Intraindividual Variation and Short-term Temporal Trend in DNA Methylation of Human Blood

Yurii B. Shvetsov1, Min-Ae Song2, Qiuyin Cai3,4, Maarit Tiirikainen1, Yong-Bing Xiang5, Xiao-Ou Shu3,4, and Herbert Yu1

Abstract

Background: Between- and within-person variation in DNA methylation levels are important parameters to be considered in epigenome-wide association studies. Temporal change is one source of within-person variation in DNA methylation that has been linked to aging and disease.

Methods: We analyzed CpG-site-specific intraindividual variation and short-term temporal trend in leukocyte DNA methylation among 24 healthy Chinese women, with blood samples drawn at study entry and after 9 months. Illumina HumanMethylation450 BeadChip was used to measure methylation. Intraclass correlation coefficients (ICC) and trend estimates were summarized by genomic location and probe type.

Results: The median ICC was 0.36 across nonsex chromosomes and 0.80 on the X chromosome. There was little difference in ICC profiles by genomic region and probe type. Among CpG loci with high variability between participants, more than 99% had ICC > 0.8. Statistically significant trend was observed in 10.9% CpG loci before adjustment for cell-type composition and in 3.4% loci after adjustment.

Conclusions: For CpG loci differentially methylated across subjects, methylation levels can be reliably assessed with one blood sample. More samples per subject are needed for low-variability and unmethylated loci. Temporal changes are largely driven by changes in cell-type composition of blood samples, but temporal trend unrelated to cell types is detected in a small percentage of CpG sites.

Impact: This study shows that one measurement can reliably assess methylation of differentially methylated CpG loci.

Introduction

Epigenetic modifications such as DNA methylation of cytosine residues at CpG dinucleotides have been extensively examined for their potential association with disease and aging in humans. DNA methylation marks are stable indicators of tissue lineage; thus, different tissues have fundamentally different methylation patterns. DNA methylation profile is established at neonatal stage and undergoes rapid change at the early age and further modifications throughout lifespan. Changes of methylation patterns have been seen in normal human aging process, and aberrant methylation changes in tissues such as blood have been linked to a number of diseases, including cancer (1–5). A link has also been established between age-related and cancer-related methylation changes. Important cancer-related genes become hypermethylated during aging, including key developmental genes and those encoding the estrogen receptor, insulin growth factor, and E-cadherin (1, 6). Furthermore, the acquisition of methylation changes during aging could underlie the development of age-related pathologic conditions (7).

Research into the role of DNA methylation in disease is affected by variability in measured methylation levels, which can be attributed to a number of sources: differences among study participants due to demographic, environmental and genetic factors; random and systematic variation between samples taken from the same individual, which may result from differences in tissue sample composition or timing of sample collection; technical variability due to measurement error and limited precision of analytic tools. The relative proportion of between- and within-person variability in the total variability is typically described by the intraclass correlation coefficient (ICC). The ICC varies between 0 and 1, with high (low) proportion of within-person variation represented by values close to 0 (respectively, 1). Low values of ICC signify lower data reliability and may affect validity of analysis results. A number of studies (8–11) have analyzed within-person variability of different blood biomarkers; however, we are unaware of any reports on that in DNA methylation, although the importance of this aspect for epigenome-wide association studies has been recognized (12).

In the present study of leukocyte DNA methylation among 24 healthy women, we examined CpG-site-specific intraindividual variation in DNA methylation and analyzed short-term temporal changes in DNA methylation using Illumina HumanMethylation450 BeadChip (13). This array has been reported to have low technical variability and high reproducibility (13), although a number of issues have been pointed out (14). The BeadChip contains 485,577 loci that include 482,421 (99.35%) CpG

1University of Hawaii Cancer Center, Honolulu, Hawaii. 2Comprehensive Cancer Center, Ohio State University and James Cancer Hospital, Columbus, Ohio. 3Division of Epidemiology and Vanderbilt Epidemiology Center, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee. 4Vanderbilt-Ingram Cancer Center, Nashville, Tennessee. 5Shanghai Cancer Institute, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Yurii B. Shvetsov, University of Hawaii Cancer Center, 701 Iwilo Street, Honolulu, HI 96813. Phone: 808-564-5825; Fax: 808-586-2982; E-mail: yshvetso@cc.hawaii.edu.

