Investigation of Global Methylation in Peripheral Blood from Breast Cancer Patients

Xue Cao1,2, Tim Holland-Letz2, Qiuqiong Tang1,2, Xiaomeng Li1,2, Melanie Günert1,2, Katarina Cuk1,2, Sarah Schott4, Jörg Heil4, Michael Golatta4, Christof Sohn1, Andreas Schneeweiss1,2 and Barbara Burwinkel1,2*

1Department of Gynecology and Obstetrics, Molecular Biology of Breast Cancer, University of Heidelberg, Heidelberg, Germany
2Division of Molecular Epidemiology (C080), German Cancer Research Center (DKFZ), Heidelberg, Germany
3Department of Biostatistics (C080), German Cancer Research Center (DKFZ), Heidelberg, Germany
4Department of Gynecology and Obstetrics, University Women’s Clinic, Heidelberg, Germany
5National Centre for Tumor Diseases, Heidelberg, Germany

Abstract

Breast cancer is the most frequent malignancy with high mortality among women around the world. Global DNA methylation has been investigated by multiple studies and suggested as a screening biomarker for cancer. DNA methylation for two repetitive elements, LINE1 and Alu, was investigated in whole blood DNA from 229 breast cancer patients and 151 controls by using MassARRAY Epityper assay. Results showed that the mean methylation level of investigated CpG sites of LINE1 in peripheral blood from breast cancer patients was lower than that in controls (P=8.78E-06), especially for one specific CpG site (LINE1_CpG_1 with P=3.64E-10). ROC curve analysis of LINE1_CpG_1 methylation and LINE1 mean methylation was used to estimate the potential clinical utility of LINE1 methylation, area under the curve (AUC) was 0.73 (95% confidence interval (CI): 0.68-0.79) and 0.68 (95% CI: 0.62-0.74), respectively. In addition, the highest increased risk was observed in the lowest quartile of LINE1_CpG_1 methylation and LINE1 mean methylation (odds ratio (OR)=38.47 and 5.94; 95% CI: 8.77-168.64 and 2.94-11.98; P for trend=1.42E-07 and 1.33E-05 respectively). For Alu, significant hypomethylation of Alu_CpG_13 and Alu_CpG_14 in peripheral blood of breast cancer cases compared to controls was observed (P=0.002 and P=0.006). For the combined analysis, the AUC of the 10 most important CpG sites of LINE1 and Alu is 0.77 (95% CI: 0.72-0.83). Our study indicates that hypomethylation of specific CpG sites in LINE1 and Alu elements in peripheral blood DNA could be potential biomarkers for breast cancer risk. Further multicenter prospective studies are needed to verify these results.

Keywords: Breast cancer; DNA methylation; LINE1; Alu

Abbreviations: BC: Breast Cancer; ROC: Receiver Operating Characteristic; OR: Odds Ratio; AUC: Area Under the Curve; IQR: Interquartile Range; CI: Confidence Interval; 450K: Infinium Human Methylation450 Bead Chip array; 5-mCdC: 5-Methyldeoxycytosine

Introduction

Breast cancer (BC) is one of the most common cancers and the leading cause of cancer death among women in the worldwide. Although screening mammography is critical for the declined mortality of breast cancer, the limitations of mammography are well recognized, especially for young women with a high mammographic density of breasts [1,2]. Therefore, other approaches are urgently needed for breast cancer early detection.

DNA methylation is a type of epigenetic alterations which plays an important role in cancer development [3]. Promoter hypermethylation of tumor suppressor genes and global hypomethylation leading to malignancy have been studied extensively in different cancer types [4]. Global DNA hypomethylation is a hallmark of most cancers [5-9], including breast cancer [10-14]. This DNA hypomethylation has been proposed to activate oncogenes [15,16], induce genomic instability [17] and promote chromosome instability [7,18-21].

