Article

Methylation profile of small breast cancer tumors

Aleksey M. Krasnyi[1]*, Alsu A. Sadekova[1], Vlada V. Kometova[1], Valeriy V. Rodionov[1] and Gennadiy.T. Sukhikh[1,2]

[1] National Medical Research Center for Obstetrics, Gynecology and Perinatology of Ministry of Healthcare of Russian Federation, Ac.Oparina str. 4, Moscow, Russia 117997
[2] I.M. Sechenov First Moscow State Medical University, Ministry of Healthcare of Russian Federation, B. Pirogovskaya str. 2-4, Russia 119991
* Correspondence: alexred@list.ru ; Tel.: +7-963-750-35-35 (A.M.K);

Abstract: The DNA methylation profile of breast cancer differs from that in healthy tissue and can be used as a diagnostic and prognostic biomarker. Aim of the study: to compare gene methylation in small malignant breast tumors less than 2 cm and in healthy tissue and fibroadenoma. Methylation of the following 15 genes was studied: *MAST1, PRDM14, ZNF177, DNM2, SSH1, AP2M1, CACNA1E, CPEB4, DLGAP2, CCDC181, GCM2, ITPRIPL1, POM121L2, KCNQ1, TIMP3.* Methods: analysis was made by our modified MS-HRM method followed confirmation of the results by pyrosequencing. The genes were selected from publications that studied DNA methylation in breast cancer with high genome coverage. The study group included 48 samples of breast cancer, the control group included 24 samples of fibroadenoma and 24 samples of healthy tissue. Results: significant differences were found in methylation of 8 genes: *CCDC181, GSM2, ITPRIPL1, ZNF177, CACNA1E, DLGAP2, TIMP3* (all $p<0.001$), and *PRDM14* ($p=0.002$). The most accurate diagnostic value, based on logistic regression, was shown with the compound of two genes – *CCDC181* and *ZNF177* (AUC=0.99) in pyrosequencing analysis. Conclusion: small breast cancer tumors have a specific DNA methylation profile that distinguishes them from healthy tissue and benign proliferative lesions.

Keywords: breast cancer, small tumors, DNA methylation, *CCDC181, ZNF177,* fibroadenoma, biomarker, MS-HRM, pyrosequencing.

1. Introduction

In 2020, 2.3 million women were diagnosed with breast cancer, making it the most common cancer in the world [1]. It is known that in the early stages of cancer, including breast cancer, there are changes in the methylation of many genes [2].

DNA methylation occurs by modification of DNA through the addition of a methyl group to the position 5' of cytosine which precedes guanine (CpG). CpG is often found in high-density in genome, forming CpG islands. The study of methylation of CpG islands in promoter regions of genes is important because hypermethylation of CpG islands can lead to the suppression of gene transcription, in particular, through downregulation of tumor suppressor genes – to the development of cancer. Aberrant DNA methylation in tumors, as an early diagnostic and prognostic marker undergo vigorous research with high genome coverage using bioinformatic approaches with a perspective of introduction into clinical practice. The gene panels proposed by researchers show high accuracy, but the sets of genes vary considerably in the panels. In this study, gene methylation in small breast cancer tumors (<2 cm) was analyzed, using gene panels that have been proposed in the studies with high genome coverage. Only genes with CpG islands in promoter regions were used in the study.

Genes included in the study: *MAST1, PRDM14, ZNF177* were selected for the study using a bioinformatics analysis approach of publicly available datasets by *Mao* et al. [3]; genes *DNM2, SSH1, AP2M1, TIMP3* were selected from a similar article by Panagopoulou et.al. [4]; genes *GCM2, ITPRIPL1, CCDC181* were selected from the studies using a bioinformatics analysis approach of publicly available datasets and results of own research by Wang et al. [5]; genes *CACNA1E, CPEB4, DLGAP2* were selected from study using a bioinformatics analysis approach of publicly available datasets with followed by

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2022 doi:10.20944/preprints202211.0513.v1

experimental confirmation of the results by Luo et al. [6]; genes *POM121L2, KCNQ1* were selected from the paper studying methylation in normal healthy breast epithelium which is a potential origin of breast cancer [7].