doi: 10.1158/1055-9965.EPI-14-0853

©2014 American Association for Cancer Research.
Materials and Methods

Study population and sample collection

Study participants were drawn from the Shanghai Physical Activity (SPA) study subcohort of the Shanghai Women’s Health Study (SWHS), a prospective cohort study of 74,943 women ages 40 to 70 years who were recruited from seven communities in Shanghai, China, between 1997 and 2000 (16). The SPA subcohort of SWHS comprises 300 women from two communities (17), who provided blood samples at study entry and after 9 months. Twenty-four women from the SPA subcohort were randomly selected for the present analysis.

For each consented participant, a 10-ml blood sample was drawn into a EDTA vacutainer tube and kept in a portable Styrofoam box with ice packs at 0°C to room temperature. The samples were processed and separated into 2-ml aliquots of plasma (4),uffy coat (2), and red blood cells (2). Samples were then stored at −80°C until DNA extraction.

Laboratory analysis

DNA extraction and bisulfite conversion. DNA was extracted fromuffy coats using a QIAamp DNA Blood Mini Kit (Qiagen Inc.). The quality and quantity of extracted DNA was examined byNanoDrop-2000 spectrometer (Thermo Scientific). Bisulfite conversion of 500 ng of DNA was performed on each sample according to the manufacturer’s recommendations for the HumanMethylation450 BeadChip using the EZ DNA Methylation Kit (Zymo Research). The treatment protocol included 16 cycles of denaturing at 95°C for 30 seconds, incubation at 50°C for 60 minutes, and holding at 4°C for a minimum of 2 hours.

Illumina methylation platform. Bisulfite-converted DNA (4 μl) was hybridized onto and analyzed using HumanMethylation450 BeadChip (Illumina). The hybridization protocol consisted of a whole-genome amplification step followed by enzymatic end-point fragmentation, precipitation, and resuspension. Resuspended samples were hybridized onto the BeadChip for 16 hours at 48°C. After hybridization, unhybridized and nonspecifically hybridized DNA was washed away, followed by single-nucleotide extension using the hybridized bisulfite-treated DNA as a template. The Illumina iScan SQ scanner was used to create images of the single arrays. Image intensities were extracted using GenomeStudio (v.2011.1) Methylation module (v.1.9.0) software.

Data quality assessment and preprocessing

Data normalization was performed in GenomeStudio using “Background Subtraction” and “Normalization to Internal Controls” methods and has been described elsewhere (18). Briefly, background subtraction values were derived from built-in negative control bead signals and subtracted from probe intensities. Normalization was performed using internal control probe pairs designed to target the same region within housekeeping genes. The methylation score for each CpG site was represented by a β value, calculated according to normalized probe fluorescence intensity ratios between methylated and unmethylated signals and varying between 0 (fully unmethylated) and 1 (fully methylated). Data quality control (QC) analyses performed on β values included principal component analysis, to assess potential batch effects, and sample histograms for signal distributions (data not shown). Probes with detection P value above 0.05 were excluded.

Intraclass Correlation Coefficients (ICCs) were then computed as ICC = σ^2 / (σ^2 + σ^2_w). To summarize patterns of variability by functional genomic location, by type of probe and for known differentially methylated regions (DMR), we computed the percentage of CpG loci with high (=0.8), midrange (0.5–0.8), and low (<0.5) ICCs for each of these groups. The estimated between- and within-person variance components were converted to the β-scale as follows:

\[ \hat{\sigma}_w = \log^{-1}(\hat{\mu} + \hat{\sigma}_w) - \log^{-1}(\hat{\mu} - \hat{\sigma}_w), \]

\[ \hat{\sigma}_w = \log^{-1}(\hat{\mu} + \hat{\sigma}_w) - \log^{-1}(\hat{\mu} - \hat{\sigma}_w), \]

where \( \hat{\mu} \) is the mean M value across all study subjects. For 14,248 CpG loci with very low mean methylation, ICCs could not be estimated and were set to 0, as the worst possible case. A short-term temporal trend in DNA methylation across all CpG loci was examined using a mixed model with time since study entry (years) as a fixed effect, and study participants as random effects:

\[ Y_i = \mu + t_i + S_i + e_i, \quad S_i \sim N(0, \sigma^2_w), \quad e_i \sim N(0, \sigma^2_u). \]

