

Methylation of *WT1*, *CA10* in peripheral blood leukocyte is associated with breast cancer risk: a case-control study

Anqi Ge

Harbin Medical University School of Public Health

Song Gao

Tumor Hospital of Harbin Medical University

Yupeng Liu

Harbin Medical University School of Public Health

Hui Zhang

Harbin Medical University School of Public Health

Xuan Wang

Harbin Medical University School of Public Health

Lei Zhang

Harbin Medical University School of Public Health

Da Pang

Tumor Hospital of Harbin Medical University

Yashuang Zhao (✉ zhao_yashuang@263.net)

Harbin Medical University <https://orcid.org/0000-0002-7425-5773>

Research article

Keywords: breast cancer; CA10; WT1; DNA methylation; leukocytes.

Posted Date: July 16th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-27678/v2>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at BMC Cancer on July 31st, 2020. See the published version at <https://doi.org/10.1186/s12885-020-07183-8>.

Abstract

Background: Studies have shown that abnormal changes of specific-gene DNA methylation in leukocytes may be associated with an elevated risk of cancer. However, associations between the methylation of the zinc-related genes, *WT1* and *CA10*, and breast cancer risk remain unknown.

Methods: The methylation of *WT1* and *CA10* was analyzed by methylation-sensitive high-resolution-melting (MS-HRM) in a case-control study with female subjects (N=959). Logistic regression was used to analyze the associations, and propensity score (PS) method was used to adjust confounders.

Results: The results showed that *WT1* hypermethylation was associated with an increased risk of breast cancer, with an odds ratio (OR) of 3.07 [95% confidence interval (CI): 1.67-5.64, $P<0.01$]. Subgroup analyses showed that *WT1* hypermethylation was specifically associated with an elevated risk of luminal A subtype (OR=2.62, 95% CI: 1.11-6.20, $P=0.03$) and luminal B subtype (OR=3.23, 95% CI: 1.34-7.80, $P=0.01$). *CA10* hypermethylation was associated with an increased risk of luminal B subtype (OR=1.80, 95% CI: 1.09-2.98, $P=0.02$).

Conclusion: The results of the present study suggest that the hypermethylation of *WT1* methylation in leukocytes is significantly associated with an increased risk of breast cancer. The hypermethylation of *WT1* is associated with an increased risk of luminal subtypes of breast cancer, and the hypermethylation of *CA10* is associated with an increased risk of luminal B subtype of breast cancer.

Background

Breast cancer is one of the most common malignancies in women worldwide [1] and presents with different molecular subtypes, including luminal A, luminal B, HER2-enriched, and basal-like that also called triple negative [2]. As a major type of epigenetic modification, DNA methylation is involved in regulating cellular processes, including chromosomal instability [3] and gene expression. The hypermethylation of CpG regions in specific genes contribute to neoplastic formation through the transcriptional silencing of tumor suppressor genes. Aberrant patterns of specific gene methylation can help identifying differences in breast cancer subtypes [2], and showing promise for utilizing in large-scale epidemiological studies. It has been suggested that leukocyte DNA methylation, as a simple non-invasive blood marker [4, 5], could serve as a surrogate for systematic methylation activity and offers great potential for predicting the increased risk of breast cancer [6].

Wilm's Tumor gene (*WT1*) is a tumor suppressor gene which involved in human cell growth and differentiation. *WT1* has been reported to be significantly different methylated in the tissues of hepatocellular carcinoma [7], lung cancer [8] and breast cancer [9]. *WT1* aberrant methylation may lead to a reduction or absence of *WT1* expression, which results in the overexpression of the insulin-like growth factor I receptor (IGF 1R) and insulin-like growth factor II (IGF II), thereby promoting breast cancer process [10-12]. *CA10* is a member of the carbonic anhydrase family, which is a large family of zinc-containing metalloenzymes that catalyze the reversible hydration of carbon dioxide and the dehydration of carbonic acid [13]. Ivanov et al. suggested that the induction or enhancement of carbonic anhydrase expression may contribute to the tumor microenvironment by maintaining an extracellular acidic pH and helping the growth and metastasis of cancer cells [14]. Studies have demonstrated that the abnormal expression of carbonic anhydrase family by aberrant methylation is related with gastric cancer and the metastasis of ovary tumors [15, 16]. Furthermore, Wojdacz et al. reported that both *WT1* and *CA10* hypermethylation were significantly different between breast cancer tumor tissues and non-malignant tissues [17]. However, how the methylation of these two genes in leukocyte DNA affects breast cancer susceptibility remains unclear.

In this study, we investigated the associations between the methylation of *WT1*, *CA10* in peripheral blood leukocyte DNA and breast cancer risk. We subsequently used an external dataset of a nested case-control cohort within the EPIC-Italy cohort study as external data to validate the association between gene methylation and breast cancer risk. We also investigated the associations between the methylation of these two genes and the risk of different molecular types of breast cancer.

