Breast Cancer Research and Treatment (2019) 178:565–572
https://doi.org/10.1007/s10549-019-05417-3

PRECLINICAL STUDY

CpG methylation signature predicts prognosis in breast cancer

Tonghua Du[1] · Bin Liu[1] · Zhenyu Wang[1] · Xiaoyu Wan[1] · Yuanyu Wu[2]

Received: 14 May 2019 / Accepted: 22 August 2019 / Published online: 13 September 2019

Abstract

Purpose DNA methylation can be used as prognostic biomarkers in various types of cancers. We aimed to identify a CpG methylation pattern for breast cancer.

Methods In this study, using the microarray data from the cancer genome atlas (TCGA) and gene expression omnibus (GEO), we profiled DNA methylation between 97 healthy control samples and 786 breast cancer samples in a training cohort (from TCGA, $n=883$) to build a gene classifier using a penalized regression model. We validated the prognostic accuracy of this gene classifier in an internal validation cohort (from GEO, $n=72$).

Results A total of 1777 differentially methylated CpGs corresponding to 1777 different methylated genes (DMGs) between breast cancer and control were chosen for this study. Subsequently, 16 CpGs were generated to classify patients into high-risk and low-risk groups in the training cohort. Patients with high-risk scores in the training cohort had shorter overall survival (hazard ratio [HR], 4.674; 95% CI 2.918 to 7.487; $P=1.678e{-}12$) than patients with low-risk scores. The prognostic accuracy was also validated in the validation cohorts. Furthermore, among patients with low-risk scores in the combined training and validation cohorts, the patients with the age > 60 years compared with the patients with the age < 60 years were associated with improved overall survival (HR 2.088, 95% CI 1.348 to 3.235; $p=7.575e{-}04$) in patients with a high-risk score but not in patients with low-risk score (HR 1.246, 95% CI 0.515 to 3.011; $p=0.625$). The patients treated with radiotherapy compared with the patients without radiotherapy were associated with improved overall survival (HR 0.418, 95% CI 0.249 to 0.703; $p=6.991e{-}04$) in patients with a high-risk score but not in patients with low-risk score (HR 2.092, 95% CI 0.574 to 7.629; $p=0.253$). For the patients with recurrence and the patients without recurrence both groups were all associated with improved overall survival (HR 7.475, 95% CI 4.333 to 12.901; $p=6.991e{-}04$) in patients with a high-risk score and in patients with low-risk score (HR 14.33, 95% CI 4.265 to 48.17; $p=4.883e{-}13$).

Conclusion The 16 CpG-based signature is useful as a biomarker in predicting prognosis for patients with breast cancer.

Keywords Breast cancer · DNA methylation · Prognosis · Overall survival · CpG sites

Introduction

Breast cancer, the most frequent cancer among women, is the leading cause of cancer mortality for women, and dramatically, the incidence is increasing among young women aged less than 40 years [1, 2]. According to the statistics of 2016, about 246,660 female were diagnosed with breast cancer, of whom an estimated 40,450 people are expected to die of breast cancer [3]. In recent years, although many advanced treatment techniques have been used to improve the survival rate, the patients still suffer from low quality of life or develop metastasis for late diagnosis [4]. As a result, the prognosis-related markers of breast cancer could increase treatment options, including surgical resection and therapeutic interventions [5].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10549-019-05417-3) contains supplementary material, which is available to authorized users.

✉ Yuanyu Wu
fangxuedcc6@163.com

[1] Department of Breast Surgery, The Second Clinical Hospital Of Jilin University, NO. 218, Ziqiang Street, Nanguan District, Changchun 130000, Jilin, China

[2] Department of Gastrointestinal and Colorectal Surgery, China-Japan Union Hospital of Jilin University, No.126 Xiantai St, Changchun, Jilin 130033, China

Biomarkers strongly associated with breast cancer risk factors will provide an opportunity to understand cancer development. Gene expression profiles have also been used for breast cancer classification and served as prognostic and therapeutic predictors. However, there are still major challenges in accurate early prediction of breast cancer incidence, detection and prognosis. Thus, it is of great need to identify sensitive biomarkers to evaluate breast cancer prognosis at an early treatable stage.

DNA methylation refers to heritable and modifiable markers that regulate gene expression without changing the underlying DNA sequence. So far, many researches have demonstrated that DNA methylation may play a critical role in carcinogenesis through downregulating tumor suppressor genes expression [6]. It has also been found that the aberrant DNA methylation could be an important factor in breast cancer. For example, Tang et al. using the DNA methylation array have also identified breast cancer-associated RPTOR, MGRN1, and RAPSN hypomethylation in peripheral blood DNA [3]. Given that DNA methylation changes are plausibly critical components of the molecular mechanisms involved in breast cancer, distinct DNA methylation profiles could be a potential biomarker to improve the accuracy of breast cancer prognosis.

