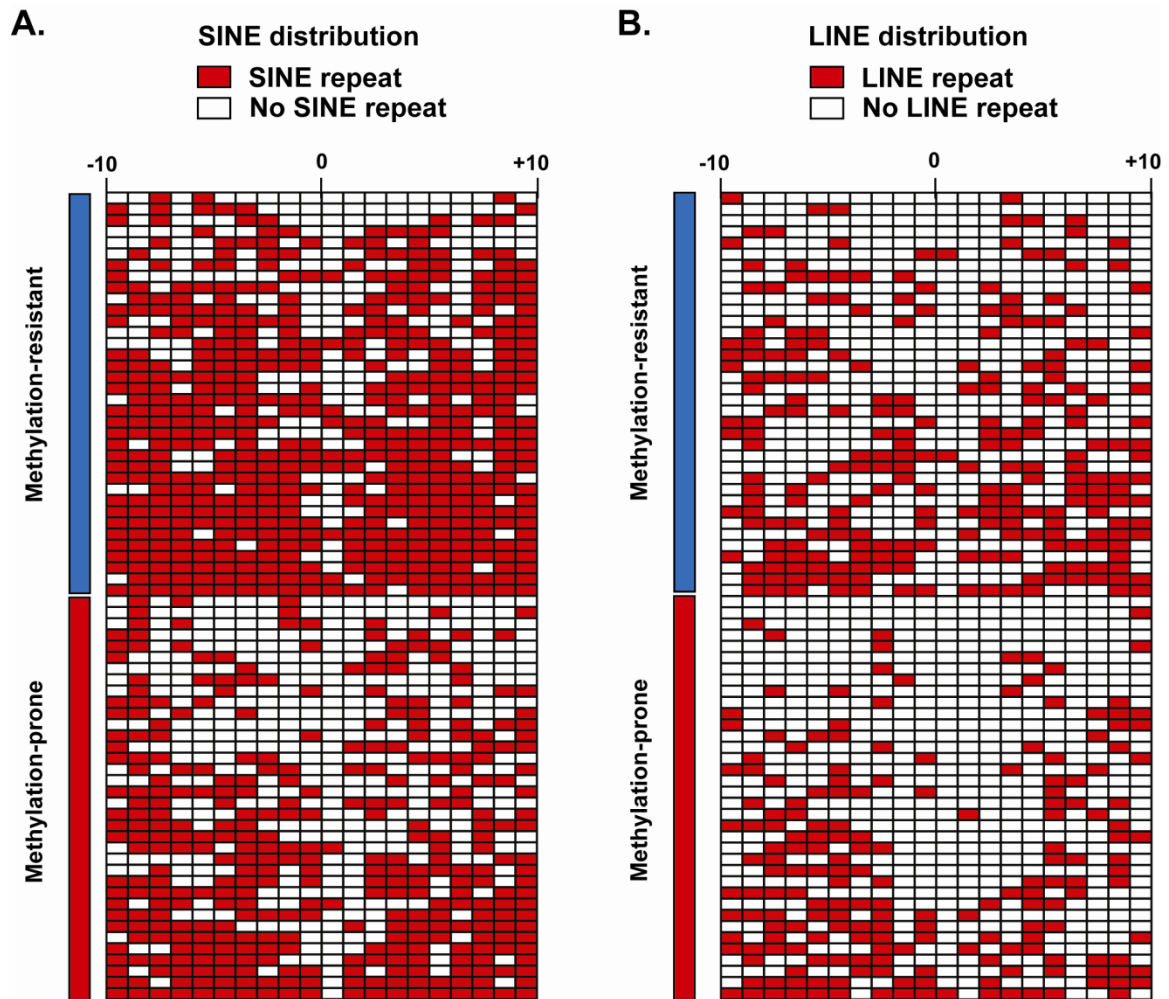
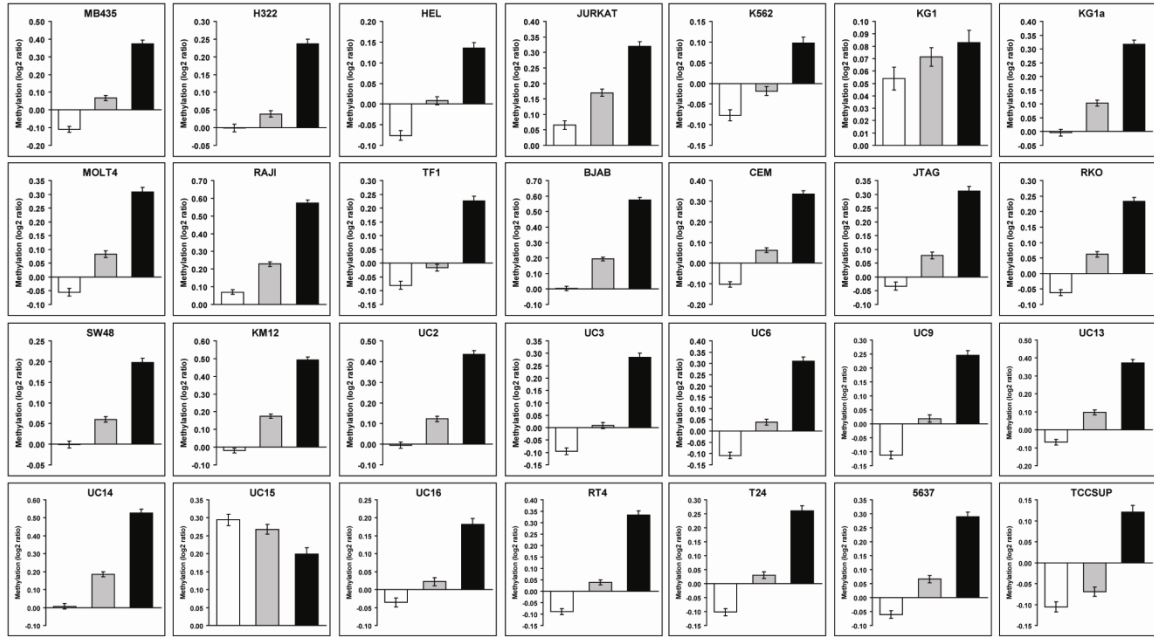