To assess the influence of quatile normalization on ICC and trend, all models were fit with untransformed and quantile normalized M values. All analyses were adjusted for cell-type composition using the cell mixture deconvolution method of Houseman and colleagues (20), which establishes DNA methylation signature panels for each cell type and uses constrained optimization to map methylation profiles of interest onto the signature panel and to predict proportions of cell types in a blood sample.

A number of recent reports define meaningful difference (14) in methylation as ΛB value (a measure of average difference in methylation level at a particular CpG site between study subjects) being above a certain threshold. A threshold of 0.2 on the β-scale is commonly used. To assess the impact of intrasubject variation on such high-variance CpG loci, we summarized ICC and trend by
Cancer Epidemiol Biomarkers Prev; 24(3) March 2015

Shvetsov et al.

Genomic location and probe type for high-variance ($\Delta p \geq 0.2$), mid-variance ($0.1 \leq \Delta p < 0.2$), and low-variance loci ($\Delta p < 0.1$). We estimated the $\Delta p$ value for every CpG site as twice the between-person SE: $\Delta p = 2\sigma_b$. All analyses were conducted using R 3.0.3 software and Bioconductor package limma 3.1-117.

### Results

The median age at study entry among the 24 study participants was 54.5 years (range, 46.5–68.8 years). Across the 483,880 CpG loci that passed data QC, the ICCs ranged from 0 to 0.999, with a median of 0.37. For the majority of the loci, the within-person variance component tended to be smaller than and uncorrelated with the between-person component (Fig. 1). The ranges of $\sigma_{b, p}$ and $\sigma_{p,b}$ were 0 to 0.42 and 0 to 0.36, respectively.

Table 1 lists proportions of low-, mid-, and high-ICC loci by genomic location, probe type, and locus variability for nonsex chromosomes. The median ICC across all CpG loci was 0.36 [interquartile range (IQR), 0.13–0.63]. More than 64% loci had low ICC, whereas 23% had mid-range ICC, and 13% had high ICC. Among the low-variability loci, the proportions of low- and mid-range ICC were comparable, whereas only 8.1% loci exhibited high ICC. On the other hand, more than 90% of moderate-variability loci ($0.1 \leq \Delta p < 0.2$) and more than 99% of high-variability loci ($\Delta p \geq 0.2$) had high ICC, whereas <1% loci in these two groups had low ICC. This latter result was observed across all genomic locations.

Among the CpG loci in known DMR (13), ICCs tended to be higher than across the entire genome, with median 0.49 (IQR, 0.24–0.72), and 51.5%, 31.8%, and 16.7% loci exhibiting low-, mid-range, and high ICCs, respectively. Across loci from CGIs, shores and shelves, the median ICC ranged from 0.34 to 0.43, with the proportion of high-ICC loci highest in CpG shores (overall and among low-variability sites) and CGIs (moderate-variability sites). The ICC profiles across different functional locations were similar, with median ICC between 0.29 and 0.36, and the percentage of high-ICC loci 7.8% to 12.4% overall, 5.2% to 8.2% among low-variability sites, and 92.0% to 96.2% among moderate-variability sites. There was also little difference in ICC profiles by probe type. CpG loci with methylation measured by Infinium I probes, compared with Infinium II probes, had lower proportion of high-ICC loci overall (11.3% vs. 13.5%), among low-variability sites (6.6% vs. 8.2%), and among moderate-variability sites (92.7% vs. 94.7%).

