

**Figure S1** A prognostic risk model constructed from the 3 methylation-driven genes. (A) The upper and lower panels represent the risk score model and heat map of gene expression, respectively. (B) Univariate Cox regression analyses between risk scores and clinicopathological characteristics in BRCA based on the TCGA database. BRCA, breast invasive carcinoma; TCGA, The Cancer Genome Atlas.

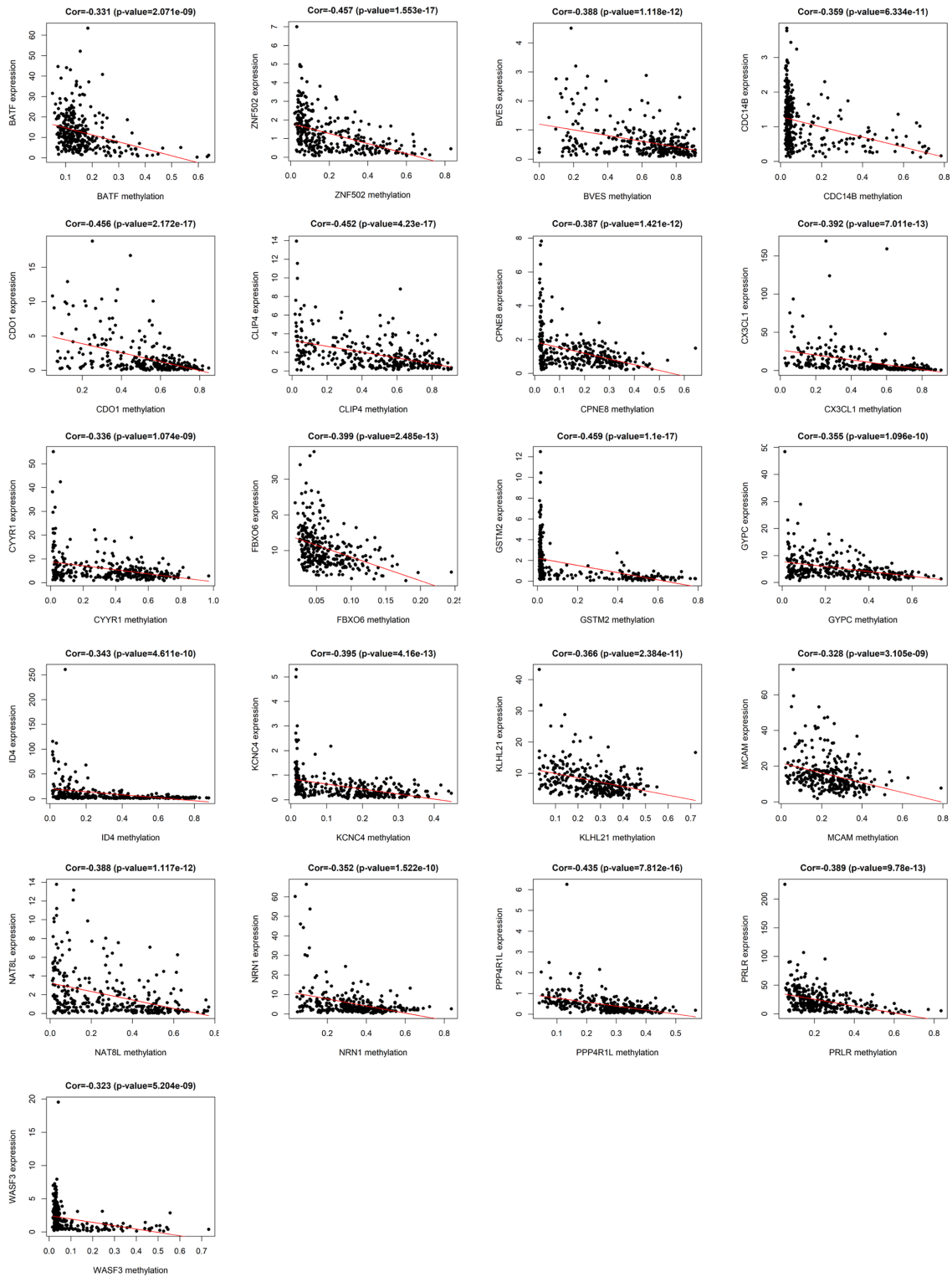
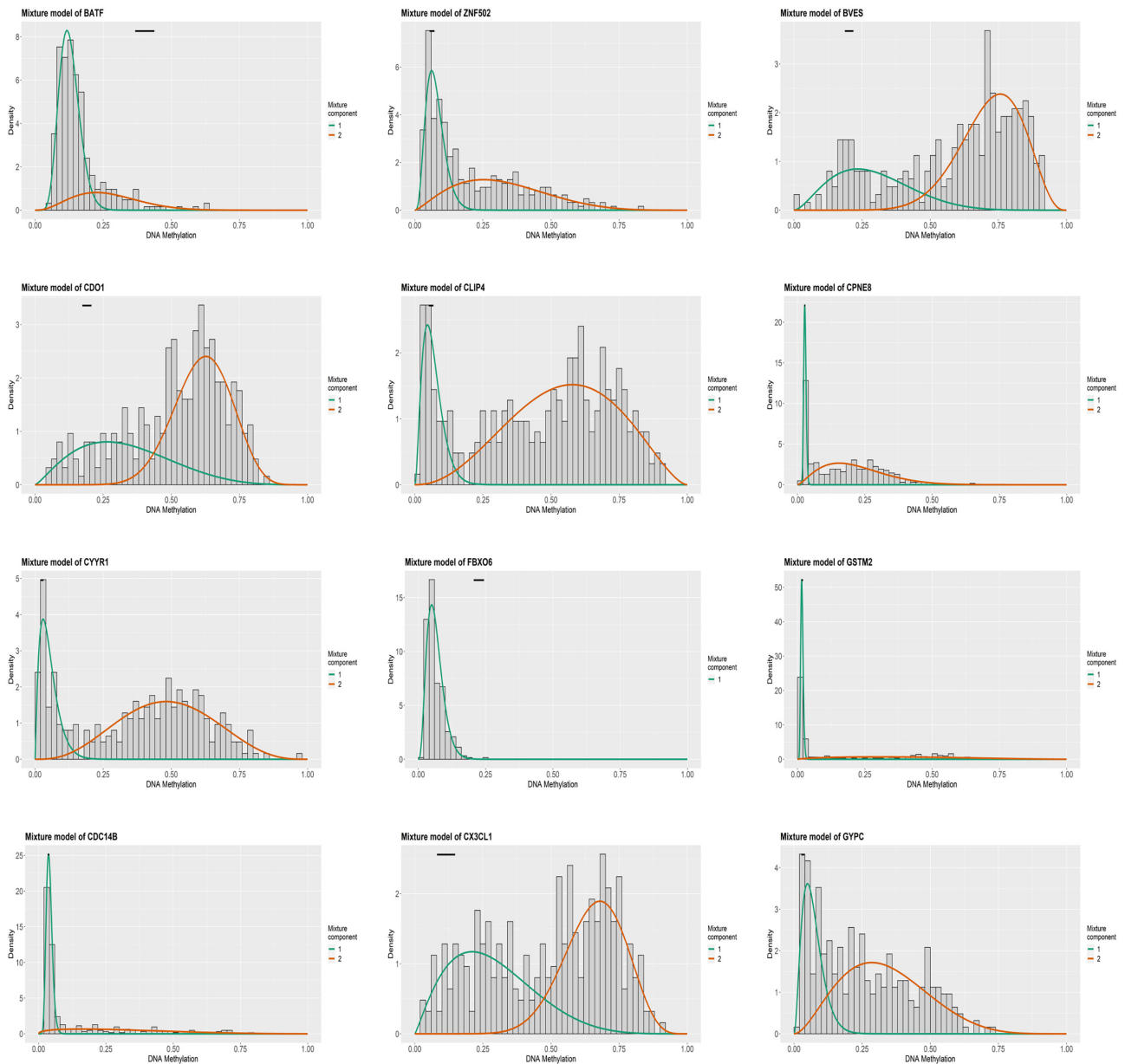
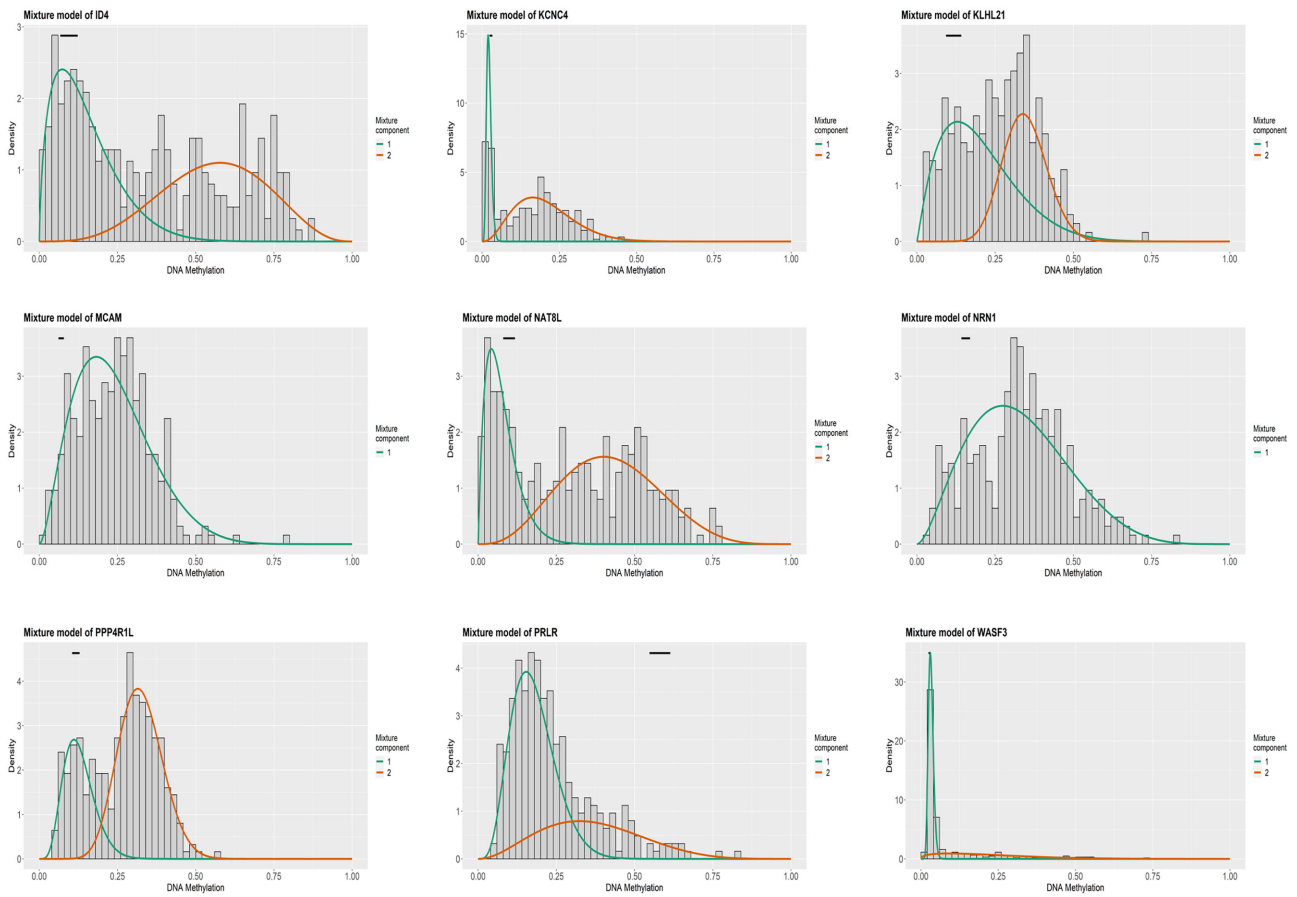


