narra j

Original Article

DNA methylation profiles for breast cancer subtype classifications: A translational study from microarray to methylation-specific PCR (MSP)

Sonar S. Panigoro[1*], Rafika I. Paramita[2,3,4*], Septelia I. Wanandi[5,6,7], Fadilah Fadilah[3,4], I GNG. Wibisana[1] and Noorwati Sutandyo[8,9]

[1] Division of Surgical Oncology, Department of Surgery, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [2]Doctoral Program in Biomedical Sciences, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [3]Department of Medical Chemistry, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [4]Bioinformatics Core Facilities - IMERI, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [5]Master Program in Biomedical Sciences, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [6]Department of Biochemistry and Molecular Biology, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [7]Molecular Biology and Proteomics Core Facilities – IMERI, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia; [8]Department of Hematology and Medical Oncology, Dharmais National Cancer Center Hospital, Jakarta, Indonesia; [9]Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia

*Corresponding authors: sonar.soni@ui.ac.id (SSP); rafikaindah@ui.ac.id (RIP)

Abstract

Breast cancer subtypes can be categorized based on their gene expression profiles using immunohistochemistry into Luminal A, Luminal B, human epidermal growth factor receptor 2-positive (HER2+), and triple-negative breast cancer (TNBC) subtypes. However, immunohistochemistry has certain limitations that can lead to misclassification. DNA methylation is an epigenetic modification, and changes in the promoter region can alter gene expression and the quantity of functional protein synthesized, disrupting gene function. The aim of this study was to identify DNA methylation biomarkers for subtype classification in breast cancer using microarray and methylation-specific polymerase chain reaction (MSP) methods. DNA samples were extracted, subjected to bisulfite conversion and then used for both the microarray and MSP methods. This study successfully identified differentially methylated CpGs (DMCs) as biomarker for each subtype classification of breast cancer: Luminal A (hypermethylation of *ADAMTSL2* gene; cg14397888), Luminal B (hypomethylation of *ADAMTSL2* gene; cg14397888), HER2+ (hypermethylation of *PTPRN2* gene; cg25910261), and TNBC (hypomethylation of *LCLAT1* gene; cg15652532). The DMC biomarker found for the HER2+ subtype, hypermethylation in the *PTPRN2* gene (cg25910261), has the potential to be used by healthcare providers to identify HER2+ patients and provide the HER2-targeted therapy to improve the patient's survival. In addition, our developed MSP method could produce an effective diagnostic tool for classifying the Luminal A and Luminal B subtypes, with accuracies of 75% and 76%, respectively.

Keywords: Breast cancer, DNA methylation, microarray, methylated-specific PCR, subtype classifications

Copyright: © 2025 by the authors.
This is an open access article distributed under the terms and conditions of the CC BY-NC 4.0.

Introduction

Breast cancer is the most prevalent cancer among women, and according to the GLOBOCAN 2022 data, breast cancer had the highest incidence rate in Indonesia, with 66,271 cases, and the

Received: August 23, 2024 - Accepted: January 6, 2025 - Published online: January 30, 2024

number increased compared to previous years [1]. In the same year, the 5-year breast cancer prevalence was 151.3 per 100,000 people in the country, with a 9.3% mortality rate, surpassing the global mortality rate of 6.9% [1]. Currently, breast cancer subtypes can be categorized based on their gene expression profile using immunohistochemistry (IHC) [2]. The standard IHC tests performed for breast cancer include estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2-positive (HER2), and Ki-67 [2]. The subtypes can be classified into Luminal A, Luminal B, HER2+, and triple-negative breast cancer (TNBC) [2]. Distinct subcategories of breast cancer encompass numerous biologically diverse entities with unique pathological characteristics and clinical implications. Current research shows that breast cancer with varying histological and biological characteristics have unique behaviors, resulting in divergent treatment responses that require different therapeutic approaches [3].

IHC is a cost-effective and easily executable technique as it allows for visualization of staining data using a conventional bright field microscope and enables assessment of protein overexpression within the tissue's morphological context. However, the IHC also has certain shortcomings that can result in misclassification [4]. Reproducibility in Ki-67 staining can contribute to the occurrence of misdiagnosis [5]. The interpretation of the IHC relies on semiquantitative evaluations, making it prone to interobserver variability and potentially resulting in significant discrepancies in IHC outcomes [6]. A study revealed a consistent misclassification trend in the alluvial diagram. For example, half of the HER2+ cases were misclassified as basal-like and Luminal B-like and 95% of Luminal B cases were classified as Luminal A [7].

DNA methylation is an epigenetic modification that involves the attachment of a methyl group to the 5th carbon of the nitrogen base cytosine (C), resulting in the formation of 5-methylcytosine (5-mC). Typically, DNA methylation takes place at C bases next to G bases (known as CpG), and these are commonly located in gene promoter regions [8]. DNA methylation in the promoter region could alter gene expression and the quantity of functional protein synthesized, leading to disruption of gene function [9]. Environmental factors, including exposure to cigarette smoke, food intake, occupational activities, and lifestyle, can affect DNA methylation [10].

There is a lack of research in Indonesia assessing the association between DNA methylation and the classification of breast cancer subtypes, in particular those using peripheral blood as a sample. There are two studies from Indonesia that have discovered the association between methylation of the *BRCA1* gene promoter in cancerous tissues and the clinical, as well as pathology of breast cancer [11,12]. Nevertheless, there has been no previous investigation conducted in Indonesia assessing the association between DNA methylation in peripheral blood and the classification of breast cancer subtypes. An examination utilizing peripheral blood offers various advantages, primarily its non-invasive nature, which makes it a relatively easier and more cost-effective procedure. Studies conducted on Polish [13], French [14,15], Arab [16], and Japanese population [17] have demonstrated an association between DNA methylation of the *BRCA1/2* promoter region in peripheral blood and susceptibility to breast cancer.

Given the present advancements in technology, the Illumina Infinium Methylation EPIC Beadchip Microarray (EPIC-array) is a widely utilized and cost-effective device for efficiently profiling DNA methylation with more than 850,000 CpG sites.[18] The use of EPIC-array for DNA methylation profiling is expected to yield DNA methylation biomarkers capable of categorizing breast cancer subtypes. Nevertheless, the utilization of this technology requires high costs and complex analysis [19], rendering it challenging to implement in regular clinical practice, particularly in healthcare facilities with more constrained resources such as in Indonesia. A translational approach towards simpler analysis is needed for larger patient benefits.