The distribution of ICCs on the X chromosome (Table 2) differed substantially from that on nonsex chromosomes, with the median ICC of 0.80 (IQR, 0.60–0.89) overall and nearly half of the X chromosome CpG loci having high ICC. The main difference from nonsex chromosomes in the distribution of ICCs was found among low-variability sites, where 27.8% to 51.1% loci had high ICC. Of all genomic locations, lineage-defining DMRs appeared the most stable, with 61.6% loci having high ICC. CGIs and shores and functional locations on the promoter side of gene coding regions contained more high-ICC loci than CGI shelves, gene body, and 3′UTR locations. The proportion of high-ICC loci measured by Infinium II probes was somewhat higher than that expected.

![Figure 1](image)

**Figure 1.** Scatter plot of within- versus between-person SE (β-scale), all CpG sites.

**Table 1.** Summary of ICCs (β-scale) of CpG locus methylation, by genomic location and probe-type, nonsex chromosomes

<table>
<thead>
<tr>
<th>Location/probe type</th>
<th>N</th>
<th>Median ICC</th>
<th>0.5–0.8</th>
<th>0.8–1</th>
<th>&lt;0.5</th>
<th>0.5–0.8</th>
<th>0.8–1</th>
<th>&lt;0.5</th>
<th>0.5–0.8</th>
<th>0.8–1</th>
<th>&lt;0.5</th>
<th>0.5–0.8</th>
<th>0.8–1</th>
</tr>
</thead>
<tbody>
<tr>
<td>All CpG loci</td>
<td>472,663</td>
<td>0.36</td>
<td>0.13</td>
<td>0.63</td>
<td>64.1</td>
<td>23.0</td>
<td>12.9</td>
<td>446,677</td>
<td>67.8</td>
<td>24.1</td>
<td>8.1</td>
<td>20,859</td>
<td>0.2</td>
</tr>
<tr>
<td>Known DMR</td>
<td>36,053</td>
<td>0.49</td>
<td>0.24</td>
<td>0.72</td>
<td>51.5</td>
<td>31.8</td>
<td>16.7</td>
<td>33,802</td>
<td>54.9</td>
<td>33.6</td>
<td>11.5</td>
<td>1,965</td>
<td>0.1</td>
</tr>
<tr>
<td>Location with respect to CpG island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG Island</td>
<td>101,663</td>
<td>0.39</td>
<td>0.15</td>
<td>0.66</td>
<td>61.1</td>
<td>24.4</td>
<td>14.5</td>
<td>95,364</td>
<td>65.1</td>
<td>25.7</td>
<td>9.2</td>
<td>5,110</td>
<td>0.2</td>
</tr>
<tr>
<td>Shoreb</td>
<td>108,839</td>
<td>0.45</td>
<td>0.18</td>
<td>0.71</td>
<td>57.2</td>
<td>25.4</td>
<td>17.4</td>
<td>101,091</td>
<td>61.6</td>
<td>26.9</td>
<td>11.5</td>
<td>5,623</td>
<td>0.2</td>
</tr>
<tr>
<td>Shelfb</td>
<td>45,790</td>
<td>0.49</td>
<td>0.24</td>
<td>0.72</td>
<td>51.5</td>
<td>31.8</td>
<td>16.7</td>
<td>33,802</td>
<td>54.9</td>
<td>33.6</td>
<td>11.5</td>
<td>1,965</td>
<td>0.1</td>
</tr>
</tbody>
</table>

*All estimates are adjusted for cell-type composition. Δ$\beta$ is estimated as twice the between-person SE (β-scale).

Shores, up to 2 kb from CpG islands; shelves, from 2 to 4 kb from CpG islands.

492 Cancer Epidemiol Biomarkers Prev; 24(3) March 2015

Cancer Epidemiology, Biomarkers & Prevention

---

Published OnlineFirst December 23, 2014; DOI: 10.1158/1055-9965.EPI-14-0853

Downloaded from cebp.aacrjournals.org on June 21, 2017. © 2015 American Association for Cancer Research.
Table 2. Summary of ICCs (β-scale) of CpG locus methylation, by genomic location and probe type. X chromosome