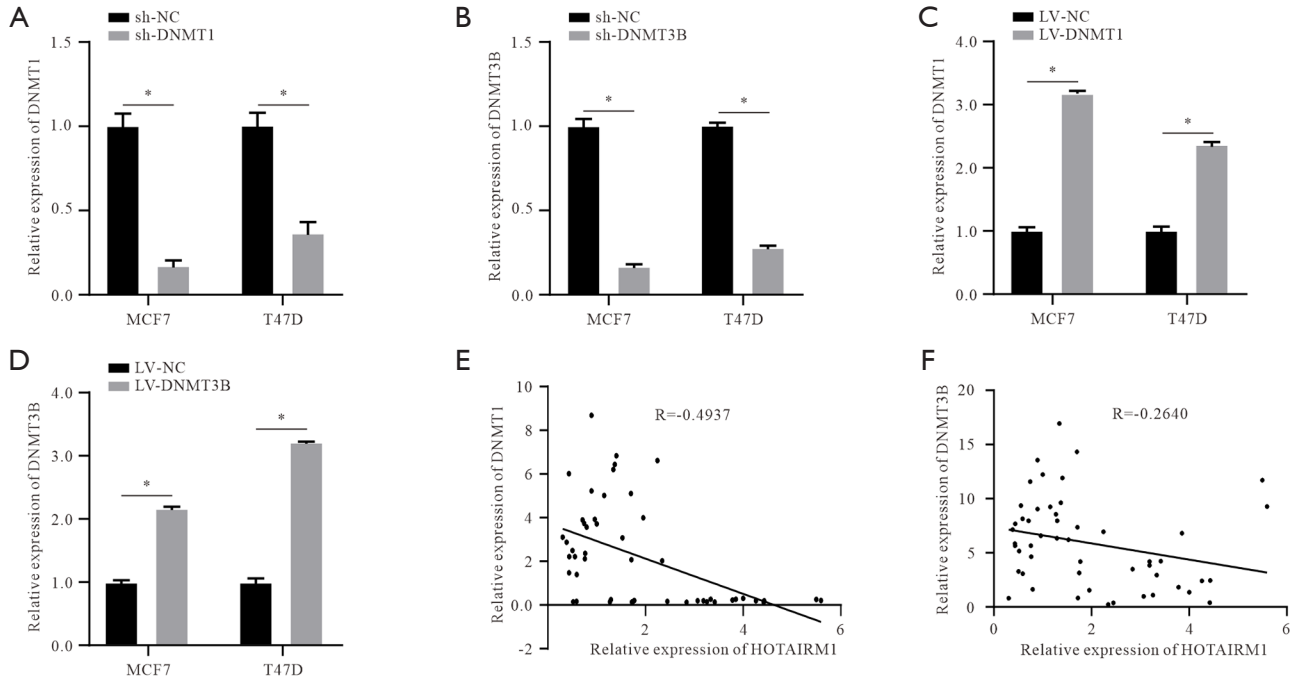
Figure S2 Correlation between methylation levels and the expression of methylation-driven genes.