In the present study, using the downloaded Illumina 450 K DNA methylation array data of breast cancer from the cancer genome atlas (TCGA) database and gene expression omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/), we aimed to screen the breast cancer-related genes with aberrant DNA methylation and constructed a model using the candidate genes in 863 training cohorts and validated those genes in 72 cohorts.

Materials and methods

Microarray data

The DNA methylation profile (Illumina Infinium Human Methylation 450 BeadChip) was downloaded from TCGA (https://gdc-portal.nci.nih.gov/) which including 97 healthy control samples and 786 breast cancer samples as the training cohorts (Table S1). In addition, the DNA methylation profile of GSE37754 (Illumina Infinium Human Methylation 450 BeadChip) was downloaded from GEO, containing 10 healthy control samples and 62 breast cancer samples as the validation cohorts.

Screening of different methylated CpGs

According to the overall survival prognosis information from TCGA, the training cohorts were divided into two groups, including group tumor and group control. Next, the methylation loci at the genes' CpG sites were further screened for analysis. Methylation at each CpG site is described as β value [β = intensity of the methylated allele (M)/(intensity of the unmethylated allele (U) + intensity of the methylated allele (M) + 100)] [7]. It is expressed as a continuous variable that ranges from 0 (no methylation) to 1 (full methylation). Limma (linear model for microarray data) package in R language was used to screen the differentially methylated genes (DMGs) [8]. Bonferroni in multi-test package was employed to adjust the p value into false discovery rate (FDR) [9]. The FDR < 0.05 and $|\log_2$ fold-change (FC)$| > 0.5$ were used as the cutoff criteria for the DMGs.

Co-methylated genes screening

WGCNA (weighted gene coexpression network analysis, version 1.63, https://cran.r-project.org/web/packages/WGCNA/index.html) package in R language was used to analyze the correlation among the DMGs [10]. Modules were defined as clusters of highly interconnected DMGs, and DMGs within the same cluster have high correlation coefficients among them.

Prognosis-related methylated genes screening

The univariate and multivariate cox regression analysis of survival package (Version 2.41-1, http://bioconductor.org/packages/survivalr/) in R language was used to identify DMGs associated to the prognosis using the DMGs with the threshold value of log-rank $p < 0.05$ [11]. In addition, the penalized package (Version 0.9-50, http://bioconductor.org/packages/penalized/) [12] was applied to identify the optimized DMGs based on the Cox-PH (Cox-proportional Hazards) model [13]. The optimized parameter "lambda" was obtained by 1000 cross-validation likelihood cycles.

Prognostic model construction

To better investigate the performance of those DMGs in predicting prognosis, a risk score was built, with the coefficients weighted by the penalized Cox model, and the risk score was calculated as per the following formula:

$$\text{Risk score} = \sum \text{coef}_{\text{gene}} \times \text{Methylation}_{\text{gene}}$$

where $\text{Coef}_{\text{gene}}$ denote the coefficients; $\text{Methylation}_{\text{gene}}$ is the methylation level of gene.

Using the median of the risk score as the cutoff point, the training cohorts were divided into high-risk group and low-risk group. Then, Kaplan–Meier plots were used to illustrate overall survival. Its performance was assessed by receiver-operating characteristic (ROC) analysis. Moreover, we also performed univariate and multivariate Cox

regression analysis using backward selection to test the independent significance of different factors; the p value threshold was 0.1 ($p > 0.1$) for removing nonsignificant variables from the analysis, and the marginally significant variables ($0.05 < p < 0.1$) remained in the final Cox model. Covariates included radiotherapy (yes vs. no), age (> 60 vs. < 60), and concurrent chemotherapy (no vs. yes).

Results

Analysis of DMGs

Based on the annotation data of the Illumina 450 K methylation platform, 10,443 unique CpG sites were obtained. Subsequently, 1777 significant DMGs were identified between the tumor group and control group in the training cohorts including 236 hypomethylation genes and 1541 hypermethylation genes. Volcano plots were used for visualization and assessment of the variation (or reproducibility) of DMGs between tumor group and control group (Fig. 1a). Two-way clustering revealed that methylation patterns between tumor group and control group were distinguishable (Fig. 1c).

Co-methylation analysis

According to the WGCNA, nine modules (617 DMGs) with different colors were identified among which six modules (brown, green, yellow, black, turquoise, and red) were significantly associated with breast cancer (Figure S1).