Genome-wide DNA hypomethylation originates from the decrease of 5-methyldeoxycytosine (5-mCdC) in dinucleotide CpG sites throughout the genome. As most 5-mCdC sites are rich in repetitive sequences that account for approximately half of the human genome and these repetitive DNA sequences are highly methylated in normal cells [7,22,23]. There are several different categories of repetitive sequences dispersed throughout the genome, such as long interspersed nuclear elements, short interspersed nuclear elements and satellite repeats. LINE1, a long interspersed nuclear element, is scattered throughout about 17% of the entire genome [24,25]. Alu is a short interspersed repetitive sequence that contributes almost 11% of the human genome [26]. Some studies found that the loss of DNA methylation for elements LINE1, Alu and Sat2 has been observed in cancer cells [22,27].

The DNA methylation of repetitive elements has been associated with global DNA methylation used as a biomarker for global methylation status by some investigators [28]. Furthermore, global hypomethylation in peripheral blood DNA has been considered as a risk factor for many tumors, such as colorectal, bladder and head and neck cancer [29-32].

In this study, we examined the methylation status of LINE1 and Alu in whole blood DNA of breast cancer cases and controls by using MassARRAY Epityper assay in order to find a potential blood-based biomarker for BC detection.
Materials and Methods

Study population

Our study was approved by the Ethics Committee of the University Hospital in Heidelberg. All samples of BC cases and healthy controls were Caucasian and obtained from the same region in southwest Germany. All enrolled patients and controls have given informed consent. Blood samples of sporadic BC patients are from the Genome study in our research group and were collected at the time-point of diagnosis before any treatment at University Hospital of Heidelberg. Clinical parameters of BC patients were confirmed according to the American Joint Committee on Cancer staging manual [33]. Detailed characteristics of BC cases are shown in Supplementary Table S1. Additionally, control blood samples come from our biomarker study, which were obtained from blood donors by the university hospital of Heidelberg. All the controls were healthy when donating blood and no one had a family history of BC. All blood samples were collected during the period from 2011 to 2014. A total of 229 sporadic BC cancer patients and 151 healthy controls were randomly selected for this study (Table 1).

DNA isolation and bisulfite conversion

DNA from whole blood samples (200 µl) was extracted by using QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s recommendations. DNA quality and quantity were measured by the Nanodrop ND-1000 UV/Vis-Spectrophotometer 3.3 (peqLab, Erlangen, Germany). Aliquots of DNA (500 ng) were bisulfite-treated with the EZ-96 DNA methylation Gold kit (Zymo Research Corporation, Orange, USA) according as the description of the manufacturer.

Primer design and PCR amplification

The online tool, “epidesigner” (http://www.epidesigner.com/start3.html), was used for the primer design. The PCR primers for LINE1 and Alu and their amplicon sequences were shown in Supplementary Tables S2. PCR reaction was performed in a total volume of 6 µl. PCR reaction components included 10 ng/ml bisulfite-treated DNA, 10 x CoralLoad Buffer(Qiagen), 10 mM dNTPs, 1 µM of each (forward and reverse) primer (Sigma), and 5 U HotStar Taq DNA polymerase (Qiagen, Valencia CA). The touch-down PCR profile was 95°C for 5 minutes, denaturation at 94°C for 30 seconds, primer annealing from 59°C to 53°C for 30 seconds, a final extension at 72°C for 1 minute, then 72°C for 5 minutes, 4°C for infinite. 1% agarose gel was used for electrophoresis to inspect PCR products and visualized under ultraviolet light.

Methylation analysis

The Sequenom MassARRAY Epityper assay was used for methylation analysis as described previously [34]. The PCR products were dealt with succeeding procedures according as the protocol of Sequenom MassARRAY Epityper Assay and cleaned by Resin. Then 8-15 nl of cleavage reaction was transferred to a 384 SpectroCHIP by using the Nanodispenser (SEQUENOM). The chip was analyzed with the MassARRAY (SEQUENOM). The mass spectra were collected from the MassARRAY (SEQUENOM). The mass spectra were collected from using the Nanodispenser (SEQUENOM). The chip was analyzed with

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sample Types</th>
<th>Group</th>
<th>Number</th>
<th>Age (y)</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINE1</td>
<td>Peripheral blood DNA</td>
<td>Sporadic BC cases</td>
<td>229</td>
<td>48.31 ± 8.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls</td>
<td>151</td>
<td>41.47 ± 10.91</td>
<td></td>
</tr>
<tr>
<td>Alu</td>
<td>Peripheral blood DNA</td>
<td>Sporadic BC cases</td>
<td>229</td>
<td>48.31 ± 8.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls</td>
<td>151</td>
<td>41.47 ± 10.91</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Sample information.