Materials And Methods

Study subjects

We investigated the relationship between *WT1* and *CA10* methylation and breast cancer risk using a case-control study. All the included breast cancer patients were newly diagnosed females and were recruited from the Tumor Hospital of Harbin Medical University from 2010 to 2014. Female breast cancer subjects were included if they diagnosed with invasive ductal carcinoma (IDC) or ductal carcinoma in situ (DCIS), other types of breast cancer (such as lipoma of the breast, metastatic breast cancer, etc.) were excluded from our study. Controls were recruited from patients admitted to the Orthopedic and Ophthalmology Department of the Second Affiliated Hospital of Harbin Medical University and volunteers from the Xiangfang community of Harbin within the same period. All controls were also female. In addition, all control participants were asked about their disease history in a questionnaire, and individuals who reported a history of any cancer were excluded from our final subjects. Finally, 402 female breast cancer cases and 557 female controls were included in our study. Blood sample (5 mL) was collected from each participant and then stored at -80°C.

Data collection

All subjects were interviewed face-to-face by trained investigators with normalized questioning methods. The questionnaire was adopted from the study by Shu et al. [18], and included information on demographic information (age, ethnicity, and others); daily dietary intake (vegetables, fruits, beverages, and snacks); behaviors (smoking, drinking, physical activity and work activity); female-specific questions involving menstruation status, breast disease history (lobular hyperplasia, cyst, and others); gynecologic surgery history (hysterectomy, ovariectomy) and family history of cancer and breast cancer. The questions involved in dietary and behavioral were about the participants' daily routine of one year prior to the interview. The basic demographic characteristics and environmental factors of the study subjects are presented in Table S1.

The study was validated with the GEO-GSE51032 (IPEC-Italy cohort) dataset with a nested case control study design to analyze the association between the methylation of *CA10* and *WT1* and breast cancer risk. The blood samples were also collected and other anthropometric measurements were taken. The sample selection criteria and the methods were reported by Riboli et al. [19]. We extracted all 232 female breast cancer cases and all 340 female controls from this nested case-control study and located the methylation probes from the Illumina 450 K array. The annotated CG sites covered by our MS-HRM sequence are illustrated in Figure 1.

DNA extraction and bisulfite conversion

DNA was extracted from peripheral blood samples using a commercial DNA extraction kit (QIAamp DNA Blood Mini Kit, Hilden, Germany). The concentration and the purity of DNA were assessed using a Nanodrop 2000 Spectrophotometer (Thermo Scientific, USA). Genomic DNA was bisulfite-modified with an EpiTect Bisulfite kit (Qiagen, Hilden, Germany). Bisulfite DNA was normalized to a concentration of 20 ng/mL and was stored at -20°C for the following experiment. DNA extraction and DNA sodium bisulfite modification were performed according to the manufacturers' instructions.

Gene methylation status analysis

We performed methylation-sensitive high-resolution melting analysis (MS-HRM) to evaluate the methylation of *WT1* and *CA10* with the LightCycler 480 system (Roche Applied Science, Mannheim, Germany) equipped with Gene Scanning software (version 2.0). The primers were adopted from a published study [17]. We used universal methylated and unmethylated DNA standards (ZYMO, USA) and mixed them at different ratios to create standards with a 0.5, 1, 2, and 5% methylation levels of *WT1* and *CA10* (Figure 2). PCR amplification and MS-HRM were optimized and performed. The conditions, reaction mixture and primer sequences used in the MS-HRM experiments are listed in Table S2-3. Each standard reaction was performed in duplicate in each run. Each plate included duplicate water blanks as negative controls. We also repeated some samples in different runs to assess the consistency of the experiment. There was a significant agreement of these samples in different runs with respect to the observed methylation status of *WT1* and *CA10*, with kappa value of 1.00 ($P<0.01$) and 0.94 ($P<0.01$), respectively (Table S4).

Definitions of different molecular subtypes of breast cancer

Four subtypes of breast cancer were defined as luminal A, luminal B, HER-2 enriched and triple negative breast cancer (TNBC) by immunohistochemical analysis based on previously validated clinicopathological criteria [20].

Statistical analysis

For the distribution of basic demographic characteristics, continuous variables such as age were analyzed by two-sample t-tests, and categorical variables were analyzed by chi-square (χ^2) tests. For missing values in the environmental factors, we applied multiple imputation to generate possible values. To measure the association between methylation of *WT1*, *CA10* and breast cancer risk and different molecular types breast cancer, we used univariate and multivariate unconditional logistic regression analyses to estimate the crude and adjusted odds ratios (ORs) and 95% confidence intervals (95% CIs). For our case-control study, we used 0% methylation as a cutoff for both *WT1* and *CA10*. We used receiver operating characteristic curve (ROC) to calculate the cut-off value of β for the validation dataset. We also applied the propensity score (PS) method to adjust covariates (involving all environmental factors in the questionnaire), in which the study outcome served as the dependent variable and PS served as the confounding variable. Kappa values were calculated to analyze the consistency between same samples in different runs. All two-sided P values <0.05 were considered statistically significant. Data were analyzed by using SPSS v.24.0 (SPSS Inc., Chicago, IL, USA).

Results

Characteristics of the cases and controls

This study included 402 female cases with a mean age of 51.75 ± 9.39 and 557 female controls with a mean age of 51.85 ± 10.31 . Other demographic information of the cases and controls is listed in Table 1. The definition of variables for environmental factors with $\leq5.8\%$ missing data were processed by the multiple imputation method are presented in Table S1.

Table 1. Demographic characteristics of breast cancer patients and controls.

