Aalborg Universitet

Reproducibility of the Infinium methylationEPIC BeadChip assay using low DNA amounts

Christiansen, Steffan Noe; Andersen, Jeppe Dyrberg; Kampmann, Marie-Louise; Liu, Jing; Andersen, Mikkel Meyer; Tfelt-Hansen, Jacob; Morling, Niels

Published in:
Epigenetics

DOI (link to publication from Publisher):
10.1080/15592294.2022.2051861

Creative Commons License
CC BY-NC-ND 4.0

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

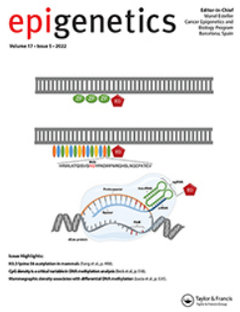
Citation for published version (APA):
Christiansen, S. N., Andersen, J. D., Kampmann, M.-L., Liu, J., Andersen, M. M., Tfelt-Hansen, J., & Morling, N. (2022). Reproducibility of the Infinium methylationEPIC BeadChip assay using low DNA amounts. *Epigenetics*, *17*(12), 1636-1645. https://doi.org/10.1080/15592294.2022.2051861

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: December 05, 2025



Epigenetics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/kepi20

Reproducibility of the Infinium methylationEPIC BeadChip assay using low DNA amounts

Steffan Noe Christiansen, Jeppe Dyrberg Andersen, Marie-Louise Kampmann, Jing Liu, Mikkel Meyer Andersen, Jacob Tfelt-Hansen & Niels Morling

To cite this article: Steffan Noe Christiansen, Jeppe Dyrberg Andersen, Marie-Louise Kampmann, Jing Liu, Mikkel Meyer Andersen, Jacob Tfelt-Hansen & Niels Morling (2022): Reproducibility of the Infinium methylationEPIC BeadChip assay using low DNA amounts, Epigenetics, DOI: 10.1080/15592294.2022.2051861

To link to this article: https://doi.org/10.1080/15592294.2022.2051861

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 31 Mar 2022.

Submit your article to this journal

Article views: 623

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=kepi20

EPIGENETICS
https://doi.org/10.1080/15592294.2022.2051861

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

OPEN ACCESS Check for updates

Reproducibility of the Infinium methylationEPIC BeadChip assay using low DNA amounts

Steffan Noe Christiansen [a], Jeppe Dyrberg Andersen [a], Marie-Louise Kampmann [a], Jing Liu [a,b], Mikkel Meyer Andersen [a,c], Jacob Tfelt-Hansen [a,d], and Niels Morling [a]

[a]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen Denmark; [b]Institute of Forensic Medicine, West China School of Basic Medical Sciences & Forensic Medicine, Sichuan University, Chengdu, Sichuan, China; [c]Department of Mathematical Sciences, Aalborg University, Aalborg Denmark; [d]The Department of Cardiology, The Heart Centre, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

ABSTRACT

The Infinium MethylationEPIC BeadChip (EPIC) is a reliable method for measuring the DNA methylation of more than 850,000 CpG positions. In clinical and forensic settings, it is critical to be able to work with low DNA amounts without risking reduced reproducibility. We evaluated the EPIC for a range of DNA amounts using two-fold serial dilutions investigated on two different days. While the β-value distributions were generally unaffected by decreasing DNA amounts, the median squared Pearson's correlation coefficient (R^2) of between-days β-value comparisons decreased from 0.994 (500 ng DNA) to 0.957 (16 ng DNA). The median standard deviation of the β-values was 0.005 and up to 0.017 (median of medians: 0.014) for β-values around 0.6–0.7. With decreasing amounts of DNA from 500 ng to 16 ng, the percentage of probes with standard deviations ≤ 0.1 decreased from 99.9% to 99.4%. This study showed that high reproducibility results are obtained with DNA amounts in the range 125–500 ng DNA, while DNA amounts equal to 63 ng or below gave less reproducible results.

ARTICLE HISTORY
Received 10 March 2021
Revised 27 February 2022
Accepted 4 March 2022

KEYWORDS
DNA methylation; illumina MethylationEPIC array; variability; correlation; reproducibility; low DNA amounts

Introduction

DNA methylation is an epigenetic mark of transcriptional regulation [1]. In humans, the chemical addition of a methyl group to the DNA most often occurs at cytosines, followed by guanines (CpG sites). There are 28 million CpG sites in the human genome [2]. Several diseases, such as imprinting disorders and different kinds of cancer, are associated with changes in CpG methylation [3–5]. Since DNA methylation alterations are associated with many diseases, there is a great interest in profiling DNA methylation to improve diagnosis, prognosis, and treatment.