Figure S2

Cancer Cell Lines



Primary Cancer Tissues

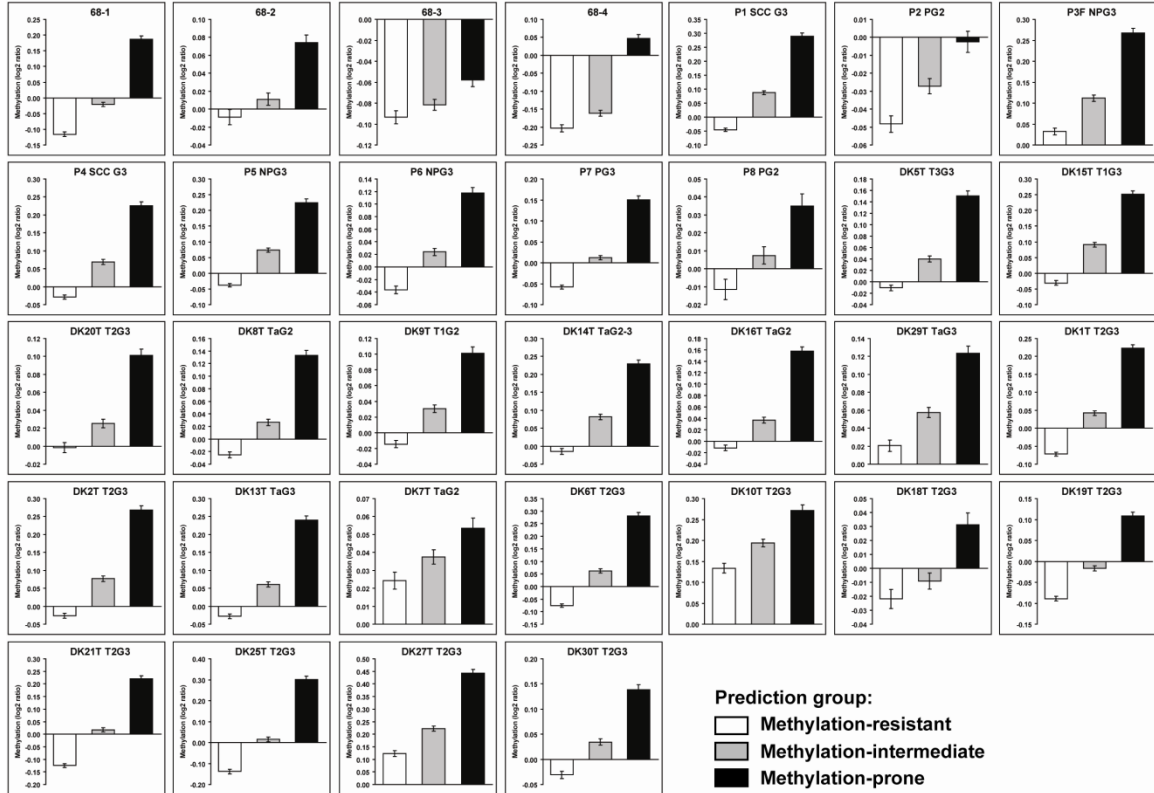


Figure S3

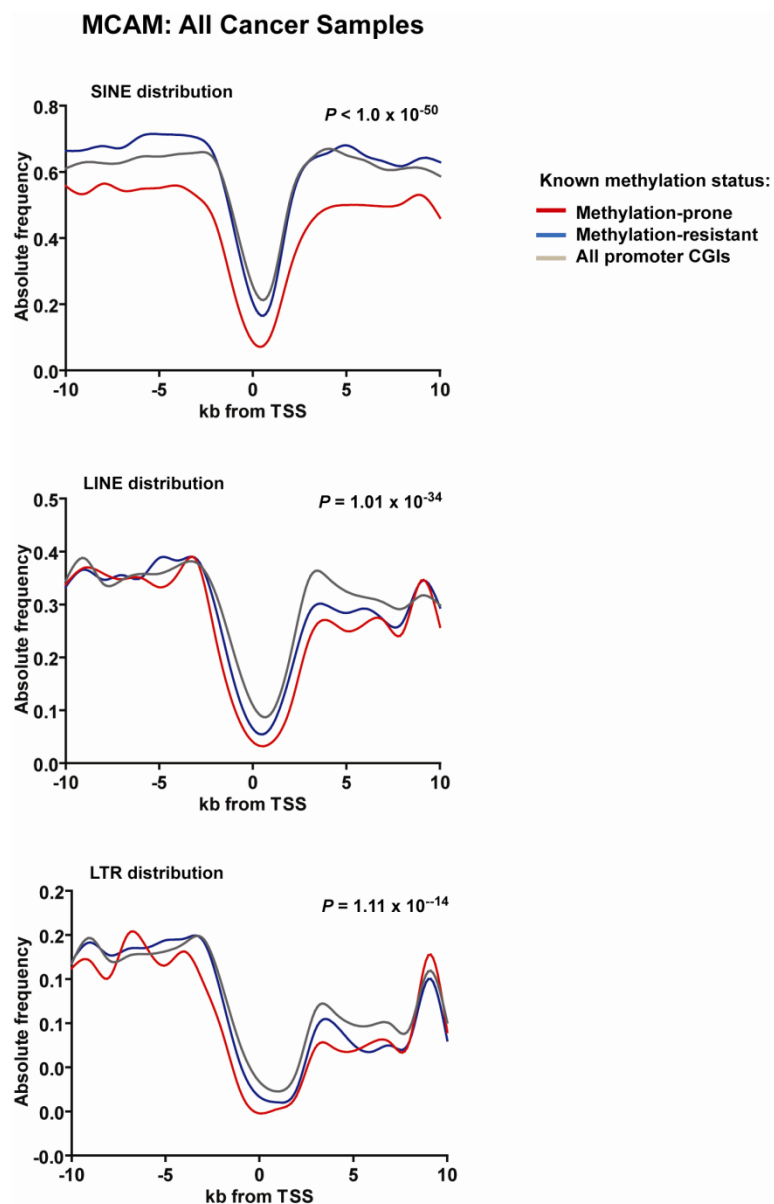


Figure S4

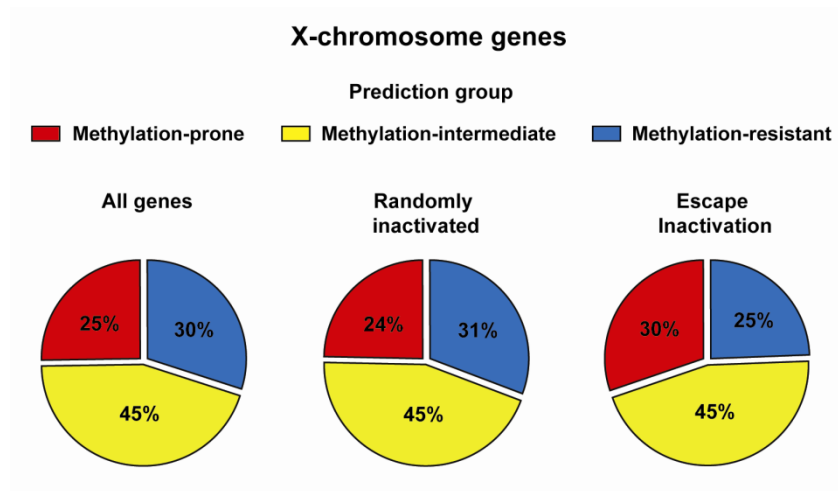


Figure S5

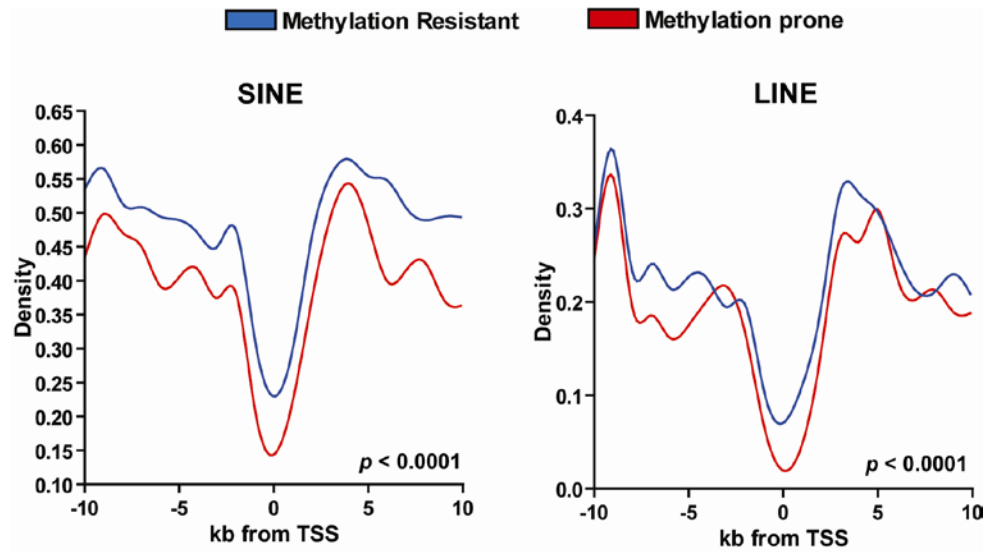


Figure S6

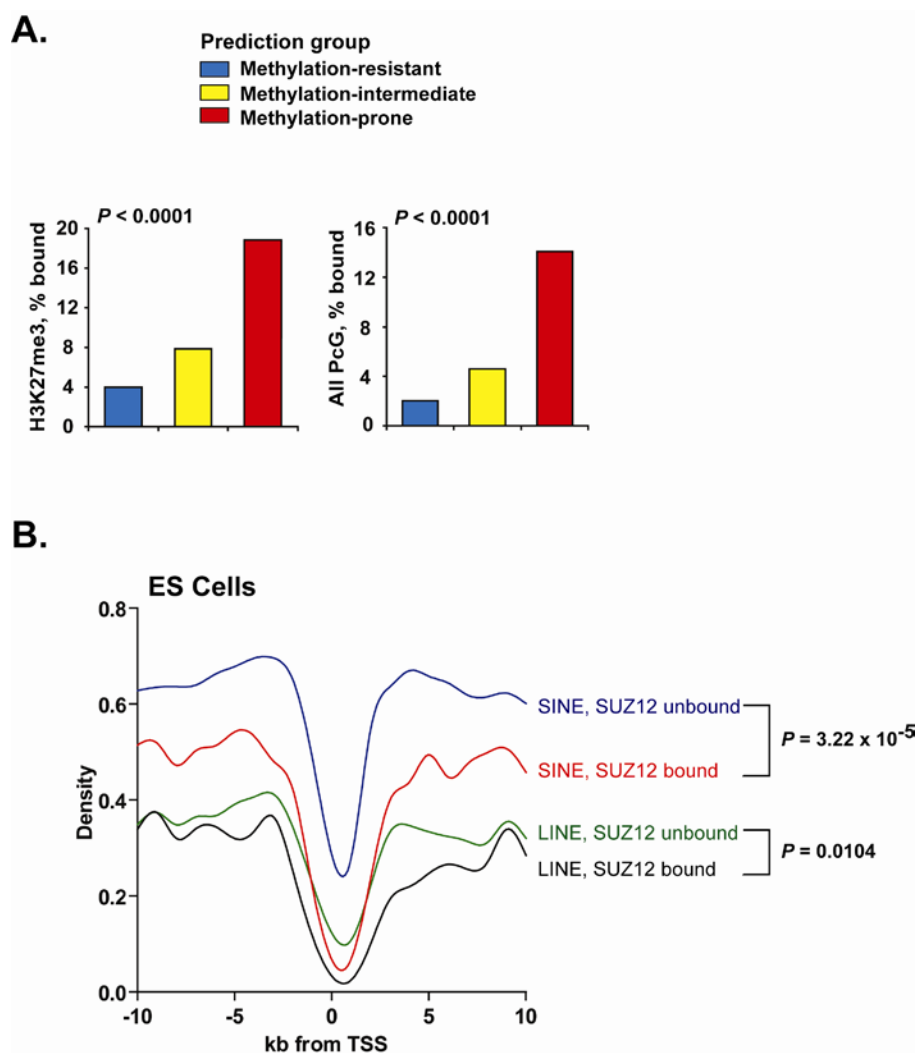


Figure S7

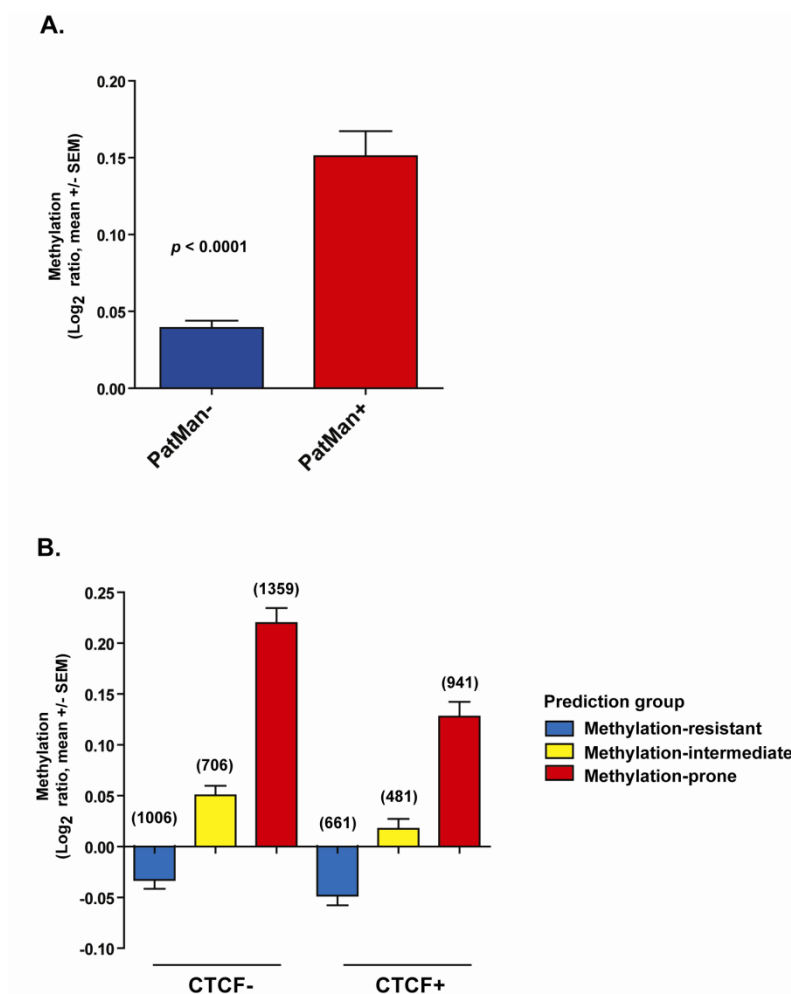


Table S1. Training set.