A translational approach that can be employed is methylation-specific polymerase chain reaction (MSP). Currently, there is no available polymerase chain reaction (PCR) detection method for identifying DNA methylation in breast cancer using population epigenomic data from Indonesia. Therefore, it is important to develop a new detection method specifically designed to identify DNA methylation. The aim of this study was to identify DNA methylations as biomarkers of breast cancer subtype classifications by using microarray and translated into MSP method. The MSP method acts as a substitute for sequencing or microarray technology that is more cost-effective and easier to use, making it suitable for regular clinical use.

Methods

Study design, setting and patients

A cross-sectional study was conducted using blood samples from female breast cancer patients who had already undergone therapy according to their subtypes (based on the IHC results). All patients were outpatients at Cipto Mangunkusumo National Hospital and Dharmais National Cancer Center Hospital, both located in Jakarta, Indonesia. The number of subjects was calculated using the cross-sectional formula, with 46 subjects determined as the minimum required sample size. In this study, 48 breast cancer patients were included, with 12 subjects allocated to each subtype group. The inclusion criteria required female patients with invasive breast carcinoma, aged at least 18 years, with comprehensive medical records and histopathological data. Patients with comorbidities were excluded. Clinical data were collected through participant surveys and computerized medical records.

DNA extractions and DNA isolates quality control

DNA samples were extracted from blood by isolating the buffy coat, which yielded a significantly higher amount of DNA, ranging from five to 10 times more. A high DNA concentration was required, as the DNA-bisulfite conversion process later required 200 ng. Whole blood was first centrifuged at 2500 × g for 10 minutes at room temperature (25°C). The buffy coat, located in the middle layer, was separated for DNA extraction. The DNA was extracted using the Genomic DNA Mini Kit (GeneAid, New Taipei City, Taiwan). DNA purity was assessed using an Implen Nanodrop P300 (Implen, Munich, Germany) by measuring the absorbance ratio at 260 nm and 280 nm. The concentration of double-stranded DNA was determined using the Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and the Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA).

Microarray preparations

A total of 200 nanograms (ng) of DNA isolates was subjected to bisulfite conversion using the Zymo EZ-96 DNA Methylation Kit (Zymo Research, Orange, CA, USA) prior to the microarray preparation procedure, which was performed using the Infinium MethylationEPIC v2.0 (Illumina, San Diego, CA, USA). The microarray scanning process was conducted on the Illumina iScan platform (Illumina, San Diego, CA, USA). The intensity of each single nucleotide polymorphism (SNP) in the samples was generated during the procedure. The scanning process produced files in the .idat format, which contained two types of intensity colors: red and green.

Microarray data analysis

The imported .idat file in R was examined for data quality using the Minfi program (https://bioconductor.org/packages/release/bioc/html/minfi.html) [20].The *p*-value parameter was utilized for each CpG probe signal in every sample. The *p*-value was determined by comparing the combined signal of methylated and unmethylated DNA (M + U) to the background signal. Prior to further analysis, the data were normalized using the "preprocessFunnorm" function from the Minfi package. Upon completion of the normalization process, the data were transformed into Beta values. The Beta value was calculated as the ratio of the intensity of methylated probes to the total intensity (the sum of the intensities of methylated and unmethylated probes), ranging from 0 (fully unmethylated) to 1 (fully methylated) [21]. The Limma tool, developed in R/Bioconductor (https://www.bioconductor.org/packages/release/bioc/html/limma.html) [22], was used to identify differentially methylated CpGs (DMCs), which are locations where methylation differences occur. DMCs were defined as CpG probes with an adjusted $p<0.05$, exhibiting hypermethylation with a log2 fold change greater than 1 or hypomethylation with a log2 fold change less than -1 [23]. A CpG probe with a Beta value greater than or equal to 0.3 was classified as methylated, while a Beta value below 0.3 was classified as unmethylated [24].

Comparison of differentially methylated CpGs (DMCs) to other datasets

Given the limited sample size for each subtype in this study, we then compared the identified DMCs with relevant research findings from Gene Expression Omnibus (GEO) dataset. Three distinct datasets were utilized: GSE141338 (consisting of 45 breast cancer subjects), GSE72245

(comprising 118 breast cancer subjects), and GSE72251 (encompassing 119 breast cancer subjects). These datasets contained the DNA methylation profiles of breast cancer patients categorized into four distinct subtypes: Luminal A, Luminal B, HER2+, and TNBC. The Illumina Infinium 450K Human Methylation BeadChip was employed for the analysis of the datasets. Shared DMCs were identified across the three GEO datasets and the study datasets were extracted using a Venn diagram tool (https://www.biotools.fr/misc/venny).

Survival analysis

Clinical outcomes related to DMCs found in each subtype were then analyzed by using Kaplan-Meier plotter database (https://kmplot.com/analysis/index.php?p=service&cancer=breast) [25]. The Kaplan–Meier curve correlates the gene expression in each subtype with time and the percent of relapse-free survival. The Kaplan-Meier plotter database used gene expression data and relapse-free survival information from GEO and European Genome-Phenome Archive [25]. Molecular subtypes were identified using the StGallen criteria, which are based on estrogen and ERBB2 expression [25]. For patient assignment, the gene expression threshold of 500 for probe set "205225_at" was used to determine estrogen positivity, and the probe set "216836_s_at" with a cutoff of 4800 was determined to assign patients into HER2 positive/negative groups [25]. The total number of breast cancer arrays was 9,423, including 7,830 distinct samples from 55 independent datasets [25]. Relapse-free survival data was available for 5268 patients [25].

In this study, the number of patients used for analysis from Luminal A, Luminal B, HER2+, and TNBC were 2,470, 2,005, 1,080, and 1,671 patients, respectively, provided from the Kaplan-Meier plotter database [25]. All the patients involved in this analysis were given systemic treatments (endocrine therapy and chemotherapy, without targeted therapy) [25]. A significant difference was defined as those with a $p<0.05$, using the median as the group cut-off. The hazard ratio (HR) was calculated with a 95% confidence interval (95%CI) to determine the statistical association between gene expression and relapse-free survival.