Characteristics	Mean \pm SD	No. of	No. of	P Value
		Controls(%)	Cases (%)	
Age	51.85 \pm 10.31	51.75 \pm 9.39		
	< 40	82(14.7)	41(10.2)	0.02
	40-	333(59.8)	274(68.2)	
BMI	\geq 60	142(25.5)	87(21.6)	
	\leq 18.5	35(6.3)	14(3.5)	0.12
	18.5-	274(49.2)	211(52.5)	
Urban and Rural Status	\geq 24.0	248(44.5)	177(44.0)	
	Rural	236(42.4)	232(57.7)	<0.01
	Urban	321(57.6)	170(42.3)	
Education Level	Primary School or Below	162(29.1)	98(24.4)	0.27
	Middle School	175(31.4)	135(33.6)	
	Senior School and Higher	220(39.5)	169(42.0)	
Occupation Type ^a	White Collar	273(49.0)	233(58.2)	0.01
	Blue Collar	284(51.0)	169(41.8)	
Ethnicity	Han	529(95.0)	386(96.0)	0.27
	Other	28(5.0)	16(4.0)	

^a The white collar occupation referred to people work that need mental rather than physical effort, such as office, doctor, accountant, business, teacher, etc.; the blue collar occupation referred to people work as manual labor, such as worker, farmer, cleaner, etc.

Associations between *WT1*, *CA10* methylation and breast cancer risk

WT1 methylation was associated with breast cancer risk both in multivariable and PS adjusted methods with ORs of 2.42 (95% CI: 1.45-4.04, $P<0.01$) and 3.07 (95% CI: 1.67-5.64, $P<0.01$), respectively. *CA10* methylation was statistically significant associated with breast cancer in the multivariable adjustment with an OR of 1.53 (95% CI: 1.14-2.05, $P<0.01$), but was only marginally associated with breast cancer after PS adjustment, with an OR of 1.35 (95% CI: 0.97-1.90, $P=0.08$) (Table 2).

Table 2. The associations between gene methylation and risk of breast cancer and different molecular types of breast cancer.

	Molecular types ^a	No. of Unmethylation(%)	No. of Methylation(%)	Crude OR (95% CI)	P Value	OR _{adjusted} ^b (95% CI)	P Value	OR _{adjusted} ^c (95% CI)	P Value
								(95% CI)	
<i>WT1</i>	Control	65(11.7)	492(88.3)	1		1		1	
	Luminal A	9(6.4)	132(93.6)	1.99(0.94-4.23)	0.07	2.61(1.18-5.74)	0.02	2.62(1.11-6.20)	0.03
	Luminal B	8(6.0)	125 (94.0)	2.12(1.50-2.99)	0.07	2.49(1.13-5.51)	0.02	3.23(1.34-7.80)	0.01
	HER-2 Enriched	5(8.9)	51(91.1)	1.34(0.51-3.50)	0.55	1.91(0.69-5.30)	0.21	1.91(0.66-5.51)	0.23
	TNBC	1(2.9)	33(27.1)	4.34(0.58-32.33)	0.15	5.63(0.73-43.63)	0.10	6.04(0.76-47.90)	0.09
	All cases	26(6.5)	376(93.5)	1.92(1.18-3.13)	0.01	2.42(1.45-4.04)	0.01	3.07(1.67-5.64)	<0.01
<i>CA10</i>	Control	209(37.5)	348(62.5)	1		1		1	
	Luminal A	40(28.4)	101(71.6)	1.52(1.00-2.26)	0.05	1.60(1.04-2.45)	0.03	1.51(0.94-2.41)	0.09
	Luminal B	34(25.6)	99(74.4)	1.79(1.17-2.74)	0.01	2.04(1.30-3.21)	<0.01	1.80(1.09-2.98)	0.02
	HER-2 Enriched	18(32.1)	38(67.9)	1.27(0.71-2.29)	0.43	1.42(0.76-2.66)	0.27	1.37(0.71-2.63)	0.35
	TNBC	14(41.1)	20(58.8)	0.86(0.43-1.74)	0.67	0.94(0.45-1.96)	0.87	1.01(0.46-2.20)	0.99
	All cases	119(29.6)	283(70.4)	1.43(1.08-1.88)	0.01	1.53(1.14-2.05)	<0.01	1.35(0.97-1.90)	0.08

^a The result excluded 38 breast cancer patients with incomplete immunohistochemical records.

^b Adjusted for age, BMI, ethnicity, urban and rural status and family history of breast cancer and cancer.

^c Adjusted by propensity score(potential confounder included age, BMI, urban and rural status, ethnicity, education level, mammography, gynecologic surgery, breast disease history, menstrual cycle, menopause, reproduction, abortion, breast feeding, oral contraceptive, female hormone intake, fruit intake, vegetable intake, tomato intake, broccoli intake, bean products, pungent food, pork, beef and lamb consumption, chicken consumption, sea-fish, egg, diary, fungus, pickles, alcohol consumption, tea consumption, cigarette, physical activity, occupation type, family history of breast cancer and cancer).