Gene Symbol	Gene Name	Chrom	Tx Start	RefSeq ID
Methylation prone				
<i>CABIN1</i>	calcineurin binding protein 1	22	22881972	NM_012295
<i>CACNA1G</i>	calcium channel, voltage-dependent, T type, alpha 1G subunit	17	45993447	NM_198397
<i>CDH1</i>	cadherin 1, type 1, E-cadherin (epithelial)	16	67328695	NM_004360
<i>CDH13</i>	cadherin 13, H-cadherin (heart)	16	81218078	NM_001257
<i>CDKN1C</i>	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	11	2863537	NM_000076
<i>CDKN2A*</i>	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	9	21965038	NM_058197
<i>DAPK1</i>	death-associated protein kinase 1	9	89302575	NM_004938
<i>DPH2</i>	DPH2 homolog (S. cerevisiae)	1	44208239	NM_001384
<i>ESR1</i>	estrogen receptor 1	6	152170378	NM_000125
<i>FGFR2</i>	fibroblast growth factor receptor 2	10	123347962	NM_022970
<i>FHIT</i>	fragile histidine triad gene	3	61212164	NM_002012
<i>FLT1</i>	fms-related tyrosine kinase 1 (vascular endothelial growth factor/vascular permeability factor receptor)	13	27967265	NM_002019
<i>FOXN3</i>	forkhead box N3	14	88953207	NM_005197
<i>GATA2</i>	GATA binding protein 2	3	129694718	NM_032638
<i>GDNF</i>	glial cell derived neurotrophic factor	5	37871350	NM_199231
<i>GPX1</i>	glutathione peroxidase 1	3	49370795	NM_000581
<i>HAND1</i>	heart and neural crest derivatives expressed 1	5	153838017	NM_004821
<i>HCN2</i>	hyperpolarization activated cyclic nucleotide-gated potassium channel 2	19	540892	NM_001194
<i>HRK</i>	harakiri, BCL2 interacting protein (contains only BH3 domain)	12	115803615	NM_003806
<i>MGMT</i>	O-6-methylguanine-DNA methyltransferase	10	131155455	NM_002412
<i>MLH3</i>	mutL homolog 3 (E. coli)	14	74587988	NM_014381
<i>MME</i>	membrane metallo-endopeptidase	3	156280129	NM_000902
<i>NKX2-3</i>	NK2 transcription factor related, locus 3 (Drosophila)	10	101282679	NM_145285
<i>PAX5</i>	paired box 5	9	37024476	NM_016734
<i>PDLIM4</i>	PDZ and LIM domain 4	5	131621285	NM_003687
<i>PPP1R1A</i>	protein phosphatase 1, regulatory (inhibitor) subunit 1A	12	53268710	NM_006741
<i>RB1</i>	retinoblastoma 1	13	47775883	NM_000321
<i>RUNX3</i>	runt-related transcription factor 3	1	25164088	NM_004350
<i>RYR1</i>	ryanodine receptor 1 (skeletal)	19	43616179	NM_000540
<i>SIN3A</i>	SIN3 homolog A, transcription regulator (yeast)	15	73530979	NM_015477
<i>TEAD2</i>	TEA domain family member 2	19	54557526	NM_003598
<i>TEC</i>	tec protein tyrosine kinase	4	47966571	NM_003215
<i>TP73</i>	tumor protein p73	1	3558988	NM_005427
<i>VCAN</i>	versican	5	82803338	NM_004385
<i>VRK2</i>	vaccinia related kinase 2	2	58127223	NM_006296
<i>ZNF160</i>	zinc finger protein 160	19	58298499	NM_198893
Methylation-resistant				
<i>ABL1</i>	c-abl oncogene 1, receptor tyrosine kinase	9	132700651	NM_005157
<i>AIP</i>	aryl hydrocarbon receptor interacting protein	11	67007096	NM_003977
<i>ARSA</i>	arylsulfatase A	22	49413473	NM_000487
<i>ARSB</i>	arylsulfatase B	5	78318113	NM_000046
<i>BNIP3</i>	BCL2/adenovirus E1B 19kDa interacting protein 3	10	133645425	NM_004052
<i>CD82</i>	CD82 molecule	11	44543716	NM_002231
<i>CDKN2A**</i>	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	9	21965038	NM_058195
<i>CLCN6</i>	chloride channel 6	1	11788793	NM_021736
<i>DCK</i>	deoxycytidine kinase	4	72078255	NM_000788
<i>FDX1</i>	ferredoxin 1	11	109805803	NM_004109
<i>FOXP2</i>	forkhead box P2	7	113842287	NM_148898
<i>GAS8</i>	growth arrest-specific 8	16	88616508	NM_001481
<i>GYS1</i>	glycogen synthase 1 (muscle)	19	54188361	NM_002103
<i>HMG2</i>	high mobility group AT-hook 2	12	64504506	NM_003483
<i>ITGAV</i>	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	2	187163044	NM_002210
<i>KCNK6</i>	potassium channel, subfamily K, member 6	19	43502323	NM_004823
<i>MKLN1</i>	muskelin 1, intracellular mediator containing kelch motifs	7	130663177	NM_013255
<i>MSH2</i>	mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli)	2	47483766	NM_000251
<i>MSH6</i>	mutS homolog 6 (E. coli)	2	47863789	NM_000179
<i>NEK4</i>	NIMA (never in mitosis gene a)-related kinase 4	3	52779991	NM_003157
<i>NSD1</i>	nuclear receptor binding SET domain protein 1	5	176494690	NM_172349
<i>PDE8A</i>	phosphodiesterase 8A	15	83326208	NM_002605
<i>PIK3R3</i>	phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	1	46370901	NM_003629
<i>PIP4K2B</i>	phosphatidylinositol-5-phosphate 4-kinase, type II, beta	17	34209684	NM_003559
<i>PXN</i>	paxillin	12	119187892	NM_002859
<i>RAD23A</i>	RAD23 homolog A (S. cerevisiae)	19	12917653	NM_005053
<i>RASA2</i>	RAS p21 protein activator 2	3	142688615	NM_006506
<i>SHFM1</i>	split hand/foot malformation (ectrodactyly) type 1	7	96177139	NM_006304
<i>STAT3</i>	signal transducer and activator of transcription 3 (acute-phase response factor)	17	37794039	NM_003150
<i>TAF11</i>	TAF11 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 28kDa	6	34963797	NM_005643
<i>TDG</i>	thymine-DNA glycosylase	12	102883746	NM_003211
<i>TGFBR2</i>	transforming growth factor, beta receptor II (70/80kDa)	3	30622997	NM_003242
<i>TOP1</i>	topoisomerase (DNA) I	20	39090875	NM_003286
<i>TOP2A</i>	topoisomerase (DNA) II alpha 170kDa	17	35827695	NM_001067
<i>TRIM27</i>	tripartite motif-containing 27	6	28999747	NM_006510
<i>VHL</i>	von Hippel-Lindau tumor suppressor	3	10158318	NM_000551

*transcript variant coding for the P16INK4A protein; **transcript variant coding for the P14ARF protein.

Table S2. Methylation profile of the training set in nine cancer cell lines from different tissue origin

Gene Symbol	NSCLC NCI-H322	PROSTATE PC3	BREAST MB 435	RENAL TK-10	CNS SNB-19	OVARIAN OVCAR-5	MELANOMA LOX IMUI	LEUK MOLT-4	COLON HCT 116
Methylation-prone									
CABIN1	100	0	0	0	0	0	0	0	0
CACNA1G	38	32	62	80	57	24	58	7	58
CDH1	0	0	100	63	68	0	100	100	0
CDH13	6	20	88	75	3	79	4	ND	97
CDKN1C	0	0	36	54	0	0	27	51	0
CDKN2A*	Deleted	89	0	81	Deleted	0	0	Deleted	57
DAPK1	59	15	74	0	8	0	64	54	10
DPH2	0	0	0	44	32	56	67	45	34
ESR1	0	43	27	44	29	66	55	66	57
FLT1	0	20	0	0	0	60	0	0	31
FOXP3	0	0	0	0	0	0	74	0	0
GATA2	0	0	68	0	0	0	0	0	53
GDNF	79	3	75	50	39	93	44	77	93
GPX1	54	0	0	0	0	0	0	55	0
HAND1	24	15	69	69	52	78	33	47	47
HRK	14	45	13	29	4	54	96	4	7
MGMT	58	3	93	64	65	0	91	0	79
MLH3	6	2	12	2	54	2	2	77	9
MME	0	0	49	44	0	11	2	81	76
PAX5	54	2	87	84	37	54	85	3	64
PDLIM4	92	0	53	0	0	0	92	76	56
PPP1R1A	0	18	27	0	0	0	20	0	28
RYR1	0	0	35	0	0	32	0	0	0
TEAD2	20	0	0	0	0	0	0	48	0
TEC	4	0	42	0	0	0	0	20	0
TP73	0	28	100	0	0	0	100	30	0
VRK2	0	0	0	0	90	5	0	15	0
ZNF160	ND	0	6	0	7	90	74	5	0
Methylation-resistant									
ABL1	0	0	0	0	0	0	0	0	0
AIP	0	0	0	0	0	0	0	0	0
ARSA	ND	3	5	ND	5	12	5	ND	10
ARSB	8	12	12	7	7	8	10	6	7
BNIP3	0	0	0	0	0	27	0	0	0
CD82	2	4	17	1	2	1	2	1	7
CDKN2A**	0	0	0	13	Deleted	0	0	0	15
CLCN6	0	0	0	0	0	0	0	0	0
DCK	4	3	3	12	2	4	2	4	11
FDX1	0	0	0	0	0	0	0	0	0
FOXP2	0	0	14	14	0	0	0	18	0
GAS8	50	0	0	0	0	0	0	0	0
GYS1	0	0	0	0	0	0	0	0	0
HMG2A	0	0	0	0	0	0	0	0	0
ITGAV	0	0	0	0	0	0	0	0	0
KCNK6	0	0	0	2	12	0	2	0	0
MKL1	14	0	0	0	0	0	0	0	0
MSH2	2	5	3	2	2	2	2	2	11
MSH6	4	6	3	2	3	2	3	2	13
NEK4	0	0	0	0	0	0	0	0	0
NSD1	0	0	0	10	0	0	0	0	0
PDE8A	0	0	0	0	0	0	0	0	0
PIK3R3	0	0	0	0	0	0	0	0	0
PIP4K2B	0	0	0	0	0	0	0	0	0
PXN	0	0	0	0	0	0	0	0	0
RAD23A	0	0	0	0	0	0	0	0	0
RASA2	10	12	7	4	3	9	2	2	14
SHFM1	0	0	0	0	0	0	0	0	0
STAT3	0	0	0	0	0	0	0	0	0
TAF11	2	3	2	2	2	2	1	2	7
TDG	0	0	0	0	0	0	0	0	0
TGFBR2	2	2	2	2	2	2	2	3	9
TOP1	0	0	0	0	0	0	0	0	0
TOP2A	1	2	2	2	1	1	1	2	6
TRIM27	6	0	14	5	5	0	0	0	0
VHL	6	8	8	5	4	4	5	8	18