Results

Hypomethylation of repetitive elements in BC patients

LINE1: DNA methylation levels in two repetitive elements, LINE1 and Alu, were compared in 229 BC patients and 151 healthy controls. 11 CpG sites were measured in the amplicon of LINE1 (Table 2). We observed significant hypomethylation of seven CpG sites, LINE1_CpG_1, LINE1_CpG_3,4,5, LINE1_CpG_9, LINE1_CpG_12, LINE1_CpG_14, in peripheral blood of BC patients. Their median methylation levels, interquartile ranges (IQRs) and the P values were shown in Table 2. The mean methylation level of all investigated CpG sites of LINE1 was also significantly lower in peripheral blood of BC cases than that in healthy controls (median of mean methylation level in BC cases=0.85 (IQR=0.84-0.86) compared to controls 0.86 (IQR=0.85-0.88), with adjusted P=8.78E-06).

In Figure 1, the methylation levels of the seven significant CpG sites of LINE1 were plotted for the BC cases and controls. Although all of these CpG sites revealed significant results, the methylation levels of cases and controls were quite similar. The biggest difference was shown in the median methylation levels between BC cases and controls for
LINE1_CpG_1 (median methylation level in BC cases=0.76 (IQR=0.74-0.79) compared to controls 0.80 (IQR=0.76-0.82), with adjusted P=3.64E-10) and LINE1_CpG_12 (median methylation level in BC cases=0.71 (IQR=0.69-0.74) compared to controls 0.74 (IQR=0.71-0.77), with adjusted P=4.09E-07). The lowest P value was observed for LINE1_CpG_1 with adjusted P=3.64E-10 (Table 2). As especially LINE1_CpG_1 showed a strong association with BC, we also calculated the diagnostic AUC for this single CpG site through a ROC analysis.

ROC curve analysis of LINE1_CpG_1 methylation and LINE1 mean methylation (Figure 2) was used to estimate the potential clinical utility of LINE1 methylation, AUC was 0.73 (95% CI: 0.68-0.79) and 0.68 (95% CI: 0.62-0.74) respectively.

Moreover, in the quartile analysis we observed that patients in the lowest methylation quartile of LINE1_CpG_1 have the highest OR of 38.3 (95% CI: 8.77-168.64) compared to the highest quartile. P for trend was 1.42E-07 for LINE1_CpG_1 methylation in Table 3.

Similarly, as shown in Table 4, an increased risk was found in the lower quartiles compared to the highest quartile of LINE1 mean methylation (OR=5.94, 4.44 and 3.90; 95% CI=2.94-11.98, 2.13-9.28 and 1.84-8.24 respectively; P for trend=1.33E-05).

**Alu**

For Alu, we investigated 11 CpG sites in the amplicon and found a significant difference in methylation of Alu_CpG_13 and Alu_CpG_14 in peripheral blood between cases and controls (Table 5) with adjusted P=0.002 and 0.006 respectively (median of methylation level for Alu_CpG_13=0.63 in cases (IQR=0.62-0.65) compared to controls 0.65 (IQR=0.62-0.67) and median of methylation level for Alu_CpG_14=0.50 (IQR=0.48-0.55) in cases compared to controls 0.54 (IQR=0.49-0.56)). All other investigated CpG sites did not show significant differences between BC cases and controls. The box plot and ROC curve of Alu_CpG_13 and Alu_CpG_14 were shown in Supplementary Figures 1 and 2 respectively with AUC=0.67 (95% CI: 0.61-0.74) and 0.67 (95% CI: 0.61-0.73).