Genome-wide DNA methylation can be investigated using different bisulphite based methods including whole-genome bisulphite sequencing (WGBS) and methylation arrays [6]. WGBS covers all CpG positions in the human genome, while methylation arrays cover a fraction of the positions for a considerably lower price per sample. The most recent methylation array by Illumina, the Infinium MethylationEPIC BeadChip (EPIC), is designed for genome-wide investigations of DNA methylation of more than 850,000 CpG positions corresponding to approximately 3% of all the CpG positions. The reproducibility of the results is high, and the results correlate well with those obtained with Illumina's HumanMethylation450 array (450 K) and WGBS [7,8].

Imprinting disorders are characterized by gene dysregulation and aberrant DNA methylation patterns at disease-specific loci [3]. While imprinting disorders may be diagnosed only by investigating specific loci, diagnostic tests of other diseases may require genome-wide methylation profiles. DNA methylation alterations in cancer cells often affect many loci throughout the genome [2,9]. EPIC and 450 K arrays are used to generate methylation profiles for a recently developed tumour classifier, which has been shown to improve treatment by

CONTACT Steffan Noe Christiansen steffan.christiansen@sund.ku.dk Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Frederik V's Vej 11, 2100, Copenhagen Denmark
Supplemental data for this article can be accessed here

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

reassigning the tumours to a different subtype in up to 12% of the cases [10].

DNA methylation has gained increased interest in forensic genetics. DNA methylation can potentially be used to identify the cellular origin of trace samples, distinguish monozygotic twins, and provide information of the suspect that left a trace sample, including age, smoking, alcohol intake, drug abuse, and body shape [11].

In clinical, forensic, and research investigations of DNA methylation, it is critical to be able to work with samples with low amounts of DNA, like trace samples in forensic genetics. Up to 550,000 single nucleotide polymorphisms have been genotyped using array technology with only a few ng of DNA [12]. However, most hybridization based methods require around 250–500 ng DNA. An important reason for the need for large amounts of DNA is that the bisulphite conversion process destroys approximately 90% of the DNA. It is rarely challenging to obtain enough DNA from peripheral blood in the clinical setting. However, sampling sufficient amounts of solid tissue may be a challenge. Therefore, it is of great interest to know the reproducibility of the assay for low DNA amounts. The use of different DNA amounts for both 450 K and EPIC has been investigated previously but with other experimental designs and analyses than those presented in this study [13,14].

We investigated the within-laboratory reproducibility of the EPIC array for a range of DNA amounts by two-fold serial dilutions of DNA starting with 500 ng DNA. The distribution of signal intensities measured as arbitrary fluorescent units (AFU) by the iScan™ system and methylation values were compared among different DNA amounts. The between-days correlations of β-values were investigated using the squared Pearson's correlation coefficient (R^2), and the between-days β-value variability was evaluated using the standard deviation (SD) and coefficient of variation (CV).

Materials and methods

Design of the study

Figure 1. shows the design of the study. The experiments were performed on two different days, which are referred to as 'Day 1' and 'Day

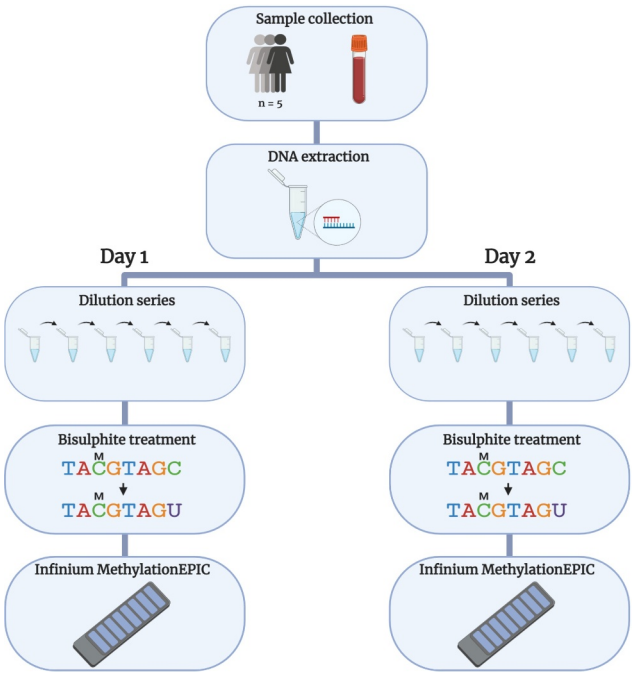


Figure 1. Study design Blood samples were collected from five female individuals. The DNA was extracted and two-fold serially diluted to 500 ng, 250 ng, 125 ng, 63 ng, 32 ng, and 16 ng DNA. The DNA was sodium bisulphite treated and investigated using the Illumina Infinium MethylationEPIC BeadChip Kit. Created with BioRender.com.

2.' The extracted DNA was stored at −80°C with approximately six months between the two investigations. The locations of the samples on the slides are shown in Figure S1.

