Original Article

A combination of transcriptome and methylation analyses reveals the role of lncRNA *HOTAIRM1* in the proliferation and metastasis of breast cancer

Gui-E Lai[1#], Jian Zhou[2#], Cui-Liu Huang[1#], Cun-Jun Mai[1], Yi-Mei Lai[1], Zhi-Qin Lin[1], Tao Peng[1], Yuan Luo[3], Feng-En Liu[1]

[1]Department of Vascular and Breast Surgery, First Affiliated Hospital of Gannan Medical University, Ganzhou, China; [2]Department of Neurosurgery Center, Zhujiang Hospital, Southern Medical University, Guangzhou, China; [3]Department of Medical Ultrasonics, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

Contributions: (I) Conception and design: GE Lai, J Zhou; (II) Administrative support: FE Liu; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: CJ Mai; (V) Data analysis and interpretation: YM Lai, ZQ Lin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Gui-E Lai. Department of Vascular and Breast Surgery, First Affiliated Hospital of Gannan Medical University, Ganzhou, China. Email: lge0919@126.com.

Background: DNA methylation status is strongly associated with the prognosis of breast invasive carcinoma (BRCA). Elucidating the mechanisms underlying DNA methylation coupled with determining its biological function is imperative to the effective development of treatment and prevention strategies for breast cancer.

Methods: We retrieved transcriptome and DNA methylation profiles of BRCA patients from The Cancer Genome Atlas (TCGA) database, then applied the "limma" package in R software to identify differentially expressed genes (DEGs) and aberrantly methylated genes. Next, we used the "MethylMix" package to screen for methylation-driven genes, and performed univariate and multivariate Cox regression analyses to determine the prognostic value of the methylation-driven genes and clinical characteristics. We validated these findings in 51 breast cancer tissues alongside 51 corresponding normal tissues. Furthermore, we used cell experiments to clarify the biological function and underlying molecular mechanisms of *HOTAIRM1 in vitro*.

Results: A total of 25 methylation-driven genes were identified in the dataset. Results from univariate and multivariate Cox regression showed that *SYN2*, *HOTAIRM1*, *BCAS1*, and *ALDOC* were significantly associated with patient prognosis. Immunohistochemistry and quantitative real-time polymerase chain reaction (qRT-PCR) results showed that the expression levels of *SYN2* and *HOTAIRM1* were negatively correlated with BRCA stage, whereas those of *BCAS1* and *ALDOC* were positively correlated with BRCA stage. Results from *in vitro* experiments showed that knockdown *HOTAIRM1* expression promoted breast cancer cells proliferation, clone formation, and invasion. Up-regulation of *HOTAIRM1* inhibited breast cancer cells proliferation, clone formation, and invasion.

Conclusions: In summary, low *HOTAIRM1* expression is a significant prognostic factor for the survival of BRCA patients and thus could be a potential therapeutic target for the treatment of BRCA.

Keywords: Breast invasive carcinoma (BRCA); DNA methylation; long noncoding RNA *HOTAIRM1* (lncRNA *HOTAIRM1*); metastasis; proliferation

Submitted Feb 21, 2022. Accepted for publication Apr 13, 2022.
doi: 10.21037/gs-22-164
View this article at: https://dx.doi.org/10.21037/gs-22-164

Introduction

The incidence of breast invasive carcinoma (BRCA), the most frequently occurring type of carcinoma among women (1,2), is currently on the rise worldwide (3). Despite the improvements in BRCA therapies over the last decade, metastasis, tumor relapse, and resistance to therapy remain the chief causes of BRCA-related deaths in patients (2,4-6). Understanding the specific molecular mechanisms underlying BRCA development and progression is imperative to the effective management of the disease. DNA methylation, an epigenetic modification process catalyzed by DNA methyltransferases such as *DNMT1*, *DNMT3A*, and *DNMT3B*, can regulate gene expression without changing the DNA sequence (7). Previous studies have shown that DNA methylation predominantly occurs in cytosine-phosphate-guanine (CpG) dinucleotides in mammals, accounting for 70–80% of the entire human genome (8-12). Additional evidence has demonstrated that aberrant DNA methylation regulates abnormal gene expression and malignant phenotypes (13,14), while gene expression levels are negatively correlated with DNA methylation factors, named methylation-driven genes (15). In fact, widespread DNA methylation alterations in normal breast tissue adjacent to cancer that become enriched with breast cancer progression have been identified, suggesting that DNA methylation alterations predate the emergence of breast cancer. Furthermore, DNA methylation status has been found to be strongly associated with the prognosis of BRCA patients (16-18). Previous studies demonstrated that silencing genes such as *PENl*, *BCSGl*, *PLAU*, *IGF*, and *CDN3* in human breast cancer cells was associated with methylation status (19,20). These genes are essential in regulating various processes, including cell proliferation, apoptosis, cell cycle progression, and metastasis (21). Although some progress has been made in the development of therapies against BRCA, there is currently no cause for optimism, mainly owing to the lack of clarity regarding the cause of the disease. Therefore, elucidating the mechanisms underlying DNA methylation coupled with understanding its biological function is imperative to guiding the development of effective treatment and prevention therapies for BRCA patients.

In recent decades, researchers have identified and elucidated the biological functions of long noncoding RNAs (lncRNAs) in various human cancers (22,23). Consequently, much attention has been paid to their functions in tumors, including cell proliferation, differentiation, chromosomal remodeling, epigenetic regulation, and transcription and post-transcriptional modification (24,25). For example, recent evidence has demonstrated that *H19* is more abundant in estrogen receptor (ER)-positive compared with ER-negative breast tumor tissues, where it promotes tamoxifen resistance (26,27).

In the present study, we combined transcriptome and methylation analyses to reveal an aberrant lncRNA, named HOX antisense intergenic RNA myeloid 1 (*HOTAIRM1*), in BRCA tissues. This lncRNA not only acts as a tumor suppressor in various human cancers, but can also regulate papillary thyroid cancer cell proliferation and invasion through the *HOTAIRM1/miR-107/TDG* axis (28). In addition, Kim *et al.* (29) found that *HOXA1* and its neighboring *HOTAIRM1* might serve as potential therapeutic targets for ER^+ breast cancer patients. To date, however, nothing is known regarding the relationship between DNA methylation of the *HOTAIRM1* gene in BRCA. Therefore, additional studies are needed to clarify the role of DNA methylation in regulating *HOTAIRM1* expression. For the first time we found DNA methylation is associated with *HOTAIRM1* using transcriptome and methylation combined analysis in breast cancer, and regulated by *DNMT1* and *DNMT3A*. Results of the present study demonstrated that *HOTAIRM1* is downregulated and hypermethylated in BRCA patients, and plays a critical role in the proliferation and metastasis of breast cancer cells. Notably, this downregulation is a significant prognostic factor for the survival of BRCA patients, and thus may be a promising therapeutic target. We present the following article in accordance with the MDAR reporting checklist (available at https://gs.amegroups.com/article/view/10.21037/gs-22-164/rc).

Methods

Data retrieval

RNA sequencing data, comprising 1,109 BRCA tumor and 113 normal tissues, were retrieved from The Cancer Genome Atlas (TCGA) database (https://tcga-data.nci.nih.gov/tcga). DNA methylation profiles, including 315 BRCA tumor tissues alongside 27 normal tissues, were also downloaded from the TCGA DNA methylation database (Illumina human methylation 27 platform).

Identification of differentially expressed and aberrantly methylated genes

Firstly, we employed the "limma" package, implemented in R software, to identify differentially expressed and aberrantly methylated genes based on the following criteria:

|$\log_2$fold change (FC)| >1, false discovery rate (FDR) <0.05. Next, we used the "MethylMix" package (30) to screen for upregulated and downregulated hypermethylated genes, with the promoter of these genes generally considered a 2,000 bp sequence upstream of the transcription start site (31). Finally, the same package was used to analyze the distribution of the methylation-driven gene promoters.

Determination of the prognostic value

To explore the prognostic value of the methylation-driven genes, we combined gene methylation and expression data with the survival status and time data of patients, then subjected them to univariate and multivariate analyses using Cox proportional hazards regression models. Multivariate analyses were conducted to determine the prognostic value of the methylation-driven genes and clinical characteristics.

Clinical specimens and approval

We collected 51 breast cancer and 51 corresponding normal adjacent tissues from patients at the First Affiliated Hospital of Gannan Medical University, China. The samples were immediately frozen in liquid nitrogen following surgical resection. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The present study was approved by the Ethics Committee of the First Affiliated Hospital of Gannan Medical University (No. LLSC-2022033101). Informed consent was obtained from patients or guardians.

Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted from the collected breast cancer and normal tissues and cells using an RNA isolation kit (Tiangen, Beijing, China) according to the manufacturer's instructions. The RNA was reverse transcribed into complementary DNA (cDNA) using a reverse transcriptase kit (TaKaRa, Dalian, China). The cDNA was subjected to qRT-PCR using the Takara SYBR® Premix Ex Taq™ II kit (TaKaRa, Dalian, China), and qRT-PCR was performed on a 7500 Real-Time PCR System (ABI) targeting genes whose primer sequences are listed in Table S1.

Cell cultures and lentiviral transfection

Human breast cancer cell lines, MCF7 and T47D, were purchased from the Chinese Academy of Sciences Cell Bank (Shanghai, China) and cultured in Dulbecco's Modified Eagle's Medium (DMEM; Corning, VA, USA) supplemented with 10% fetal bovine serum (FBS; Corning, VA, USA).