The following genes were classified as methylation-prone based on methylation analysis of RKO (colon cancer cell line) and Raji (leukemia cell line): *VCAN*, *FGFR2*, *FHIT*, *HCN2*, *NKX2-3*, *RB1*, *RUNX3* and *SIN3A*. *CABIN1* was classified as methylation-prone due to its high methylation density. Bold: methylation values equal or above 15%. *transcript variant coding for the P16INK4A protein; **transcript variant coding for the P14ARF protein.

Table S3. Methylation-prone genes in cancer used as test set.

Gene Symbol	Gene Name	RefSeq	Chrom	Tx Start	S score	L score	Prediction	Source
ANKHD1-EIF4EBP3	ANKHD1-EIF4EBP3 readthrough	NM_020690	5	139761612	-5.322	2.160	intermediate	MCAM
BARHL1	BarH-like homeobox 1	NM_020064	9	134447813	3.776	0.379	prone	Bis-PCR
BIN3	bridging integrator 3	NM_018688	8	22582606	1.374	-1.956	intermediate	MCAM
BMP7	bone morphogenetic protein 7	NM_001719	20	55274708	4.903	0.701	prone	18
C19orf70	chromosome 19 open reading frame 70	NM_205767	19	5631911	-3.435	0.235	intermediate	MCAM
CALCA	calcitonin-related polypeptide alpha	NM_001741	11	14950408	2.834	0.680	prone	2
CEBPD	CCAAT/enhancer binding protein (C/EBP), delta	NM_005195	8	48813279	1.265	2.055	prone	23
CHFR	checkpoint with forkhead and ring finger domains	NM_018223	12	131974257	-2.538	-0.969	resistant	5
CLDN3	claudin 3	NM_001306	7	72822512	-5.427	-1.544	resistant	13
CNNM1	cyclin M1	NM_020348	10	101080022	-1.270	-2.175	resistant	10
COL2A1	collagen, type II, alpha 1	NM_001844	12	46684552	6.333	1.245	prone	10
CPT1C	carnitine palmitoyltransferase 1C	NM_152359	19	54886218	-1.354	0.124	intermediate	10
DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	NM_001343	5	39461092	2.058	0.533	prone	4
DDIT4L	DNA-damage-inducible transcript 4-like	NM_145244	4	101330636	4.524	-1.492	intermediate	10
DERL3	Der1-like domain family, member 3	NM_198440	22	22511201	2.592	0.052	prone	10
DHRS3	dehydrogenase/reductase (SDR family) member 3	NM_004753	1	12600407	0.700	1.099	prone	10
DKK1	dickkopf homolog 1 (Xenopus laevis)	NM_012242	10	53744046	5.597	1.351	prone	1
DLG1	deleted in liver cancer 1	NM_024767	8	13416766	1.425	0.132	prone	28
DLEC1	deleted in lung and esophageal cancer 1	NM_007337	3	38055699	-1.228	-2.988	resistant	20
DPYS	dihydropyrimidinase	NM_001385	8	105548453	4.745	1.418	prone	10
DSC3	desmocollin 3	NM_024423	18	26876779	3.530	-0.565	intermediate	19
EDIL3	EGF-like repeats and discoidin I-like domains 3	NM_005711	5	83716367	4.375	2.502	prone	MCAM
EDNRB	endothelin receptor type B	NM_000115	13	77447665	6.416	-1.733	intermediate	6
EFEMP2	EGF-containing fibulin-like extracellular matrix protein 2	NM_016938	11	65396852	2.029	0.292	prone	10
EMX2	empty spiracles homeobox 2	NM_004098	10	119291945	7.077	1.310	prone	MCAM
EPHX3	epoxide hydrolase 3	NM_024794	19	15204231	-0.863	1.271	intermediate	10
ERBB4	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	NM_005235	2	213111597	7.454	1.702	prone	MCAM
ESYT3	extended synaptotagmin-like protein 3	NM_031913	3	139636108	2.109	-1.271	intermediate	10
FAM19A2	family with sequence similarity 19 (chemokine (C-C motif)-like), member A2	NM_178539	12	60872818	8.355	1.006	prone	MCAM
GDNF	glial cell derived neurotrophic factor	NM_000514	5	37875539	2.242	2.919	prone	MCAM
GSTP1	glutathione S-transferase pi 1	NM_000852	11	67107861	-1.541	0.471	intermediate	29
HNF1B	HNF1 homeobox B	NM_000458	17	33179209	-0.911	-0.152	resistant	24
HPSE2	heparanase 2	NM_021828	10	100985609	5.000	2.973	prone	MCAM
KCNAB1	potassium voltage-gated channel, shaker-related subfamily, beta member 1	NM_172159	3	157491519	3.711	-2.544	intermediate	MCAM
LHX6	LIM homeobox 6	NM_199160	9	124030840	-0.761	1.858	intermediate	Bis-PCR
LHX9	LIM homeobox 9	NM_020204	1	196153139	8.240	1.702	prone	MCAM
LMX1A	LIM homeobox transcription factor 1, alpha	NM_177399	1	163591641	2.291	0.904	prone	MCAM
LRP2	low density lipoprotein receptor-related protein 2	NM_004525	2	169927368	2.017	1.707	prone	Bis-PCR
MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	NM_000249	3	37009982	2.962	-1.882	intermediate	Bis-PCR
MLPH	melanophilin	NM_024101	2	238060616	1.683	0.264	prone	MCAM
MT1G	metallothionein 1G	NM_005950	16	55259478	-0.934	1.702	intermediate	11
MT3	metallothionein 3	NM_005954	16	55180767	-3.033	-2.058	resistant	21
MYOD1	myogenic differentiation 1	NM_002478	11	17697685	6.576	-1.652	intermediate	12
NFIX	nuclear factor I/X (CCAAT-binding transcription factor)	NM_002501	19	12967583	1.490	1.702	prone	MCAM
ONECUT2	one cut homeobox 2	NM_004852	18	53253914	9.237	1.702	prone	MCAM
OTP	orthopedia homeobox	NM_032109	5	76970278	9.237	2.299	prone	MCAM
PRDX2	peroxiredoxin 2	NM_181738	19	12773694	-3.191	1.863	intermediate	10
PRRG1	proline rich Gla (G-carboxyglutamic acid) 1	NM_000950	X	37093545	4.674	-0.296	intermediate	MCAM
PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	NM_000963	1	184916179	2.147	0.992	prone	26
PTPRG	protein tyrosine phosphatase, receptor type, G	NM_002841	3	61522284	-1.424	-0.229	resistant	10
PVRL2	poliovirus receptor-related 2 (herpesvirus entry mediator B)	NM_002856	19	50041232	-4.724	-1.748	resistant	MCAM
RASD1	RAS, dexamethasone-induced 1	NM_016084	17	17340432	3.615	2.434	prone	10
RASGRF2	Ras protein-specific guanine nucleotide-releasing factor 2	NM_006909	5	80292313	4.645	-0.079	intermediate	8
RASSF2	Ras association (RalGDS/AF-6) domain family member 2	NM_014737	20	4752291	1.156	-0.063	intermediate	14
REC8L1	REC8 homolog (yeast)	NM_005132	14	23711073	1.261	1.208	prone	10
RECK	reversion-inducing-cysteine-rich protein with kazal motifs	NM_021111	9	36026915	1.046	-5.119	intermediate	7
RPS5	ribosomal protein S5	NM_001009	19	63590447	-1.227	-0.579	resistant	MCAM
SFRP1	secreted frizzled-related protein 1	NM_003012	8	41286137	5.634	2.585	prone	22
SMAD4	SMAD family member 4	NM_005359	18	46810610	0.255	1.258	prone	25
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	NM_003070	9	2005341	5.634	1.702	prone	MCAM
SOC1	suppressor of cytokine signaling 1	NM_003745	16	11257540	-3.038	-1.196	resistant	16
STAT1	signal transducer and activator of transcription 1, 91kDa	NM_139266	2	191587181	4.775	-1.399	intermediate	27
TBX5	T-box 5	NM_080717	12	113330630	5.412	1.841	prone	MCAM
TERT	telomerase reverse transcriptase	NM_198253	5	1348162	2.149	-0.394	intermediate	9
TEX9	testis expressed 9	NM_198524	15	54444935	4.942	-5.345	intermediate	MCAM
THBS1	thrombospondin 1	NM_003246	15	37660571	5.911	1.763	prone	Bis-PCR
UCHL1	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	NM_004181	4	40953685	-1.778	-1.969	resistant	17
UNC5C	unc-5 homolog C (C. elegans)	NM_003728	4	96689185	5.060	1.889	prone	10
VAX2	ventral anterior homeobox 2	NM_012476	2	70981227	5.211	-0.533	intermediate	MCAM
WIF1	WNT inhibitory factor 1	NM_007191	12	63801383	6.206	1.060	prone	3
WNT10A	wingless-type MMTV integration site family, member 10A	NM_025216	2	219453498	2.172	0.974	prone	18
WNT6	wingless-type MMTV integration site family, member 6	NM_006522	2	219432789	3.192	1.409	prone	18
ZGPAT	zinc finger, CCH-type with G patch domain	NM_032527	20	61809834	0.756	-3.336	intermediate	18
ZHX2	zinc fingers and homeoboxes 2	NM_014943	8	123863081	2.024	0.310	prone	15

Source of methylation data: 1-29, reported in the literature (see Supplemental References); Bis-PCR, bisulfite-PCR followed by pyrosequencing or COBRA assays; MCAM, methylated CpG island amplification and microarray.