In addition, as shown in Table 6, we found an increased risk in the first two lower quartiles compared with the highest quartile of the
methylation in \(Alu\) \_CpG\_13 (OR=2.50 and 2.10; 95% CI: 1.25-5.02 and 1.05-4.19 respectively; \(P\) for trend=0.002). Contrarily, as shown in Supplementary Tables S3 and S4, no increased risk was found between the lower quartiles and the highest quartile of \(Alu\) \_CpG\_14 methylation with \(P\) for trend=0.005.

**Combination**

Finally, we did a multiple logistic regression analysis including all CpG sites from both repetitive elements plus age (Supplementary Table S5). For this calculation, \(Alu\) \_CpG\_5 and \(Alu\) \_CpG\_15,16,17 were excluded because of the large number of missing values. Both of these CpGs were not noticeable in the univariate analysis. Again, \(LINE1\) \_CpG\_1 was dominant, but different other CpGs contributed additional information. Furthermore, the effect of \(Alu\) \_CpG\_13 was no longer significant. Performing a ROC analysis on this model, an AUC of 0.78 (95% CI: 0.72-0.83) was observed (Supplementary Figure 3). The 0.6324+ corrected AUC was 0.74.

If only the two most promising univariate candidates \(LINE1\) \_CpG\_1 and \(Alu\) \_CpG\_13 plus age were included (Supplementary Table S6), \(Alu\) \_CpG\_13 still lost significance and an AUC of 0.73 (95% CI: 0.68-0.79, corrected AUC=0.73) was observed (Supplementary Figure 4).

We also calculated a model with the 10 most important CpG sites plus age (Table 7). The AUC was 0.77 (95% CI=0.72-0.83, Cross-Validation AUC=0.75) (Figure 3).

**Correlation of \(LINE1\) and \(Alu\) methylation with clinical characteristics of BC patients**

We evaluated the association of the methylation levels of these two repetitive elements with clinical features of breast cancer patients. Overall, there were no significant associations between the methylation levels of most CpG sites in \(LINE1\) and \(Alu\) and the different clinical characteristics. Further studies are needed to discover the association between blood DNA methylation and clinicopathological parameters (Supplementary Tables S7 and S8).

**Comparison of the results from this study with the results of Infinium Human Methylation450 BeadChip array and with literature**

As the methylation level of repetitive elements is thought to reflect the average methylation level of genomic DNA, we compared the mean methylation of \(LINE1\) and \(Alu\) in Sequenom MassARRAY EpiTyper data with the mean of all the CpG sites in 450 K methylation array. In the epigenome-wide 450 K methylation data, we observed the mean methylation level was lower in peripheral blood of patients compared to controls but the \(P\) value was not statistically significant (0.523 ± 0.28 and 0.524 ± 0.24 for cases and controls, respectively, \(P\)=0.09 (Table 8)). In our here presented study, the mean methylation level of \(LINE1\) was significantly lower in peripheral blood DNA of BC cases than that in controls (median of mean methylation level in BC cases=0.85 (IQR=0.84-0.86) compared to controls 0.86 (IQR=0.85-0.88), with \(P\)=2.14E-06). However, the mean methylation of \(Alu\) did not show a significant difference between BC patients and healthy controls (Table 8).

**Discussion**

In this study, we found statistically significant \(LINE1\) hypomethylation in the peripheral blood DNA of sporadic breast cancer patients compared with healthy controls, especially for \(LINE1\) \_CpG\_1. We identified \(LINE1\) \_CpG\_1 methylation to be strongly associated with breast cancer. For \(Alu\), we observed that one single CpG sites was significantly hypomethylated in the peripheral blood DNA from breast cancer cases compared to controls. Furthermore, our results showed that the decreased methylation level of \(LINE1\) \_CpG\_1 was associated with an increased breast cancer risk. Similar to the discovery studied by DeRoo et al, quartiles of \(LINE1\) methylation levels were associated with the risk of breast cancer, with an increased risk observed in the lowest quartile compared with those in the highest quartile [35].