Constructs for overexpression of *HOTAIRM1*, *DNMT1*, and *DNMT3B* (LV-HOTAIRM1, LV-DNMT1, and LV-DNMT3B) were constructed and transfected in MCF7 and T47D cells using lentiviral overexpression plasmids Ubi-MCS-SV40-puromycin. Cells transfected with LV-NC lentivirus served as a negative control. Next, we used the hU6-MCS-CMV-puromycin plasmid to knock-down *HOTAIRM1*, *DNMT1*, and *DNMT3B* in cells (sh-HOTAIRM1, sh-DNMT1, and sh-DNMT3B) before transfection in MCF7 and T47D cells, with sh-NC cells used as a negative control. All cells were cultured in a humidified incubator with 5% CO_2 and 95% air at 37 ℃ (Thermo Scientific, Waltham, MA, USA). Target sequences for RNA interference are listed in Table S2.

Immunohistochemistry and Western blot (WB) assays

Immunohistochemical experiments were performed using an immunohistochemistry kit (Solarbio, Beijing, China), according to the manufacturer's instructions. Stained specimens were observed and images were captured under a Leica microscope (Leica, Germany). Total proteins were extracted using a total protein extraction kit (KeyGen Biotech, Nanjing, China), and concentrations were determined using the BCA-100 Protein Quantitative Analysis Kit. WB was performed according to standard protocols. WB transfer solution, WB electrophoresis solution, primary antibody diluent, and secondary antibody diluent were purchased from Beyotime Biotechnology (China). All antibodies used for WB are listed in Table S3.

Determination of cell viability and migration

Cell viability was determined using the cell counting kit 8 (CCK8) assay kit, according to the manufacturer's instructions. Briefly, 1×10^3 breast cancer cells were first seeded into 96-well plates after cell transfection, then the CCK8 reagent (Dojindo, Japan) was added to each well followed by a 1 h incubation at 37 ℃ at 5 different time points (0, 1, 2, 3, and 4 days). Cell viability was measured by detecting absorbance at 450 nm and an optical density (OD) value was obtained. For the cell cloning assay, after cell transfection, breast cancer cells clone formation assay

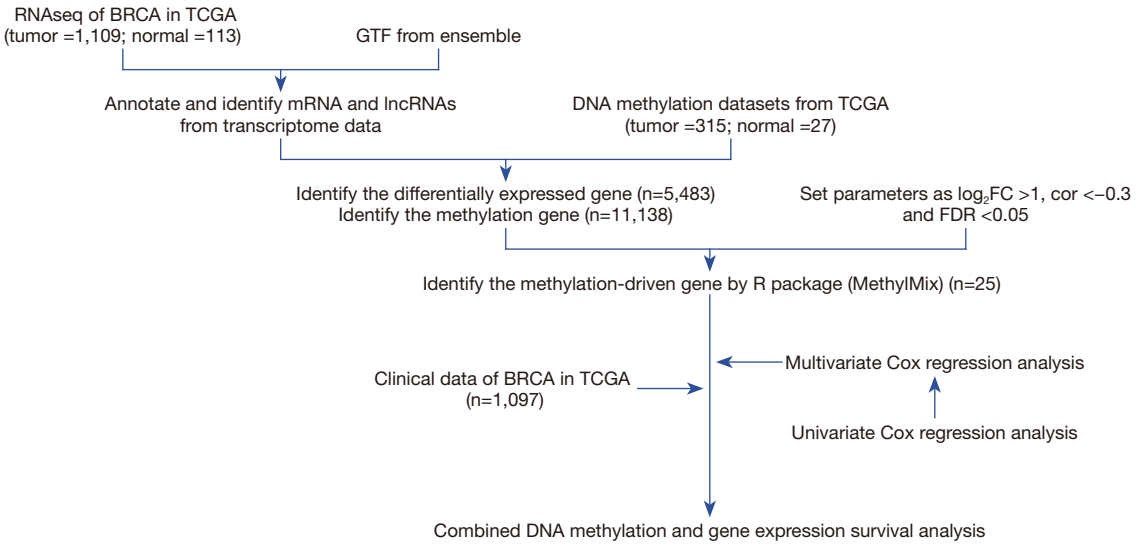


Figure 1 Flowchart showing the process for the identification of differentially methylated genes. BRCA, breast invasive carcinoma; TCGA, The Cancer Genome Atlas; GTF, gene transfer format; mRNA, messenger RNA; lncRNA, long noncoding RNA; FC, fold change; FDR, false discovery rate.

were performed through limiting dilution. Briefly, breast cancer cells were seeded into a 6-well plate and cultured for 14 days in DMEM supplemented with 10% FBS at 37 ℃ in 5% CO_2 humidified air incubators. The cells were then stained with crystal violet and counted. For the cell migration assay, 5×10^4 breast cancer cells were seeded into a transwell chamber, with culture medium containing 20% serum DMEM used in the lower layer and serum-free DMEM used in the upper chamber. Cells were stained with crystal violet and counted after a 48-h incubation.

Statistical analysis

All bioinformatics analyses were performed using packages implemented in R software (unless stated otherwise). Statistical analysis of all experimental data, including Student's *t*-test and Spearman correlations, were performed in GraphPad Prism 8, with P<0.05 considered statistically significant.

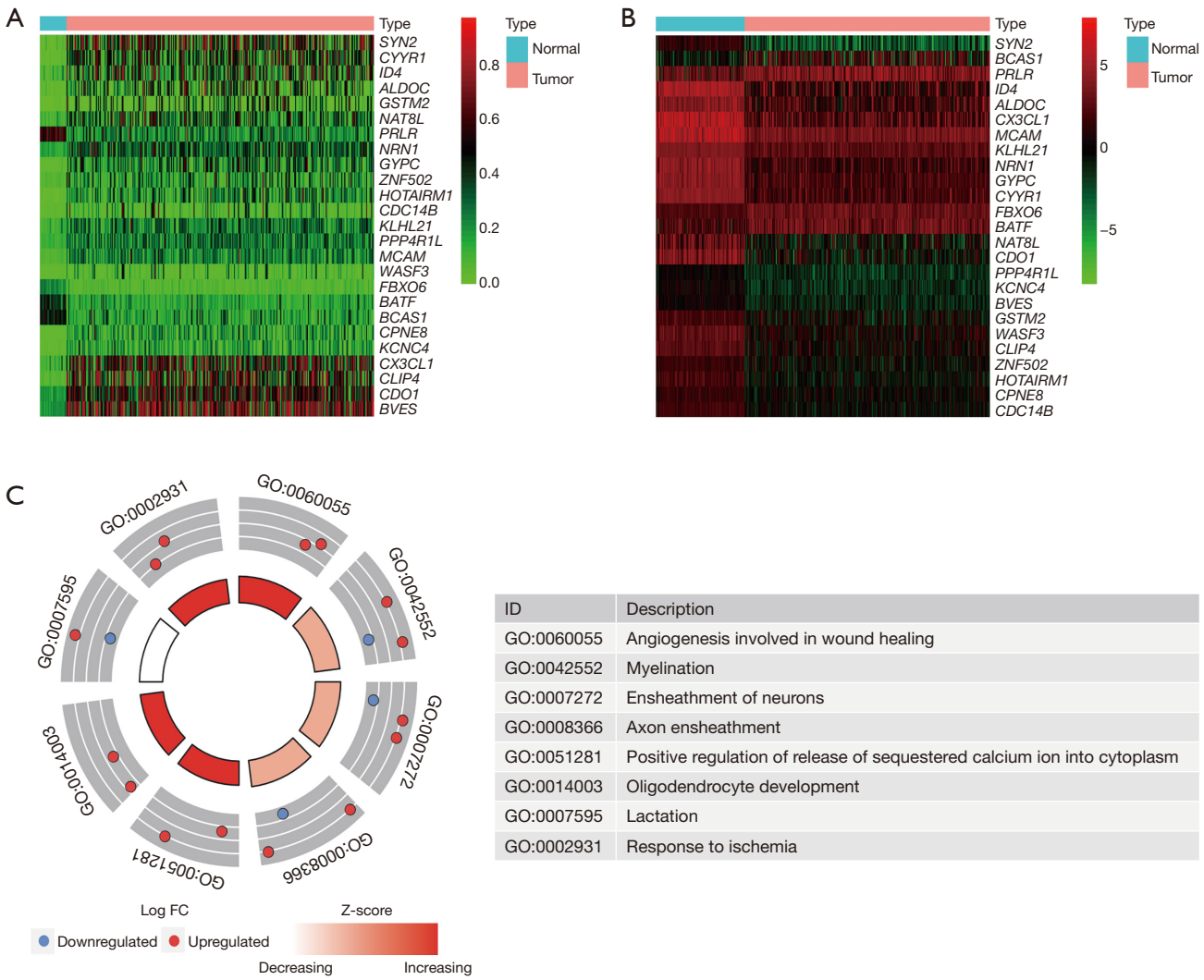
Results

Identification of methylation-driven genes in BRCA

Profiles of methylation-driven genes in BRCA are shown using a flow diagram in *Figure 1*. A total of 5483 differentially expressed genes (DEGs) were identified based on FDR <0.05, $\log_2$FC >1, of which 3,216 and 2,267 were upregulated and downregulated in BRCA, respectively (available online: https://cdn.amegroups.cn/static/public/gs-22-164-01.xls). Since previous studies have demonstrated that overall promoter methylation levels are inversely correlated with gene expression (9-12), we next explored methylation-driven gene patterns by analyzing the DNA methylation data with DEG expression profiling, and identified a total of 25 methylation-driven genes based on FDR <0.05, $\log_2$FC >1, and cor <–0.3 (Table S4). Profiles of differential methylation levels and DEGs in the methylation-driven genes are shown in *Figure 2A,2B*. Next, we employed the "cluster profile", "org.Hs.eg.db", "enrichplot", "ggplot2", and "GOplot" packages in R to further investigate enriched pathways. Results from pathway analysis revealed significant enrichment of 8 Gene Ontology (GO) terms (GO:0060055, GO:0042552, GO:0007272, GO:0008366, GO:0051281, GO:0014003, GO:0007595, GO:0002931) (FDR <0.05) (*Figure 2C*) and 38 Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathways (Table S5). Collectively, these results suggested that gene expression levels were not only closely correlated with DNA methylation levels, but also play crucial roles in BRCA biology.