Blood samples and DNA preparation

Peripheral blood from five females was collected in EDTA and stored at −80°C until DNA extraction. The DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's recommendations for the purification of total DNA from whole blood. The DNA concentrations were measured using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). Two-fold serial dilutions of DNA were performed twice using nuclease-free water to obtain samples with the following theoretical DNA amounts: 500 ng, 250 ng, 125 ng, 63 ng, 32 ng, and 16 ng. The bisulphite conversion of DNA was performed using the EZ DNA Methylation™ Kit (Zymo Research) following the manufacturer's recommendations with a final elution volume of 10 µl.

Measurement of DNA methylation

The samples were analysed using the Infinium MethylationEPIC Kit (Illumina) following the manufacturer's recommendations. Briefly, the bisulphite treated DNA was amplified using whole-genome amplification before fragmentation and precipitation. The resuspended DNA was subsequently hybridized to probes attached to the BeadChips, and the unhybridised DNA was removed. Thereafter, the attached probes were subject to single-base extension and stained. The BeadChips were scanned using the iScan™ system (Illumina) following the manufacturer's recommendations.

Data analysis

The data analysis was performed in the statistical environment R (version 3.6.0) using the *tidyverse* packages (version 1.3.0) 15. Bead counts were obtained using the *illuminaio* package (version 0.34.0) [16]. Raw .idat files were imported into R using the *minfi* package (version 1.32.0) [17,18]. *MethylAid* (version 1.20) [19] was used to identify poorly performing samples and analyse the results of the sample-independent controls (staining, extension, target removal, and hybridization controls) and sample-dependent controls (bisulphite conversion of type I and II probes, specificity of type I and II probes, nonpolymorphic, and negative controls) as described in the Infinium® HD Assay Methylation Protocol Guide (15019519 v01).

The default threshold values of *MethylAid* were used to evaluate the performance of the sample-dependent and sample-independent controls positioned on the BeadChip. The threshold values were 10 for the median methylated vs. unmethylated signals, 12 for the overall quality control, 11.75 for the bisulphite conversion efficiency, 12.75 for the hybridization value, and 0.95 for the mean detection *p*-value. All quality checks had to be accepted to allow sample inclusion. The identities of the samples were checked using the 59 identification SNPs of the EPIC array. The heatmap and dendrograms (Figure S2) were visualized using *pheatmap* (v. 1.0.12) using the Euclidian distance and complete clustering.

The red and green signal intensity measures from the iScan™ were converted into unmethylated and methylated signals using the *preprocessRaw()* followed by the *getUnmeth()* and *getMeth()* functions in *minfi*. The data analysis of the signal intensities was performed for all 866,091 positions. The β-values were calculated as indicated in equation 1 with α = 0 except in the test of using α = 100.

$$\beta = \frac{methylated\,signal}{unmethylated\,signal + methylated\,signal + \alpha} \tag{1}$$

The α-value is an arbitrarily selected value that is often set to 100 to avoid data that are difficult to handle. By using an α-value of 100, the β-value is, however reduced. β-values of positions with zero intensities of methylated or unmethylated signals were excluded for further analysis. Results of 4,534 probes were removed due to the lack of signal. The β-values of the remaining 861,557 methylation positions were analysed.

Pearson's correlation coefficient was calculated to estimate the correlation between the technical replicates and among the different DNA amounts. R^2 was reported in the analyses since this value explains the fraction of systematic variation

between two variables. The between-days variability of the β-value of each position was evaluated using both the SD and the CV. The methylation data analyses were visualized using *ggplot2* (v. 3.3.2). The density plots were generated with density estimates based on a Gaussian kernel with default settings and plotted using *geom_density()*. The overlapping histogram was plotted using *geom_histogram()* with the density at the ordinate. Scatter plots were visualized using *geom_bin2d()* with the counts represented in each bin among 500 × 500 bins. Dot plots were visualized using *geom_point()*. Violin plots overlapped by 25th percentiles, medians, and 75th percentiles were made using *geom_violin()*. The distribution of the variability was visualized using *geom_line()*.

Results

Quality control

All samples except one (Individual D, 500 ng, Day 2) passed the *MethylAid* quality control. For that particular sample, the calculated values were below the threshold values for all sample dependent and sample-independent controls except for the hybridization control. All remaining samples passed the sample identity control using the 59 identification SNPs (Figure S2).

Distribution of signal intensities

The methylated and unmethylated signal intensities showed two slightly different distributions with two sets of peaks (Figure S3). The first set of peaks at approximately 1,000 AFU was highest for small DNA amounts, while the second set of broader peaks around 8,000 AFU showed the highest densities for high DNA amounts. The density distribution of the unmethylated signal intensities was similar except for a plateau instead of a peak at 8,000–18,000 AU. There was a relationship between the DNA amount and the signal distributions with a sequential decrease in the density of the highest signal intensities with each reduction of the DNA amount.