Primer design of methylation-specific PCR (MSP)

The MSP required the use of two primer sets to detect regions containing CpG sites in bisulfite-treated DNA and to generate distinct PCR results. To achieve this, the main primer pair of MSP was designed to meet two criteria: (1) it had to be specific for bisulfite-converted DNA to prevent false-positive results caused by partial bisulfite conversion; and (2) it needed to clearly differentiate between methylated and unmethylated DNA strands following bisulfite treatment [26]. The identification of the specific genomic region was prioritized based on the results obtained from the EPIC-methylation array. After the desired genomic region was identified, MSP primer pairs were designed using MethPrimer, a web-based program (http://www.urogene.org/cgi-bin/methprimer/meetprimer.cgi). MethPrimer, a freely accessible tool, provided resources and databases for DNA methylation analysis, including the ability to design MSP primers. The design of the MSP primers adhered to specific criteria: (a) a product length ranging from 100 to 300 base pairs; (b) a primary melting temperature (Tm) between 50°C and 60°C; (c) primer lengths ranging from 20 to 30 base pairs; and (d) a 5°C difference in Tm between the forward and reverse primers [26].

DNA methylation detection methods with methylation-specific PCR (MSP)

The PCR was performed at 95°C for 15 minutes, followed by 38 cycles of 95°C for 50 seconds, 59°C for 50 seconds, and 72°C for 50 seconds, with a final extension at 72°C for 10 minutes. The PCR mixture was prepared by combining 1 µL of bisulfite-converted DNA at a concentration of 15 ng/µL, 10 µL of 2× Vazyme SYBR Green PCR Master Mix (Vazyme, Nanjing, People's Republic of China), 0.8 µL of forward and reverse primers at a concentration of 100 µM, and ddH_2O to achieve a total volume of 20 µL. Each sample was processed in duplicate to ensure reproducibility. The PCR products were analyzed by electrophoresis using the BlueGel system (miniPCR bio, Cambridge, MA, USA) with 0.4 g of agarose gel (Invitrogen, Waltham, MA, USA) dissolved in 20 mL of 1× TAE buffer. The PCR product bands were visualized under 254 nm ultraviolet light.

Results

Subject characteristics

A total of 48 breast cancer patients, aged between 32 and 66 years, were included in this study. Among them, eight patients were identified as having familial breast cancer, while the remaining 40 patients were classified as having non-familial breast cancer. The clinical characteristics of the patients are presented in **Table 1**. Of the total number of patients, 26 individuals (54.2%) were found to have a positive metastatic status.

Table 1. Characteristics of breast cancer patients included in the study (n=48)

Characteristics	Frequency (percentage)
Age (year), mean±SD	49±9.1
Breast cancer family history	7 (14.6)
Molecular subtype	
Triple-negative breast cancer (TNBC)	12 (25.0)
HER2+	12 (25.0)
Luminal B	12 (25.0)
Luminal A	12 (25.0)
Metastasis status	10 (20.9)
Breast cancer stadium	
I	4 (8.3)
II	11 (22.9)
III	23 (47.9)
IV	10 (20.9)

According to the IHC data (**Table 1**), 12 patients were assigned to each of the four subtype groups: Luminal A, Luminal B, HER2+, and TNBC. The IHC results were used to classify the subtypes as follows: ER/PR positivity (>10%) and Ki67 negativity (<20%) were categorized as the Luminal A group; ER/PR positivity (>10%) and Ki67 positivity (>20%) were categorized as the Luminal B group; HER2+ positivity (+3) was classified as the HER2+ group; and negativity for ER/PR/HER2 was classified as the TNBC group following the guidelines [2,3].

Quality control of DNA samples and microarray results

The isolated DNA were subjected to quality control by assessing their purity through the absorbance ratio of 260/280 nm using a Nanodrop and measuring their quantitative concentration with a Qubit 4.0 fluorometer. Based on the quality control results, it was confirmed that all DNA isolate samples exhibited excellent purity (with ratio of absorbance A260/A280 nm ranging from 1.8 to 2.0) and concentrations suitable for bisulfite conversion and microarray analysis (at least 300 ng/μL).

The quality control process of the .idat files, conducted using the Minfi program, revealed that all samples exhibited a *p*-value of less than 0.01, indicating excellent data quality. CpG probes containing single nucleotide polymorphisms (SNPs) were excluded. Of the 1,105,209 CpG probes initially examined, only 903,232 CpG probes met the quality control criteria. As part of the quality control process, the Beta values of each CpG probe were also assessed. The Beta values were observed to fall within the low range (0–0.2) and the high range (0.8–1.0) (**Supplementary 1**). A CpG probe with a Beta value of ≥0.3 was classified as methylated, whereas a Beta value below 0.3 was classified as unmethylated [24].

Analysis of differentially methylated CpGs (DMC)

In this step, a comprehensive analysis was conducted to identify probes exhibiting differential methylation in each molecular subtype. This analysis, called paired analysis, involved calculating the differences between data groups within each individual, and then averaging these differences across individuals to determine whether there was an overall significant difference in the mean methylation levels for each CpG site. CpG sites showing significant differences in methylation levels between two different groups were referred to as differentially methylated CpGs (DMCs).

Original Article

Table 2. Differentially methylated CpG (DMCs) identified in each subtype group which has the potential to be subtype biomarkers

Molecular subtype	Methylation status	CpG locus	Methylation positions	CpG regions	Gene	Log2FC
Luminal A (Luminal A vs non-Luminal A)	Hypermethylation	cg11799593	chr12: 68452089	OpenSea/promoter	*LINC02384*	1.49
		cg14397888	chr9: 133554506	Island/gene body	*ADAMTSL2*	1.59
Luminal B (Luminal B vs non-Luminal B)	Hypomethylation	cg14397888	chr9: 133554506	Island/gene body	*ADAMTSL2*	-3.06
HER2+ (HER2+ vs non-HER2+)	Hypermethylation	cg25910261	chr7: 157613273	Island/promoter	*PTPRN2*	1.06
TNBC (TNBC vs non-TNBC)	Hypomethylation	cg15652532	chr2: 30446893	Shore/promoter	*LCLAT1*	-1.5
		cg26371957	chr12: 630114	OpenSea/promoter	*NINJ2*	-1.48

Log2FC: log2 fold change

In this study, the DMCs were analyzed in four different groups: Luminal A vs non-Luminal A (Luminal B, HER2+, and TNBC); Luminal B vs non-Luminal B (Luminal A, HER2+, and TNBC); HER2+ vs non-HER2+ (Luminal A, Luminal B, and TNBC); TNBC vs non-TNBC (Luminal A, Luminal B, and HER2+). The Limma software, developed in R/Bioconductor, was used to detect the DMCs. DMCs were defined as CpG probes with an adjusted $p<0.05$. Hypermethylation was defined as DMC with log2 fold change (log2FC) > 1 and a Beta value ≥0.3, while hypomethylation was defined as DMC with log2FC <-1 and a Beta value <0.3. The volcano plot of DMCs in each subtype is presented in **Figure 1**. The DMC list identified from this study is presented in **Table 2**.