Selection of Candidate DMGs

The univariate cox regression analysis obtained 123 DMGs which were significantly related to the prognosis. Subsequently, the multivariate cox regression analysis selected 20 DMGs. Finally, 16 optimized DMGs including ZNF462, SYCN, PCSK1, CHL1, CORIN, RAB13, USP4, KCNA1, ZNF274, BHLHA9, HBM, PXN, GRP, SYCE1L, COL4A2, and CRHR1 were identified using the Cox-PH model (Table 1).

Building a predictive signature

A risk score was built, with the coefficients were weighted by the penalized Cox-PH model in the training cohort. The risk score was calculated as follows:

$$\text{Risk score} = (-0.9271 \times \text{methylation level of ZNF462}) + (0.5220 \times \text{methylation}$$

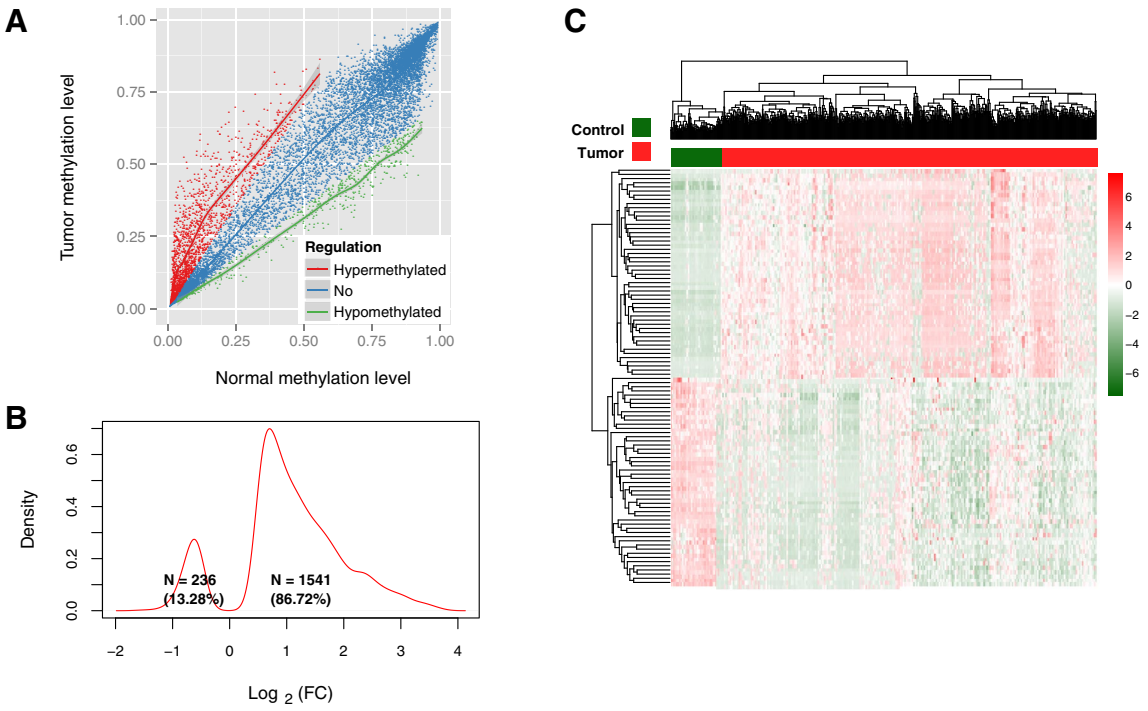

Fig. 1 Analysis of the significantly different methylation CpGs. **a** The volcano plot shows the significantly different methylation CpG-related DMGs; red points represent the hypermethylated DMGs, blue points represent the hypomethylated DMGs; **b** Log2 Kernel density curve based on DMGs; **c** The two-way clustering of DMGs; the colors indicate methylation levels, and from green to red indicates methylation levels from low to high, respectively