β-value distribution

The β-values of CpG positions with low signal intensities should, in theory, be relatively more affected by α-values above 0 than positions with high signal intensities (Equation 1). We compared the β-values when using α equal to 0 and 100, respectively, for high and low DNA amounts (Figure S4). The maximal β-values for positions with total signal intensities below approximately 12,500 AFU were reduced for both samples with high (250 ng) and low (16 ng) amounts of DNA when α = 100 was used instead of α = 0. We decided to calculate the β-values with α = 0 since the use of α = 100 artificially reduced the maximal β-values. In total, 4,534 positions were removed in this part of the analysis because no methylated or unmethylated signal was measured.

Figure 2 shows the density plots of the β-values with different amounts of DNA from five individuals. The plots are supplemented with a histogram of the β-values and the corresponding density plot of the β-values (Individual A, 500 ng, Day 1). The density plots revealed no systematic difference between the β-values with different DNA amounts for any individual.

Between-days correlations of β-values with different DNA amounts

All samples were investigated twice on different days to evaluate the between-days correlation of the β-values. The results obtained for individual A is shown as a representative example (Figure 3; Figure S5-8). Decreasing DNA amounts were associated with increased dispersion around the identity line (45°). This was most pronounced for high β-values. A gradual decrease in median R^2 from 0.994 to 0.957 was observed with decreasing DNA amounts from 500 ng to 16 ng (Figure 4).

The β-value correlations among different DNA amounts investigated on the same day were analysed (Figure S9; Table S1). The correlations between the β-values obtained from samples with high amounts of DNA (125 ng, 250 ng, and 500 ng DNA) were, in general, high with median R^2 of 0.994, 0.991, and 0.991, respectively, for the mutual comparisons. The correlation coefficients

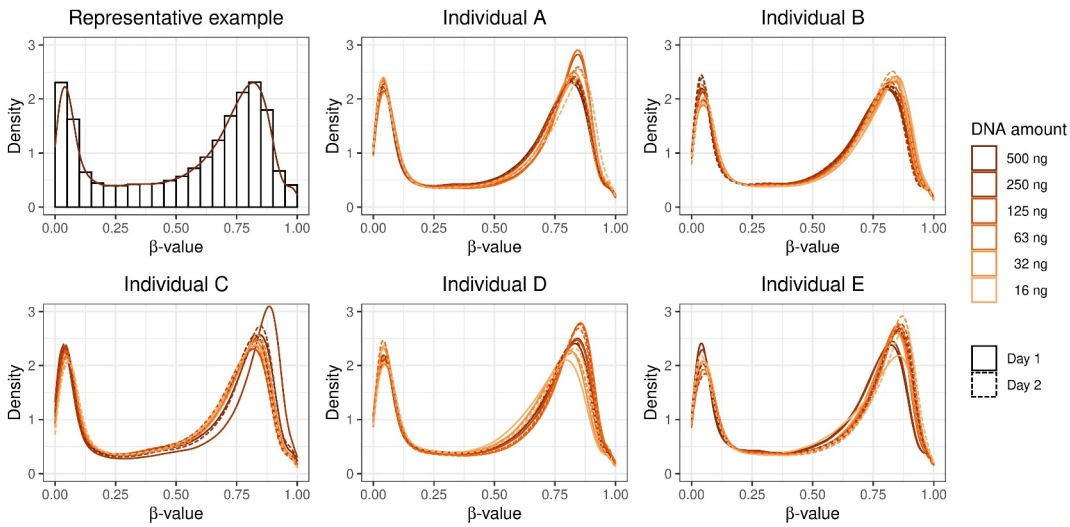


Figure 2. Histogram and density plots of β-values The representative example shows a histogram of β-values overlaid with a density plot for a single sample with 500 ng from Individual A. The remaining density plots were obtained using various amounts of DNA from five individuals investigated twice on different days. All density estimates were based on a Gaussian kernel.

were lower among samples with low amounts of DNA (16 ng, 32 ng, and 63 ng DNA) with median correlation coefficients of 0.979, 0.962, and 0.961 for the mutual comparisons, respectively. There was a clear association between decreasing amounts of DNA and the decrease in R^2 (Table S1).

Variabilities of β-values with various DNA amounts

The between-days variabilities of β-values were calculated as the SD and the CV. The decrease in DNA amount was associated with an increase in the SD and CV (Figure 5). The percentage of probes with SD ≤ 0.1 was decreased from 99.9%