Construction of a prognostic model based on methylation-driven genes

Next, we analyzed the relationship between gene expression



ID	Description
GO:0060055	Angiogenesis involved in wound healing
GO:0042552	Myelination
GO:0007272	Ensheathment of neurons
GO:0008366	Axon ensheathment
GO:0051281	Positive regulation of release of sequestered calcium ion into cytoplasm
GO:0014003	Oligodendrocyte development
GO:0007595	Lactation
GO:0002931	Response to ischemia

Figure 2 Profiles of methylation-driven genes in BRCA. (A) Heat map showing the differentially methylated genes (P≤0.05 and $\log_2$FC ≥1). (B) Heat map showing DEGs (P≤0.05 and $\log_2$FC ≥1). (C) Top enriched GO terms of peaks in each cluster. Terms with P≤0.05 were considered statistically significant. FC, fold change; GO, Gene Ontology; BRCA, breast invasive carcinoma; DEGs, differentially expressed genes.

and the overall survival of breast cancer patients to understand the clinical relevance of these 25 methylation-driven genes. Firstly, we performed univariate Cox regression analysis on the training set and identified 6 prognosis-related genes (FDR <0.05) (Table S6). Next, we stratified patients into two groups, then subjected them to univariate and multivariate Cox proportional hazard regression analysis (Figure S1A). The resulting Kaplan-Meier survival curves suggested that the model had good performance, with individuals in the high-risk group associated with worse survival outcomes relative to those in the low-risk group (*Figure 3A*). Next, we performed univariate and multivariate Cox regression analyses to determine the relationship between clinicopathological features and the risk scores, and found that risk scores and each clinicopathological feature had good risk predictive ability (Figure S1B). However, results from multivariate analyses indicated that only the risk scores and age had excellent predictive ability for different clinical features (*Figure 3B*).

Further assessment of the association of methylation and expression with patient prognosis, based on Kaplan-Meier

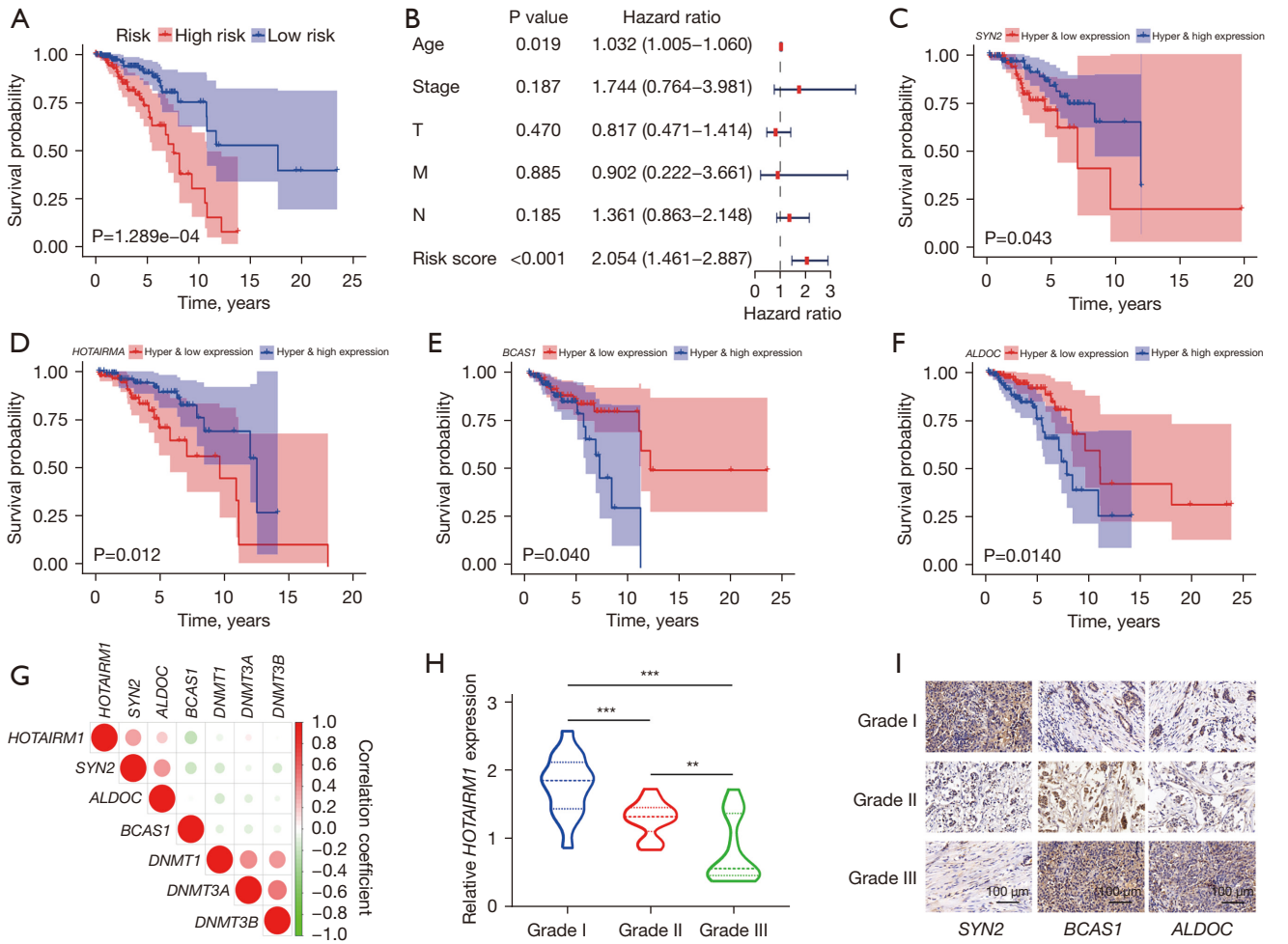


Figure 3 Prognostic model constructed based on the identified methylation-driven genes. (A) Kaplan-Meier curves for patients in the high and low risk groups. (B) Multivariate Cox regression analyses between risk scores and clinicopathological characteristics based on the BRCA database in TCGA. (C-F) Kaplan-Meier curves of the methylation-driven genes. (G) Correlation analysis of the expression of methylation-driven genes and DNA methylation genes. (H) qRT-PCR results targeting the expression of *HOTAIRM1* across different BRCA grades. Statistical significance was considered at **P<0.01, ***P<0.0001. (I) Immunohistochemical staining results showing the expression levels of 3 methylation-driven genes across different BRCA grades. BRCA, breast invasive carcinoma; TCGA, The Cancer Genome Atlas; qRT-PCR, quantitative real-time polymerase chain reaction.

survival analysis, indicated that hypomethylation/high-expression genes (*SYN2* and *HOTAIRM1*) (*Figure 3C,3D*) as well as hypermethylation/low-expression genes (*BCAS1* and *ALDOC*) were significantly associated with patient prognosis (*Figure 3E,3F*). Moreover, DNA methylation transferases were negatively correlated with gene expression (*Figure 3G*). Validation of the 4 methylation-driven genes in clinical BRCA samples, based on qRT-PCR and immunohistochemical analyses, showed that the levels of *SYN2* and *HOTAIRM1* expression were negatively correlated with BRCA stage, whereas those of *BCAS1* and *ALDOC* expression were positively correlated with BRCA stage (*Figure 3H,3I*). Collectively, these results indicated the successful construction of a methylation-driven gene risk prediction model for estimating the risk of BRCA.