Table 1 The optimized CpG data

ID	Gene	Chr.	Position	Location	Coef	Hazard Ratio	95% CI	*P* value
cg13804575	ZNF462	chr9	108666820	5′UTR	−0.9271	0.1110	0.0221–0.557	7.560E−03
cg22290648	SYCN	chr19	44386549	1stExon	0.5220	6.7910	1.281–9.199	2.437E−02
cg23187653	PCSK1	chr5	95794764	TSS200	−0.1412	0.1053	0.0164–0.674	1.750E−02
cg25482786	CHL1	chr3	213496	TSS200	2.4676	6.5910	4.865–9.012	4.280E−05
cg26232715	CORIN	chr4	47533975	promoter	1.8749	2.8250	1.769–5.412	1.811E−02
cg26336059	RAB13	chr1	152225601	TSS200	−2.1655	0.0224	0.00825–0.0606	7.930E−03
cg26353296	USP4	chr3	49353128	promoter	−0.5395	0.0944	0.0212–0.419	1.920E−03
cg26590537	KCNA1	chr12	4890463	5′UTR	−0.5816	0.0870	0.00879–0.860	3.673E−02
cg26698460	ZNF274	chr19	63407816	Body	−0.7784	0.0776	0.00671–0.896	4.058E−02
cg26953640	BHLHA9	chr17	1121223	1stExon	−0.6012	0.1043	0.0115–0.945	4.449E−02
cg26976732	HBM	chr16	156100	Body	−0.5501	0.0854	0.0121–0.602	1.358E−02
cg27168573	PXN	chr12	119185312	Body	−1.5005	0.0423	0.00303–0.589	1.862E−02
cg27338487	GRP	chr18	55038384	1stExon	1.5255	2.7970	1.004–7.793	4.974E−02
cg27430961	SYCE1L	chr16	75804158	promoter	1.5883	4.6790	2.976–6.364	2.500E−04
cg27546237	COL4A2	chr13	109758453	TSS1500	0.3422	1.6660	1.586–2.175	1.905E−02
cg27551605	CRHR1	chr17	41218673	Body	−2.6558	0.0025	0.000195–0.0318	4.010E−06

ID CpGs ID, *Gene* the CpG-related genes, *Chr* chromosome, *Location* the location of the methylated site on the gene, *Coef* cox regression coefficient

level of SYCN)—(0.1412 × methylation level of PCSK1) + (2.4676 × methylation level of CHL1) + (1.8749 × methylation level of CORIN)—(2.1655 × methylation level of RAB13)—(0.5395 × methylation level of USP4)—(0.5816 × methylation level of KCNA1)—(0.7784 × methylation level of ZNF274)—(0.6012 × methylation level of BHLHA9)—(0.5501 × methylation level of HBM)—(1.5005 × methylation level of PXN) + (1.5255 × methylation level of GRP) + (1.5883 × methylation level of SYCE1L) + (0.3422 × methylation level of COL4A2)—(2.6558 × methylation level of CRHR1).

Using the median of the risk score as the cutoff point, 393 high-risk patients (50%) had poorer RFS (hazard ratio [HR], 4.674; 95% CI 2.918 to 7.487; $P = 1.678e{-}12$; Fig. 2a) than did the 393 low-risk patients (50%) in the training cohort. The validation analyses were performed in an internal validation cohort. A total of 31 patients were categorized (50%) into the high-risk group and 31 patients (50%) into the low-risk groups (HR 2.679; 95% CI 1.224 to 5.863; $P = 0.0104$; Fig. 2b). The AUC and corresponding 95% confidence intervals for training and validation datasets were 0.972 and 0.942, respectively.

The univariate and multivariate analysis of overall survival by clinical factors in the training is shown in Table 2. The stratified analysis of clinical characteristics of age, radiotherapy, and recurrence illustrated that the patients with the age > 60 years compared with the patients with the age < 60 years were associated with improved overall survival (HR 2.088, 95% CI 1.348 to 3.235; $p = 7.575e{-}04$) in patients with a high-risk score but not in patients with low-risk score (HR: 1.246, 95% CI 0.515 to 3.011; $p = 0.625$). The patients with radiotherapy compared with the patients without radiotherapy were associated with improved overall survival (HR: 0.418, 95% CI 0.249 to 0.703; $p = 6.991e{-}04$) in patients with a high-risk score but not in patients with low-risk score (HR: 2.092, 95% CI 0.574 to 7.629; $p = 0.253$). With regard to the association between the patients with recurrence and the patients without recurrence they were all associated with improved overall survival (HR: 7.475, 95% CI 4.333 to 12.901; $p = 6.991e{-}04$) in patients with a high-risk score and in patients with low-risk score (HR: 14.33, 95% CI 4.265 to 48.17; $p = 4.883e{-}13$) (Fig. 3).

Discussion

In this study, we used methylation microarray data from TCGA and GEO to develop and validate a novel prognostic tool based on 16 CpGs corresponding to 16 genes (ZNF462, SYCN, PCSK1, CHL1, CORIN, RAB13, USP4, KCNA1, ZNF274, BHLHA9, HBM, PXN, GRP, SYCE1L, COL4A2, and CRHR1) referred as DMGs that are compared with clinical risk factors, which has the improved ability to predict prognosis in patients with breast cancer. Our results indicated that the 16 DMGs identified in this study could be used to categorize patients into high-risk and low-risk groups of patients who had significantly different overall survival.