2. Results

Analysis of clinical data of the study group (n=48) and the control group (n=24) showed no significant differences in age: 51 (40.7; 60.2) and 47 (42.2; 52) years, respectively (p=0.22); BMI also had no statistically significant differences: in the study group BMI was 27.6 (24.7; 32.3), in the control group – 26.1 (23.2; 27.7) kg/m² (p=0.06). In the study group, the tumor size was 1.5 (1.3: 1.8) cm. The following types of breast cancer were found in the study group: Lum A – 16(33.3%), Lum B- – 17(35.4%), Lum B+ – 7(14.6%), HER2+ – 4(8.3%), TH – 4(8.3%). 17 (35.4%) patients of the study group had metastases in lymph nodes.

The methylation level of 15 genes *MAST1, PRDM14, ZNF177, DNM2, SSH1, AP2M1, CACNA1E, CPEB4, DLGAP2, CCDC181, GCM2, ITPRIPL1, POM121L2, KCNQ1, TIMP3* was studied by modified MS-HRM in small malignant breast tumor (n=48), fibroadenoma (n=24) and healthy tissue from patients with fibroadenoma (n=24). No statistically significant differences in the methylation level of the studied genes were found between the samples of fibroadenoma and of healthy tissue. Therefore, samples of fibroadenoma and healthy tissue were included in the control group. Significant differences between the study group and the control group were found in 8 genes: *CCDC181, GSM2, ITPRIPL1, ZNF177, CACNA1E, DLGAP2, TIMP3* (all p<0.001), and *PRDM14* (p=0.002) (Figure 1).

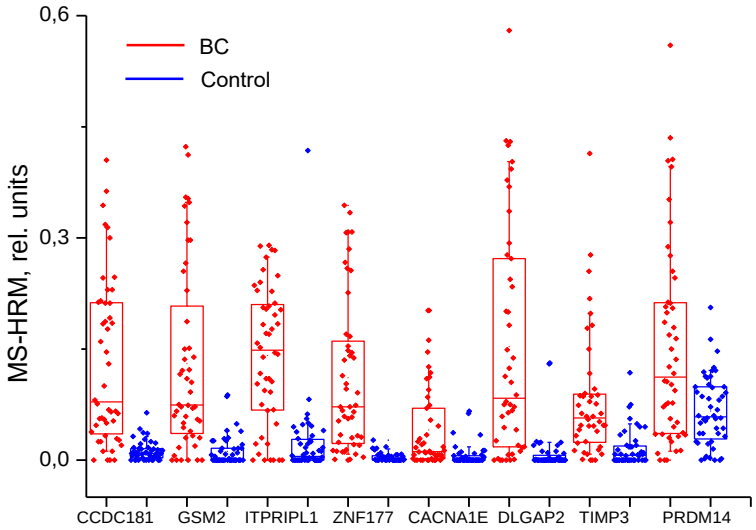

Figure 1. The levels of methylation of the studied genes in breast cancer compared to the control samples (healthy breast tissue, fibroadenoma).

CCDC181 and *ZNF177* genes showed the highest diagnostic value, according to ROC analysis (Table 1).

Table 1. Diagnostic value of genes according to ROC analysis.

Gene	AUC (95% CI)
CCDC181	0.89 (0.82-0.96)
GSM2	0.87 (0.8-0.94)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2022 doi:10.20944/preprints202211.0513.v1

ITPRIPL1	0.85(0.78-0.93)
ZNF177	0.93(0.88-0.99)
CACNA1E	0.76(0.67-0.86)
DLGAP2	0.88(0.81-0.95)
TIMP3	0.88(0.81-0.95)
PRDM14	0.68(0.57-0.78)

Pyrosequencing of *CCDC181* and *ZNF177* genes was used to confirm the results obtained by MS-HRM. The amplicons of these genes were shortened by NEST-PCR to a length suitable for pyrosequencing methylation analysis. The results of comparison between the average methylation levels obtained by pyrosequencing (%) and MS-HRM (in relative units) for *CCDC181* and *ZNF177* genes are shown in Figure 2 and Figure 3:

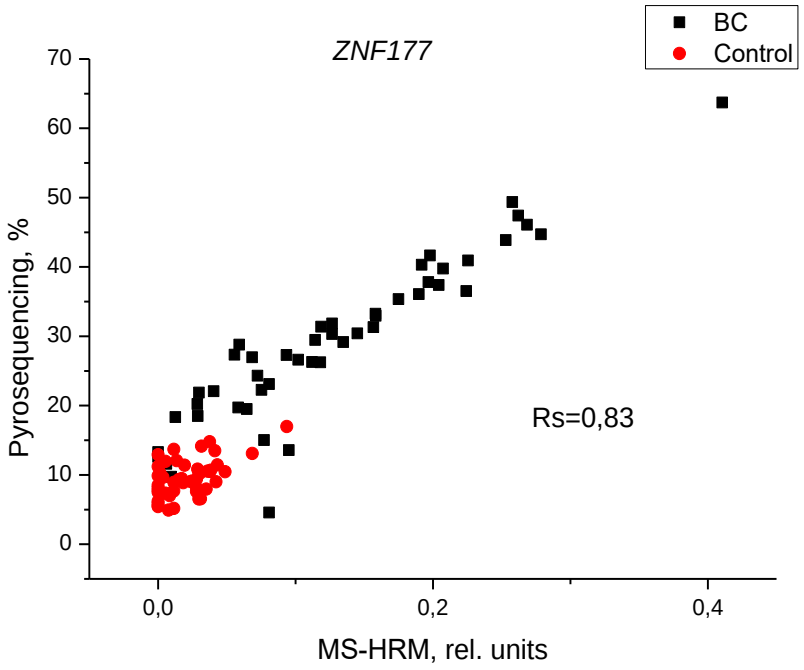

Figure 2. Comparison between the average methylation level of *ZNF177* gene obtained by pyrosequencing (%) and methylation level obtained by MS-HRM (in relative units).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2022 doi:10.20944/preprints202211.0513.v1

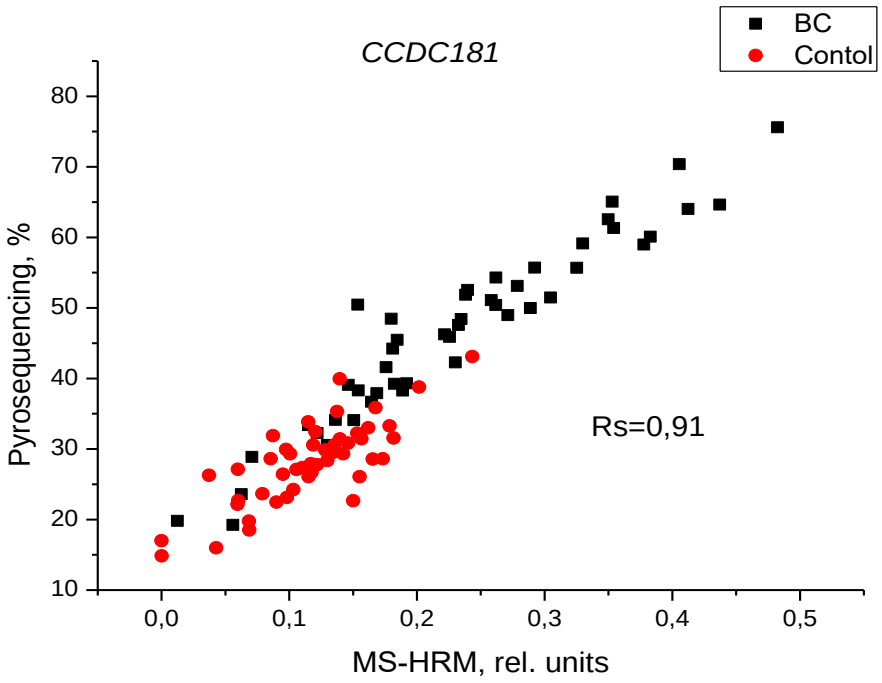

Figure 3. Comparison between the average methylation level of *CCDC181* gene obtained by pyrosequencing (%) and the methylation level obtained by MS-HRM (in relative units).

The Spearman's rank correlation coefficient (r_s) for *CCDC181* gene was 0.91 and for *ZNF177* gene – 0.83, $p<0.001$. These results confirm the reliability of MS-HRM to assess DNA methylation level. The diagnostic value (according to ROC analysis) of determination of methylation by pyrosequencing and MS-HRM of the DNA fragment obtained using NEST PCR for *CCDC181* and *ZNF177* genes is shown in Table 2.