Relationship between methylation-driven gene expression and DNA methylation levels

The correlation between the DNA methylation levels and expression of 25 methylation-driven genes in BRCA revealed a negative relationship between them

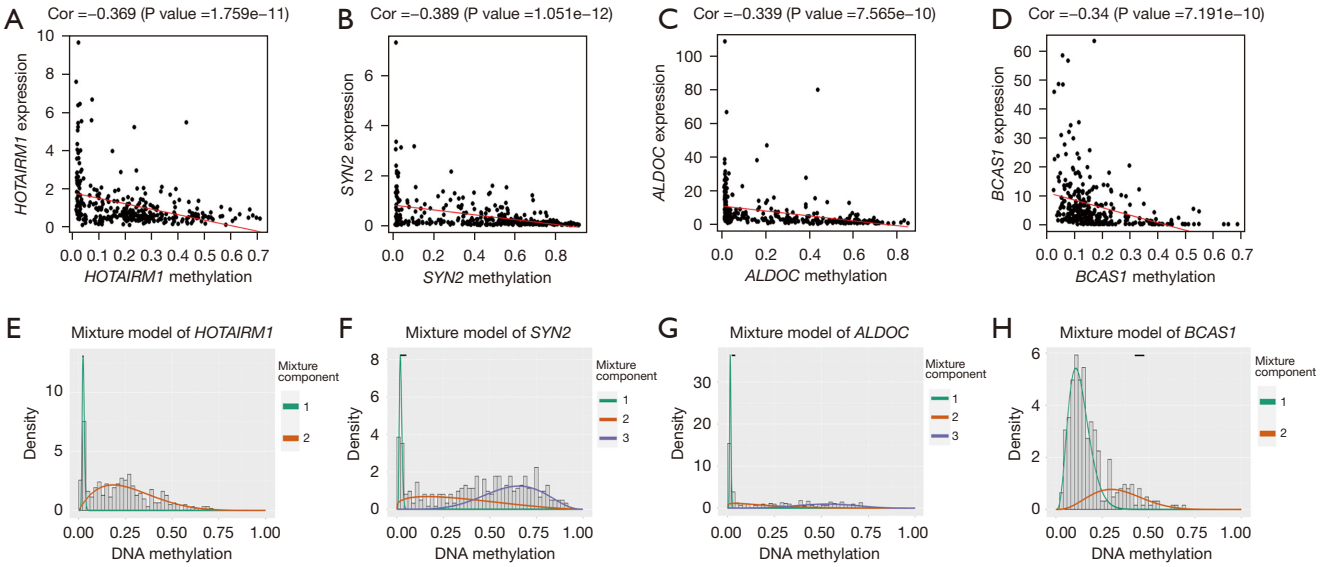


Figure 4 The relationship between gene expression and DNA methylation levels. (A-D) Correlation between methylation levels and gene expression of methylation-driven genes. (E-H) Degree of methylation between BRCA patients and normal subjects. Curves 1, 2, and 3 indicate the degree of methylation in promoter regions, while the black line above the figure denotes the distribution of methylation levels in normal patients. BRCA, breast invasive carcinoma.

(*Figure 4A-4D*, Figure S2). Moreover, results from the analysis of methylation levels of promoters of the 25 methylation-driven genes in BRCA and normal tissues indicated that most of the methylation-driven genes were hypermethylated in BRCA patients and hypomethylated in the normal group. On the other hand, *FBXO6*, *BATF*, *BCAS1*, and *PRLR* were hypomethylated in BRCA patients but hypermethylated in the normal group (*Figure 4E-4H*, Figures S3,S4). Collectively, these results suggested that BRCA patients were associated with high DNA methylation of methylation-driven genes.

Downregulation of HOTAIRM1 promotes the proliferation and migration of BRCA cells

We investigated the role of *HOTAIRM1* in BRCA, owing to the fact that lncRNAs play an essential role in the proliferation and invasion of tumor cells. Firstly, we knocked down the expression of *HOTAIRM1* in MCF7 and T47D cells, and verified the knockdown efficiency via qRT-PCR (*Figure 5A*). Next, we performed CCK8 and cell clonogenic found that downregulation of *HOTAIRM1* promoted both the proliferation and clonal formation ability (*Figure 5B-5D*). Furthermore, we overexpressed *HOTAIRM1* in MCF7 and T47D cells to determine its effect in BRCA, and verified the overexpression efficiency via qRT-PCR (*Figure 5E*). Results suggested that upregulation of *HOTAIRM1* suppressed the proliferation and clonal formation ability of tumor cells (*Figure 5F-5H*). Parallelly, the invasion of MCF7 and T47D cells was promoted by down-regulation of *HOTAIRM1*, also inhibited by up-regulation of *HOTAIRM1* expression (*Figure 5I,5J*). Finally, we performed a WB assay to determine the levels of proliferation-related (*Cyclin E1* and *Cyclin D1*) and invasion-related (Vimentin, N-cadherin, E-cadherin) proteins (*Figure 5K,5L*). These findings suggested that *HOTAIRM1* expression may be closely correlated with the proliferation and invasion of BRCA cells.

HOTAIRM1 expression is regulated by DNMT1 and DNMT3B

Next, we analyzed the correlation between methylation-driven gene expression and 3 methylation transferases (*DNMT1*, *DNMT3A*, and *DNMT3B*) and found that *HOTAIRM1* was negatively correlated with methylation transferases *DNMT1* and *DNMT3B* in BRCA (*Figure 3G*). To understand whether *DNMT1* or *DNMT3B* also regulated *HOTAIRM1* expression, we performed qRT-PCR and WB analyses in MCF7 and T47D cells transfection efficiency of silencing *DNMT1* and *DNMT3B* (*Figure 6A*,

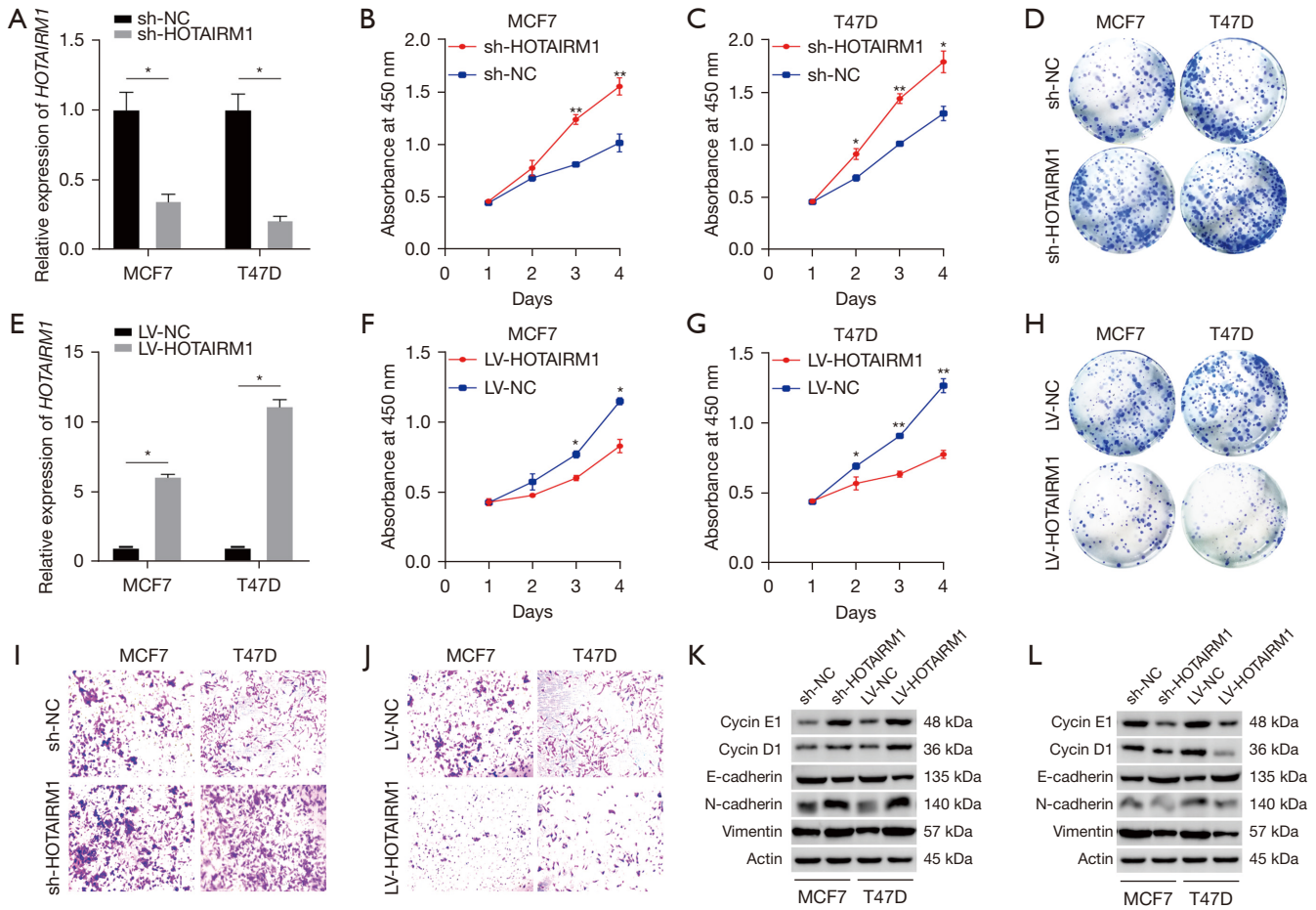


Figure 5 *HOTAIRM1* downregulation promotes the proliferation and migration of BRCA cells. (A) Validation of transfection efficiency after knockdown of *HOTAIRM1* expression via qRT-PCR. (B,C) Profiles of viability of BRCA cells after knockdown of *HOTAIRM1* expression via the CCK8 assay. (D) Clone formation capacity of BRCA cells assessed after knockdown of *HOTAIRM1* expression by the clone formation assay. The cells were stained with crystal violet and the plates were photographed. (E) qRT-PCR-based validation of transfection efficiency following *HOTAIRM1* overexpression. (F,G) CCK8 assay results showing the viability of BRCA cells after *HOTAIRM1* overexpression. (H) Clone formation capacity of BRCA cells was assessed after *HOTAIRM1* overexpression. The cells were stained with crystal violet and the plates were photographed. (I,J) Migration capacity of BRCA cells assessed after changes in *HOTAIRM1* expression by the transwell migration assay. Crystal violet staining and cells were photographed at 200× magnification. (K,L) WB showing the levels of cell cycle- and migration-associated proteins. Data with P<0.05 were considered statistically significant (*P<0.05, **P<0.01). NC, negative control; BRCA, breast invasive carcinoma; qRT-PCR, quantitative real-time polymerase chain reaction; CCK8, cell counting kit 8; WB, western blot.