Table S4. Methylation-resistant genes in cancer used as test set.

Gene Symbol	Gene Name	RefSeq	Chrom	Tx Start	S score	L score	Prediction	Source
ABCB10	ATP-binding cassette, sub-family B (MDR/TAP), member 10	NM_012089	1	227761065	-1.740	-0.251	resistant	Bis-PCR
ADAT1	adenosine deaminase, tRNA-specific 1	NM_012091	16	74214655	-0.550	-1.042	resistant	MCAM
ADRB1	adrenergic, beta-1-, receptor	NM_000684	10	115793795	1.304	1.930	prone	Bis-PCR
ALDH7A1	aldehyde dehydrogenase 7 family, member A1	NM_001182	5	125958839	-4.027	0.168	intermediate	MCAM
ALG12	asparagine-linked glycosylation 12, alpha-1,6-mannosyltransferase homolog (S. cerevisiae)	NM_024105	22	48698110	0.421	-1.304	intermediate	Bis-PCR
ATP6V1E1	ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E1	NM_001696	22	16491588	-5.621	-0.730	resistant	Bis-PCR
ATP6V1G1	ATPase, H+ transporting, lysosomal 13kDa, V1 subunit G1	NM_004888	9	116389814	-5.650	-1.336	resistant	MCAM
ATXN10	ataxin 10	NM_013236	22	44446350	-3.514	-6.224	resistant	Bis-PCR
BAG2	BCL2-associated athanogene 2	NM_004282	6	57145292	3.993	-0.241	intermediate	MCAM
BCAT2	branched chain amino-acid transaminase 2, mitochondrial	NM_001190	19	54006113	-1.196	-0.463	resistant	MCAM
BCL2	B-cell CLL/lymphoma 2	NM_000657	18	59137593	4.734	1.163	prone	Bis-PCR
BID	BH3 interacting domain death agonist	NM_197966	22	16637258	0.358	0.980	prone	Bis-PCR
C12orf10	chromosome 12 open reading frame 10	NM_021640	12	51979736	0.841	0.584	prone	MCAM
C14orf21	chromosome 14 open reading frame 21	NM_174913	14	23838937	4.985	-1.985	intermediate	MCAM
COL9A1	collagen, type IX, alpha 1	NM_001851	6	71069494	0.597	-0.290	intermediate	MCAM
CYBASC3	cytochrome b, ascorbate dependent 3	NM_153611	11	60886305	-2.257	-0.154	resistant	MCAM
DHRS1	dehydrogenase/reductase (SDR family) member 1	NM_138452	14	23838506	2.535	-2.490	intermediate	MCAM
DUS3L	dihydrouridine synthase 3-like (S. cerevisiae)	NM_020175	19	5742190	-1.711	-0.880	resistant	MCAM
FAM53C	family with sequence similarity 53, member C	NM_016605	5	137701866	-1.716	-0.313	resistant	Bis-PCR
FIBP	fibroblast growth factor (acidic) intracellular binding protein	NM_198897	11	65412586	-3.148	-1.851	resistant	MCAM
FNBP1	formin binding protein 1	NM_015033	9	131845294	-3.395	0.517	intermediate	MCAM
FRAT1	frequently rearranged in advanced T-cell lymphomas	NM_005479	10	99069011	-3.925	2.607	intermediate	MCAM
GABPB1	GA binding protein transcription factor, beta subunit 1	NM_181427	15	48434687	-1.762	-1.130	resistant	Bis-PCR
GNPAT	glyceronephosphate O-acyltransferase	NM_014236	1	229443604	-1.568	-3.546	resistant	MCAM
GPM6B	glycoprotein M6B	NM_005278	X	13866752	4.910	1.898	prone	MCAM
GTSE1	G-2 and S-phase expressed 1	NM_016426	22	45071475	-2.325	-5.270	resistant	Bis-PCR
HNRPA0	heterogeneous nuclear ribonucleoprotein A0	NM_006805	5	137117938	3.743	0.267	prone	Bis-PCR
HTT	huntingtin	NM_002111	4	3046205	-1.297	0.502	intermediate	MCAM
KIAA0494	KIAA0494	NM_014774	1	46957323	-2.563	-2.687	resistant	MCAM
KIF20A	kinesin family member 20A	NM_005733	5	137543247	-0.370	0.670	intermediate	Bis-PCR
KPNA3	karyopherin alpha 3 (importin alpha 4)	NM_002267	13	49265058	-2.690	-0.376	resistant	Bis-PCR
MCEE	methylmalonyl CoA epimerase	NM_032601	2	71210867	-2.576	-5.401	resistant	MCAM
MFAP1	microfibrillar-associated protein 1	NM_005926	15	41904243	-6.891	-1.775	resistant	MCAM
MOBP	myelin-associated oligodendrocyte basic protein	NM_182935	3	39484073	1.464	-2.316	intermediate	MCAM
MPHOSPH10	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)	NM_005791	2	71210951	-2.939	-0.845	resistant	MCAM
MTX2	metaxin 2	NM_006554	2	176842382	1.586	-0.246	intermediate	MCAM
NUDT15	nudix (nucleoside diphosphate linked moiety X)-type motif 15	NM_018283	13	47509703	-0.803	-3.181	resistant	Bis-PCR
PARK2	Parkinson disease (autosomal recessive, juvenile) 2, parkin	NM_013988	6	163068790	7.919	1.222	prone	MCAM
PEX3	peroxisomal biogenesis factor 3	NM_003630	6	143813809	-1.002	0.629	intermediate	Bis-PCR
PGD	phosphogluconate dehydrogenase	NM_002631	1	10381671	-5.650	0.555	intermediate	MCAM
POLR2E	polymerase (RNA) II (DNA directed) polypeptide E, 25kDa	NM_002695	19	1046336	-3.333	0.002	intermediate	MCAM
PPAN	peter pan homolog (Drosophila)	NM_020230	19	10077964	-2.877	0.866	intermediate	Bis-PCR
PSMB7	proteasome (prosome, macropain) subunit, beta type, 7	NM_002799	9	126217542	-4.979	-3.440	resistant	MCAM
RBM17	RNA binding motif protein 17	NM_032905	10	6171012	-0.992	-0.437	resistant	Bis-PCR
RCBTB1	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1	NM_018191	13	49057720	-0.533	-2.394	resistant	Bis-PCR
SART3	squamous cell carcinoma antigen recognized by T cells 3	NM_014706	12	107479295	0.367	-1.122	intermediate	MCAM
SCFD2	sec1 family domain containing 2	NM_152540	4	53926999	0.882	-3.513	intermediate	MCAM
SELO	selenoprotein O	NM_031454	22	48981534	3.141	-1.841	intermediate	Bis-PCR
SETDB2	SET domain, bifurcated 2	NM_031915	13	48916510	0.037	-2.806	intermediate	Bis-PCR
SIL1	SIL1 homolog, endoplasmic reticulum chaperone (S. cerevisiae)	NM_022464	5	138561964	-2.659	0.866	intermediate	Bis-PCR
SMURF2	SMAD specific E3 ubiquitin protein ligase 2	NM_022739	17	60088848	-2.694	-1.538	resistant	MCAM
SOX7	SRY (sex determining region Y)-box 7	NM_031439	8	10625432	1.781	1.267	prone	MCAM
SPTLC2	serine palmitoyltransferase, long chain base subunit 2	NM_004863	14	77152863	-2.979	-2.403	resistant	MCAM
ST8SIA1	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	NM_003034	12	22378915	2.606	-0.710	intermediate	MCAM
STMN1	stathmin 1	NM_203401	1	26105231	-2.580	-0.032	resistant	Bis-PCR
SUCLA2	succinate-CoA ligase, ADP-forming, beta subunit	NM_003850	13	47473463	-2.072	-3.282	resistant	Bis-PCR
TCF21	transcription factor 21	NM_198392	6	134251969	3.646	0.787	prone	MCAM
TMEM33	transmembrane protein 33	NM_018126	4	41631925	1.263	-2.671	intermediate	MCAM
TMEM45A	transmembrane protein 45A	NM_018004	3	101694152	0.954	-1.974	intermediate	MCAM
TRIB2	tribbles homolog 2 (Drosophila)	NM_021643	2	12774658	-1.211	-0.279	resistant	MCAM
TRIP4	thyroid hormone receptor interactor 4	NM_016213	15	62467072	-6.451	0.081	intermediate	MCAM
UBQLN1	ubiquilin 1	NM_013438	9	85512773	3.277	-3.856	intermediate	Bis-PCR
USP18	ubiquitin specific peptidase 18	NM_017414	22	17012757	-1.142	-0.969	resistant	Bis-PCR
ZAK	leucine zipper- and sterile alpha motif-containing kinase	NM_016653	2	173648810	-0.846	-2.494	resistant	MCAM
ZBED4	zinc finger, BED-type containing 4	NM_014838	22	48633500	-0.388	-1.704	resistant	Bis-PCR
ZNF224	zinc finger protein 224	NM_013398	19	49290336	-2.214	-5.826	resistant	MCAM
ZNF418	zinc finger protein 418	NM_133460	19	63138552	-1.253	0.639	intermediate	MCAM
ZNF786	zinc finger protein 786	NM_152411	7	148418720	-4.792	-1.125	resistant	Bis-PCR