So far, several studies have reported the DNA methylation profiles in BC. For example, by means of bisulfite

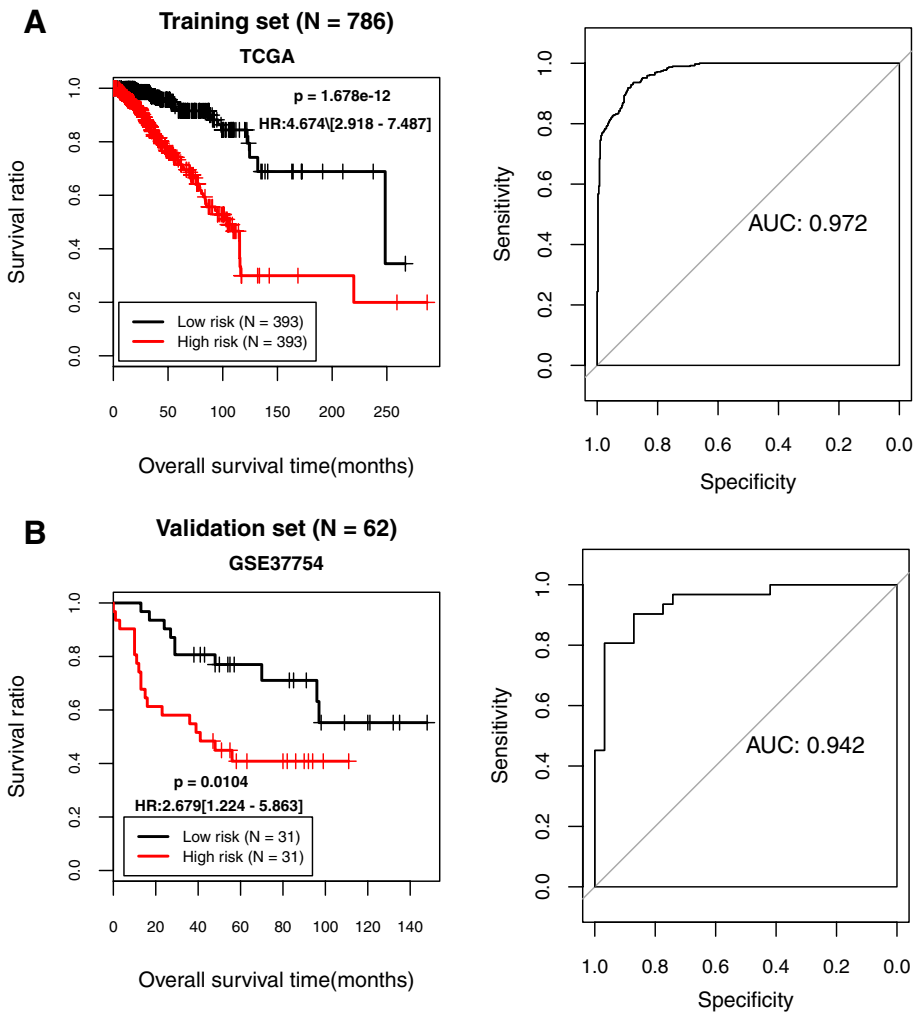

Fig. 2 The performance of 16 CpG Methylation Signatures in predicting recurrence in the training and the validation of cohorts. **a** For the training cohort; **b** For the validation cohort

Table 2 Clinical factor statistics and cox regression analysis

Clinical characteristics	TCGA(N = 786)	Uni-variables cox			Multi-variables cox		
		HR	95%CI	*P*	HR	95%CI	*P*
Age (years, mean ± SD)	57.99 ± 13.19	1.03	1.015–1.045	6.492E−05	1.043	1.0194–1.068	3.350E−04
ER status (positive/negative/–)	574/170/42	0.648	0.424–0.992	4.400E−02	0.577	0.301 –1.107	9.802E−02
PR status (positive/negative/–)	501/240/45	0.76	0.505–1.144	1.870E−01	–	–	–
Her2 status (positive/negative/–)	58/406/322	0.936	0.425–2.06	8.690E−01	–	–	–
Histological type (Basal/Her2/LumA/LumB/ Normal/–)	88/32/278/128/19/241	1.164	0.937–1.445	1.690E−01	–	–	–
Pathologic_M (M0/M1/–)	617/13/156	5.245	2.664–10.33	1.229E−07	2.168	0.689–6.819	1.858E−01
Pathologic_N (N0/N1/N2/N3/–)	349/272/95/58/12	1.6	1.297–1.973	8.374E−06	1.135	0.709–1.817	5.953E−01
Pathologic_T (T1/T2/T3/T4/–)	200/451/109/23/3	1.484	1.171–1.882	1.050E−03	1.387	0.848–2.268	1.923E−01
Pathologic_stage (I/II/III/IV/–)	127/440/199/11/9	2.262	1.685–3.038	3.963E−08	2.315	0.919–5.829	7.467E−02
Radiotherapy (yes/no/–)	413/314/59	0.563	0.355–0.892	1.320E−02	0.481	0.247–0.937	3.149E−02
Recurrence (yes/no/–)	74/623/89	9.173	5.570–15.11	2.000E−16	6.602	3.521–12.379	4.010E−09
Risk status (high/low)	393/393	4.674	2.918–7.487	1.678E−12	8.601	3.662–20.202	7.860E−07
Dead (death/alive)	105/681	–	–	–	–	–	–
Overall survival time (months, mean ± SD)	42.03 ± 38.71	–	–	–	–	–	–