**Figure S3** Degree of methylation *BATF*, *ZNF502*, *BVES*, *CDO1*, *CLIP4*, *CPNE8*, *CYR1*, *FBXO6*, *GSTM2*, *CDC14B*, *CX3CL1*, and *GYPC* between BRCA patients and normal subjects. Curves 1, and 2 indicate the methylation degree in promoter regions, while the black line above the figure denotes the distribution of methylation levels in normal subjects. BRCA, breast invasive carcinoma.



**Figure S4** Degree of methylation *ID4*, *KCNC4*, *KLHL21*, *MCAM*, *NAT8L*, *NRN1*, *PPP4R1L*, *PRLR*, and *WASF3* between BRCA patients and normal subjects. Curves 1, and 2 indicate the methylation degree in promoter regions, while the black line above the figure denotes the distribution of methylation levels in normal subjects. BRCA, breast invasive carcinoma.



**Figure S5** *HOTAIRM1* expression is regulated by *DNMT1* and *DNMT3B*. (A-D) qRT-PCR-based validation of transfection efficiency after knockdown or overexpression of *DNMT1* and *DNMT3B*. (E) qRT-PCR results showing the correlation of expression patterns between *DNMT1* and *HOTAIRM1*. (F) qRT-PCR results showing the correlation between the expression of *DNMT3B* and *HOTAIRM1*. Values followed by  $P < 0.05$  were considered statistically significant ( $*P < 0.05$ ). NC, negative control; qRT-PCR, quantitative real-time polymerase chain reaction.

**Table S1** List of primers used for qRT-PCR

Genes	Forward primers (5'-3')	Reverse primers (5'-3')
<i>HOTAIRM1</i>	CCCACCGTTC AATGAAAG	GTTTCAAACACCCACATTC
<i>DNMT1</i>	AGGCGGCTCAAAGATTTGGAA	GCAGAAATTCGTGCAAGAGATTC
<i>DNMT3B</i>	AGGGAAGACTCGATCCTCGTC	GTGTGTAGCTTAGCAGACTGG
<i>GAPDH</i>	CTGGGCTACACTGAGCACC	AAGTGGTCGTTGAGGGCAATG
<i>U6</i>	ATTGGAACGATACAGAGAAGATT	GGAACGCTTCACGAATTTG

qRT-PCR, quantitative real-time polymerase chain reaction.

**Table S2** Target sequences of shRNAs used for gene knockdown in breast cancer cells

shRNA	Sequence (5'-3')
<i>HOTAIRM1</i> -shRNA	UCAAUGAAAGAUGAACUGGTT
<i>DNMT1</i> -shRNA	GAAGAGACGTAGAGTTACA
<i>DNMT3B</i> -shRNA	AATTAATAAAGATGACGGATGCCTAGAGTCTCTTGAAGTCTAGGCATCCGTCATCTCG

shRNA, short hairpin RNA.

**Table S3** List of antibodies used for immunohistochemistry and western blotting in breast cancer

Reagent	Source	Identifier
Anti- <i>DNMT1</i>	Abcam	Cat#ab188453
Anti- <i>DNMT3B</i>	Abcam	Cat#ab79822
Anti- <i>SYN2</i>	Abcam	Cat#ab76494
Anti- <i>BCAS1</i>	Abcam	Cat#ab106661
Anti- <i>ALDOC</i>	Abcam	Cat#ab115212
Anti- <i>cyclin E1</i>	Abcam	Cat#ab238081
Anti- <i>cyclin D1</i>	Abcam	Cat#ab16663
Anti- <i>E-cadherin</i>	Abcam	Cat#ab40772
Anti- <i>Vimentin</i>	Abcam	Cat#ab92547
Anti- <i>N-cadherin</i>	Abcam	Cat#ab98952
Anti- <i>Actin</i>	Abcam	Cat#ab8226