Figure S5A-S5D). Results showed that *HOTAIRM1* was significantly upregulated in cells that had *DNMT1* or *DNMT3B* knocked down (*Figure 6B,6C*), We also performed qRT-PCR and WB analyses in MCF7 and T47D cells transfection efficiency of overexpressing *DNMT1* and *DNMT3B* (*Figure 6D*), but downregulated in cells overexpressing *DNMT1* and *DNMT3B* (*Figure 6E,6F*). Further evaluation of the association between *HOTAIRM1* and *DNMT1* or *DNMT3B* revealed a negative correlation in BRCA (Figure S5E,S5F). Collectively, these results demonstrated that *DNMT1* and *DNMT3B* regulate the expression of *HOTAIRM1*.

Discussion

BRCA is the most common type of carcinoma among women (1,2,18). The current annual increase in BRCA incidence has necessitated the exploration of the specific

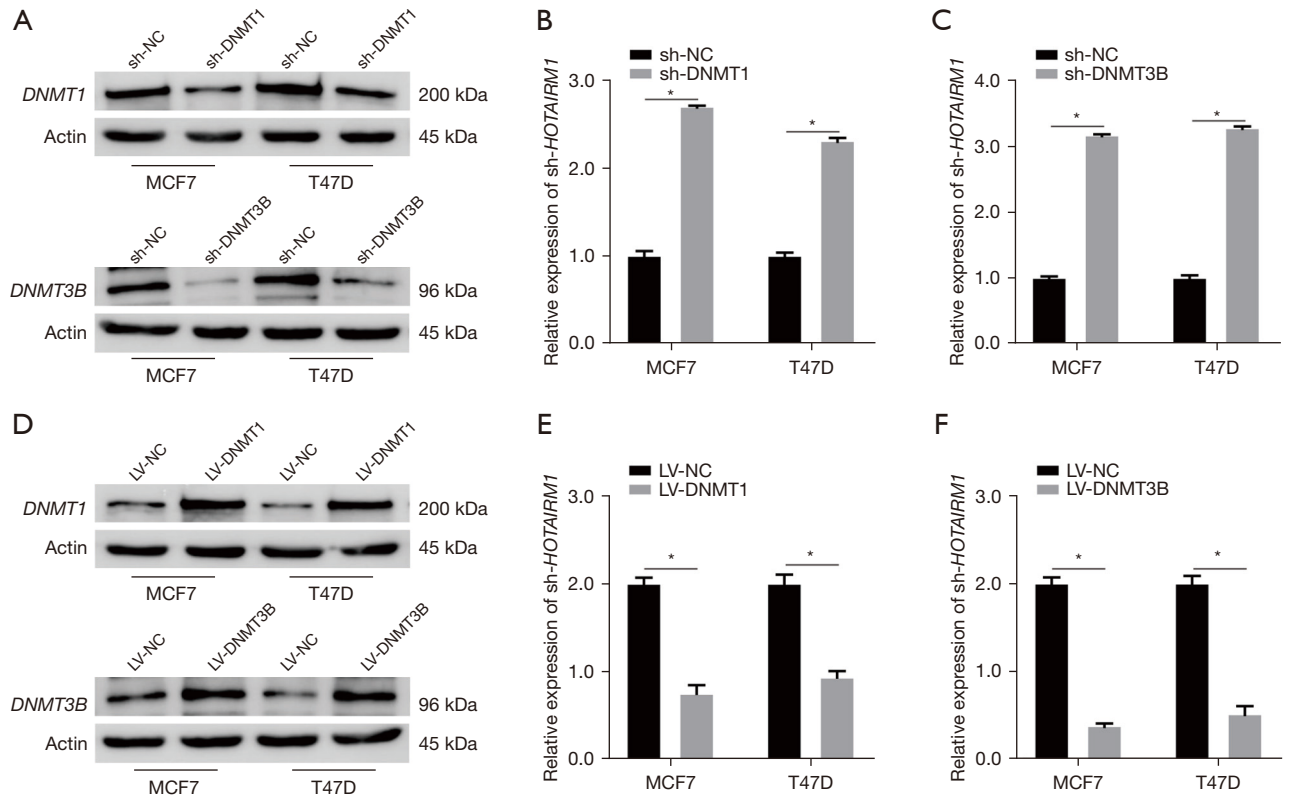


Figure 6 *HOTAIRM1* expression is regulated by *DNMT1* and *DNMT3B*. (A) WBs used to validate transfection efficiency after knockdown of *DNMT1* and *DNMT3B* expression. (B,C) qRT-PCR results showing *HOTAIRM1* expression after knockdown of *DNMT1* and *DNMT3B*. (D) WBs used to validate transfection efficiency after overexpression of *DNMT1* and *DNMT3B*. (E,F) qRT-PCR results showing *HOTAIRM1* expression after overexpressing *DNMT1* and *DNMT3B*. Values followed by P<0.05 were considered statistically significant (*P<0.05). NC, negative control; WB, western blot; qRT-PCR, quantitative real-time polymerase chain reaction.

molecular mechanisms underlying disease development and progression, as this will aid in the identification of effective prognostic biomarkers for predicting the survival rates of patients.

Previous studies have shown that DNA methylation plays a crucial role in regulating gene expression, while abnormal distribution of DNA methylation has been observed in many malignancies, including BRCA (13,14). Moreover, DNA methylation status has been strongly associated with the prognosis of BRCA patients (16,18). In the present study, we used a combination of methylation and transcriptome analyses to identify BRCA-specific diagnostic biomarkers for predicting survival rates.

We identified a total of 25 methylation-driven genes, then explored their clinical relevance by correlating their expression profiles with the overall survival of BRCA patients. Results from univariate and multivariate Cox regression analyses showed that hypomethylation/high-expression genes (*SYN2* and *HOTAIRM1*) as well as hypermethylation/low-expression genes (*BCAS1* and *ALDOC*) were significantly associated with patient prognosis. Additionally, results from distribution analysis of the degree of methylation indicated that most of the methylation-driven genes were hypermethylated in BRCA patients but hypomethylated in subjects in the normal group, while *FBXO6*, *BATF*, *BCAS1*, and *PRLR* were hypomethylated in BRCA patients but hypermethylated in subjects in the normal group. Among the identified methylation-driven genes, only *HOTAIRM1* was a lncRNA. Much research evidence has described the roles played by lncRNAs in tumors, including cell proliferation, differentiation, chromosomal remodeling, epigenetic regulation, and transcription and post-transcriptional modification (22,23,25). For example, Kim *et al.* (29) demonstrated that

HOTAIRM1 promotes tamoxifen resistance by mediating *HOXA1* expression in ER$^+$ breast cancer cells, indicating that it plays a vital role in the breast cancer cells. To date, however, no study has described the relationship between *HOTAIRM1* and DNA methylation in BRCA.

For this reason, much attention has been directed towards *HOTAIRM1*'s role as a prognostic factor. In this research, we found that knockdown *HOTAIRM1* expression promoted breast cancer cells proliferation, clone formation, and invasion. Up-regulation of *HOTAIRM1* inhibited breast cancer cells proliferation, clone formation, and invasion. The expression of *HOTAIRM1* is regulated by diverse and sophisticated mechanisms. Such as, inhibiting KDM6A demethylase represses *HOTAIRM1* transcription, in addition *IRF4* transcriptionally activate *HOTAIRM1* (32,33). We also analyzed the relationship between methylation-driven gene expression and 3 methylation transferases (*DNMT1*, *DNMT3A*, and *DNMT3B*) and found that *HOTAIRM1* was negatively correlated with *DNMT1* and *DNMT3B* in BRCA. Overall, our results indicated that *HOTAIRM1* was not only downregulated and hypermethylated in BRCA, but also plays a critical role in the proliferation and metastasis of breast cancer cells.

In conclusion, downregulation of *HOTAIRM1* is a significant prognostic factor for the survival of BRCA patients, and thus may be a potential therapeutic target for the management of the disease.

Acknowledgments

We thank all of the participants in the present research. We also thank the reviewers for their valuable advice.
Funding: None.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://gs.amegroups.com/article/view/10.21037/gs-22-164/rc

Data Sharing Statement: Available at https://gs.amegroups.com/article/view/10.21037/gs-22-164/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://gs.amegroups.com/article/view/10.21037/gs-22-164/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The present study was approved by the Ethics Committee of the First Affiliated Hospital of Gannan Medical University (No. LLSC-2022033101). Informed consent was obtained from patients or guardians.