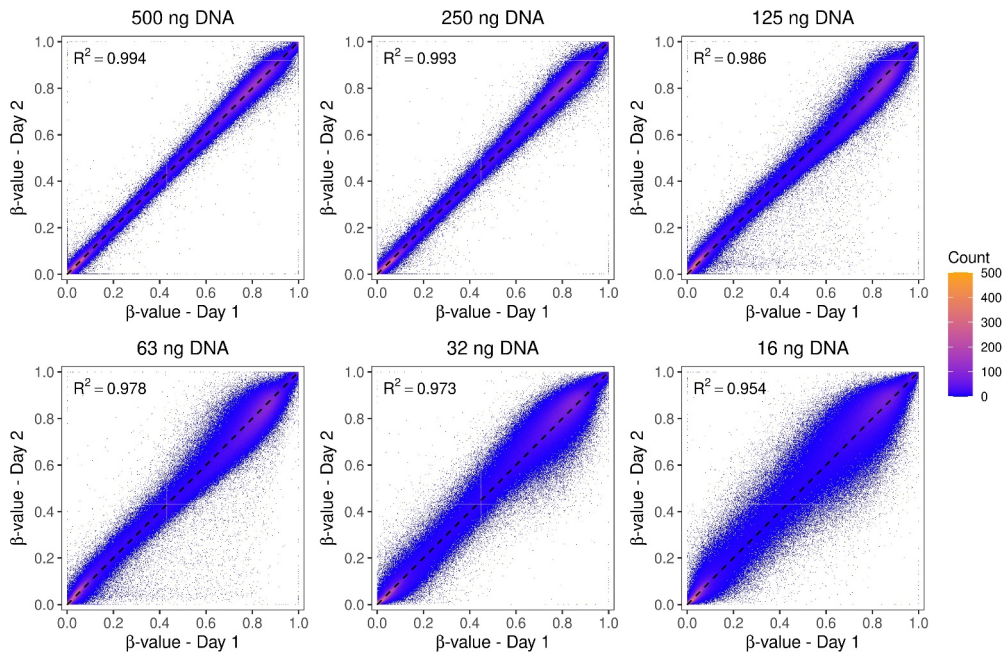


Figure 3. Between-days correlations of β-values of a representative individual. Scatter plots of β-values obtained on two different days with 500 ng (a), 250 ng (b), 125 ng (c), 63 ng (d), 32 ng (e), and 16 ng (f) DNA, respectively, from individual A.

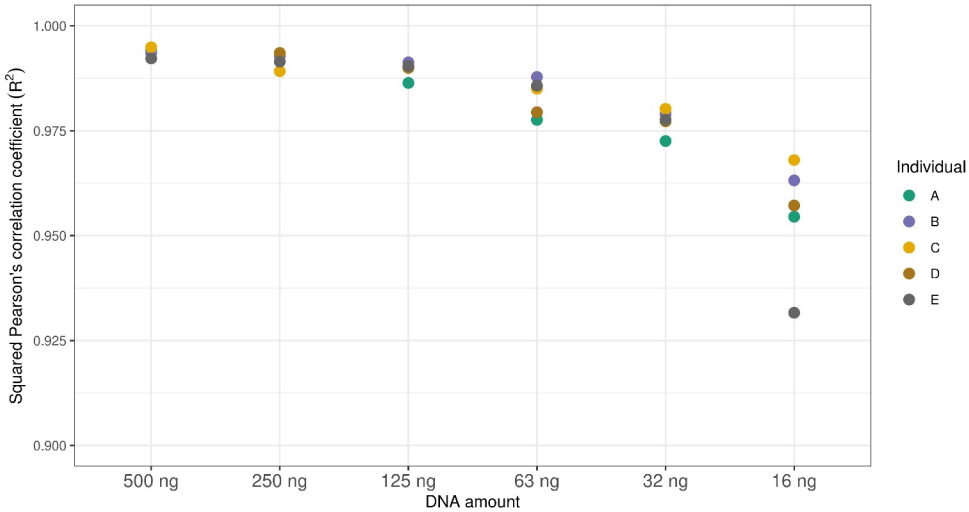


Figure 4. Between-days correlations of β-values of all individuals. The squared Pearson's correlation coefficient between β-values obtained with various amounts of DNA from five individuals, respectively, investigated twice on different days.

to 99.4% with decreasing amounts of DNA from 500 ng to 16 ng. The percentage of probes with CV ≤ 0.1 was decreased from 84.6% to 67.2% with decreasing amounts of DNA from 500 ng to 16 ng.

The variability measures, SD and CV, of β-the values varied with the β-values. Hence, the β-values were grouped into intervals of 0.10 to evaluate the variabilities (Figure 6). The SDs were lowest in the intervals [0.0–0.1] (median of medians: 0.006; range of medians: 0.005–0.006) and the interval [0.9–1.0] (median of medians: 0.011; range of medians: 0.009–0.013) (Figure 6a; Table S2) and highest in the interval [0.6–0.7] (median of medians: 0.019; range of medians: 0.014–0.039). Among all intervals, the median SDs ranged from 0.005 to 0.017 (median of medians: 0.014) in samples with 125–500 ng DNA, while the median SDs ranged from 0.006 to 0.039 (median of medians: 0.021) in samples with 16–63 ng DNA (median of medians: 0.021) (Table S2).