Table S5. Multi clusters of predicted methylation-prone genes

Chrom	Start	End	Size (Mb)	Location	Number of genes (Prone/Resistant or Intermediate)	Hyperconserved	Multi Cluster
1	850983	1437417	0.6	tel	17/10	no	no
1	1812355	6182266	4.4	tel	16/14	no	no
2	175578361	176702722	1.1	q31	9/2	<i>HOXD13</i>	<i>HOXD</i>
2	219432789	223433975	4.0	q35	21/16	no	no
2	233633279	238060616	4.4	q37	9/2	<i>GBX2</i>	<i>UGT1</i>
3	49019829	49289456	0.3	p21	9/3	no	no
3	73756762	99795131	26.0	centr	12/15	no	<i>OR5</i>
5	496210	9599233	9.1	tel	18/14	<i>IRX1, IRX2</i>	no
5	140146059	140835752	0.7	q31	15/11	no	<i>PCDHA, PCDHB</i>
6	31938662	33486771	1.5	p21	16/13	no	<i>HLA</i>
7	23111926	30688665	7.6	p15	24/20	<i>IGF2BP3, CBX3, HOXA</i>	<i>HOXA</i>
7	78920826	87343604	8.4	q11	9/3	no	no
8	142080514	145988533	3.9	tel	42/29	no	<i>LY6</i>
9	135204793	139633070	4.4	tel	43/27	no	no
10	18469611	25504295	7.0	p12	13/8	<i>MLLT10, BMI1</i>	no
10	128584012	135333668	6.7	tel	19/13	<i>EBF3</i>	no
11	497273	3833508	3.3	tel	31/23	no	no
11	19219083	20341978	1.1	p15	7/0	<i>NAV2</i>	no
11	127897371	133628643	5.7	tel	13/7	<i>FLI1</i>	no
12	52179538	54497773	2.3	q13	13/8	<i>HOXC</i>	<i>HOXC, OR</i>
12	55808548	62524807	6.7	q13	17/11	no	no
14	19999477	20636648	0.6	q11	7/0	<i>uncharacterized</i>	<i>RNASE</i>
16	43625	1908232	1.9	tel	43/24	no	<i>HB</i>
17	77110011	78009348	0.9	tel	21/19	no	no
18	57711284	76106388	18.4	q21-q23	22/13	<i>SALL3</i>	<i>SERPINB</i>
20	56848189	61676036	4.8	tel	28/21	no	no
21	44351635	46568213	2.2	tel	13/7	no	<i>KRTAP10, 12</i>
22	48698347	49569090	0.9	tel	16/10	no	no

Table S6. Predicted methylation status of known imprinted genes

Gene Symbol	Gene Name	Expressed allele	Chrom	Tx start	RefSeq Name	S score	L score	Prediction
<i>ATP10A</i>	ATPase, class V, type 10A	Maternal	15	23659963	NM_024490	2.1525	-1.2103	Intermediate
<i>CDKN1C</i>	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	Maternal	11	2863537	NM_000076	2.2754	3.0559	Prone
<i>DIRAS3</i>	DIRAS family, GTP-binding RAS-like 3	Paternal	1	68289048	NM_004675	0.3746	1.0786	Prone
<i>DLK1</i>	delta-like 1 homolog (Drosophila)	Paternal	14	100263005	NM_003836	3.133	0.6564	Prone
<i>DLX5</i>	distal-less homeobox 5	Maternal	7	96492079	NM_005221	7.5135	1.9779	Prone
<i>GNAS</i>	GNAS complex locus	Maternal	20	56848189	NM_016592	4.6449	-1.0696	Intermediate
<i>HYMAI</i>	hydatidiform mole associated and imprinted (non-protein coding)	Paternal	6	144371560	NM_001080954	3.6932	-6.0203	Intermediate
<i>IGF2</i>	insulin-like growth factor 2 (somatomedin A)	Paternal	11	2116780	NM_000612	10.0794	1.9779	Prone
<i>KCNK9</i>	potassium channel, subfamily K, member 9	Maternal	8	140784481	NM_016601	1.4654	-0.8115	Intermediate
<i>KLF14</i>	Kruppel-like factor 14	Maternal	7	130069400	NM_138693	4.9041	-1.9255	Intermediate
<i>L3MBTL</i>	l(3)mbt-like (Drosophila)	Paternal	20	41576466	NM_032107	-1.599	2.4102	Intermediate
<i>MEST</i>	mesoderm specific transcript homolog (mouse)	Paternal	7	129919173	NM_002402	2.979	-2.2642	Intermediate
<i>NNAT</i>	neuronatin	Paternal	20	35583020	NM_005386	6.5027	-0.6502	Intermediate
<i>OSBPL5</i>	oxysterol binding protein-like 5	Maternal	11	3143116	NM_020896	0.1943	0.6564	Prone
<i>PEG10</i>	paternally expressed 10	Paternal	7	94123572	NM_001040152	9.0299	1.6288	Prone
<i>PEG3</i>	paternally expressed 3	Paternal	19	62043876	NM_006210	6.4992	-0.741	Intermediate
<i>PHLDA2</i>	pleckstrin homology-like domain, family A, member 2	Maternal	11	2907226	NM_003311	0.9173	0.3317	Prone
<i>PLAGL1</i>	pleiomorphic adenoma gene-like 1	Paternal	6	144371234	NM_002656	3.6932	-6.0203	Intermediate
<i>PPP1R9A</i>	protein phosphatase 1, regulatory (inhibitor) subunit 9A	Maternal	7	94374884	NM_017650	8.1646	1.8538	Prone
<i>SGCE</i>	sarcoglycan, epsilon	Paternal	7	94123457	NM_003919	9.7476	0.5498	Prone
<i>SLC22A2</i>	solute carrier family 22 (organic cation transporter), member 2	Maternal	6	160599949	NM_003058	-1.219	-0.182	Resistant
<i>SLC22A3</i>	solute carrier family 22 (extraneuronal monoamine transporter), member 3	Maternal	6	160689414	NM_021977	1.8184	-3.0299	Intermediate
<i>SNRPN</i>	small nuclear ribonucleoprotein polypeptide N	Paternal	15	22751227	NM_003097	-2.4611	2.4025	Intermediate
<i>SNURF</i>	SNRPN upstream reading frame	Paternal	15	22751227	NM_003097	-2.4611	2.4025	Intermediate
<i>TCEB3C</i>	transcription elongation factor B polypeptide 3C (elongin A3)	Maternal	18	42804532	NM_145653	3.9506	1.4642	Prone
<i>TFPI2</i>	tissue factor pathway inhibitor 2	Maternal	7	93358001	NM_006528	3.175	-2.7699	Intermediate
<i>UBE3A</i>	ubiquitin protein ligase E3A	Maternal	15	23235221	NM_130838	0.5189	-4.1931	Intermediate
<i>WT1</i>	Wilms tumor 1	Paternal	11	32413663	NM_000378	5.4481	1.3196	Prone
<i>ZIM2</i>	zinc finger, imprinted 2	Paternal	19	62043887	NM_015363	6.4992	-0.741	Intermediate
<i>ZNF264</i>	zinc finger protein 264	Maternal	19	62394680	NM_003417	-1.8817	1.3543	Intermediate