In the subgroup analyses, after PS adjustment, *WT1* methylation was associated with breast cancer risk in both the younger (<60-years-old) and older (≥ 60 -years-old) groups, with ORs of 2.64 (95% CI: 1.31-5.32, $P=0.01$) and 4.72 (95% CI: 1.31-16.97, $P=0.01$), respectively. *CA10* methylation was associated with breast cancer risk in younger age group (<60-years-old) before PS adjustment, with OR of 1.56 (95% CI: 1.15-2.11, $P=0.01$); However, the association was not statistically significant after PS adjustment (Table 3). We also analyzed the combination and interaction of age and the methylation of *WT1*, *CA10* on the risk of breast cancer. The P values for the interactions between age and the methylation of *WT1* and *CA10* on the risk of breast cancer were 0.40 and 0.73, respectively. The results are presented in Table 4.

Table 3. The subgroup analysis of the associations between methylation of genes and the risk of breast cancer based on different age.

			Crude OR 95% CI	P Value	OR adjusted ^a 95% CI	P Value
<i>WT1</i>	≤60	Unmethylation	1		1	
		Methylation	1.64(0.95-2.84)	0.08	2.64(1.31-5.32)	0.01
	≥60	Unmethylation	1		1	
		Methylation	3.16(1.05-9.5)	0.04	4.72(1.31-16.97)	0.01
<i>CA10</i>	≤60	Unmethylation	1		1	
		Methylation	1.56(1.15-2.11)	0.05	1.32(0.90-1.96)	0.15
	≥60	Unmethylation	1		1	
		Methylation	1.20(0.61-2.37)	0.60	1.52(0.69-3.37)	0.30

^a Adjusted by propensity score.

Table 4. The interaction between age and gene methylations on the risk of breast cancer.

		Age			P
		≥60		<60	
		OReg _{adjusted} ^a (95% CI)	ORi _{adjusted} ^a (95% CI)		
<i>WT1</i>	Unmethylation	1	1.70(0.40-6.84)		1
	Methylation	4.90(1.36-17.67)	4.44(1.29-15.34)	0.53(0.13-2.28)	0.40
<i>CA10</i>	Unmethylation	1	1.17(0.54-2.54)		1
	Methylation	1.55(0.70-3.45)	1.55(0.74-3.27)	0.86(0.35-2.09)	0.73

^a Adjusted for propensity score.

Associations between methylation of *WT1*, *CA10* and risk of different molecular types of breast cancer

WT1 methylation was significantly associated with the risk of luminal A subtype of breast cancer with multivariable adjusted OR of 2.61 (95% CI: 1.18-5.74, *P*=0.02), and PS adjusted OR of 2.62 (95% CI: 1.11-6.20, *P*=0.03). *WT1* methylation was also significantly associated with the risk of luminal B subtype breast cancer with ORs of 2.49 (95% CI: 1.13-5.51, *P*=0.02) and 3.23 (95% CI: 1.34-7.80, *P*=0.01) after multivariable and PS adjustment. However, *WT1* methylation was not significantly associated with the risk of HER-2 enriched and TNBC subtypes (Table 2).

The associations between *CA10* methylation and the risk of luminal B subtype breast cancer with multivariable adjusted and PS adjusted ORs were 2.04 (95% CI: 1.30-3.21, *P*<0.01) and 1.80 (95% CI: 1.09-2.98, *P*=0.02), respectively. However, *CA10* methylation had no significant associations with the risk of luminal A, HER-2 enriched and TNBC subtypes after the adjustment of PS. The association between the methylation of *WT1*, *CA10* and other clinicopathological characteristics of breast cancer patients were analyzed are showed in Table S5.

Association between *WT1*, *CA10* methylation and breast cancer risk in GEO dataset

The GSE51032 dataset is a nested case control study that includes 233 female breast cancer cases and 340 female cancer-free controls. After the data extraction from the 450K array, we identified two CG loci each in our targeted *WT1* and *CA10* sequences (Figure 1). ROC curves were used to calculate the cut-off values of β , which were 0.057 and 0.226 for

average β of probes in *WT1* and *CA10*. The average methylation level of Cg14657517 and Cg19074340, which are located within the *WT1* targeted sequence, was associated with breast cancer with OR of 1.88 (95% CI: 1.25-2.83, $P=0.03$). However, the average methylation level of Cg14054928 and Cg20405017, which are located within the targeted *CA10* sequence, was not statistically significant associated breast cancer risk (OR=0.76, 95% CI: 0.54-1.06, $P=0.11$) (Table 5).

Table 5. The association between gene average CpG sites methylation and risk of female breast cancer in GEO51032.

	Hypomethylation [%]	Hypermethylation [%]	Crude OR (95% CI)	<i>P</i> Value
<i>WT1</i>				
Control	285(83.8)	55(16.2)		1
Case	171(73.4)	62(26.6)	1.88(1.25-2.83)	0.03
<i>CA10</i>				
Control	146(42.9)	194(57.1)		1
Case	116(49.8)	117(50.2)	0.76(0.54-1.06)	0.11

Discussion

This is the first case-control study to investigate the associations between the methylation of *WT1*, *CA10* in leukocyte DNA and breast cancer risk, and the risk of different molecular subtypes of breast cancer in a Chinese female population. After PS adjustment, we observed that methylation of *WT1* was significantly elevated breast cancer risk by 2.07-fold, *CA10* methylation was marginally associated with breast cancer risk with OR of 1.35. Women with *WT1* methylation presented a 1.62 higher risk of luminal A and 2.23 higher risk of luminal B subtype of breast cancer than those without methylation. *CA10* methylation was significantly associated with the risk of luminal B subtype with OR of 1.80. We subsequently used GEO-GSE51032 dataset, a nested case control study with clear temporal relationship between methylation changes and breast cancer, as an external dataset to validate our retrospective study. The nested case control study's results showed a lower but still significant association between *WT1* methylation and breast cancer risk, but the association between *CA10* methylation and breast cancer risk was not statistically significant.