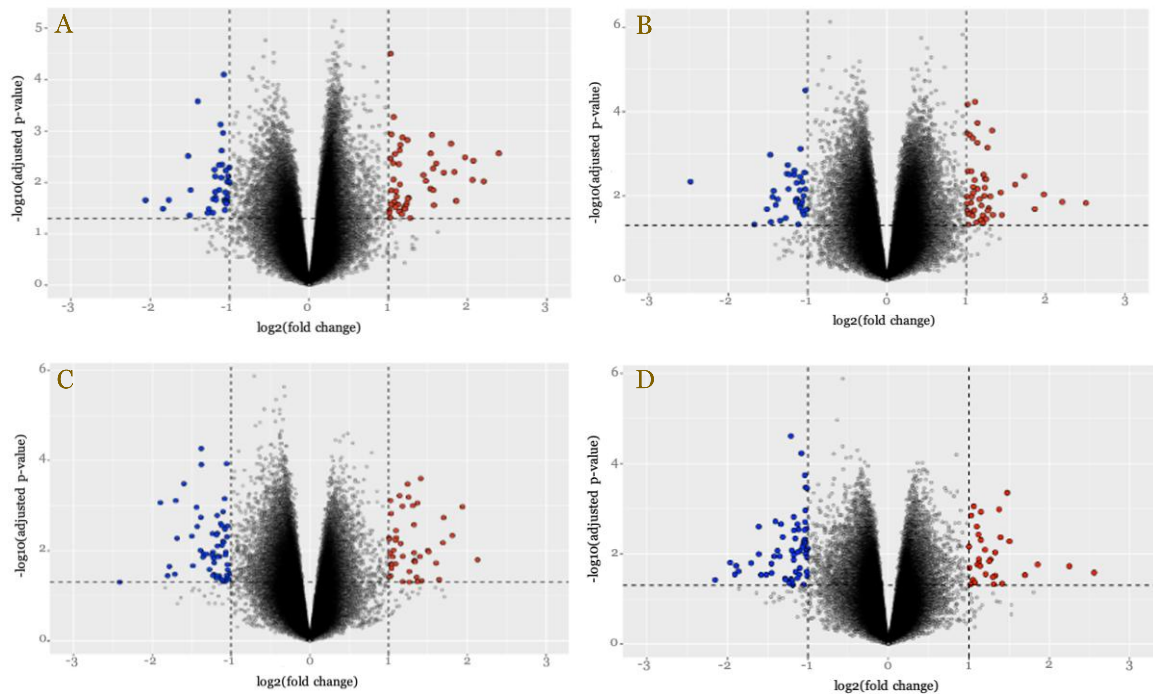

Figure 1. Volcano plot of differentially methylated CpG (DMC): (A) Luminal A, (B) Luminal B, HER2+ and (D) triple-negative breast cancer (TNBC). Hypomethylations are shown in blue dots (adjusted $p<0.05$; log2 fold changes (log2FC) <-1), hypermethylations are shown in red dots (adjusted $p< 0.05$; log2FC >1).

To assess potential biomarkers for the subtypes, a comparison was made between our identified DMCs and those reported in other datasets (using tissue samples) that employed a similar methodology. By examining the Venn diagram (**Figure 2**), the cg15652532 biomarker was found to overlap in the DMCs identified in our dataset for the TNBC subtype and in the other three datasets as well. This indicated that DNA methylation in TNBC was found not only in germinal tissue but also in somatic tissue. Unfortunately, no overlaps were found in the Luminal A, Luminal B, and HER2+ subtypes. The cg14397888 biomarker, identified as potentially useful for distinguishing between Luminal A and Luminal B subtypes, was not included in the Illumina Infinium 450K Human Methylation manifest. The absence of cg14397888 in the manifest probes used across the other three datasets may have caused this discrepancy.

Based on our analysis, DNA methylation biomarkers were identified for each subtype: cg14397888 (Luminal A and Luminal B), cg25910261 (HER2+), and cg15652532 (TNBC). It was shown that the Beta value scores between the two distinct groups were sufficiently different, indicating methylated and unmethylated DNA (**Figure 3**). These DMCs have the potential to be further analyzed, as they could serve as biomarkers for classifying the molecular subtypes of breast cancer.

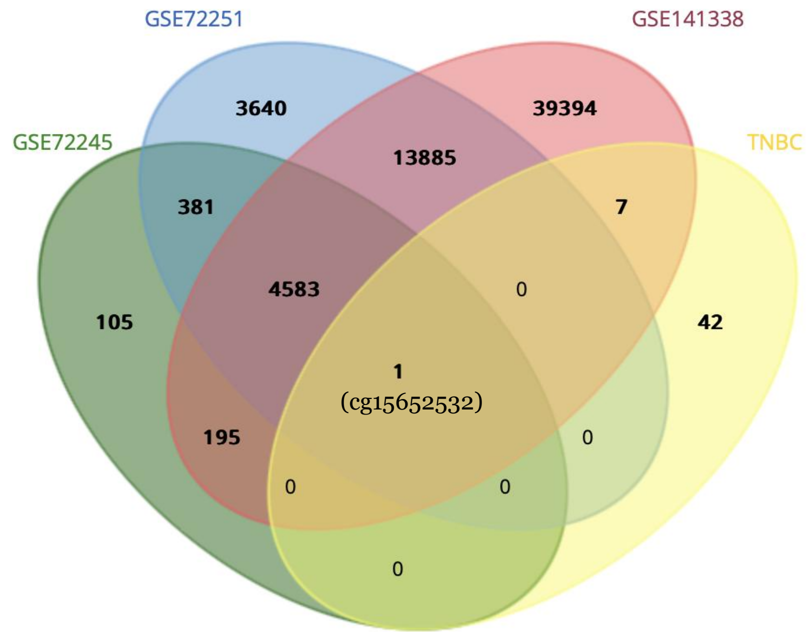

Figure 2. Venn diagram of triple-negative breast cancer (TNBC) subtype's biomarkers from the present study dataset (TNBC) and other three datasets from Gene Expression Omnibus (GEO) databases (GSE141338, GSE72245, and GSE72251).