TCGA the cancer genome atlas, *HR* hazard ratio

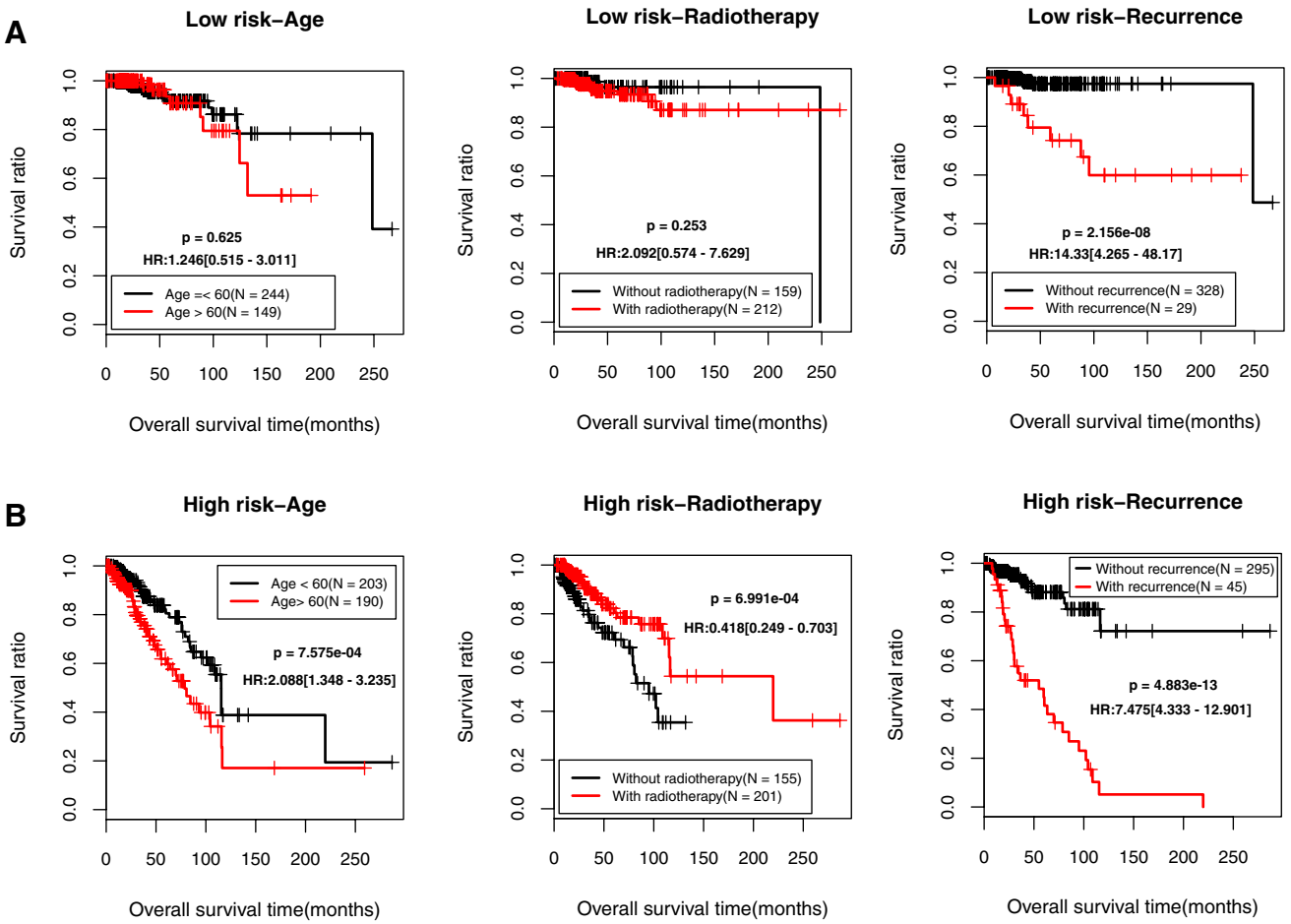

Fig. 3 Kaplan–Meier curves of overall survival according to low-risk or high-risk scores stratified by age, administration of radiotherapy, and recurrence. **a** For low risk—age, radiotherapy, and recurrence; **b** For high risk—age, radiotherapy, and recurrence