Breast cancer is a heterogeneous disease with different molecular subtypes, which may present different genetic and epigenetic susceptibilities. Previous studies predominantly focused on the aberrant methylation in tissue samples and its association with the risk of different molecular types of breast cancer [21, 22], with few studies having focused on the gene-specific methylation in leukocyte DNA. The methylation alteration in leukocyte DNA presented a response of the hematopoietic system [23]. Leukocyte DNA methylation can represent germline methylation, which can be used to analyze the association with cancer risk [24]. It was further reported that *BRCA1* hypermethylation in peripheral blood DNA was associated with TNBC with an OR of 5.0 [25]. The results of our study indicated that after PS adjustment, *WT1* methylation was associated with the risk of luminal A and luminal B subtypes of breast cancer with ORs of 2.62 and 3.23, and *CA10* methylation was significantly associated with luminal B subtype of breast cancer with OR of 1.80.

WT1 is a zinc finger transcription factor located on 11p13, which was first identified as a tumor suppressor gene. *WT1* exon displayed significantly increased methylation in cancer tissue compared to nonmalignant breast tissue [17]. *WT1* methylation in the promoter and first exon region was shown to be associated with the silencing of *WT1* mRNA expression in MCF-7 and MDA-MB-231 breast cancer cells [9]. Our investigated sequence was 160 bp downstream of the

Laux et al. sequence position. Here, we observed methylation of the CpG island in the first exon of *WT1* in blood leukocyte DNA, which contains 11 CpGs in the CpG island. Furthermore, we used external data from an IPEC- Italy cohort (GEO-GSE51032) with a nested case control study design and found the significant association between *WT1* methylation and breast cancer risk, with two CpG probes inside our sequence, with OR of 1.88.

A previous study showed *CA10* can undergo methylation during breast carcinogenesis in tumor tissue [17]. *CA10* was reported to be hypermethylated among a panel of genes in urine, which may contribute to the highly accurate and early detection of bladder cancer [26]. The result of our study suggested that *CA10* methylation in leukocyte DNA was marginally associated with an elevated breast cancer risk after PS adjustment. The amplified sequence contained 7 CpGs and located in the second exon of *CA10*. The external validation dataset of GEO-GSE51032 only included 2 CpG probes and did not exhibit a significant association between *CA10* hypermethylation and breast cancer risk.

To further investigate the functional relevance of the observed associations, it would be important to test whether methylation of the specific CpGs in *WT1* and *CA10* associated with the alteration of their expression. Therefore, we investigated the correlations between methylation probes and expression using TCGA (<http://cancergenome.nih.gov/>) and Mexpress (<https://mexpress.be/>) databases. The results showed that *WT1* hypermethylation was also negatively correlated with its expression (Cg14657517, $r=-0.204$, $P<0.001$; Cg19074340, $r=-0.201$, $P<0.001$), and *CA10* hypermethylation was negatively related to its mRNA expression as well (Cg14054928, $r=-0.182$, $P<0.001$; Cg20405017, $r=-0.162$, $P<0.001$). Although discounted by different sample-derived DNA, the significant negative correlations between *WT1*, *CA10* methylation and gene expression were consistent with our study and indicated promising potential in breast cancer risk assessment.

In our previous study, we tested the accuracy of MS-HRM by assessing the *WT1* methylation level with both MS-HRM and pyrosequencing, and the results were highly correlated between these two methods [27]. However, the methylation level of leukocyte DNA is relatively low and the limitation of pyrosequencing is 2%. As a reliable and highly sensitive technique, MS-HRM can be used to assess the methylation level of targeted CpGs as low as 0.1% [28]. The high consistency of our results for different runs which making the non-misclassification of methylation level between case and control and the probability of higher sensitivity of MS-HRM comparing pyrosequencing can make our study results more conserved [29].

The limitations of this study should be taken into consideration before drawing a conclusions: first, as in all retrospective analyses, our study may have some recall bias when collecting information on environmental factors. Second, the sample size of our study is not large enough for subgroup analysis, including the subgroup analyses of low frequency environmental factors, such as smoking behavior, therefore their associations with DNA methylation of *WT1*, *CA10* could not be analyzed. Third, selection bias may have occurred, since we recruited the breast cancer patient subjects at the Tumor Hospital of Harbin Medical University, which might not be representative of the distribution of breast cancer patients to some extent.

List Of Abbreviations

WT1: Wilm's Tumor 1; *CA10*: Carbonic Anhydrase 10; IDC: Invasive ductal carcinoma; DCIS: ductal carcinoma in situ; MS-HRM: Methylation-Sensitive High-Resolution Melting; TNBC: Triple Negative Breast Cancer; CIs: Confidence Intervals; OR: Odds Ratio; ROC: receiver operating characteristic curve; ER: Estrogen Receptor; PR: Progesterone Receptor; HER-2: Human Epidermal growth factor Receptor-2; BMI: Body Mass Index

Declarations

Ethics approval and consent to participate

This study was approved by the Human Research and Ethics Committee of Harbin Medical University. Informed written consent was provided by all the subjects.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the National Natural Science Foundation of China (Yashuang Zhao, grant no. 81172743) The funder had no role in the design of the study, analysis, interpretation of data and manuscript writing.