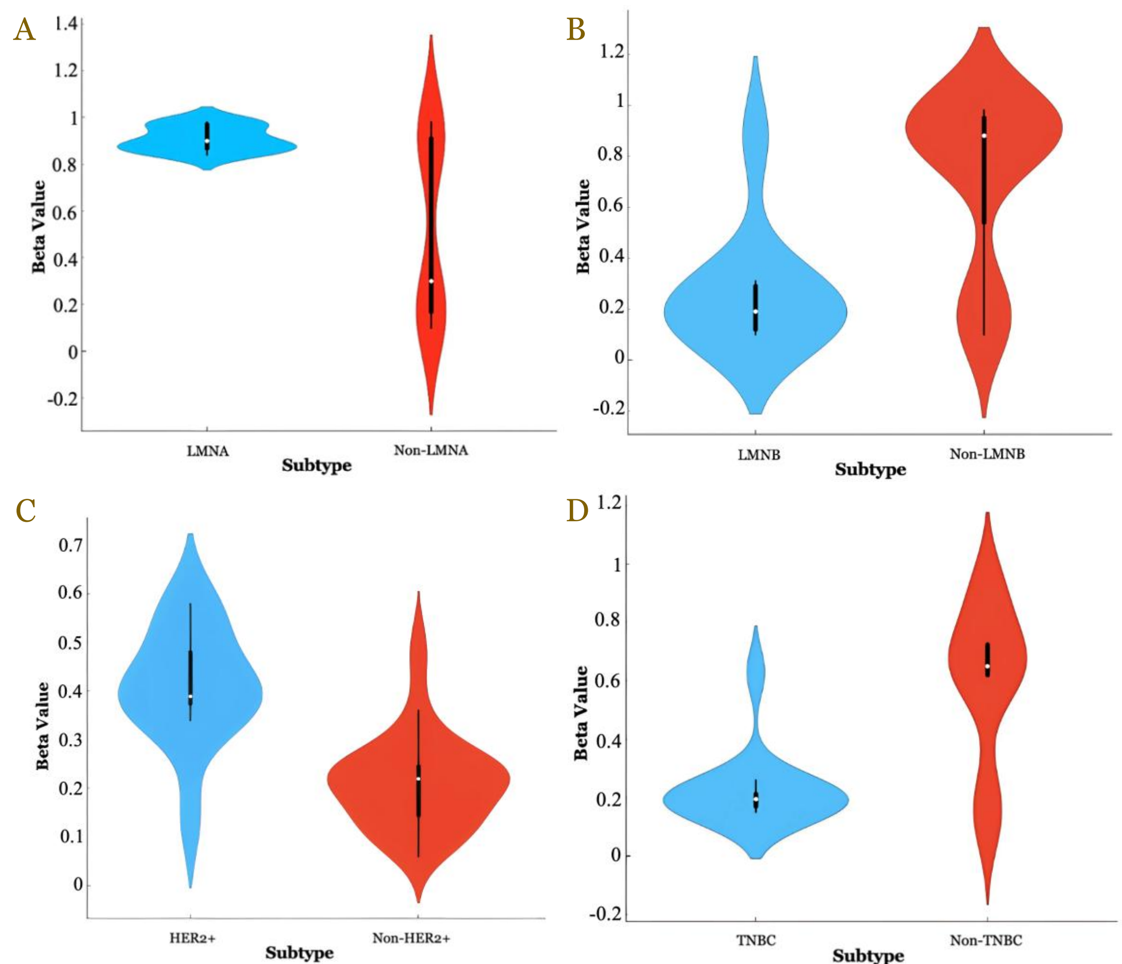

Figure 3. Violin plot of Beta value for each differentially methylated CpG (DMC): (A) cg14397888 (Luminal A vs non-Luminal A); (B) cg14397888 (Luminal B vs non-Luminal B); (C) cg25910261 (HER2 + vs non-HER2+); and (D) cg15652532 (triple-negative breast cancer (TNBC)) vs non-TNBC).

Survival analysis

We used relapse-free survival analysis with the Kaplan-Meier plotter database to analyze the prognostic outcome due to DMCs in *ADAMTSL2*, *PTPRN2*, *and LCLAT1* genes (**Figure 4**). It was informed that relapse-free survival in the HER2+ subtype was significantly different between low and high expression of *PTPRN2* expression (p=0.0096). Hypermethylation in *PTPRN2* promoter gene (cg25910261) found in HER2+ subtype could downregulate the *PTPRN2* expression. Low expression of *PTPRN2* showed lower relapse-free survival probability than high expression of *PTPRN2*. Therefore, hypermethylation in *PTPRN2* (cg25910261) could be proposed as HER2+ identifier and relapse-free survival predictor as well.