Table 2. Comparison of diagnostic value according to ROC analysis of the studied genes methylation by pyrosequencing and MS-HRM.

Gene	AUC (95% CI)	
	Pyrosequencing	MS-HRM
CCDC181	0.9 (0.84-0.97)	0.84 (0.76-0.92)
ZNF177	0.98 (0.95-1)	0.89 (0.83-0.96)

Logistic regression showed that compound analysis of *CCDC181* and *ZNF177* gene methylation resulted in AUC=0.99 (95% CI 0.98-1) by pyrosequencing and AUC=0.96 (95% CI 0.92-1) by MS-HRM for the same fragments, and AUC=0.97 (95% CI 0.93-1) by MS-HRM for long fragments.

3. Discussion

Breast cancer, like other types of cancer, is associated with epigenetic changes. Assessment of epigenetic changes, such as DNA methylation, can provide important diagnostic and prognostic information. However, biomarkers of DNA methylation are not applied in routine medical practice. This can be explained by the absence of simple, generally accepted and clinically applicable methods for assessment of methylation that can be used in clinical practice. In the same time, an optimal combination of genes, which methylation could provide important clinical data, has not been defined yet. The results of our study of methylation in small breast cancer tumors can be useful for cancer

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2022 doi:10.20944/preprints202211.0513.v1

identification in small samples obtained by needle biopsy, where cancer is sometimes missed [8]. An important aspect of the study of methylation in small tumors is selection of potential markers for effective early diagnosis of breast cancer by detection of circulating cancer DNA in plasma [9]. A lot of data has been collected on the DNA methylation profile in breast cancer by means of chips such as the Infinium Methylation EPIC BeadChip (Illumina, USA), and through the whole genome methyl-sensitive sequencing. A number of authors have proposed the breast cancer methylation profiles, which they consider to be effective for tumor diagnosis. In this study, we showed that our simple and non-expensive modified MS-HRM method can be used for preliminary assessment of the results obtained by whole genome methylation analysis of breast cancer. A possibility of long fragments assessment increases the MS-HRM diagnostic value. The results obtained by MS-HRM were confirmed by subsequent pyrosequencing. However, pyrosequencing provides higher diagnostic accuracy than MS-HRM in at least two genes studied in our research. Selection of patients for the control group is an important aspect. In many studies healthy tissue from patients with breast cancer were used as a control. However, a possibility of migration of cancer cells to the adjacent healthy tissue should not be ignored. Our control group comprised the patients with fibroadenoma; this allowed us to compare methylation in both healthy tissue and a benign tumours. We found that a combination of two genes, *CCDC181* and *ZNF177*, showed the highest diagnostic value for small malignant breast tumors. Together, both genes, as the most methylated in breast cancer tumors, are noted in the publication by Wang et al. among other eight genes [5]. There are also publications where the methylation of these genes in breast cancer is mentioned separately [3, 10-12].

4. Materials and Methods

72 patients were included in the study: 48 patients with small breast cancer tumors and 24 patients with fibroadenoma. Breast cancer tissues were obtained from the patients of the study group, and fibroadenoma tissue and healthy breast tissue were obtained from the patients of the control group. All samples were stabilized in RNAlater and stored at -80°C. DNA was isolated from the samples using QIAamp DNA Mini Kit spin columns (Qiagen, USA). DNA was modified by bisulfite conversion using EpiJET Bisulfite Conversion Kit (Thermo Scientific, USA). To assess the relative level of methylation, we used our modified MS-HRM (Methylation Sensitive High Resolution Melting) method, which has a number of benefits compared to the one proposed by Wojdacz et al. (2008) [13]. Modification of MS-HRM was aimed to elaboration of a protocol for assessment of methylation of long fragments of CpG islands (500-600 nucleotides) containing a large number of CpG sites. The fragment lengths and the number of CpG sites are presented in Table 3.

Table 3. Primers for MS-HRM.