The CVs of the β-value intervals decreased from the interval [0.0–0.1] (median of medians: 0.148; range of medians: 0.135–0.166) to the interval [0.9–1.0] (median of medians: 0.012; range of medians: 0.010–0.014) (Figure 6b;

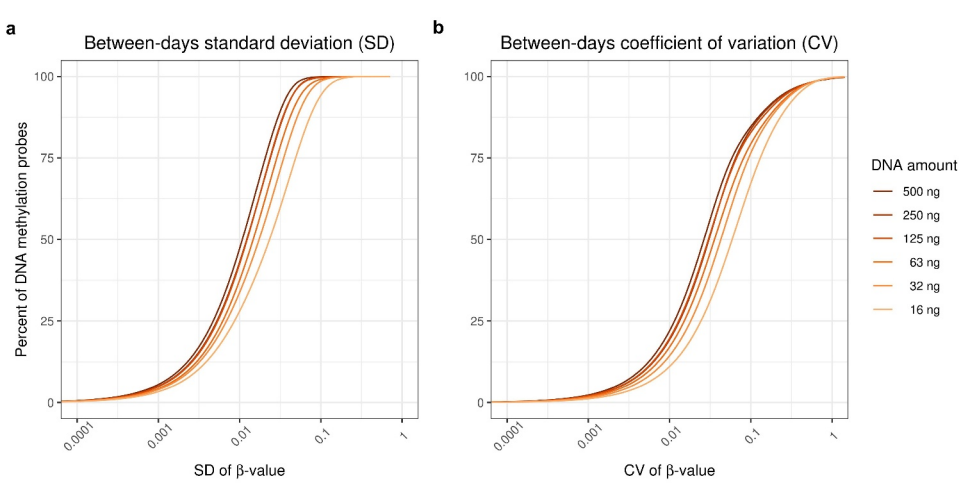


Figure 5. Distributions of the variabilities of the β-values of the probes investigated with various amounts of DNA on two different days.

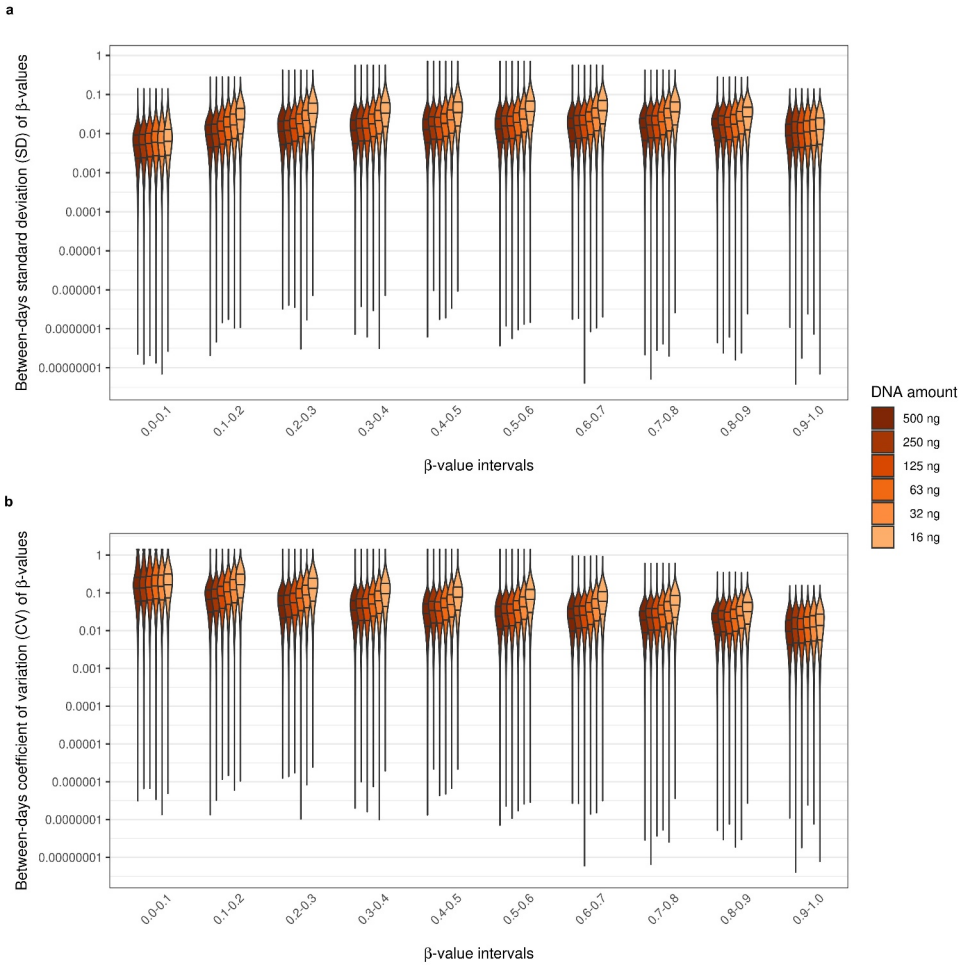


Figure 6. Between-days variabilities of β-values Violin plots of the between-days standard deviations (SD) (a) and coefficients of variation (CV) (b) of all β-values of DNA from five individuals. The SD and CV values were calculated for β-value intervals of 0.10. The horizontal lines within the violins indicate the upper quartiles, medians, and lower quartiles, respectively.