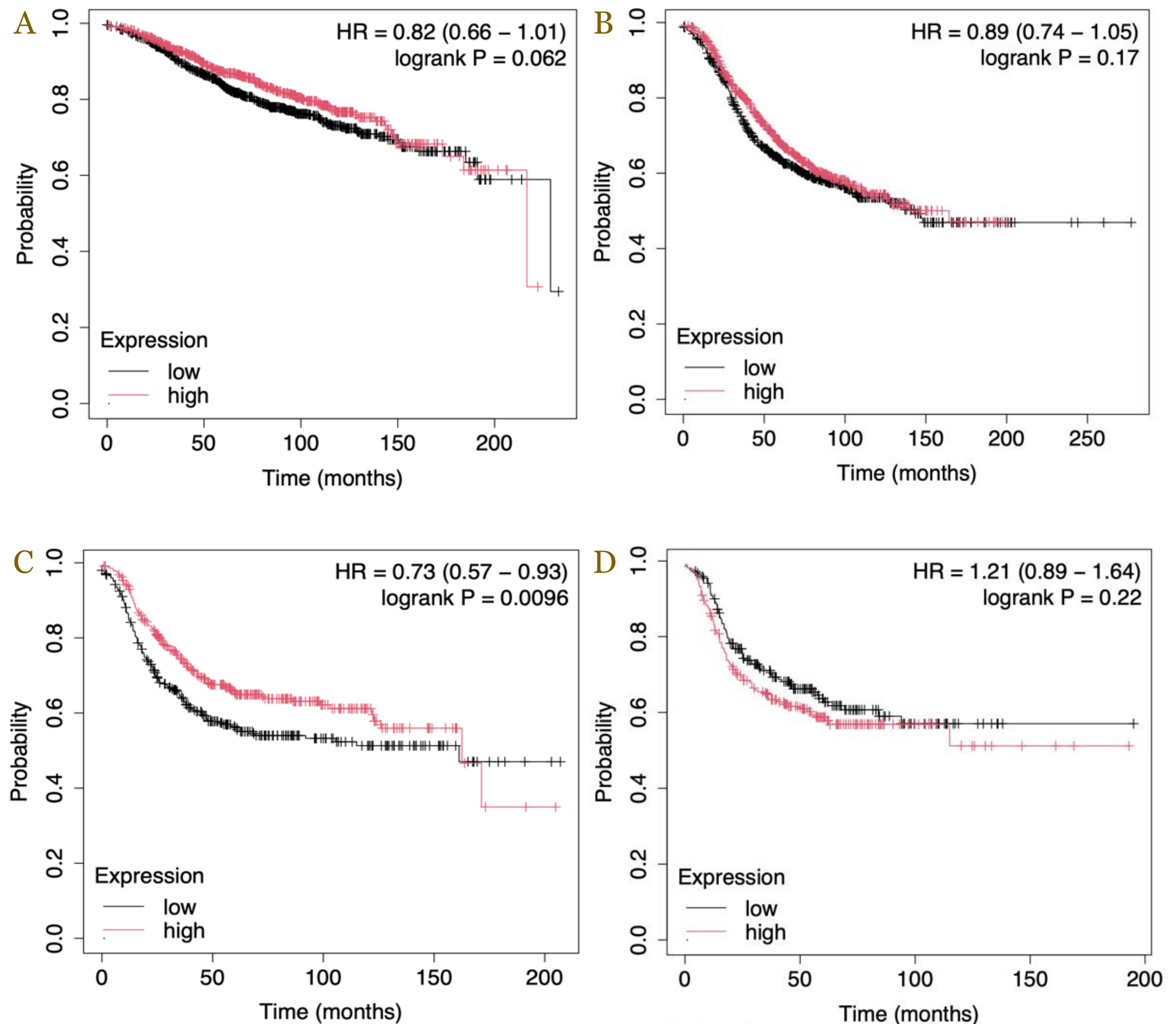

Figure 4. Kaplan-Meier plot: (A) *ADAMTSL2* expression in Luminal A patients; (B) *ADAMTSL2* expression in Luminal B patients; (C) *PTPRN2* expression in HER2+ patients; and (D) *LCLAT1* expression in triple-negative breast cancer (TNBC) patients. Significant differences are defined with a $p<0.05$; the hazard ratio (HR) with a 95% confidence interval.

DNA methylation analysis of methylation-specific PCR (MSP)

The chosen biomarkers were derived from the findings of the microarray and were subsequently translated using PCR. Epigenomic biomarkers, specifically DNA methylation indicators, were utilized through the application of MSP method. The initial step undertaken was the design of main MSP primers for each group (**Table 3**). Using MethPrimer software, primers were designed for each DMC biomarker for MSP analysis, consisting of methylated and unmethylated primers.

Table 3. Designed primers for each differentially methylated CpG biomarker

Group	Gene	cgID	Primer sequences	Product size (bp)
Luminal A / Luminal B	*ADAMTSL2*	cg14397888	Methylated: F: TTTTTTTTAGGTGTGGAATTAGAAC (25) R: CGAATAACCGAACAACCGAT (20)	238
			Unmethylated: F: TTTTTTTAGGTGTGGAATTAGAATG (25) R: CCAAATAACCAAACAACCAAT (21)	239
HER2+	*PTPRN2*	cg25910261	Methylated: F: TAGTCGTTTTTGTTAGAATTATCGG (25) R: TAATTCCTTCAACTCCTAAAACGAT (25)	211
			Unmethylated: F: GTAGTTGTTTTTGTTAGAATTATTGG (26) R: TAATTCCTTCAACTCCTAAAACAAT (25)	212
TNBC	*LCLAT1*	cg15652532	Methylated: F: TTGATTTGTTTTGTGATTTCGC (22) R: TAAAAACTCCTTTAAAAAAACCTCG (25)	239
			Unmethylated: F: TTATTGATTTGTTTTGTGATTTTGT (25) R: AAAACTCCTTTAAAAAAACCTCACT (25)	240

In MSP analysis, it is advisable to initially examine the dissociation curve and determine if there is a distinct peak indicating primary methylation at 81°C. If a dissociation curve has a distinct peak at 81°C for methylated primers, it indicates the presence of methylated DNA. If not, then the sample does not include any methylation DNA. The PCR products showing a peak at 81°C (methylated) and below 81°C (unmethylated) were further analyzed by electrophoresis. The electrophoresis bands also confirmed that the methylated and unmethylated products had the same size as originally designed (**Supplementary 2**).

The DNA methylation and unmethylation analysis results obtained using MSP were compared with the microarray results using accuracy, sensitivity, and specificity scores. When comparing the MSP analysis to the microarray analysis, the test accuracy for the Luminal A and Luminal B subtypes was determined to be 75% and 76%, respectively (**Table 4**). The accuracy for the HER2+ and TNBC subtypes was only 49% and 54%, respectively (**Table 4**). Hence, the utilization of the MSP approach was demonstrated to have promising testing capabilities in the Luminal A and Luminal B subtype groups.