Gene	Annealing temperature	Amplicon length	Number of CpG sites	Forward primer	Revers primer
CCDC181	60	432	26	GGAGTGAGGTGTTTTGGGGTTTA	CTAATATATAAATTTCCTTCATCTTAT
GCM2	60	528	36	GTTTTTGGGATTGTGGTGGAG	AATACCATTCTCCCTCCTTC
ITRIPL1	63.5	210	13	GGATAAGTATGGTTTATAATTTAGG	CCAAAAACCCATCCTACCTCGAAATATCTCC

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2022 doi:10.20944/preprints202211.0513.v1

POM121L2	58	828	67	GGAAATTTTTAATTAGTTGTT	CCTCCTCACAAATCTATACC
KCNQ1	60	494	51	GTAGGAGTAAGTAGGGGAGATGTAGA	AACAACCACTACTACCAAC
MAST1	63	605	75	TTTATGGGGGTATTAGGAGGT	ACCCCAACCCCATCCCCCTA
PRDM14	58	654	41	TTATTTAGTTAAGAGGAAGTAG	ACCTTCTAAAACAAACAATATTAC
DNM2	61	401	47	GGTTAGAGTTGTTATTTGGATTTGT	ACCAACCAAATAACAAACTTCACC
ZNF177	63	345	23	GGGTAGTTTATTTTTTTAGTTGTTGGT	CACATAAACCCACTTACCTCCTC
AP2M1	59	691	63	GTGTATGTTGGAATGTGTAT	ACACAAAAAAATATCACTATCCTAC
CPEB4	57	226	23	GGTTTTGAGGAGAAGGATTTAGT	AACTTTTATTTCTCCTCA
DLGAP2	57	199	34	GGAGGTTTTGTTTTTAGTATTTAAG	AACAAAATCAACCTTTCTAAAC
CACNA1E	62	380	29	TAGAGTTGGAGTTTAGGAAGGGGTTAT	CATACATACCCACCACCC
SSH1	62	341	40	AGTTTTTTAAAGTGTTGGGATTATAGG	CACTACATACACAATCCCTACAA
TIMP3	60	165	15	GGTTAGAGATATTTAGTGGTTTAGGTGG	TTCAAATCCTTATAAAAAATAATACC

A modified MS-HRM protocol:

1. Amplification of fragments of CpG islands of the studied genes was carried out according to the following protocol: 95° - 5 min; (95°C - 15 s, 60°C - 30 s, 72° - 45 s)x30; (95°C - 15 s, 50°C - 30 s, 72°C - 45 s)x25. Reducing the annealing temperature from 60°C to 50°C after the 30th cycle gives an increased product yield compared to single-step amplification. Primers are shown in Table 3. PCR was performed using the **CFX96** system (Bio-Rad, Hercules, CA, USA);

2. Intercalating dye EVAGreen (Syntol, Moscow, Russia) was added to the derived product. Supplementation of EVAGreen on step 1 can lead to PCR inhibition;

3. The melting curve was generated according to the following protocol: 95° - 30 s; 60° - 10 min; melting analysis was performed in the temperature interval from 65° to 90° with increments of 0.2°;

4. After the construction of melting curve, the quality of the PCR product was checked by electrophoresis in 1.5% agarose gel.

Methylation levels were assessed by melting curves using Precision Melt Analysis Software, version 3 (Bio-Rad, Hercules, CA, USA). We did not use standards of different methylation levels when assessing HRM results, since the standards proposed by Wojdacz et al. [13] provide accurate results of methylation levels only if one CpG site is present in the PCR product. To compare the data, the curves were normalized according to the sample with the lowest level of methylation. Methylation was quantified with relative fluorescence units (RFU) at the temperature of the maximum peaks of the HRM curves. The temperature of the maximum peaks was manually determined for each

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2022 doi:10.20944/preprints202211.0513.v1

gene by the researchers based on the appearance of the melting curves (Figure 4). To confirm the validity of the results of modified MS-HRM, the results of pyrosequencing and MS-HRM were compared for two genes which had shown the best diagnostic significance – *CCDC181* and *ZNF177*.