sequencing PCR, they confirmed that RASSF1a, P16, and PCDHGB7 displayed significant sensitivity and specificity as diagnostic biomarkers for BC ($P < 0.001$) [14]. Based on a quantitative methylation-specific real-time PCR analysis, it has been found that both the higher methylation of genes *PER1, 2, 3* and the lower methylation of *CLOCK, BMAL1*, and *CRY2* could be potential indicators of BC [15]. Yang et al. have discovered 45 CpGs residing in 18 genomic regions which have not previously been associated with breast cancer risk from 228,951 women of European descent [16]. Different studies showed the different markers, which illustrated that there are no exact or specific markers for the diagnosis of BC. Unlike in the previous studies, we used in this study the WGCNA, penalized Cox-PH model, the univariate and multivariate analysis of overall survival by clinical factors to screen the markers for BC. We believed that analyses based on differential statistical tests may result in different outcomes.

Cancer is a heterogeneous disease, and exploring the dysmethylated genes involved in carcinogenesis and development might help to improve prognostic and therapeutic strategies. In the present study, we identified a group of 16 different methylated CpGs corresponding to 16 genes that effectively predict prognosis of breast cancer. Among these genes, the CHL1 gene which encodes a cell-adhesion molecule belongs to the L1 family of CAMs, and CHL1 is involved in human breast tumorigenesis and progression [17]. It has been found CHL1 hypermethylation as a potential biomarker of poor prognosis in breast cancer [18]. Besides, ubiquitin-specific protease 4 (USP4), which is a deubiquitinating enzyme with key roles in the regulation of TGF-β1, is also identified in our study. Many studies have revealed the important role of USP4 in breast cancer. For example, USP4 could promote invasion of breast cancer cells via Relaxin/TGF-β1/Smad2/MMP-9 signal [19], and USP4 could inhibit breast cancer cell growth through the upregulation of PDCD4 [20]. In addition, collagen type IV alpha 2 (COL4A2), is also detected. COL4A2 is the major structural component of basement membranes. The C-terminal portion of the protein, known as canstatin, is an inhibitor

of angiogenesis and tumor growth [21]. The suppression of COL4A2 mRNA inhibits triple-negative breast cancer cell proliferation and migration [22]. Compared to our study, the commercial kit used for the BC diagnosis showed main fucus in the expression of genes related to cell proliferation, estrogen, HER2, and invasion [23], which were different from our markers. Therefore, different genes corresponding to different functions could show the different mechanism of BC. However, further studies are needed to validate their functions.

Aberrant DNA methylation is a critical mechanism in carcinogenesis [24]. Although DNA methylation profiles are often tissue- and cell-type specific, recent data indicate that epigenetic changes in blood cell DNA are potential markers for solid tumors [25]. In our study, we used the tissue samples from TCGA to screen the markers and the blood samples from GEO to validate the markers. And the markers screened in tissue samples showed a good performance in the blood samples, which illustrated the abnormal methylation of blood could be used to diagnose BC. However, these results should be validated through further studies.

The study's limitations should be noted. First, the methods for screening the biomarkers was based on the statistical method rather than the biological experiment. Therefore, the biologic mechanisms of the candidate markers are still unknown. Second, additionally no experiments such as pyrosequencing were conducted to validate the methylation levels of the marker genes. Third, further validation in prospective studies and multicenter clinical trials are needed. Fourth, we did not analyze the relevance of chemotherapy or endocrine therapy to the overall survival of BC, which could, to some degree, illustrate the key factors related to the overall survival of BC. Fifth, there existed probe design bias in Illumina Infinium 450 k DNA methylation data, and a key statistical method should be used to adjust for the two different probe designs.