<table>
<thead>
<tr>
<th>Location/probe type</th>
<th>N</th>
<th>Median</th>
<th>IQR</th>
<th>ICC range</th>
<th>Percentage of loci by ICC range</th>
<th>All β &lt; 0.1</th>
<th>Percentage of loci by ICC range</th>
<th>All β &lt; 0.2</th>
<th>Percentage of loci by ICC range</th>
<th>All β &gt; 0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>All CpG loci</td>
<td>11,215</td>
<td>0.80</td>
<td>0.60</td>
<td>0.89</td>
<td>18.5</td>
<td>31.8</td>
<td>49.7</td>
<td>9,051</td>
<td>36.6</td>
<td>38.6</td>
</tr>
<tr>
<td>Enhancer</td>
<td>629</td>
<td>0.84</td>
<td>0.71</td>
<td>0.91</td>
<td>9.9</td>
<td>31.5</td>
<td>58.7</td>
<td>478</td>
<td>40.0</td>
<td>46.9</td>
</tr>
<tr>
<td>Known DMR</td>
<td>1,231</td>
<td>0.85</td>
<td>0.72</td>
<td>0.91</td>
<td>10.6</td>
<td>27.8</td>
<td>61.6</td>
<td>936</td>
<td>40.0</td>
<td>51.1</td>
</tr>
<tr>
<td>Location with respect to CpG island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG Island</td>
<td>4,272</td>
<td>0.82</td>
<td>0.67</td>
<td>0.90</td>
<td>13.2</td>
<td>32.0</td>
<td>54.8</td>
<td>3,391</td>
<td>16.6</td>
<td>38.9</td>
</tr>
<tr>
<td>Shoreb</td>
<td>2,837</td>
<td>0.82</td>
<td>0.63</td>
<td>0.90</td>
<td>13.5</td>
<td>31.4</td>
<td>55.1</td>
<td>1,866</td>
<td>22.5</td>
<td>39.3</td>
</tr>
<tr>
<td>Shelfb</td>
<td>1138</td>
<td>0.71</td>
<td>0.46</td>
<td>0.85</td>
<td>28.0</td>
<td>36.3</td>
<td>35.7</td>
<td>1,003</td>
<td>31.8</td>
<td>40.4</td>
</tr>
<tr>
<td>Functional location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSS5000</td>
<td>2,777</td>
<td>0.80</td>
<td>0.60</td>
<td>0.90</td>
<td>18.5</td>
<td>31.3</td>
<td>50.2</td>
<td>2,176</td>
<td>31.8</td>
<td>37.2</td>
</tr>
<tr>
<td>TSS2000</td>
<td>2,441</td>
<td>0.82</td>
<td>0.67</td>
<td>0.90</td>
<td>13.5</td>
<td>31.4</td>
<td>55.1</td>
<td>1,873</td>
<td>16.6</td>
<td>39.6</td>
</tr>
<tr>
<td>5'UTR</td>
<td>2,401</td>
<td>0.82</td>
<td>0.67</td>
<td>0.90</td>
<td>13.2</td>
<td>31.9</td>
<td>54.8</td>
<td>1,993</td>
<td>15.5</td>
<td>40.7</td>
</tr>
<tr>
<td>Body</td>
<td>2,893</td>
<td>0.76</td>
<td>0.52</td>
<td>0.88</td>
<td>23.3</td>
<td>32.5</td>
<td>44.2</td>
<td>2,407</td>
<td>31.8</td>
<td>40.4</td>
</tr>
<tr>
<td>Location by probe type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron I</td>
<td>3,042</td>
<td>0.78</td>
<td>0.60</td>
<td>0.87</td>
<td>16.8</td>
<td>38.3</td>
<td>44.8</td>
<td>2,444</td>
<td>20.9</td>
<td>45.7</td>
</tr>
<tr>
<td>Intron II</td>
<td>8,372</td>
<td>0.81</td>
<td>0.60</td>
<td>0.90</td>
<td>19.1</td>
<td>29.4</td>
<td>51.5</td>
<td>6,570</td>
<td>23.7</td>
<td>35.7</td>
</tr>
</tbody>
</table>

Table 3. Summary of temporal trend (β-scale) in CpG site methylation, by genomic location and probe typea