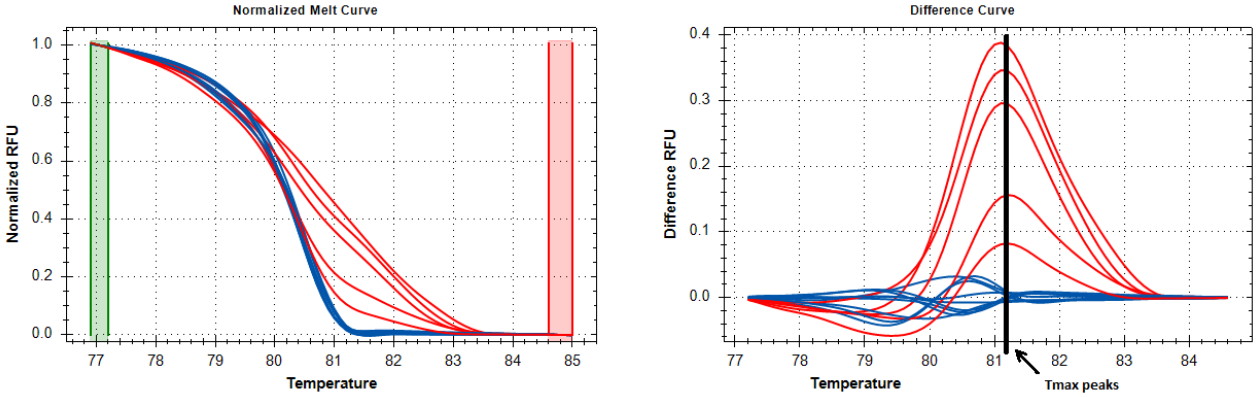

Figure 4. Determination the temperature of the maximum peaks, at which the RFU values were compared.

A comparative analysis was performed according to the following protocol:

1. 0.3 µl of the PCR product of the studied gene was taken;
2. NEST-PCR was performed with the primers shown in Table 4. The amplification program was as follows: 95°C for 5 min: (95°C for 15 s., 50°C for 30 s., 72°C for 45 s.)x12. NEST-PCR was required to shorten the length of the product to a size suitable for pyrosequencing;
3. A melting curve of the product obtained in step 2 was generated, and MS-HRM results were processed according to the protocol mentioned above;
4. Sequencing of the product obtained in step 2 was performed on a PyroMark Q48 (Qiagen, Hilden, Germany) in accordance with manufacturer's instructions;
5. The results of MS-HRM and pyrosequencing were compared using Spearman's rank correlation coefficient.

Table 4. Primers for pyrosequencing.

Gene	Annealing temperature	Amplicon length	Number of CpG sites	Forward primer	Revers primer
CCDC181	50	107	6	TTTTGGGGTTTAATTTGTG	ACCTACTTCCAATCTTCAAC-biotin
ZNF177	50	100	7	GGGTAGTTTATTTTTTTAGTTGTTGGT	AACAACCCTTTCTCAACTACA-biotin

Statistical analysis and graphs construction were performed using SPSS 17 (IBM, Armonk, NY, USA) and OriginPro 8.5 (OriginLab Corporation, Northampton, MA, USA software. The Mann–Whitney U test was used to determine the significance of differences for continuous values. Correlation was assessed using Spearman's rank correlation coefficient (r_s). Data are presented as median (upper quartile; lower quartile). Qualitative characteristics are presented as percentages. We applied logistic regression and receiver operating characteristic (ROC) curve analysis separately to assess diagnostic value of the studied parameters. Differences were considered statistically significant at $p<0.05$.

5. Conclusions

The analysis of gene methylation performed by a modified MS-HRM of breast cancer determined an optimal combination of *CCDC181* and *ZNF177* genes, allowing to accurately distinguish breast cancer from fibroadenoma and healthy tissue.

Author Contributions: Conceptualization, methodology, formal analysis, visualization and writing—original draft preparation, Aleksey M. Krasnyi; formal analysis, investigation, methodology, visualization and writing—review and editing, Alsu A. Sadekova; data curation, formal analysis and resources, Vlada V. Kometova; project administration, Valeriy V. Rodionov; supervision, Gennadiy.T. Sukhikh.

All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by research project «Study of gene methylation in blood plasma for early diagnosis of breast cancer » № 121032500098-3.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of National Medical Research Center for Obstetrics, Gynecology and Perinatology of Ministry of Healthcare of Russian Federation (Protocol № 1, 04.02.202)

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, upon request.