In conclusion, 16-CpG methylation, is a potential prognostic tool for predicting prognosis for patients with breast cancer and might help clinicians in directing personalized therapeutic regimen selection for patients with breast cancer.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Johnson RH, Chien FL, Bleyer A (2013) Incidence of breast cancer with distant involvement among women in the United States, 1976 to 2009. JAMA 309(8):800–805. https://doi.org/10.1001/jama.2013.7761656255
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 136(5):E359–386. https://doi.org/10.1002/ijc.29210
3. Tang Q, Holland-Letz T, Slynko A, Cuk K, Marme F, Schott S, Heil J, Qu B, Golatta M, Bewerunge-Hudler M, Sutter C, Surowy H, Wappenschmidt B, Schmutzler R, Hoth M, Bugert P, Bartram CR, Sohn C, Schneeweiss A, Yang R, Burwinkel B (2016) DNA methylation array analysis identifies breast cancer associated RPTOR, MGRN1 and RAPSN hypomethylation in peripheral blood DNA. Oncotarget 7(39):64191–64202. https://doi.org/10.18632/oncotarget.1164011640
4. Cady B (2007) Local therapy and survival in breast cancer. N Engl J Med 357(10):1051–1052 **author reply 1052**
5. Hudis CA (2007) Trastuzumab–mechanism of action and use in clinical practice. N Engl J Med 357(1):39–51
6. Jones PA, Baylin SB (2007) The epigenomics of cancer. Cell 128(4):683–692
7. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL, Fan JB, Shen R (2011) High density DNA methylation array with single CpG site resolution. Genomics 98(4):288–295
8. Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S (2005) Bioinformatics and computational biology solutions using R and Bioconductor. Springer, New York
9. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodological) 57(1):289–300
10. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559
11. Wang P, Wang Y, Hang B, Zou X, Mao JH (2016) A novel gene expression-based prognostic scoring system to predict survival in gastric cancer. Oncotarget 7(34):55343–55351
12. Goeman JJ (2010) L1 penalized estimation in the Cox proportional hazards model. Biom J 52(1):70–84. https://doi.org/10.1002/bimj.200900028
13. Tibshirani R (1997) The lasso method for variable selection in the Cox model. Stat Med 16(4):385–395
14. Shan M, Zhang L, Liu Y, Gao C, Kang W, Yang W, He Y, Zhang G (2019) DNA methylation profiles and their diagnostic utility in BC. Dis Mark 2019:6328503. https://doi.org/10.1155/2019/6328503
15. Lesicka M, Jablonska E, Wieczorek E, Seroczynska B, Kalinowski L, Skokowski J, Reszka E (2019) A different methylation profile of circadian genes promoter in breast cancer patients according to clinicopathological features. Chronobiol Int. https://doi.org/10.1080/07420528.2019.1617732
16. Yang Y, Wu L, Shu XO, Cai Q, Shu X, Li B, Guo X, Ye F, Michailidou K, Bolla MK, Wang Q, Dennis J, Andrulis IL, Brenner H, Chenevix-Trench G, Campa D, Castelao JE, Gago-Dominguez M, Dork T, Hollestelle A, Lophatananon A, Muir K, Neuhausen SL, Olsson H, Sandler DP, Simard J, Kraft P, Pharoah PDP, Easton DF, Zheng W, Long J (2019) Genetically predicted levels of DNA methylation biomarkers and breast cancer risk: data from 228,951 women of European descent. J Natl Cancer Inst. https://doi.org/10.1093/jnci/djz109
17. He LH, Ma Q, Shi YH, Ge J, Zhao HM, Li SF, Tong ZS (2013) CHL1 is involved in human breast tumorigenesis and progression. Biochem Biophys Res Commun 438(2):433–438
18. Martin-Sanchez E, Mendaza S, Ulazia-Garmendia A, Monreal-Santesteban I, Blanco-Luquin I, Cordoba A, Vicente-Garcia F, Perez-Janices N, Escors D, Megias D, Lopez-Serra P, Esteller M, Illarramendi JJ, Guerrero-Setas D (2017) CHL1

hypermethylation as a potential biomarker of poor prognosis in breast cancer. Oncotarget 8(9):15789–15801
19. Cao WH, Liu XP, Meng SL, Gao YW, Wang Y, Ma ZL, Wang XG, Wang HB (2016) USP4 promotes invasion of breast cancer cells via Relaxin/TGF-beta1/Smad2/MMP-9 signal. Eur Rev Med Pharmacol Sci 20(6):1115–1122
20. Li Y, Jiang D, Zhang Q, Liu X, Cai Z (2016) Ubiquitin-specific protease 4 inhibits breast cancer cell growth through the upregulation of PDCD4. Int J Mol Med 38(3):803–811. https://doi.org/10.3892/ijmm.2016.2685
21. Turner AW, Nikpay M, Silva A, Lau P, Martinuk A, Linseman TA, Soubeyrand S, McPherson R (2015) Functional interaction between COL4A1/COL4A2 and SMAD3 risk loci for coronary artery disease. Atherosclerosis 242(2):543–552
22. JingSong H, Hong G, Yang J, Duo Z, Li F, WeiCai C, XueYing L, YouSheng M, YiWen O, Yue P, Zou C (2017) siRNA-mediated suppression of collagen type IV alpha 2 (COL4A2) mRNA inhibits triple-negative breast cancer cell proliferation and migration. Oncotarget 8(2):2585–2593
23. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE Jr, Wickerham DL, Wolmark N (2006) Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 24(23):3726–3734
24. Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9(6):465–476
25. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S, Jones A, Lechner M, Beck S, Jacobs IJ, Widschwendter M (2009) An epigenetic signature in peripheral blood predicts active ovarian cancer. PLoS ONE 4(12):e8274. https://doi.org/10.1371/journal.pone.0008274