<table>
<thead>
<tr>
<th>Location/probe type</th>
<th>N</th>
<th>Percentage of loci with significant trend</th>
<th>X chromosome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Nonsex chromosomes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative trend</td>
<td>Positive trend</td>
</tr>
<tr>
<td>All CpG loci</td>
<td>472,665</td>
<td>1.86</td>
<td>1.54</td>
</tr>
<tr>
<td>Enhancer</td>
<td>101,653</td>
<td>1.83</td>
<td>1.79</td>
</tr>
<tr>
<td>Known DMR</td>
<td>36,053</td>
<td>1.91</td>
<td>1.29</td>
</tr>
<tr>
<td>Location with respect to CpG island</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG Island</td>
<td>145,571</td>
<td>1.91</td>
<td>1.31</td>
</tr>
<tr>
<td>Shoreb</td>
<td>108,839</td>
<td>1.92</td>
<td>1.50</td>
</tr>
<tr>
<td>Shelfb</td>
<td>45,790</td>
<td>1.86</td>
<td>1.70</td>
</tr>
<tr>
<td>Functional location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSS5000</td>
<td>81,267</td>
<td>1.94</td>
<td>1.44</td>
</tr>
<tr>
<td>TSS2000</td>
<td>60,005</td>
<td>1.88</td>
<td>1.38</td>
</tr>
<tr>
<td>5'UTR</td>
<td>62,975</td>
<td>1.88</td>
<td>1.42</td>
</tr>
<tr>
<td>1st Exon</td>
<td>37,656</td>
<td>1.93</td>
<td>1.34</td>
</tr>
<tr>
<td>Body</td>
<td>172,068</td>
<td>1.86</td>
<td>1.58</td>
</tr>
<tr>
<td>3'UTR</td>
<td>19,174</td>
<td>1.83</td>
<td>1.60</td>
</tr>
<tr>
<td>Location by probe type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron I</td>
<td>132,039</td>
<td>1.55</td>
<td>1.59</td>
</tr>
<tr>
<td>Intron II</td>
<td>340,626</td>
<td>1.99</td>
<td>1.52</td>
</tr>
</tbody>
</table>

aAll estimates are adjusted for cell-type composition.

5Shores, up to 2 kb from CpG islands; shelves, from 2 to 4 kb from CpG islands.
adjustment alone left temporal effects nearly unaffected (Fig. 3B). Compared with cell-type adjusted model, quantile normalization resulted in inconsistent changes in the temporal effect magnitude, with larger changes for some loci and smaller changes for others (Fig. 3C and D). Neither cell-type adjustment nor quantile normalization, alone or combined, changed the direction of the observed temporal effects.

Discussion

To our knowledge, this is the first study to look at within- and between-person variability profiles of CpG sites from the Human-Methylation450 BeadChip. Our results show that on average, ICC < 0.5 is common across the genome, and that ICC profiles are similar across all functional locations and probe types. This implies that genomic location at or near CGIs or functional regions has little effect on inter- and intraindividual variability of a CpG locus. We also found that within-person variation tended not to exceed a certain threshold (0.36 for $\hat{\sigma}_w$). Thus, larger between-person variation generally implies a higher ICC. In particular, among CpG loci with between-person variation above the threshold of meaningfulness ($\Delta \hat{\sigma} \geq 0.2$), more than 99% loci had high ICC. Therefore, for differentially methylated loci with sizeable differences in methylation across subjects, that is, loci of interest to most association studies, one measurement may be sufficient for a reliable estimate of methylation level.

We note that CpG loci within DMRs had better ICC profiles: even among those with low between-person variation, the percentage of high-ICC loci was higher than in other genomic locations. This observation is expected, as methylation of tissue-specific lineage-defining regions is stable in blood DNA.

The low ICC values may result in substantial attenuation (reduction in magnitude compared with the true value) in estimated parameters, so repeated measurements would be required per study participant to keep attenuation within some acceptable level. The average ICC across all CpG sites in our study was under 0.5, with the largest proportion of low- and mid-range ICC values observed among low-variability CpG loci. More than one
measurement would be needed for these loci to adequately assess their methylation and to limit attenuation in the estimates; however, because their variability is well below the threshold of meaningful difference, they will likely not be primary targets of an association study.