**Table S4** Differential methylation-driven gene expression screening

Gene	Normal mean	Tumor mean	Log <sub>2</sub> FC	P value	Cor	Cor P value
<i>FBXO6</i>	0.224784454	0.063448768	-1.824878214	5.16E-17	-0.398688477	2.49E-13
<i>PRLR</i>	0.581113147	0.238593848	-1.284262251	1.58E-15	-0.389371718	9.78E-13
<i>BATF</i>	0.402611351	0.160279088	-1.328801636	3.78E-15	-0.330957916	2.07E-09
<i>GYPE</i>	0.032675096	0.254327294	2.960422899	7.30E-15	-0.354802761	1.10E-10
<i>BCAS1</i>	0.45092556	0.183273106	-1.298894192	1.32E-14	-0.339775751	7.19E-10
<i>CYYR1</i>	0.025197102	0.366317786	3.861766013	6.91E-14	-0.336465455	1.07E-09
<i>MCAM</i>	0.07000091	0.238288463	1.767263521	7.83E-14	-0.327507098	3.11E-09
<i>CX3CL1</i>	0.114725528	0.487511945	2.087251117	9.96E-14	-0.391662689	7.01E-13
<i>HOTAIRM1</i>	0.022892996	0.221642758	3.275258023	2.51E-13	-0.368699673	1.76E-11
<i>CDO1</i>	0.190193575	0.514537485	1.435807674	5.76E-13	-0.455526134	2.17E-17
<i>SYN2</i>	0.036116546	0.433624697	3.585715077	2.24E-12	-0.388872907	1.05E-12
<i>BVES</i>	0.202469899	0.58967092	1.542202609	4.70E-12	-0.388446322	1.12E-12
<i>CLIP4</i>	0.059932548	0.458103315	2.934261391	5.16E-12	-0.451805598	4.23E-17
<i>CPNE8</i>	0.026749376	0.168588893	2.655932341	6.42E-11	-0.386778431	1.42E-12
<i>PPP4R1L</i>	0.11688142	0.261411253	1.161275637	8.88E-10	-0.434978657	7.81E-16
<i>NRN1</i>	0.155825902	0.330357012	1.084090902	2.33E-09	-0.352237685	1.52E-10
<i>KLHL21</i>	0.117275066	0.258511945	1.140334635	2.37E-08	-0.366433384	2.38E-11
<i>ID4</i>	0.095035194	0.34655469	1.866549256	2.82E-08	-0.343398505	4.61E-10
<i>WASF3</i>	0.026023665	0.086954199	1.740431524	1.26E-07	-0.323044123	5.20E-09
<i>KCNC4</i>	0.028492988	0.153199585	2.426733589	1.73E-07	-0.39521981	4.16E-13
<i>ZNF502</i>	0.063274687	0.206485276	1.706338544	3.08E-07	-0.457379614	1.55E-17
<i>NAT8L</i>	0.098869887	0.296370985	1.58380113	1.42E-05	-0.388454334	1.12E-12
<i>ALDOC</i>	0.031806323	0.255893594	3.008158533	0.000457881	-0.339359412	7.57E-10
<i>CDC14B</i>	0.036848676	0.125814535	1.771613927	0.000705105	-0.359038265	6.33E-11
<i>GSTM2</i>	0.017708769	0.180491308	3.349393547	0.00158673	-0.459272787	1.10E-17

FC, fold change.