Table S7. Primers for bisulfite-pyrosequencing method used in this study

Primer Name*	Primer Sequence**	Sequence to analyze***
ARSA/F1	GTTGTTGGTTGAGTTTTGAA	GYGTTYGGGTTG
ARSA/F2	TTGAAAAGGGTTTTGGGAGTT	
ARSA/R	CAACCCCTCCTCCTCTAAAAAA	
ARSA/RU	GGGACACCGCTGATCGTTTCAACCCCTCCTCCTCTAAAAAA	
ARSA/S1	GGTAGGTTTTGGTTGAG	
ARSB/F1	GGTAGTTTGTAGTAAGAAGATTAGGTGG	GTGYTYGAGGTTYGGGTTT
ARSB/F2	GGAGGAAGAAGGGGAGGAGTA	
ARSB/R	ACAACCCAATTCTCATTCTATCA	
ARSB/RU	GGGACACCGCTGATCGTTTACAACCCAATTCTCATTCTATCA	
ARSB/S1	GAAGAAGGGGAGGAGTA	
CD82/F1	GAGGAGTTTTTAGGTTGTAGTAGGGAGG	YGAYGYGATTGAGGTA
CD82/F2	GGAGGGGCTAGAGGAGTGATTGA	
CD82/R	CCTTACTCACCTACCAATTCTACTTTCTACTCC	
CD82/RU	GGGACACCGCTGATCGTTTACCTTACTCACCTACCAATTCTACTTTCTACTCC	
CD82/S1	GAGTTTAGGGTTGTGTAGTT	
CDH1/F	GAGTTATAGGAGGGGTTGGGATT	TATYGYGTTTATGYGAGGTYGGGT
CDH1/RU	GGGACACCGCTGATCGTTTACTACCCACCACAACCAATCAA	
CDH1/S	AGTAATTTTAGGTTAGAGGGT	
CDH13/F	AGTTTGTTTTTAAGGAAAATATGTTTAGT	GTGYGTGTATGAATGAAAYG
CDH13/RU	GGGACACCGCTGATCGTTTAAACCAAAATCTCCACTACATTTTATCC	
CDH13/S	AGGAAAATATGTTTAGTGTA	
VCAN/F	TAGAGGGGTTTTAATAATTTTA	GGYGYGGGAGAGAAT
VCAN/RU	GGGACACCGCTGATCGTTTCTCTTTAACCACTTAAAACTA	
VCAN/S	GTGTAGGATAAAGAATTTAGG	
DCK/F1	GTAAGAGGGAATTAGAGTTGGAGG	GGGYGGGTYGTTT
DCK/F2	AGAGGCTAGAGGGTAGTTAGGGA	
DCK/R	CCCCCCTAAACCAATCCTAA	
DCK/RU	GGGACACCGCTGATCGTTTACCCCCCTAAACCAATCCTAA	
DCK/S1	GGAGAGAGGTTTGAGGA	
FGFR2/F	GGGTTTTAAGAGTATTTTAGGTTG	GTTGYGGAGGAGYGYGGGTATGAYG
FGFR2/RU	GGGACACCGCTGATCGTTTCAACTCCCCAACAAATTTAATAA	
FGFR2/S	GGGTTTTAAGAGTATTTAG	
FHIT/F	ATTGTTTTAGTTTTGTGTTTAGATA	TTYGTTTTTATYGTGGGGTTTTTTGGGA
FHIT/RU	GGGACACCGCTGATCGTTTACAAAAACAAAAATCCTATAA	
FHIT/S	GATAAGYGGTTATTTTGGGT	
HAND1/FU2	GGGACACCGCTGATCGTTTAGTAGGGTAAGGTTGAAATGAG	CRACRACCTCRCCRCRAC
HAND1/R2	TCCCCCTTTTCTTATCC	
HAND1/S2	AAATTAACAAATCTTTAACTT	
HCN2/F	GGAGAGTAAGGTTAAGGGTAG	TTYGAAYGGYAGT
HCN2/RU	GGGACACCGCTGATCGTTTACAAAACACACCTTAAAC	
HCN2/S	GAGTAAGGTTAAGGGTAG	
HRK/F	TAGGGGTAAGGGTATATGAT	YGTGTTTGTGTTTTYGT
HRK/RU	GGGACACCGCTGATCGTTTCCACAACACAAAAAACTTAATATC	
HRK/S	GGGTAAGGGTATATGAT	
MLH3/F1	GGGTTTTTTTGGTGTTAGAAGAGATT	YGTATGYGTATTAGTTGTTTTAGTTA
MLH3/F2	AGAAGAGAGTATTTGGATTTGAGGT	
MLH3/R	ACCCATAAATACCAATCCTTAACATAAACAC	
MLH3/RU	GGGACACCGCTGATCGTTTACCCATAAATACCAATCCTTAACATAAACAC	
MLH3/S1	GTTAGGTAGAGAGATATTAGTG	
MSH2/F1	TTAGGGGGAAGTGGGAGGG	TYGATATGGYGGTGTAGT
MSH2/F2	GAGGGGAGGAGGGAAATAGTTT	
MSH2/R	ACCCCCCTCTCCAACACAA	
MSH2/RU	GGGACACCGCTGATCGTTTACCCCCCTCTCCAACACAA	
MSH2/S1	AGGAGGTGAGGAGGT	
MSH6/F1	GTTTGTTGATTGTTATTGGGGTT	TTTYGTTYGATAGAA
MSH6/F2	GAGGTAGATGAGGTGTTTTAGGA	
MSH6/R	CCTAACCCAAACCTTATTAACATCACT	
MSH6/RU	GGGACACCGCTGATCGTTTACCTAACCCAAACCTTATTAACATCACT	
MSH6/S1	GATGAGGTGTTTTAGGAG	
NKX2-3/F	GGGGAGTGTTTATAGTTA	ATYGTGTTTG
NKX2-3/RU	GGGACACCGCTGATCGTTTACTACCAATCCAACCTTTTAACAAA	
NKX2-3/S	GGAGGAGATTAGATTGG	
TP73/F	AGAAGGTTGTAGGTAGGAGGTTTAGG	GGGTYGGGGGGYGGTTYGG
TP73/RU	GGGACACCGCTGATCGTTTCAACCCCAATCCCAAAA	
TP73/S	GTTGTAGGTAGGAGGTTTA	
PAX5/F	GTTAGGAGAGAGGAAAGGGTAGG	GGGTAGGYGTTGYGYGGTYGGG
PAX5/RU	GGGACACCGCTGATCGTTTACCCACCCAAACTAAACCT	
PAX5/S	GTTAGGAGAGAGGAAA	

Table S7. continuation

RAS2/F1	TAGGATAGTAGAGAGGTTAGAGTGTGTAGAG	YGTTTTGGYGG
RAS2/F2	GTTGTAGAGTTTGAGGGGTAAGATTTGTAA	
RAS2/R	AACACACCCCTTCCAACCT	
RAS2/RU	GGGACACCGCTGATCGTTA AACACACCCCTTCCAACCT	
RAS2/S1	GTTGGGTTGAGGGGA	
RB1/F	TTAGGGGAAGTTGAAATTTATTTGT	TTTTTGTAAYGGGAGTYGGG
RB1/RU	GGGACACCGCTGATCGTTA ATCCTACTCTAAATCCTCCTCAAAA	
RB1/S	TAGGGGAAGTTGAAATTA	
RUNX3-F	GGGTATTTTTATTTTATTGT	YGAYGGGTAGYGTTTTG
RUNX3-R	ACAACCCCAACTTCTCTA	
RUNX3-RU	GGGACACCGCTGATCGTTA CAACCCCAACTTCTCTA	
RUNX3-S	GTATTTATTTTGAAGG	
SIN3A/F	GGGAGGGAAAATTATAAGATAGGTT	GAYGGGYGGGTATTGYTTTTGTATTGGAA
SIN3A/RU	GGGACACCGCTGATCGTTA TCAACTTCTTCCAATACAAA	
SIN3A/S	TGGTTGGTTGAGTTTTT	
TAF11/F1	GAGGGAGATTAGTGGGTATAGTTTATTA	GAGATYGYGGAGATGTTTGAGGTAGAAGTT
TAF11/F2	ATAGGATTGGAGGGGAGAGGA	
TAF11/R	TCCCAACTTCTACCTCAAAACATCT	
TAF11/RU	GGGACACCGCTGATCGTTA TCCCAACTTCTACCTCAAAACATCT	
TAF11/S1	GATTGGAGGGGAGAG	
TGFBR2/F1	TTGGAAGTAGAGGAGAGGGAGAAG	TYGYGYGTAYGGAG
TGFBR2/F2	GGGAGGAGAGAGGTTTTGTTAGTTG	
TGFBR2/R	CACCCAAAAAACTACACAAAAATCC	
TGFBR2/RU	GGGACACCGCTGATCGTTA CACCCAAAAAACTACACAAAAATCC	
TGFBR2/S1	GAGTTGAAGTTGAGTGAGTTAT	
TOP2A/F1	GAGATGTTAGGGTAGAGAGGAGTTTT	GTGAYGGTYGTGAAGGGGTTTAAG
TOP2A/F2	AGTAGATTTGTAATGGTGATTTTTATGGT	
TOP2A/R	TTTCAAAATCTTAAACCCCTTCA	
TOP2A/RU	GGGACACCGCTGATCGTTA TTTCAAAATCTTAAACCCCTTCA	
TOP2A/S1	TTGTAATGGTGATTTTTATG	
VHL/FU	GGGACACCGCTGATCGTTA AGGTTAAGGTTGTAGTGAGTTAAGTT	ACTCTATCRCCRAACT
VHL/R	CCAACTCCTTACTATATTTCCACTATAAAA	
VHL/S	TTTTTTTAAACAAAATCTC	

*F, forward primer; R, reverse primer; S, sequencing primer.

**sequence in blue is biotinylated universal adaptor primer.

*** Each investigated cytosine is represented by Y (C/T, forward strand) or R (G/A, reverse strand) polymorphisms.

Table S8. Primers for COBRA method used in this study

Primer Name*	Primer Sequence	Restriction enzyme
ABL1/F	TTAATAAAGGGTTYGGAGAG	TaqI
ABL1/R	CTAAAATAAAAATAAACAAACCTACA	
AIP/F	TTTTTTTGATAGGTTGGGGTTTA	HpyCH4IV
AIP/R	ACTTCRACAACCTCCTAACACCTC	
BNIP3/F	TTYGGTYGGAGGAATTTATAGGGTAG	RsaI
BNIP3/R	CCCTCRCCCACCRCAAAAC	
CABIN1/F	TYGGTGTGTTGTAAGGGTTGGAAT	RsaI
CABIN1/R	CCCRAACTAAAACTAACTCRTAA	
CACNA1G/F	GAYGGYGTAGTAGTTATTTTGTT	BstU I
CACNA1G/R	CATCACCACCCCTCACTTTAC	
MME/F	TTYGGTTTATGTTTGAGTT	BstUI
MME/R	CCCTTTAAACCTTTCTCCCT	
FOXN3/F	GAGTAGATTAGGAGGTAGGAAATGAG	HpyCH4IV
FOXN3/R	CTAAAACCCAACAAAACCTCAA	
CLCN6/F	AGGGGGTTTTTGTTGTTGT	TaqI
CLCN6/R	CTAACTACRATCCCCAAACCC	
DAPK1/F	TAGATTTGTGTTGYGAGTTGT	BstUI
DAPK1/R	ATCCCCATTAACCRCTACC	
DPH2/F	TGGTTTGGGATTTAGGTTGAA	Hpy188I
DPH2/R	TCCCTTCAACCATCCTAAC	
ESR1/F	GGTTTTGAGTTTTTGTGTTT	BstUI
ESR1/R	AACTTACTACTATCCAAATACACCTC	
FDX1/F	GTTGGTTTYGTTTTATGGGAT	Hpy188I
FDX1/R	ACRAAAAAACCTACTAAAAACAC	
FLT1/F	AAAATCAATCCTCTATTAATAAAAA	HpyCH4IV
FLT1/R	AAAAATAATTATCTCTAACTAAAACC	
FOXP2/F	GGAAGGAGAGAGATGGAAAGG	HpyCH4IV
FOXP2/R	CAAACCCACCCACAAACAC	
GAS8/F	TGGYGTTATGGTGAGTAGGG	Hpy188I
GAS8/R	ACTCTCAATTCCAAAAACCCAA	
GATA2/F	GTGGGGAGGGGAGGGATT	HpyCH4IV
GATA2/R	CCCAACTCTACCCTATAAACCC	
GNDF/F	GTTTAATAGGTTAGGGGTAAGAGTT	Fnu4HI
GNDF/R	CAACCCTATCAACCCCCC	
GPX1/F	AAGTATTGTAAGGGGAGGTTAGTAGG	Hpy188I
GPX1/R	CACTCTCCAACCTTTTCCTAA	
GYS1/F	GGGGTYGAGTAGTTGGTGTT	BstUI
GYS1/R	AAAAACRCCAATAAACTCAATAA	
HMGA2/F	AGGATAATTGTTAGGGGTGATTT	BstBI
HMGA2/R	ACRCCTAACTACTAACCAAAAAACA	
ITGAV/F	TTTAGGTTTAGGAGTTGGGGG	RsaI
ITGAV/R	ACRCCCTTCTACAAAAAAAAC	
KCNK6/F	GAAGGAATGAGGGGAAGAGAG	TaqI
KCNK6/R	ACAACAAATACCCCAAAAAA	
ZNF160/F	GGYGGTTTTGTTTATGTT	BstUI
ZNF160/R	ACAAAAAACACTCATACRCCA	
MGMT/F	TTGGTAAATTAAGGTATAGAGTTTT	TaqI
MGMT/R	CTAAAACAATCTACRCATCCTC	
MKLN1/F	GTGGGAAATTTTGAGATAGGATA	BstUI
MKLN1/R	AACRTAACCTACTCTTAAAAAACCC	
NEK4/F	GGGTTTTATTTTGTAGATTGTTTG	Hpy99I
NEK4/R	ACRCTAAAAACATACCCCTAACCC	