Table 4. Evaluation criteria of the DNA methylation test utilizing the methylation-specific PCR (MSP) technique compared to microarray

Criteria	Luminal A (cg14397888/ *ADAMTSL2*)	Luminal B (cg14397888/ *ADAMTSL2*)	HER2+ (cg25910261/ *PTPRN2*)	TNBC (cg15652532/ *LCLAT1*)
Accuracy	0.75	0.76	0.49	0.54
Precision	0.76	0.69	0.36	0.46
Sensitivity	0.86	0.56	0.80	1.00
Specificity	0.58	0.86	0.34	0.24

Discussion

This study sought to identify DNA methylation biomarkers that could be used to differentiate the breast cancer subtypes using MSP method. Several DMCs in each subtype group, which had the potential as subtype biomarkers, were identified (**Table 2**). The cg14397888/*ADAMTSL2* was associated with Luminal A and Luminal B subtype but had different methylation status: hypermethylated in Luminal A and hypomethylated in Luminal B. The downregulation or upregulation of *ADAMTSL2* may be associated with the transforming growth factor beta (TGFβ) signaling pathway [27]. A study found that *ADAMTSL2* has the ability to inhibit the function of TGFβ by attaching itself to fibrillin-1 and latent TGFβ-1, hence hindering the release of unbound TGFβ [27]. The expression of *ADAMTSL2*, TGF-β, and phospho-SMAD2 plays a crucial role in the development of breast cancer, as they are part of the canonical TGF-β signaling pathway [27-29].

We also found a methylation biomarker in the HER2+ subtype, cg25910261/*PTPRN2* (**Table 2**). Hypermethylation of the *PTPRN2* promoter gene in the HER2+ subtype can cause gene silencing, that might downregulate the *PTPRN2* expression. Previous investigations have found that *PTPRN2* promoter hypermethylation is present in both lung glioblastoma and adenocarcinoma, indicating a tumor suppressor function [30,31]. A previous study showed that *PTPRN2* expression in breast cancer was higher than in healthy population, but no further data about the *PTPRN2* expression in HER2+ subjects [31].

In our study, we found that downregulation of *PTPRN2* expression had lower relapse-free survival probability in HER2+ subjects without targeted therapy. A previous study showed that 12 weeks of weekly paclitaxel plus dual HER2 targeted therapy may be an effective de-escalated neoadjuvant regimen in early HER2+ breast cancer patients, with high pathological complete response rates and favorable 5-year survival [32]. Therefore, the DMC biomarker found for HER2+ subtype, hypermethylation in *PTPRN2* (cg25910261), can help clinicians identify HER2+ patients early and give HER2-targeted therapy for better patient survival.

Hypomethylation of the *LCLAT1* promoter was observed in the TNBC subtype group, specifically in the TNBC patient group (**Table 2**). This hypomethylation may have led to an increase in the expression of the *LCLAT1* gene. A study [28] demonstrated that the acyltransferase *LCLAT1/LYCAT* was necessary for EGF-induced phosphatidylinositol-3,4,5-trisphosphate and Akt activation in MDA-MB-231 cells. The upregulation of *LCLAT1* expression may enhance EGF-induced Akt signaling and play a role in the PI3/Akt signaling pathway.

The present study successfully developed MSP method for classifying Luminal A and Luminal B subtypes (accuracy of 75% and 76%, respectively). Alongside with microarray approach, the MSP is a widely utilized technique that relies on the conversion of modified DNA samples. This method involves subjecting the samples to PCR amplification using specific primers for methylated and unmethylated DNA. MSP has been developed as a minimally invasive method for early diagnosis or prognosis of various types of cancer and is commonly employed for this purpose [33]. The present study successfully produced effective diagnostic tools for classifying Luminal A and Luminal B subtypes by MSP method. Nevertheless, this result is anticipated to have advantages in identifying individuals with low-risk breast cancer (Luminal A), who only require hormone therapy [2] by using less invasive techniques (without biopsy) and at a cost-effective price.

Despite the findings, this study has some limitations. Biomarker testing has not been conducted on other breast cancer patient populations outside the test sample, and cross-validation has not yet been performed. The accuracy, precision, specificity, and selectivity of the MSP method were relatively low for other types of breast cancer and this could be improved by incorporating more data, defining better inclusion or exclusion criteria, and addressing potential sources of bias. Further refinement of the MSP analysis is therefore necessary to achieve more satisfactory accuracy.

Conclusion

This study successfully identified DMCs as biomarkers for each subtype classification of breast cancer: Luminal A (hypermethylation; cg14397888/*ADAMTSL2*), Luminal B (hypomethylation; cg14397888/*ADAMTSL2*), HER2+ (hypermethylation; cg25910261/*PTPRN2*), and TNBC (hypomethylation; cg15652532/LCLAT1). The DMC biomarker found for HER2+ subtype, hypermethylation in *PTPRN2* (cg25910261), had the potential to help clinicians identify HER2+ patients early and prescribe HER2-targeted therapy for a better survival rate. For the translational approach, MSP method produced effective diagnostic tools for classifying the Luminal A and Luminal B subtypes, with accuracies of 75% and 76%, respectively.

Ethics approval

The study received clearance from the Ethical Committee of the Faculty of Medicine, Universitas Indonesia, on September 4, 2023 (approval number: KET-1140/UN2.F1/ETIK/PPM.00.02/2023). The procedures were conducted in accordance with the relevant norms and regulations.

Prior to the collection of the patients' samples, all individuals provided their informed consent for participation in this study.

Acknowledgments

None.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This research was funded by RIIM-BRIN/LPDP Research Grant (Grant number: 34/IV/KS/05/2023; 22/PKS/WRIII-DISTP/UI/2023; Grant receiver name: Sonar Soni Panigoro)

Underlying data

The full data of Beta values of each CpG probe (Supplementary 1) are available from: https://doi.org/10.6084/m9.figshare.27634503.v1 and the results of electrophoresis bands during the methylation-specific PCR (MSP) (Supplementary 2) are available from: https://doi.org/10.6084/m9.figshare.27634506

Declaration of artificial intelligence use

We hereby confirm that no artificial intelligence (AI) tools or methodologies were utilized at any stage of this study, including during data collection, analysis, visualization, or manuscript preparation. All work presented in this study was conducted manually by the authors without the assistance of AI-based tools or systems.

How to cite

Panigoro SS, Paramita RI, Wanandi SI, *et al*. DNA methylation profiles for breast cancer subtype classifications: A translational study from microarray to methylation-specific PCR (MSP). Narra J 2025; 5 (1): e1364 - http://doi.org/10.52225/narra.v5i1.1364.