Authors' contributions

GA and GS have been involved in drafting the manuscript. LY and ZH performed subsequent data analysis. WX, ZL and above researchers together completed experiment part of this research. Dr. ZY and Dr. PD revised the manuscript for important intellectual content. All authors have read and approved the manuscript.

Acknowledgements

The authors thank all the patients and healthy volunteers for providing blood samples and all the research staff for their contributions to this project.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA Cancer J Clin* 2018, **68**(6):394-424.
2. Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, Riggs M, Banerjee N, Synnestvedt M, Schlichting E: **DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables.** *Molecular Oncology* 2011, **5**(1):77-92.
3. Esteller M: **Cancer epigenomics: DNA methylomes and histone-modification maps.** *Nature Reviews Genetics* 2007, **8**(4):286-298.
4. Kang HJ, Kim JM, Kim SY, Kim SW, Shin IS, Kim HR, Park MH, Shin MG, Yoon JH, Yoon JS: **A Longitudinal Study of BDNF Promoter Methylation and Depression in Breast Cancer.** *Psychiatry Investigation* 2015, **12**(4):523-531.
5. Woo HD, Kim J: **Global DNA Hypomethylation in Peripheral Blood Leukocytes as a Biomarker for Cancer Risk: A Meta-Analysis.** *Plos One* 2012, **7**(4):e34615.
6. Guan Z, Yu H, Cuk K, Zhang Y, Brenner H: **Whole-Blood DNA Methylation Markers in Early Detection of Breast Cancer: A Systematic Literature Review.** *Cancer Epidemiol Biomarkers Prev* 2019, **28**(3):496-505.
7. Mžik M, Chmelařová M, John S, Laco J, Slabý O, Kiss I, Bohovicová L, Palička V, Nekvindová J: **Aberrant methylation of tumour suppressor genes WT1, GATA5 and PAX5 in hepatocellular carcinoma.** *Clinical Chemistry & Laboratory*

Medicine 2016, **54**(12):1971-1980.

8. Bruno P, Gentile G, Mancini R, Vitis CD, Esposito MC, Scozzi D, Mastrangelo M, Ricci A, Mohsen I, Ciliberto G: **WT1 CpG islands methylation in human lung cancer: A pilot study.** *Biochemical & Biophysical Research Communications* 2012, **426**(3):306-309.
9. Laux DE, Curran EM, Welshons WV, Lubahn DB, Huang THM: **Hypermethylation of the Wilms' tumor suppressor gene CpG island in human breast carcinomas.** *Breast Cancer Research & Treatment* 1999, **56**(1):35-43.
10. Werner H, Re GG, Drummond IA, Sukhatme VP, Rd RF, Sens DA, Garvin AJ, Leroith D, Jr RC: **Increased expression of the insulin-like growth factor I receptor gene, IGF1R, in Wilms tumor is correlated with modulation of IGF1R promoter activity by the WT1 Wilms tumor gene product.** *Proceedings of the National Academy of Sciences of the United States of America* 1993, **90**(12):5828-5832.
11. Paik S: **Expression of IGF-I and IGF-II mRNA in breast tissue.** *Breast Cancer Research & Treatment* 1992, **22**(1):31-38.
12. Silberstein GB, Horn KV, Strickland P, Roberts CT, Daniel CW: **Altered expression of the WT1 Wilms tumor suppressor gene in human breast cancer.** *Proceedings of the National Academy of Sciences of the United States of America* 1997, **94**(15):8132-8137.
13. Nakamura J, Kitajima Y, Kai K, Hashiguchi K, Hiraki M, Noshiro H, Miyazaki K: **Expression of Hypoxic Marker CA IX Is Regulated by Site-Specific DNA Methylation and Is Associated with the Histology of Gastric Cancer.** *American Journal of Pathology* 2011, **178**(2):515.
14. Ivanov S, Liao SY, Ivanova A, Danilkovitch-Miagkova A, Tarasova N, Weirich G, Merrill MJ, Proescholdt MA, Oldfield EH, Lee J et al: **Expression of hypoxia-inducible cell-surface transmembrane carbonic anhydrases in human cancer.** *Am J Pathol* 2001, **158**(3):905-919.
15. Nakamura J, Kitajima Y, Kai K, Hashiguchi K, Hiraki M, Noshiro H, Miyazaki K: **Expression of hypoxic marker CA IX is regulated by site-specific DNA methylation and is associated with the histology of gastric cancer.** *Am J Pathol* 2011, **178**(2):515-524.
16. Sung HY, Ju W, Ahn JH: **DNA hypomethylation-mediated overexpression of carbonic anhydrase 9 induces an aggressive phenotype in ovarian cancer cells.** *Yonsei Med J* 2014, **55**(6):1656-1663.
17. Wojdacz TK, Windeløv JA, Thestrup BB, Damsgaard TE, Overgaard J, Hansen LL: **Identification and characterization of locus-specific methylation patterns within novel loci undergoing hypermethylation during breast cancer pathogenesis.** *Breast Cancer Research* 2014, **16**(1):R17.
18. Shu XO, Yang G, Jin F, Liu D, Kushi L, Wen W, Gao YT, Zheng W: **Validity and reproducibility of the food frequency questionnaire used in the Shanghai Women's Health Study.** *European Journal of Clinical Nutrition* 2004, **58**(1):17-23.
19. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J et al: **European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection.** *Public Health Nutr* 2002, **5**(6B):1113-1124.
20. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel M: **Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015.** *Ann Oncol* 2015, **26**(8):1533-1546.
21. Conway K, Edmiston SN, May R, Pei FK, Chu H, Bryant C, Tse CK, Swift-Scanlan T, Geraerts J, Troester MA: **DNA methylation profiling in the Carolina Breast Cancer Study defines cancer subclasses differing in clinicopathologic characteristics and survival.** *Breast Cancer Research Bcr* 2014, **16**(5):450.
22. Holm K, Hegardt C, Staaf J, Vallonchristersson J, Jönsson G, Olsson H, Borg Å, Ringnér M: **Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns.** *Breast Cancer Research : BCR* 2010, **12**(3):R36.