Table S3). The CV was increased with decreasing DNA amounts. This tendency was most pronounced in the midrange intervals.

Discussion

The lack of adequate amounts of DNA from samples is a common challenge in forensic genetic investigations and diagnostic laboratories using scarce amounts of tissue and blood. In this study, we investigated the use of small DNA amounts extracted from blood for examination with the EPIC array to evaluate the performance when only limited DNA amounts are available. Illumina recommends using ≥ 250 ng DNA (*Infinium HD Methylation assay workflow for fresh/frozen versus FFPE samples*, 2020.02.18). We decided to use twice that amount since many studies were performed with higher DNA amounts [7,8,20,21]. Our results showed that decreasing DNA amounts affected the raw signal intensities but barely changed the β-value distributions. In contrast, the between-days correlation and variability were affected to a great extent.

The investigated dilution series were made from the same DNA extractions. The experimental procedures were conducted six months apart [Figure 1]. Bulla et al. [22] showed that the difference in DNA methylation of EDTA stabilized blood samples stored for up to one year at −80°C was < 1%. We assumed that the storage effect on DNA methylation in our experimental setup would be at a similar, low level.

Normalization and probe removal during methylation array data processing is a common procedure to address, e.g., unbalanced gender distribution, common genetic variation, or poorly performing probes [23]. We applied a simple data processing approach without normalization and probe filtration in contrast to previous investigations of different DNA amounts [13,14].

We used an α-value equal to 0 for the β-value calculation since the α-value may be unnecessary in many preprocessing pipelines [24]. Additionally, an α-value equal to 100 may spuriously reduce the β-values of positions with low signal intensities (Figure S4). Simple data processing was performed to conduct a transparent and unbiased evaluation of the kit. However, we are also well aware that the methylation measurements for each individual would most likely have been more correlated if normalization and probe filtering had been performed. The simple data processing may explain the relative low β-value correlation between the samples with the highest and lowest DNA amounts (Figure S9; Table S1) compared to the previous comparison of high and low DNA amounts on the 450 K array [14].

Y-chromosome probes were not expected to give any signal, since only women were investigated. However, 270 out of 537 Y-chromosome probes gave signals. The median signal intensity from Y-chromosome probes was 16.2–34 times lower than that for the remaining chromosomes. Signals with Y-chromosome probes in females have been observed and discussed before [25,26]. The signals are most likely caused by Y-chromosome probes binding to X-chromosome sites with high target sequence similarities with Y-chromosome sites.

While the densities of the signal intensities were affected by decreasing DNA amounts (Figure S3), we did not observe systematic changes in the β-value distributions with decreasing DNA amounts (Figure 2). β-values are ratios of methylated and unmethylated signal intensities, as shown in Equation 1. If the signal intensities were equally decreased for both signals, the β-value-distributions would not be affected. We expected that decreased signal intensities would increase the variance of the ratio. This was indeed what we observed when we investigated the correlation between the β-values of sample duplicates. Thus, the between-days R^2 decreased with decreasing DNA amounts (Figure 4). The ranges of the R^2 for the between-days correlations were 0.992–0.995 (500 ng), 0.989–0.993 (250 ng), 0.986–0.991 (125 ng), 0.978–0.988 (63 ng), 0.973–0.980 (32 ng), and 0.932–0.968 (16 ng DNA).

Investigations of different individuals on the same day with 500 ng DNA showed inter-individual R^2 values of 0.974–0.982. The high inter-individual correlations suggest that the methylation degree of relatively few CpG positions vary from one individual to another. Hence, inter-individual differences will be very difficult to detect using a low DNA amount, where the between-days R^2 in some cases is even lower.

Our results showed that decreased DNA amounts lead to increased variability (Figure 6; Table S2-3). This tendency was present in all β-value intervals except for the most extreme ones ([0.0–0.1] and [0.9-1-0]). This increase in the variability of β-values was evident for the lower and upper quartiles and the median values within each interval. The relative CV change ranged from only 1.5 to 20.9% for the different β-value intervals using 500 ng and 250 ng DNA. In comparison, the relative CV change ranged from 7.7 to 58.1% between 32 ng and 16 ng suggesting an increased susceptibility to deviation in low-amount samples. This increase in variability emphasizes that the results of especially low-amount samples should be interpreted very cautiously. The variability depended on both the β-values and the DNA amount (Table S2-3). The low SD of β-values close to 0 and 1 was previously demonstrated for the HumanMethylation27 BeadChip [27]. Here, we showed that even though the SD changed almost three-fold among the β-value intervals using the recommended DNA amounts (250 ng: 0.005–0.017), the SD barely changed from 125 ng to 500 ng within each β-value interval.