References

1. Bray F, Laversanne M, Sung H, *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229–263.
2. Kementerian Kesehatan Republik Indonesia. Pedoman nasional pelayanan kedokteran tata laksana kanker payudara. Indonesia: 2018. p. 58–60.
3. Dai X, Li T, Bai Z, *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res 2015;5(10):2929–2943.
4. Furrer D, Sanschagrin F, Jacob S, *et al.* Advantages and disadvantages of technologies for HER2 Testing in Breast Cancer Specimens. Am J Clin Pathol 2015;144(5):686–703.
5. Nielsen TO, Leung SCY, Rimm DL, *et al.* Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. J Natl Cancer Inst 2021;113(7):808–819.
6. Holm J, Yu NYL, Johansson A, *et al.* Concordance of Immunohistochemistry-Based and Gene Expression-Based Subtyping in Breast Cancer. JNCI Cancer Spectr 2021;5(1):1–9.
7. Allott EH, Geradts J, Cohen SM, *et al.* Frequency of breast cancer subtypes among African American women in the AMBER consortium. Breast Cancer Res 2018;20(1):1–9.

8. Stirzaker C, Taberlay PC, Statham AL, *et al.* Mining cancer methylomes: Prospects and challenges. Trends Genet 2014;30(2):75–84.
9. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology 2013;38(1):23–38.
10. Fragou D, Pakkidi E, Aschner M, *et al.* Smoking and DNA methylation: Correlation of methylation with smoking behavior and association with diseases and fetus development following prenatal exposure. Food Chem Toxicol 2019;129(April):312–327.
11. Harahap WA, Sudji IR, Nindrea RD. BRCA1 Promoter Methylation and Clinicopathological Characteristics in Sporadic Breast Cancer Patients in Indonesia. Asian Pac J Cancer Prev 2018;19(9):2643–2649.
12. Sugiharto S, Massi MuhN, Wahid S, *et al.* The Correlation of BRCA 1 Promoter Methylation and Clinicopathological Appearance in Breast Cancer. Int J Sci Basic Appl Res IJSBAR 2018;41(1 SE-Articles):51–60.
13. Prajzendanc K, Domagała P, Hybiak J, *et al.* BRCA1 promoter methylation in peripheral blood is associated with the risk of triple-negative breast cancer. Int J Cancer 2020;146(5):1293–1298.
14. Bosviel R, Garcia S, Lavediaux G, *et al.* BRCA1 promoter methylation in peripheral blood DNA was identified in sporadic breast cancer and controls. Cancer Epidemiol 2012;36(3):e177–e182.
15. Bosviel R, Durif J, Guo J, *et al.* BRCA2 Promoter Hypermethylation in Sporadic Breast Cancer. Https://HomeLiebertpubCom/Omi 2012;16(12):707–710.
16. Al-Moghrabi N, Al-Qasem AJS, Aboussekhra A. Methylation-related mutations in the BRCA1 promoter in peripheral blood cells from cancer-free women. Int J Oncol 2011;39(1):129–135.
17. Iwamoto T, Yamamoto N, Taguchi T, *et al.* BRCA1 promoter methylation in peripheral blood cells is associated with increased risk of breast cancer with BRCA1 promoter methylation. Breast Cancer Res Treat 2010 1291 2010;129(1):69–77.
18. Zhang W, Young JI, Gomez L, *et al.* Critical evaluation of the reliability of DNA methylation probes on the Illumina MethylationEPIC v1.0 BeadChip microarrays. Epigenetics 2024;19(1):0–13.
19. Pruneri G, De Braud F, Sapino A, *et al.* Next-Generation Sequencing in Clinical Practice: Is It a Cost-Saving Alternative to a Single-Gene Testing Approach? PharmacoEconomics - Open 2021;5(2):285–298.
20. Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al.* Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30(10):1363–1369.
21. Cappetta M, Fernandez L, Brignoni L, *et al.* Discovery of novel DNA methylation biomarkers for non-invasive sporadic breast cancer detection in the Latino population. Mol Oncol 2021;15(2):473–486.
22. Ritchie ME, Phipson B, Wu D, *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47.
23. Fan Y, Gao D, Zhang Y, *et al.* Genome-Wide Differentially Methylated Region Analysis to Reveal Epigenetic Differences of Articular Cartilage in Kashin–Beck Disease and Osteoarthritis. Front Cell Dev Biol 2021;9(March):1–10.
24. Warden CD, Lee H, Tompkins JD, *et al.* COHCAP: An integrative genomic pipeline for single-nucleotide resolution DNA methylation analysis. Nucleic Acids Res 2013;41(11):1–11.
25. Győrffy B. Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. Comput Struct Biotechnol J 2021;19:4101–4109.
26. Casadio V, Salvi S. Cell-free DNA as Diagnostic Markers: Methods and PRotocols. 2019.
27. Id ZH, Hu X, Wei Y, *et al.* ADAMTSL2 is a potential prognostic biomarker and immunotherapeutic target for colorectal cancer : Bioinformatic analysis and experimental verification 2024:1–20.
28. Figueroa JD, Flanders KC, Garcia-Closas M, *et al.* Expression of TGF-β signaling factors in invasive breast cancers: relationships with age at diagnosis and tumor characteristics. Breast Cancer Res Treat 2009;121(3):727.
29. Chan MKK, Chan ELY, Ji ZZ, *et al.* Transforming growth factor-β signaling: from tumor microenvironment to anticancer therapy. Explor Target Anti-Tumor Ther 2023;4(2):316.
30. Bourgonje AM, Verrijp K, Schepens JTG, *et al.* Comprehensive protein tyrosine phosphatase mRNA profiling identifies new regulators in the progression of glioma. Acta Neuropathol Commun 2016;4(1):96.
31. Sengelaub CA, Navrazhina K, Ross JB, *et al.* PTPRN 2 and PLC β1 promote metastatic breast cancer cell migration through PI (4,5)P 2 -dependent actin remodeling . EMBO J 2016;35(1):62–76.

32. Nitz U, Gluz O, Graeser M, *et al.* De-escalated neoadjuvant pertuzumab plus trastuzumab therapy with or without weekly paclitaxel in HER2-positive, hormone receptor-negative, early breast cancer (WSG-ADAPT-HER2+/HR–): survival outcomes from a multicentre, open-label, randomised, phase 2 trial. Lancet Oncol 2022;23(5):625–635.

33. Kundu S, Das R, Laskar S, *et al.* Principles of bi-sulfite conversion of DNA and methylation-specific PCR (MSP) in biological research. Epigenetics Methods. Elsevier; 2020. p. 17–36.