23. Koestler DC, Marsit CJ, Christensen BC, Accomando W, Langevin SM, Houseman EA, Nelson HH, Karagas MR, Wiencke JK, Kelsey KT: **Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers.** *Cancer Epidemiol Biomarkers Prev* 2012, **21**(8):1293-1302.
24. Wang Y, Li D, Li X, Teng C, Zhu L, Cui B, Zhao Y, Hu F: **Prognostic significance of hMLH1/hMSH2 gene mutations and hMLH1 promoter methylation in sporadic colorectal cancer.** *Medical Oncology* 2014, **31**(7):39.
25. Gupta S, Jaworska-Bieniek K, Narod SA, Lubinski J, Wojdacz TK, Jakubowska A: **Methylation of the BRCA1 promoter in peripheral blood DNA is associated with triple-negative and medullary breast cancer.** *Breast Cancer Research & Treatment* 2014, **148**(3):615-622.
26. Chung W, Bondaruk J, Jelinek J, Lotan Y, Liang S, Czerniak B, Issa JP: **Detection of bladder cancer using novel DNA methylation biomarkers in urine sediments.** *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2011, **20**(7):1483.
27. Gao HL, Wang X, Sun HR, Zhou JD, Lin SQ, Xing YH, Zhu L, Zhou HB, Zhao YS, Chi Q: **Methylation Status of Transcriptional Modulatory Genes Associated with Colorectal Cancer in Northeast China.** *Gut & Liver* 2018, **12**(2).
28. Wojdacz TK, Dobrovic A: **Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation.** *Nucleic Acids Res* 2007, **35**(6):e41.
29. Copeland KT, Checkoway H, McMichael AJ, Holbrook RH: **Bias due to misclassification in the estimation of relative risk.** *Am J Epidemiol* 1977, **105**(5):488-495.