References

1. Mannu GS, Wang Z, Broggio J, et al. Invasive breast cancer and breast cancer mortality after ductal carcinoma in situ in women attending for breast screening in England, 1988-2014: population based observational cohort study. BMJ 2020;369:m1570.
2. Markham MJ, Wachter K, Agarwal N, et al. Clinical Cancer Advances 2020: Annual Report on Progress Against Cancer From the American Society of Clinical Oncology. J Clin Oncol 2020;38:1081.
3. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
4. Cuzick J, Sestak I, Forbes JF, et al. Use of anastrozole for breast cancer prevention (IBIS-II): long-term results of a randomised controlled trial. Lancet 2020;395:117-22.
5. Padmanaban V, Krol I, Suhail Y, et al. E-cadherin is required for metastasis in multiple models of breast cancer. Nature 2019;573:439-44.
6. Ahern TP, Broe A, Lash TL, et al. Phthalate Exposure and Breast Cancer Incidence: A Danish Nationwide Cohort Study. J Clin Oncol 2019;37:1800-9.
7. Schmitz RJ, Lewis ZA, Goll MG. DNA Methylation: Shared and Divergent Features across Eukaryotes. Trends Genet 2019;35:818-27.

8. Zhao S, Allis CD, Wang GG. The language of chromatin modification in human cancers. Nat Rev Cancer 2021;21:413-30.
9. Zhou S, Zeng H, Huang J, et al. Epigenetic regulation of melanogenesis. Ageing Res Rev 2021;69:101349.
10. Li Y, Chen X, Lu C. The interplay between DNA and histone methylation: molecular mechanisms and disease implications. EMBO Rep 2021;22:e51803.
11. Martisova A, Holcakova J, Izadi N, et al. DNA Methylation in Solid Tumors: Functions and Methods of Detection. Int J Mol Sci 2021;22:4247.
12. Do WL, Gohar J, McCullough LE, et al. Examining the association between adiposity and DNA methylation: A systematic review and meta-analysis. Obes Rev 2021;22:e13319.
13. Parry A, Rulands S, Reik W. Active turnover of DNA methylation during cell fate decisions. Nat Rev Genet 2021;22:59-66.
14. Bates SE. Epigenetic Therapies for Cancer. N Engl J Med 2020;383:650-63.
15. Sun X, Yi J, Yang J, et al. An integrated epigenomic-transcriptomic landscape of lung cancer reveals novel methylation driver genes of diagnostic and therapeutic relevance. Theranostics 2021;11:5346-64.
16. Győrffy B, Bottai G, Fleischer T, et al. Aberrant DNA methylation impacts gene expression and prognosis in breast cancer subtypes. Int J Cancer 2016;138:87-97.
17. Xu Z, Sandler DP, Taylor JA. Blood DNA Methylation and Breast Cancer: A Prospective Case-Cohort Analysis in the Sister Study. J Natl Cancer Inst 2020;112:87-94.
18. Boyne DJ, O'Sullivan DE, Olij BF, et al. Physical Activity, Global DNA Methylation, and Breast Cancer Risk: A Systematic Literature Review and Meta-analysis. Cancer Epidemiol Biomarkers Prev 2018;27:1320-31.
19. Dietrich D, Lesche R, Tetzner R, et al. Analysis of DNA methylation of multiple genes in microdissected cells from formalin-fixed and paraffin-embedded tissues. J Histochem Cytochem 2009;57:477-89.
20. Yballe CM, Vu TH, Hoffman AR. Imprinting and expression of insulin-like growth factor-II and H19 in normal breast tissue and breast tumor. J Clin Endocrinol Metab 1996;81:1607-12.
21. Li S, Rong M, Iacopetta B. DNA hypermethylation in breast cancer and its association with clinicopathological features. Cancer Lett 2006;237:272-80.
22. Statello L, Guo CJ, Chen LL, et al. Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol 2021;22:96-118.
23. Rinn JL, Chang HY. Long Noncoding RNAs: Molecular Modalities to Organismal Functions. Annu Rev Biochem 2020;89:283-308.
24. Gil N, Ulitsky I. Regulation of gene expression by cis-acting long non-coding RNAs. Nat Rev Genet 2020;21:102-17.
25. Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. Nat Cell Biol 2019;21:542-51.
26. Sun H, Wang G, Peng Y, et al. H19 lncRNA mediates 17β-estradiol-induced cell proliferation in MCF-7 breast cancer cells. Oncol Rep 2015;33:3045-52.
27. Wang J, Xie S, Yang J, et al. The long noncoding RNA H19 promotes tamoxifen resistance in breast cancer via autophagy. J Hematol Oncol 2019;12:81.
28. Li D, Chai L, Yu X, et al. The HOTAIRM1/miR-107/TDG axis regulates papillary thyroid cancer cell proliferation and invasion. Cell Death Dis 2020;11:227.
29. Kim CY, Oh JH, Lee JY, et al. The LncRNA HOTAIRM1 Promotes Tamoxifen Resistance by Mediating HOXA1 Expression in ER+ Breast Cancer Cells. J Cancer 2020;11:3416-23.
30. Cedoz PL, Prunello M, Brennan K, et al. MethylMix 2.0: an R package for identifying DNA methylation genes. Bioinformatics 2018;34:3044-6.
31. Gao C, Zhuang J, Li H, et al. Exploration of methylation-driven genes for monitoring and prognosis of patients with lung adenocarcinoma. Cancer Cell Int 2018;18:194.
32. Li L, Deng J, Huang T, et al. IRF4 transcriptionally activate HOTAIRM1, which in turn regulates IRF4 expression, thereby affecting Th9 cell differentiation and involved in allergic rhinitis. Gene 2022;813:146118.
33. Bah I, Youssef D, Yao ZQ, et al. Inhibiting KDM6A Demethylase Represses Long Non-Coding RNA Hotairm1 Transcription in MDSC During Sepsis. Front Immunol 2022;13:823660.

(English Language Editor: C. Betlazar-Maseh)

Cite this article as: Lai GE, Zhou J, Huang CL, Mai CJ, Lai YM, Lin ZQ, Peng T, Luo Y, Liu FE. A combination of transcriptome and methylation analyses reveals the role of lncRNA *HOTAIRM1* in the proliferation and metastasis of breast cancer. Gland Surg 2022;11(5):826-836. doi: 10.21037/gs-22-164

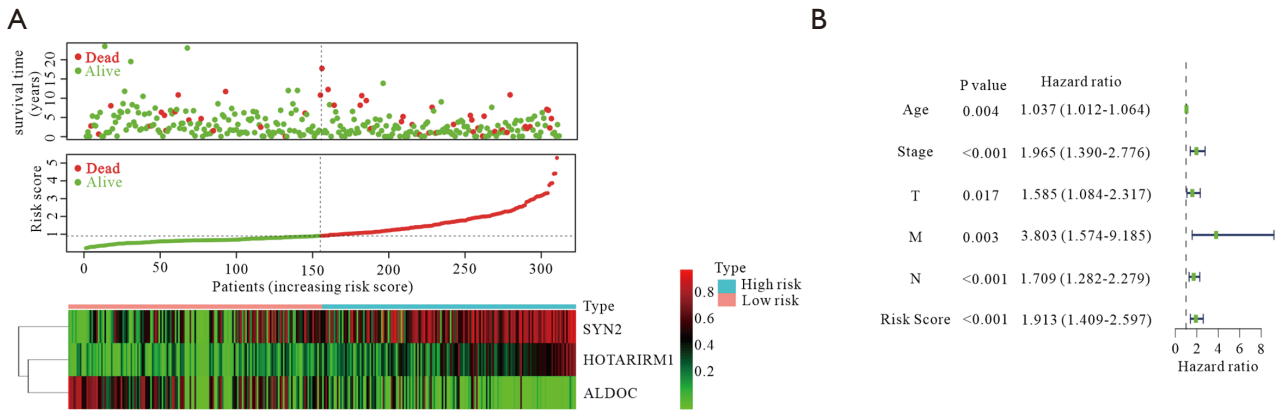


Figure S1 A prognostic risk model constructed from the 3 methylation-driven genes. (A) The upper and lower panels represent the risk score model and heat map of gene expression, respectively. (B) Univariate Cox regression analyses between risk scores and clinicopathological characteristics in BRCA based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