However, other studies found that there were no significant differences in the methylation levels of \(LINE1\) and \(Alu\) in peripheral blood DNA between breast cancer cases and healthy controls, measured by different detection methods including pyrosequencing, methylLight and COBRA assays [36-41]. But some studies assessed \(LINE1\) in only a small sample size with the number of cases smaller than 50 [38, 40]. Partly in contrast to our observations, Wu et al. reported that there was no association between breast cancer risk and \(LINE1\) and \(Alu\) methylation [41]. Even though as mentioned in our previous review [42], \(LINE1\) as a surrogate for global DNA methylation level might be limited, it is difficult to make a comparison among studies because of the multiple experimental assays, various study designs, different sample sizes, distinct blood fractions and lack of an agreed standard for global methylation measurement. Although the results are inconclusive, studies have provided some clues for investigating global methylation in blood DNA. Our study contributes to a growing literature showing global hypomethylation in peripheral blood DNA of breast cancer cases compared with controls.

One possible concern is that methylation status may be cell type specific. However, one study did not find a significant difference in \(LINE1\) measurements between purified granulocyte and lymphocyte...
samples [43]. Generally, peripheral blood samples from epidemiological studies always comprise mixed cell populations, and it is not feasible or practical for fractionation of cell populations in an epidemiological study setting [37].

Conclusion and Limitations

Our study had some potential limitations. Many factors such as lifestyle, environmental exposures [44, 45] and even genetic susceptibility can affect methylation signatures. In one study, LINE1 methylation in leukocyte DNA was lower in persons who smoked blond tobacco than nonsmokers and was inversely associated with arsenic toenail concentration [43]. These factors were not investigated in our study and should be taken into account in future studies. In addition, we did not investigate the whole LINE1 and was inversely associated with arsenic toenail concentration [43]. In summary, this study indicates that hypomethylation of specific CpG sites in LINE1 and especially in 4. Mutirangura A (2007) Quantitative PCR analysis for methylation level of 5. Wilson AS, Power BE, Molloy PL (2007) DNA hypomethylation and human cancer. Nat Rev Genet 7: 21-33.

Acknowledgement

We thank Sabine Serick and Yanxiang Jiang for their technical support. This study was supported by the German Cancer Research Center (DKFZ) and University Women’s Clinic, Heidelberg. Xue Cao was financed by the China Scholarship Council (CSC).

References


Table 8: Comparison of LINE1 and Alu methylation levels between different analysis methods.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Literature#</th>
<th>450K results</th>
<th>Sequenom Mass ARRAY Epityper assay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Author, year</td>
<td>BC Cases No./ Controls No.</td>
<td>Meth (%) (cases)</td>
</tr>
<tr>
<td>Kikumunthom [40]</td>
<td>36/144</td>
<td>40</td>
<td>42</td>
</tr>
<tr>
<td>Xu [38]</td>
<td>1064/1100</td>
<td>78.8</td>
<td>78.8</td>
</tr>
<tr>
<td>Brennan</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BGS cohort</td>
<td>241/242</td>
<td>79.0 (78.1-79.9)</td>
<td>79.0 (77.9-80.1)</td>
</tr>
<tr>
<td>EPIC cohort</td>
<td>232/263</td>
<td>75.1 (73.9-76.3)</td>
<td>75.1 (73.9-76.3)</td>
</tr>
<tr>
<td>KConFab cohort</td>
<td>153/218</td>
<td>76.6 (75.2-77.6)</td>
<td>76.0 (74.3-78.0)</td>
</tr>
<tr>
<td>Wu [42]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MethyLight</td>
<td>265/333</td>
<td>107.4+0.50</td>
<td>108.5+0.50</td>
</tr>
<tr>
<td>pyro sequencing (mean)</td>
<td>279/340</td>
<td>74.5+0.50</td>
<td>74.5+0.50</td>
</tr>
<tr>
<td>Cho [41]</td>
<td>40/40</td>
<td>70</td>
<td>78</td>
</tr>
<tr>
<td>Choi [39]</td>
<td>19/18</td>
<td>74.7</td>
<td>73.9</td>
</tr>
<tr>
<td>Alu</td>
<td>229/151</td>
<td>0.51 (0.50-0.53)</td>
<td>0.52 (0.50-0.53)</td>
</tr>
</tbody>
</table>

Note: *P value was calculated by logistic regression and adjusted by age. Significant P values are in bold (P<0.05); # literature based on review by Tang et al. in 2016.