Figures

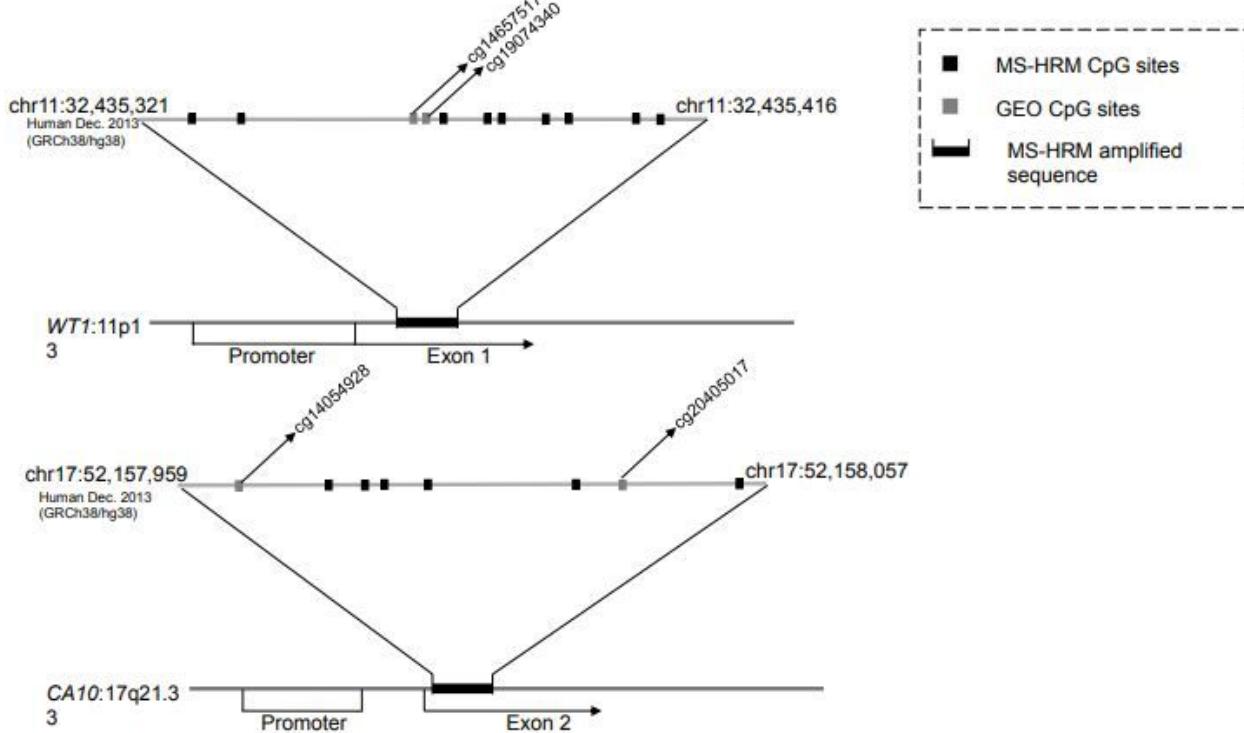


Figure 1

MS-HRM amplified sequence of WT1 and CA10 and the validated Cg sites in GSE51032

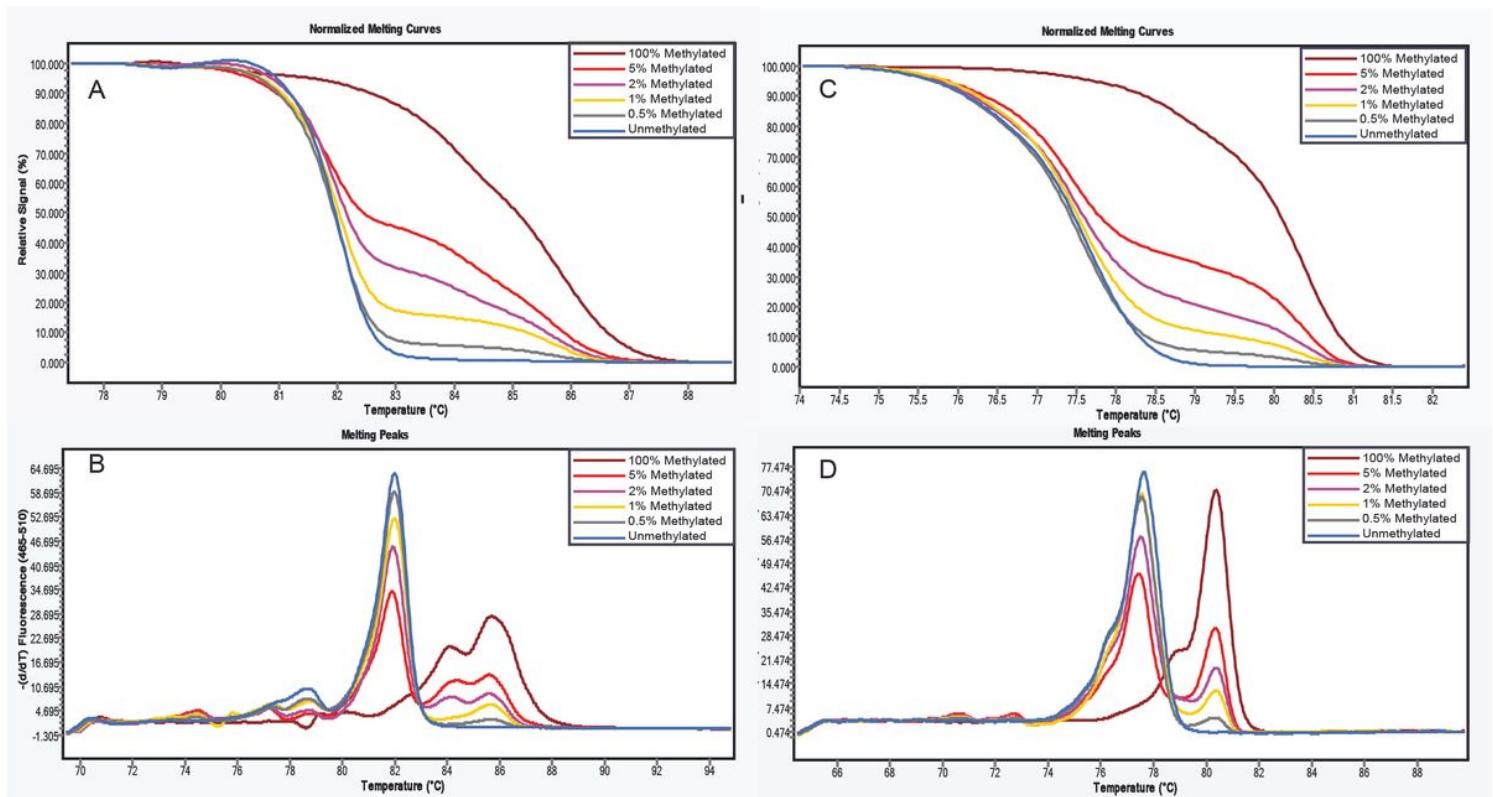


Figure 2

The MS-HRM based method for WT1 and CA10 methylation detection. The figures showed normalized melting curves and melting peaks for standards methylation level and of WT1(A)(B) and CA10(C)(D).The methylation status of the standards were 0%, 0.5%, 1%, 2%, 5%, 100%, respectively

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.docx
- AdditionalFile2.docx
- AdditionalFile3.docx
- AdditionalFile4.docx