Table S8. continuation

NSD1/F	GTTYGTATTTTAGTGTAATGAGGGTT	Hpy188III
NSD1/R	CTTACRACCACCACCATCTTC	
P14/F	TTTYGGGGYGGAGATGGGT	TaqI
P14/R	ATCACCAAAAACCTACRCACCATATTC	
P16/F	GGTTGTTTTYGGTTGGTGTITT	EcoRV
P16/R	ACCCTATCCCCTCAAATCCTCTAAAA	
CDKN1C/F	GGTTGGGYGTTTTATAGGTTA	EcoRI
CDKN1C/R	CCTAACTATCCRATAATAAACTCTTC	
PDE8A/F	GGTTTGGGGGGGTTTTTT	NruI
PDE8A/R	TACRCTCTCCTTACCCAAAAAC	
PIK3R3/F	AGGTGGGTGGTGATTGGTTG	BstUI
PIK3R3/R	TCTTTTCCCTTCTACCTTTATCC	
PIP4K2B/F	TTAGAAGGGTTGGAGGGAAGG	BstUI
PIP4K2B/R	TACAACCCTAAACCCTTACCCC	
PPP1R1A/F	TTGTTAAGGGAGGATTAGGTAGG	Hpy188III
PPP1R1A/R	CATAAAACAAAACAACACCCC	
PXN/F	GTTGGTTTTGGGGTTTTTT	Hpy188III
PXN/R	CRACCTCAACTCCTATCCCTAC	
RAD23A/F	TGTTTATAAAGGTTAAGGAGGGAGT	HpyCH4IV
RAD23A/R	AATCCTCACACAACATACAACCTCA	
TRIM27/F	GGTTCGTTGTTTTTGAGAGGTAT	HpyCH4IV
TRIM27/R	CCTAACCCTAAAACCTAACCCTC	
PDLIM4/F	GTTTATTAGGYGGAAGTTTTAGG	HpyCH4IV
PDLIM4/R	AACCAATCCAAACRCACAA	
RYR1/F	GGAGTTTTGGTTAATGGGGT	BstUI
RYR1/R	ATCGAAAACCTCTAAAAACACCA	
SHFM1/F	YGGTTGTTTTTTTTTGATATT	HhaI
SHFM1/R	CTCCAATATCCCAACATTCTCT	
STAT3/F	GGGGATTGTTTTTTATTTTGT	HhaI
STAT3/R	CAAAAAACCTCTAAAACCTTAAAAA	
TDG/F	GAAAATATTGGAGATAAAAGTTGA	BstBI
TDG/R	ATACCCAAACAATAACTAAACTCA	
TEAD2/F	AGAGTTTTTATTTTATTGAGGTT	RsaI
TEAD2/R	AATCRAAAATCCTACTTTAAAAATAT	
TEC/F	GGYGTTTTTATTTGGTTGAAG	HpyCH4IV
TEC/R	ACCTACCACCCAAAACCCA	
TOP1/F	GAGAGGAAGGGGAGGTTGT	HpyCH4IV
TOP1/R	CCAAACRAACCTATCTTTACATC	
VRK2/F	AGTTAGGTAGGTTTATGGGAGGG	TaqI
VRK2/R	CCGAACCTCAAATCCCC	

*F, forward primer; R, reverse primer.

SUPPLEMENTAL REFERENCES

The following papers were cited in the column “Source” of Table S4 and are numbered as presented.

1. Aguilera O, et al. 2006. Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 DKK-1 gene in human colorectal cancer. *Oncogene* **25**: 4116-4121.
2. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP. 1998. Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Res* **58**: 5489-5494.
3. Ai L, et al. 2006. Inactivation of Wnt inhibitory factor-1 WIF1 expression by epigenetic silencing is a common event in breast cancer. *Carcinogenesis* **27**: 1341-1348.
4. Anupam K, Tusharkant C, Gupta SD, Ranju R. 2006. Loss of disabled-2 expression is an early event in esophageal squamous tumorigenesis. *World J Gastroenterol* **12**: 6041-6045.
5. Banno K, et al. 2007. Epigenetic inactivation of the CHFR gene in cervical cancer contributes to sensitivity to taxanes. *Int J Oncol* **31**: 713-720.
6. Bastian PJ, et al. 2007. Prognostic value of CpG island hypermethylation at PTGS2, RAR-beta, EDNRB, and other gene loci in patients undergoing radical prostatectomy. *Eur Urol* **51**: 665-674.
7. Chang HC, Cho CY, Hung WC. 2007. Downregulation of RECK by promoter methylation correlates with lymph node metastasis in non-small cell lung cancer. *Cancer Sci* **98**: 169-173.
8. Chen H, et al. 2006. Aberrant methylation of RASGRF2 and RASSF1A in human non-small cell lung cancer. *Oncol Rep* **15**: 1281-1285.
9. Choi JH, et al. 2007. Site-specific methylation of CpG nucleotides in the hTERT promoter region can control the expression of hTERT during malignant progression of colorectal carcinoma. *Biochem Biophys Res Commun* **361**: 615-620.
10. Furuta J, et al. 2006. Silencing of Peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. *Cancer Res* **66**: 6080-6086.

11. Henrique R, et al. 2005. MT1G hypermethylation is associated with higher tumor stage in prostate cancer. *Cancer Epidemiol Biomarkers Prev* **14**: 1274-1278.
12. Hiranuma C, et al. 2004. Hypermethylation of the MYOD1 gene is a novel prognostic factor in patients with colorectal cancer. *Int J Mol Med* **13**: 413-417.
13. Honda H, Pazin MJ, D'Souza T, Ji H, Morin PJ. 2007. Regulation of the CLDN3 gene in ovarian cancer cells. *Cancer Biol Ther* **6**: 1733-1742.
14. Imai T, et al. 2008. Epigenetic inactivation of RASSF2 in oral squamous cell carcinoma. *Cancer Sci* **99**: 958-966.
15. Lv Z, et al. 2006. Promoter hypermethylation of a novel gene, ZHX2, in hepatocellular carcinoma. *Am J Clin Pathol* **125**: 740-746.
16. Ogino S, et al. 2008. LINE-1 hypomethylation is inversely associated with microsatellite instability and CpG island methylator phenotype in colorectal cancer. *Int J Cancer* **122**: 2767-2773.
17. Okochi-Takada E, et al. 2006. Silencing of the UCHL1 gene in human colorectal and ovarian cancers. *Int J Cancer* **119**: 1338-1344.
18. Ordway JM, et al. 2006. Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis* **27**: 2409-2423.
19. Oshiro MM, et al. 2005. Epigenetic silencing of DSC3 is a common event in human breast cancer. *Breast Cancer Res* **7**: R669-680.
20. Qiu GH, et al. 2008. The tumor suppressor gene DLEC1 is frequently silenced by DNA methylation in hepatocellular carcinoma and induces G1 arrest in cell cycle. *J Hepatol* **48**: 433-441.
21. Smith E, et al. 2005. Metallothionien 3 expression is frequently down-regulated in oesophageal squamous cell carcinoma by DNA methylation. *Mol Cancer* **4**: 42.
22. Sogabe Y, et al. 2008. Epigenetic inactivation of SFRP genes in oral squamous cell carcinoma. *Int J Oncol* **32**: 1253-1261.

23. Tang D, Sivko GS, DeWille JW. 2006. Promoter methylation reduces C/EBPdelta CEBPD gene expression in the SUM-52PE human breast cancer cell line and in primary breast tumors. *Breast Cancer Res Treat* **95**: 161-170.
24. Terasawa K, et al. 2006. Epigenetic inactivation of TCF2 in ovarian cancer and various cancer cell lines. *Br J Cancer* **94**: 914-921.
25. Wang LH, et al. 2007. Inactivation of SMAD4 tumor suppressor gene during gastric carcinoma progression. *Clin Cancer Res* **13**: 102-110.
26. Woodson K, et al. 2006. CD44 and PTGS2 methylation are independent prognostic markers for biochemical recurrence among prostate cancer patients with clinically localized disease. *Epigenetics* **1**: 183-186.
27. Xi S, et al. 2006. Decreased STAT1 expression by promoter methylation in squamous cell carcinogenesis. *J Natl Cancer Inst* **98**: 181-189.
28. Ying J, et al. 2007. Tumor-specific methylation of the 8p22 tumor suppressor gene DLC1 is an epigenetic biomarker for Hodgkin, nasal NK/T-cell and other types of lymphomas. *Epigenetics* **2**: 15-21.
29. Yuan Y, et al. 2008. Reduction of GSTP1 expression by DNA methylation correlates with clinicopathological features in pituitary adenomas. *Mod Pathol* **21**: 856-865.