Conflicts of Interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71(3)*, 209-249. doi: 10.3322/caac.21660
2. Esteller M. Epigenetics in cancer. N *Engl J Med*. **2008**, *358(11)*, 1148-59. doi: 10.1056/NEJMra072067
3. Mao, X-H; Ye, Q.; Zhang, G-B.; Jiang, J-Y.; Zhao, H-Y.,; Shao, Y-F.; Ye, Z-Q.; Xuan, Z-X.; Huang, P. Identification of differentially methylated genes as diagnostic and prognostic biomarkers of breast cancer. *World J. Surg. Oncol*. **2021,** *19(1)*, 29. doi: 10.1186/s12957-021-02124-6
4. Panagopoulou, M.; Karaglani, M.; Manolopoulos, V.G.; Iliopoulos, I.; Tsamardinos, I.; Chatzaki, E. Deciphering the Methylation Landscape in Breast Cancer: Diagnostic and Prognostic Biosignatures through Automated Machine Learning. *Cancers (Basel).* **2021,** *13(7)*, 1677. doi: 10.3390/cancers13071677
5. Wang, S.C.; Liao, L-M.; Ansar, M.; Lin, S-Y.; Hsu, W-W.; Su, C-M.; Chung, Y-M.; Liu, C-C.; Hung, C-S.; Lin, R-K. Automatic Detection of the Circulating Cell-Free Methylated DNA Pattern of GCM2, ITPRIPL1 and CCDC181 for Detection of Early Breast Cancer and Surgical Treatment Response. *Cancers (Basel).* **2021,** *13(6)*, 1375. doi: 10.3390/cancers13061375
6. Luo, C.; Huang, J.; Guo, Z.; Guo, J.; Zeng, X.; Li, Y,; Liu, M. Methylated biomarkers for breast cancer identified through public database analysis and plasma target capture sequencing. *Ann. Transl. Med.* **2021**, *9(8)*, 683. doi: 10.21037/atm-21-1128
7. Ennour-Idrissi, K.; Dragic, D.; Issa, E.; Michaud, A.; Chang, S-L.; Provencher, L.; Durocher, F.; Diorio, C. DNA Methylation and Breast Cancer Risk: An Epigenome-Wide Study of Normal Breast Tissue and Blood. *Cancers (Basel).* **2020**, *12(11)*, 3088. doi: 10.3390/cancers12113088
8. Masood, S.; El-Gabry, E.; Zhang, C.; Wang, Z. The potential of identification of a malignancy-associated biomarker in breast cancer diagnosis and research: hTERT gene DNA methylation. *Diagn. Cytopathol.* **2016**, *44(8)*, 670-675. doi: 10.1002/dc.23505
9. Huiyan, L.; Wei, W.; Ye, Z.; Zheng, J.; Xu, R-H. Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. *Trends Mol. Med*. **2021**, *27(5)*, 482-500. doi: 10.1016/j.molmed.2020.12.01
10. de Almeida, B.P.; Apolónio, J.D.; Binnie, A.; Castelo-Branco, P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. *BMC Cancer*. **2019**, *19(1)*, 219. doi: 10.1186/s12885-019-5403-0
11. Ishihara, H.; Yamashita, S.; Fujii, S.; Tanabe, K.; Mukai, H.; Ushijima, T. DNA methylation marker to estimate the breast cancer cell fraction in DNA samples. *Med. Oncol.* **2018**, *35(11)*, 117. doi: 10.1007/s12032-018-1207-3
12. Karaglani, M.; Panagopoulou, M.; Baltsavia, I.; Apalaki, P.; Theodosiou, T.; Iliopoulos, I.; Tsamardinos, I.; Chatzaki, E. Tissue-Specific Methylation Biosignatures for Monitoring Diseases: An In Silico Approach. *Int. J. Mol. Sci*. **2022**, *23(6)*, 2959. doi: 10.3390/ijms23062959
13. Wojdacz, T.K.; Dobrovic, A.; Hansen, L.L. Methylation-sensitive high-resolution melting. *Nat. Protoc*. **2008**, *3(12)*, 1903-8. doi: 10.1038/nprot.2008.191