https://dx.doi.org/10.21037/gs-22-164

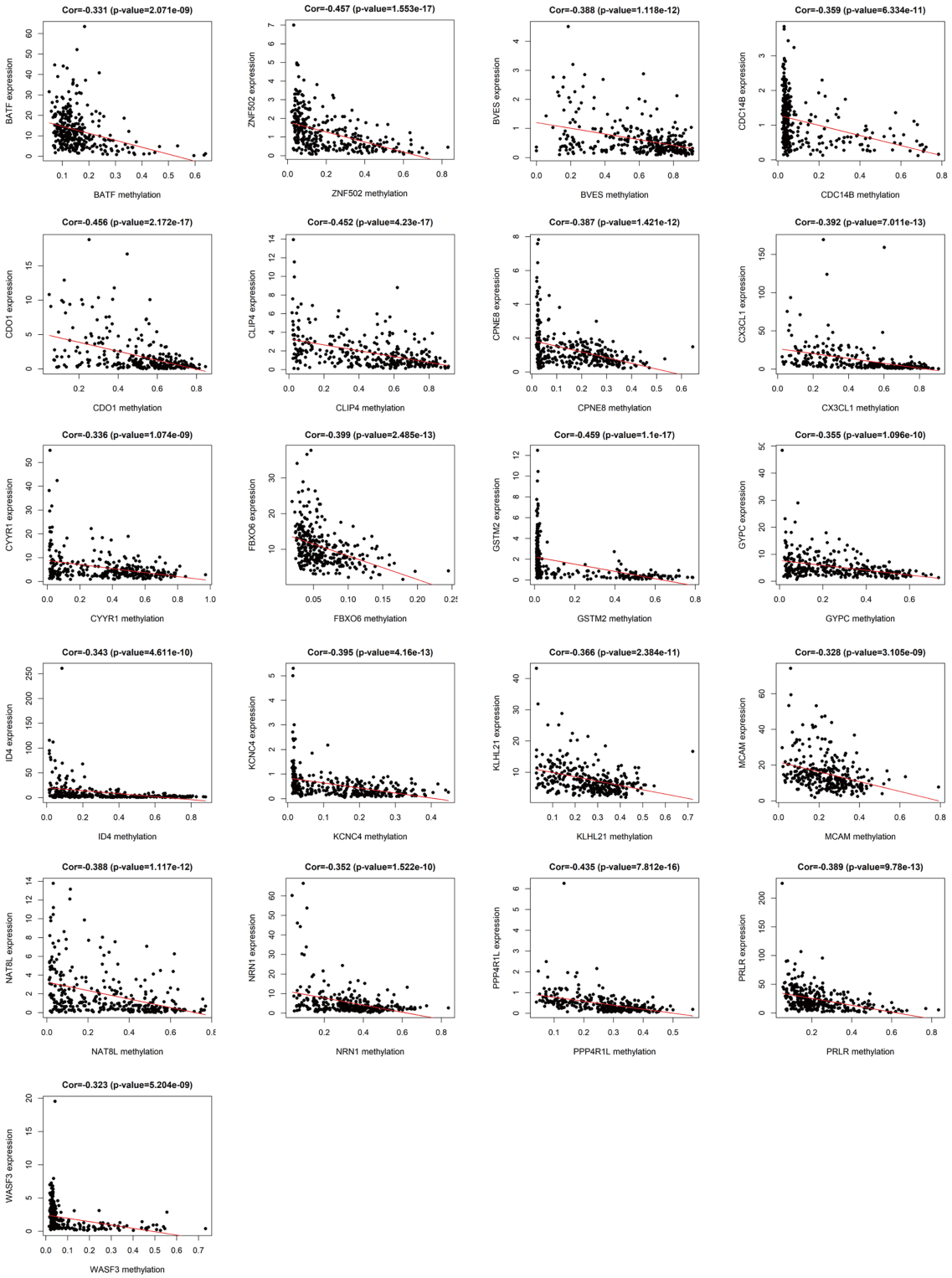


Figure S2 Correlation between methylation levels and the expression of methylation-driven genes.

https://dx.doi.org/10.21037/gs-22-164

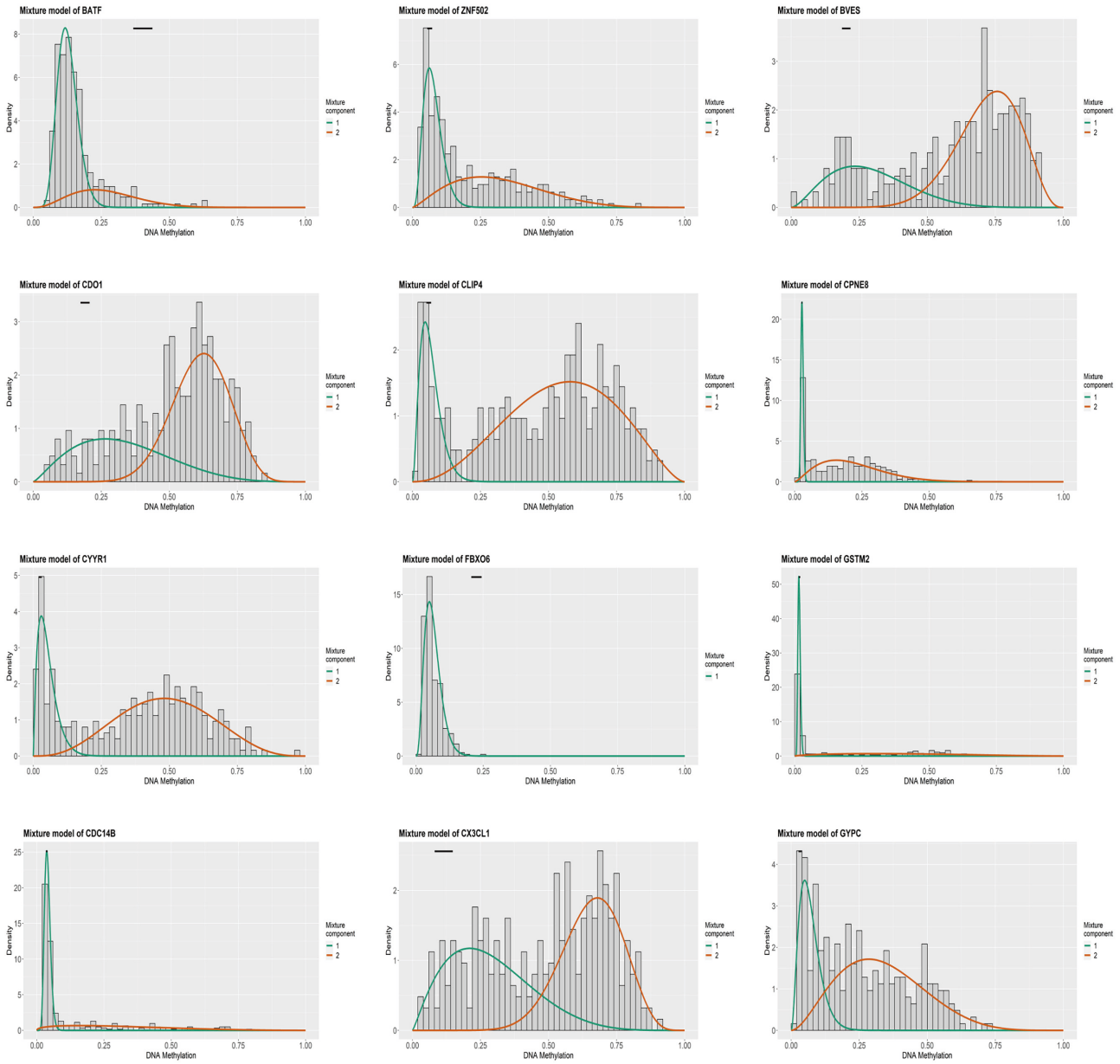


Figure S3 Degree of methylation *BATF*, *ZNF502*, *BVES*, *CDO1*, *CLIP4*, *CPNE8*, *CYYR1*, *FBXO6*, *GSTM2*, *CDC14B*, *CX3CL1*, and *GYPC* between BRCA patients and normal subjects. Curves 1, and 2 indicate the methylation degree in promoter regions, while the black line above the figure denotes the distribution of methylation levels in normal subjects. BRCA, breast invasive carcinoma.