It should be noted that the low-variability CpG locus group would include most of unmethylated loci with mean methylation close to 0, as such averages can only be attained with low variation. Because of the nature of the logit transformation, this translates to negative M values of very large magnitude, whereby even small differences in the β values may translate to M values that are wide apart. As a result, any estimates for such loci are very unstable and should be treated cautiously. For these loci, it may be preferable to use β values rather than M values, with appropriate adjustments in analytic methods. For example, beta regression techniques (21) could be used for comparison of DNA methylation between two groups.

We have found that the percentage of high-ICC loci is much larger on the X chromosome, implying more stability in methylation patterns. Prior studies have shown that X chromosome inactivation is accompanied by methylation increase at CGIs and at the promoters of genes silenced by X chromosome inactivation (22). Furthermore, methylated promoter CGIs are usually associated with genes in a stable long-term silenced state (15). Although we were unable to distinguish between active and inactive X chromosomes, our observation of more stable methylation patterns at CGIs and near the start of gene coding regions on the X chromosome appears consistent with the effect of gene silencing in X chromosome inactivation.

In our study, we have found that most CpG sites did not exhibit temporal trend that is sufficiently strong to be detected in a small window of <1 year. Among a rather small percentage of sites that did exhibit a trend, negative trend was somewhat more common than positive trend. This finding is in agreement with several prior studies of age-related methylation changes. Using HumanMethylation450 platform on 421 individuals ages 14 to 94 years, Johansson and colleagues (23) reported 29% of CpG sites significantly associated with age, of which 60.5% exhibited decrease.
and 39.5% increase in methylation. Heyn and colleagues (24) showed that most of the genome undergoes age-associated hypermethylation. At the same time, a number of studies found that high CpG density promoters of key developmental genes tend to exhibit age-associated hypermethylation, during both early and late stages of life (6, 25, 26). In two longitudinal studies of newborns followed for 1.5 to 5 years, Martino and colleagues (26, 27) found clear distinction in methylation profile between samples collected at birth and at subsequent clinic visits. They also observed increase in methylation across all classes of annotated genomic regions, with intergenic regions most likely to undergo such changes. A unique aspect of our study is its longitudinal nature combined with a focus on short-term temporal changes. We have been able to show that in a number of CpG loci, temporal changes in methylation that occur in mature adulthood are detectable over a 1-year period of time.

In addressing intragene variation in DNA methylation, adjustment for known sources of such variability could improve the ICC. Temporal change may be one such source of within-person variance for CpG loci that exhibit methylation changes over time. Removing a temporal trend may potentially reduce intragene variation and improve statistical power. However, because age-related methylation change may be an important contributor to carcinogenesis or other pathogenic process, temporal trend removal may sometimes obscure a true association between DNA methylation and disease risk.

Different cell mixture composition between samples taken from the same individual may also confound measured DNA methylation and contribute to the observed within-person variability. Reinious and colleagues (28) established that CpG methylation differs between cell types, such as mononuclear cells, granulocytes, natural killer cells, B cells, and T cells. Jacoby and colleagues (29) analyzed methylation of 58 CpG sites and observed differences in interindividual variability across blood cell types. We adjusted our analyses for cell-type composition using the method of Houseman and colleagues (20). Koestler and colleagues (30) further tested this algorithm on data from the Gene Expression Omnibus database, reporting moderate to high agreement between predicted cell-type composition and that from complete blood counts. In our analyses, there was little difference in ICCs before and after cell-type adjustment. It has also been suggested that changes in cell composition of human blood across a person’s lifespan may largely explain age-associated methylation change (1, 31). In our study, we found that a significant temporal trend disappears after cell-type adjustment in the majority of CpG sites with such trend (Table 3, Supplementary Table S3). We also observed an age-related decrease in the proportion of CD8+ T cells (3.5%/year; \( P = 0.015 \)) and B cells (1.2%/year; \( P = 0.056 \)) and a corresponding increase (4.7%/year; \( P = 0.044 \)) in the proportion of granulocytes (data not shown). Thus, our results support the hypothesis that cell composition change largely accounts for temporal changes in DNA methylation. Although two of the aforementioned studies of methylation change did not account for cell-type composition of whole blood (23, 25), others either considered cell-type composition (27) or focused on specific cell types (6, 24). Thus, associations reported by this latter group of studies are independent of age-related changes in cell composition.