**Table S5** KEGG signaling pathways

ID	Description	Bg ratio	P value	P.adjust	Q value	Gene ID	Count
hsa00430	Taurine and hypotaurine metabolism	11/8,102	0.01751977	0.257701576	0.256987721	CDO1	1
hsa00030	Pentose phosphate pathway	30/8,102	0.047115348	0.257701576	0.256987721	ALDOC	1
hsa00051	Fructose and mannose metabolism	33/8,102	0.051712384	0.257701576	0.256987721	ALDOC	1
hsa00250	Alanine, aspartate and glutamate metabolism	37/8,102	0.057809923	0.257701576	0.256987721	NAT8L	1
hsa00270	Cysteine and methionine metabolism	50/8,102	0.077377873	0.257701576	0.256987721	CDO1	1
hsa05144	Malaria	50/8,102	0.077377873	0.257701576	0.256987721	GYPC	1
hsa04060	Cytokine-cytokine receptor interaction	295/8,102	0.079125031	0.257701576	0.256987721	PRLR/CX3CL1	2
hsa00480	Glutathione metabolism	57/8,102	0.087758405	0.257701576	0.256987721	GSTM2	1
hsa00010	Glycolysis/gluconeogenesis	67/8,102	0.102400815	0.257701576	0.256987721	ALDOC	1
hsa04917	Prolactin signaling pathway	70/8,102	0.106751046	0.257701576	0.256987721	PRLR	1
hsa04520	Adherens junction	71/8,102	0.108196793	0.257701576	0.256987721	WASF3	1
hsa00982	Drug metabolism—cytochrome P450	72/8,102	0.109640379	0.257701576	0.256987721	GSTM2	1
hsa01524	Platinum drug resistance	73/8,102	0.111081808	0.257701576	0.256987721	GSTM2	1
hsa01230	Biosynthesis of amino acids	75/8,102	0.113958206	0.257701576	0.256987721	ALDOC	1
hsa00980	Metabolism of xenobiotics by cytochrome P450	78/8,102	0.118256697	0.257701576	0.256987721	GSTM2	1
hsa00983	Drug metabolism - other enzymes	80/8,102	0.121111655	0.257701576	0.256987721	GSTM2	1
hsa05204	Chemical carcinogenesis	83/8,102	0.125378096	0.257701576	0.256987721	GSTM2	1
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	89/8,102	0.133853682	0.257701576	0.256987721	BATF	1
hsa04350	TGF-beta signaling pathway	94/8,102	0.140858686	0.257701576	0.256987721	ID4	1
hsa04666	Fc gamma R-mediated phagocytosis	97/8,102	0.145036548	0.257701576	0.256987721	WASF3	1
hsa05231	Choline metabolism in cancer	98/8,102	0.146424995	0.257701576	0.256987721	WASF3	1
hsa04061	Viral protein interaction with cytokine and cytokine receptor	100/8,102	0.149195649	0.257701576	0.256987721	CX3CL1	1
hsa04066	HIF-1 signaling pathway	109/8,102	0.161561162	0.262272819	0.261546301	ALDOC	1
hsa04668	TNF signaling pathway	112/8,102	0.165645991	0.262272819	0.261546301	CX3CL1	1
hsa01200	Carbon metabolism	118/8,102	0.173760594	0.264116102	0.263384479	ALDOC	1
hsa04110	Cell cycle	124/8,102	0.181802293	0.265711044	0.264975002	CDC14B	1
hsa05418	Fluid shear stress and atherosclerosis	139/8,102	0.201591522	0.280648341	0.279870922	GSTM2	1
hsa04550	Signaling pathways regulating pluripotency of stem cells	143/8,102	0.206793514	0.280648341	0.279870922	ID4	1
hsa04630	JAK-STAT signaling pathway	162/8,102	0.231078452	0.287779491	0.286982318	PRLR	1
hsa05225	Hepatocellular carcinoma	168/8,102	0.238603598	0.287779491	0.286982318	GSTM2	1
hsa04530	Tight junction	169/8,102	0.23985116	0.287779491	0.286982318	BVES	1
hsa04141	Protein processing in endoplasmic reticulum	171/8,102	0.242340624	0.287779491	0.286982318	FBXO6	1
hsa04062	Chemokine signaling pathway	192/8,102	0.268029485	0.306264561	0.305416183	CX3CL1	1
hsa05130	Pathogenic Escherichia coli infection	197/8,102	0.274026186	0.306264561	0.305416183	WASF3	1
hsa05163	Human cytomegalovirus infection	225/8,102	0.306778007	0.333073264	0.332150624	CX3CL1	1
hsa05132	Salmonella infection	249/8,102	0.333759619	0.35230182	0.351325915	WASF3	1
hsa04080	Neuroactive ligand-receptor interaction	341/8,102	0.428459671	0.440039662	0.438820716	PRLR	1
hsa04151	PI3K-Akt signaling pathway	354/8,102	0.440790397	0.440790397	0.439569371	PRLR	1

KEGG, Kyoto Encyclopedia of Genes and Genomes.



**Table S6** Univariate Cox regression analysis

ID	HR	HR.95L	HR.95H	P value
<i>BCAS1</i>	0.044789404	0.002040809	0.982987878	0.048740769
<i>HOTAIRM1</i>	14.76450385	2.924127423	74.54893115	0.001119041
<i>CDO1</i>	5.278474895	1.054323328	26.42671037	0.042935975
<i>SYN2</i>	4.014728163	1.225469881	13.15254048	0.021687793
<i>NRN1</i>	18.277134	2.725685216	122.5576693	0.002764883
<i>ALDOC</i>	0.214274623	0.056818032	0.808081737	0.022929781

HR, hazard ratio; HR.95L, lower 95% confidence interval; HR.95H, higher 95% confidence interval.