https://dx.doi.org/10.21037/gs-22-164

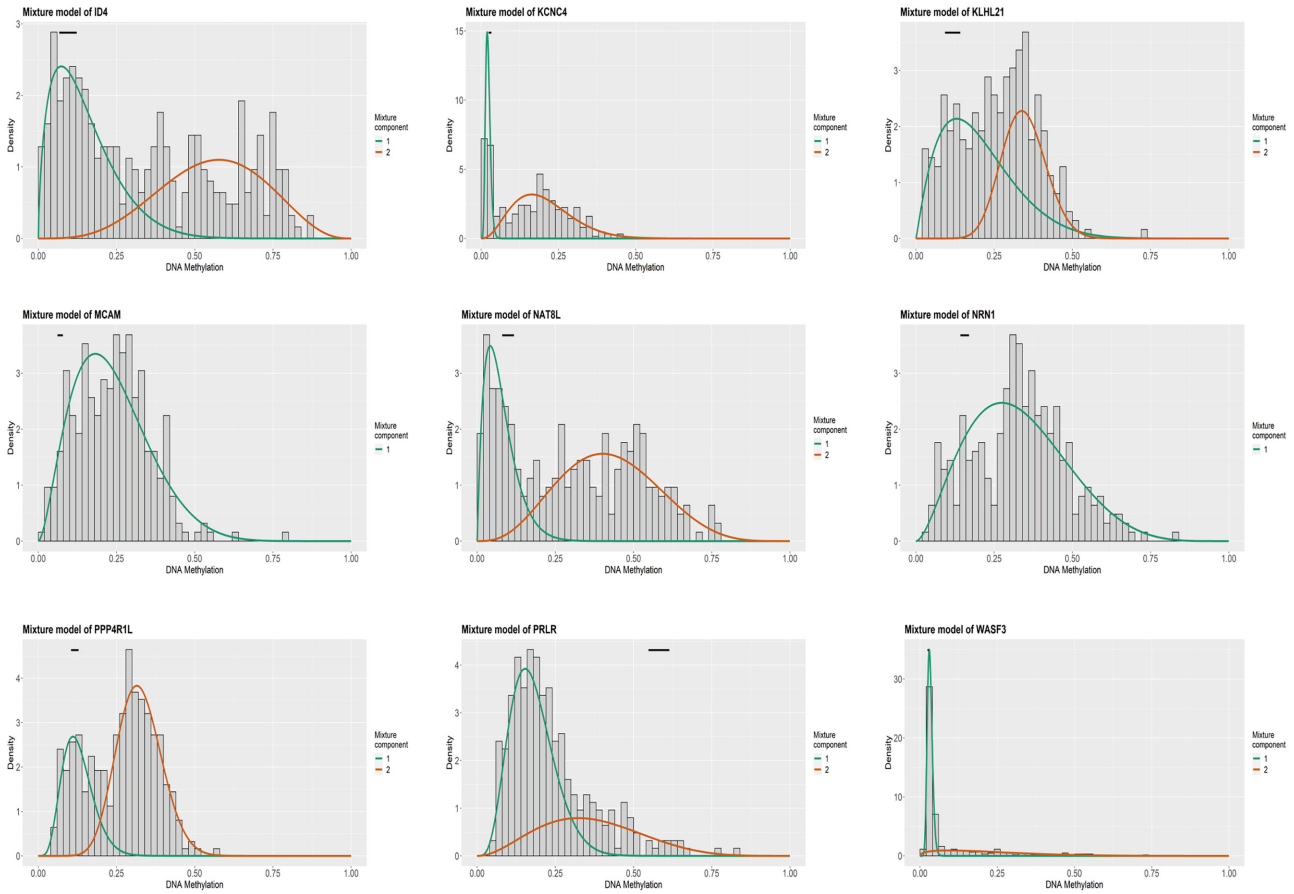


Figure S4 Degree of methylation *ID4*, *KCNC4*, *KLHL21*, *MCAM*, *NAT8L*, *NRN1*, *PPP4R1L*, *PRLR*, and *WASF3* between BRCA patients and normal subjects. Curves 1, and 2 indicate the methylation degree in promoter regions, while the black line above the figure denotes the distribution of methylation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

https://dx.doi.org/10.21037/gs-22-164

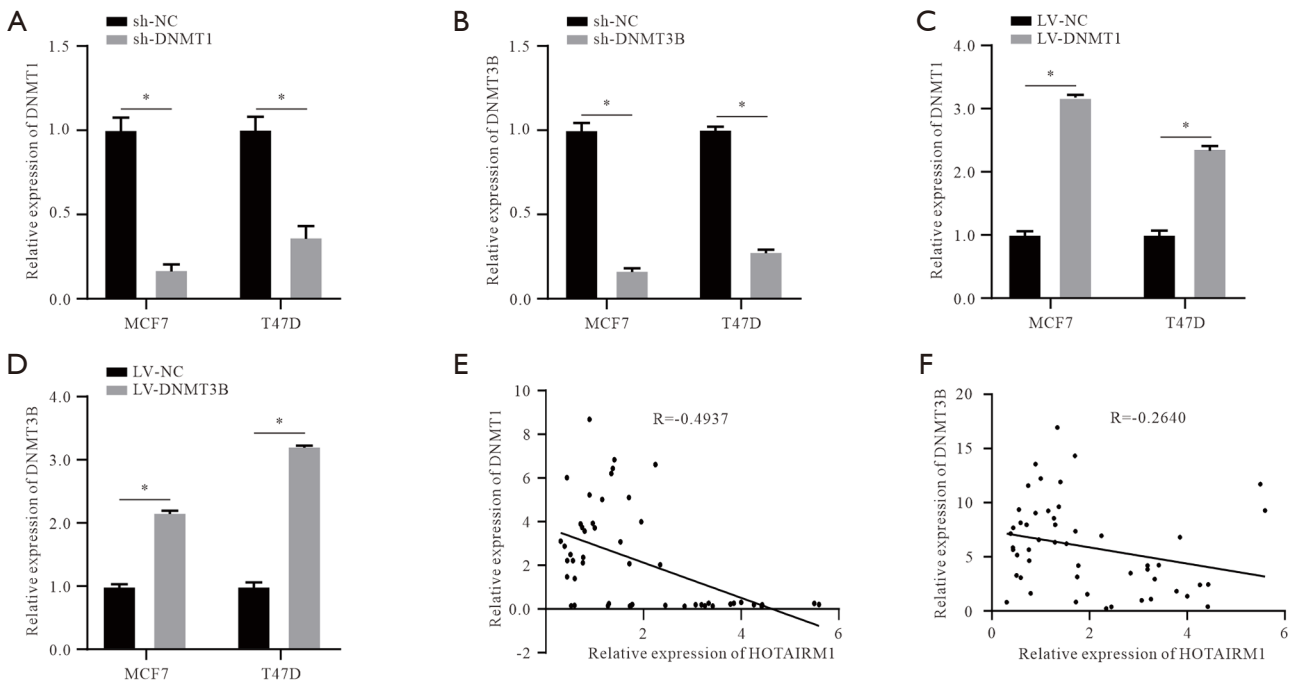


Figure S5 *HOTAIRM1* expression is regulated by *DNMT1* and *DNMT3B*. (A-D) qRT-PCR-based validation of transfection efficiency after knockdown or overexpression of *DNMT1* and *DNMT3B*. (E) qRT-PCR results showing the correlation of expression patterns between *DNMT1* and *HOTAIRM1*. (F) qRT-PCR results 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')
HOTAIRM1	CCCACCGTTCAATGAAAG	GTTTCAAACACCCACATTTC
DNMT1	AGGCGGCTCAAAGATTTGGAA	GCAGAAATTCGTGCAAGAGATTC
DNMT3B	AGGGAAGACTCGATCCTCGTC	GTGTGTAGCTTAGCAGACTGG
GAPDH	CTGGGCTACACTGAGCACC	AAGTGGTCGTTGAGGGCAATG
U6	ATTGGAACGATACAGAGAAGATT	GGAACGCTTCACGAATTTG

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

https://dx.doi.org/10.21037/gs-22-164

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

shRNA	Sequence (5'-3')
HOTAIRM1-shRNA	UCAAUGAAAGAUGAACUGGTT
DNMT1-shRNA	GAAGAGACGTAGAGTTACA
DNMT3B-shRNA	AATTAAAAAAAGATGACGGATGCCTAGAGTCTCTTGAACTCTAGGCATCCGTCATCTCG

shRNA, short hairpin RNA.

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

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

https://dx.doi.org/10.21037/gs-22-164

Table S4 Differential methylation-driven gene expression screening

Gene	Normal mean	Tumor mean	Log_2FC	P value	Cor	Cor P value
FBXO6	0.224784454	0.063448768	−1.824878214	5.16E-17	−0.398688477	2.49E-13
PRLR	0.581113147	0.238593848	−1.284262251	1.58E-15	−0.389371718	9.78E-13
BATF	0.402611351	0.160279088	−1.328801636	3.78E-15	−0.330957916	2.07E-09
GYPC	0.032675096	0.254327294	2.960422899	7.30E-15	−0.354802761	1.10E-10
BCAS1	0.45092556	0.183273106	−1.298894192	1.32E-14	−0.339775751	7.19E-10
CYYR1	0.025197102	0.366317786	3.861766013	6.91E-14	−0.336465455	1.07E-09
MCAM	0.07000091	0.238288463	1.767263521	7.83E-14	−0.327507098	3.11E-09
CX3CL1	0.114725528	0.487511945	2.087251117	9.96E-14	−0.391662689	7.01E-13
HOTAIRM1	0.022892996	0.221642758	3.275258023	2.51E-13	−0.368699673	1.76E-11
CDO1	0.190193575	0.514537485	1.435807674	5.76E-13	−0.455526134	2.17E-17
SYN2	0.036116546	0.433624697	3.585715077	2.24E-12	−0.388872907	1.05E-12
BVES	0.202469899	0.58967092	1.542202609	4.70E-12	−0.388446322	1.12E-12
CLIP4	0.059932548	0.458103315	2.934261391	5.16E-12	−0.451805598	4.23E-17
CPNE8	0.026749376	0.168588893	2.655932341	6.42E-11	−0.386778431	1.42E-12
PPP4R1L	0.11688142	0.261411253	1.161275637	8.88E-10	−0.434978657	7.81E-16
NRN1	0.155825902	0.330357012	1.084090902	2.33E-09	−0.352237685	1.52E-10
KLHL21	0.117275066	0.258511945	1.140334635	2.37E-08	−0.366433384	2.38E-11
ID4	0.095035194	0.34655469	1.866549256	2.82E-08	−0.343398505	4.61E-10
WASF3	0.026023665	0.086954199	1.740431524	1.26E-07	−0.323044123	5.20E-09
KCNC4	0.028492988	0.153199585	2.426733589	1.73E-07	−0.39521981	4.16E-13
ZNF502	0.063274687	0.206485276	1.706338544	3.08E-07	−0.457379614	1.55E-17
NAT8L	0.098869887	0.296370985	1.58380113	1.42E-05	−0.388454334	1.12E-12
ALDOC	0.031806323	0.255893594	3.008158533	0.000457881	−0.339359412	7.57E-10
CDC14B	0.036848676	0.125814535	1.771613927	0.000705105	−0.359038265	6.33E-11
GSTM2	0.017708769	0.180491308	3.349393547	0.00158673	−0.459272787	1.10E-17

FC, fold change.

https://dx.doi.org/10.21037/gs-22-164

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.

https://dx.doi.org/10.21037/gs-22-164

Table S6 Univariate Cox regression analysis

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

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

https://dx.doi.org/10.21037/gs-22-164