Quantile normalization is a widely used technique to correct for batch effects in DNA methylation and gene expression data. As this technique changes the data, it can affect variability patterns and trends in methylation levels. In our study, we have observed that, although quantile normalization has little effect on the ICC of differentially methylated CpG sites, it often lowers the ICC of low-variability loci. In addition, it changes the magnitude of the temporal trend estimates for CpG loci with significant trend, but does not change its direction. These observations suggest that for most likely targets of a DNA methylation study, such as differentially methylated loci or those with significant trend, applying quantile normalization as a batch correction technique is unlikely to significantly alter analysis results. However, for studies interested in the magnitude of temporal effects in methylation, this normalization technique can have a profound effect on the results and should be used with caution. Other batch correction techniques should also be considered (32).

Limitations of the present study include its modest sample size, which, nonetheless, is not unusual for epigenome-wide DNA methylation studies to date due to the cost constraint. Second, the absence of samples from men may restrict generalizability of our results. Although a recent study by Lam and colleagues (33) found only subtle sex differences in methylation of a small subset of CpG loci, one cannot discount the possibility that variability and temporal changes in some CpG loci may differ between men and women. Besides, there is conflicting evidence on the effect of sex on age-related changes in the methylome outside of sex chromosomes (23, 34). Third, in the absence of duplicate contemporaneous samples from the same subject and of technical replicates, it is unclear how much of the detected within-person variation is due to the temporal trend or other nontemporal factors, such as cell fraction differences. Also for this reason, we could not separate technical variability as part of the total variability in DNA methylation. Despite these limitations, we were able to examine variability patterns and detect short-term temporal methylation changes in a substantial number of CpG sites.

In summary, for CpG loci with differences in methylation between people, methylation levels can be reliably assessed with one blood sample; however, more samples and possibly special statistical methods are needed for low-variability and unmethylated loci. The X chromosome exhibits more stable methylation patterns, especially in CGIs and gene promoters, which is consistent with the effects of X chromosome inactivation. Although short-term temporal changes are largely driven by changes in the cell type composition of blood, trend unrelated to cell type was also detected in a small fraction of CpG sites. Further studies are needed to examine whether CpG loci with short-term temporal trend undergo similar methylation changes throughout lifespan and to what extent such changes are related to the onset or progression of disease.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: Y.B. Shvetsov, M.-A. Song, X.-O. Shu, H. Yu

Development of methodology: Y.B. Shvetsov, M. Tiirikainen

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Q. Cai, M. Tiirikainen, Y.-B. Xiang, H. Yu

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y.B. Shvetsov, M. Tiirikainen, H. Yu

Writing, review, and/or revision of the manuscript: Y.B. Shvetsov, M.-A. Song, Q. Cai, M. Tiirikainen, Y.-B. Xiang, X.-O. Shu, H. Yu

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y.-B. Xiang, X.-O. Shu
Intraindividual Variation in DNA Methylation

Grant Support

This study was supported by NIH grants and contracts P30 CA068455 (Vanderbilt-Ingram Cancer Center), R37 CA070867 (to W. Zheng), and NO2-CP1010-66 (to X.-O. Shu).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 25, 2014; revised December 3, 2014; accepted December 15, 2014; published OnlineFirst December 23, 2014.

S.-A. Song

Acknowledgments

The authors thank Wei Zheng for his leadership and support, Dale Liu, Gong Yang, and Charles Matthews for their contribution to the SPA, Regina Courtney and Ji Wu for assistance in DNA extraction, and Hui Cai for data preparation. Sample preparation was performed at the Survey and Biospecimen Shared Resource, Vanderbilt-Ingram Cancer Center.

References