The increased variability with decreased DNA amounts is mainly caused by a reduced number of probes bound to DNA, leading to increased stochastic variation and an increased proportion of

failing probes, as shown by Abbassi et al. [13), who explored the analytical sensitivity and other technical details of the Illumina 450HM array.

The low number of repeated investigations is a limitation of the study of the between-days variability. However, the homogeneity of the results obtained with 850,000 methylation positions makes it most likely that more extensive studies will result in similar estimates. The handling, including storage temperature and freezing-thawing of the tissue samples, may also affect the amount needed for the EPIC analysis as seen with miRNA in blood samples [28].

The sample location setup was similar on the two investigation days. The setup optimized the analysis of the technical reproducibility. However, the setup may have affected the comparisons of signal intensities and β-values between the various positions on the slides, DNA amounts, and individuals because positional effects are present with the Illumina BeadChips [29,30].

The last limitation is the unknown amount of DNA degraded by bisulphite treatment, where most DNA is usually destroyed. It is recommended, by Illumina, to use the entire eluate from the last step of the bisulphite conversion as input for the EPIC array protocol. Hence, it was not possible to measure the DNA amount after the conversion to evaluate the DNA amounts.

Conclusions

This study showed that 125–500 ng DNA from peripheral blood gives reproducible results with the Infinium MethylationEPIC Kit with median SDs of the β-value intervals of 0.005 and up to 0.017 (median of medians: 0.014). DNA amounts equal to 63 ng or below showed less reproducible results.

Acknowlegdements

We thank Anja Ladegaard Jørgensen and Stine Bøttcher Jacobsen for assistance with this work.

Data Availability Statement

Data cannot be shared publicly because of data protection legislations (Danish Data Protection Agency no. 2011-54-1262). Data are available from the University of Copenhagen, Department of Forensic Medicine Institutional Data Access/Ethics Committee (contact: retsmedicinsk.institut@sund.ku.dk, University of Copenhagen, Department of Forensic Medicine) for researchers, who meet the criteria for access to confidential data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Disclosure of interest

The authors report no conflict of interest.

Ethics

All procedures performed in the study were in accordance with the ethical standards of the Committees on Health Research Ethics in the Capital Region of Denmark (H-20039524) and the Danish Data Protection Agency (2011-54-1262). Written informed consent was provided by all study participants.

Funding

This work was supported by the Ellen and Aage Andersen's Foundation, the China Scholarship Council, and the Novo Nordic foundation No. NFF18OC0031634.

ORCID

Steffan Noe Christiansen http://orcid.org/0000-0002-8655-3452
Jeppe Dyrberg Andersen http://orcid.org/0000-0002-4395-5232
Marie-Louise Kampmann http://orcid.org/0000-0001-5403-5960
Mikkel Meyer Andersen http://orcid.org/0000-0002-0234-0266
Jacob Tfelt-Hansen http://orcid.org/0000-0003-3895-9316
Niels Morling http://orcid.org/0000-0002-9463-5087

References

[1] Schübeler D. Function and information content of DNA methylation. Nature. 2015;517(7534):321–326.
[2] Stirzaker C, Taberlay PC, Statham AL, et al. Mining cancer methylomes: prospects and challenges. Trends Genet. 2014;30(2):75–84.
[3] Eggermann T, Perez de Nanclares G, Maher ER, et al. Imprinting disorders: a group of congenital disorders with overlapping patterns of molecular changes affecting imprinted loci. Clin Epigenetics. 2015;7(1):123.

[4] Ernst C, McGowan PO, Deleva V, et al. The effects of pH on DNA methylation state: in vitro and post-mortem brain studies. J Neurosci Methods. 2008;174(1):123–125.
[5] Feinberg AP, Longo DL. The key role of epigenetics in human disease prevention and mitigation. N Engl J Med. 2018;378(14):1323–1334.
[6] Bock C. Analysing and interpreting DNA methylation data. Nat Rev Genet. 2012;13(10):705–719.
[7] Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8(3):389–399.
[8] Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17(1):208.
[9] Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet. 2016;17(5):284–299.
[10] Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. Nature. 2018;555(7697):469–474.
[11] Vidaki A, Kayser M. From forensic epigenetics to forensic epigenomics: broadening DNA investigative intelligence. Genome Biol. 2017;18(1):238.
[12] Pan X, Urban AE, Palejev D, et al. A procedure for highly specific, sensitive, and unbiased whole-genome amplification. Proc Nat Acad Sci. 2008;105(40):15499–15504.
[13] Abbasi M, Smith AD, Swaminathan H, et al. Establishing a stable, repeatable platform for measuring changes in sperm DNA methylation. Clin Epigenetics. 2018;10(1):119.
[14] Whalley C, Payne K, Domingo E, et al. Ultra-low DNA input into whole genome methylation assays and detection of oncogenic methylation and copy number variants in circulating tumour DNA. Epigenomes. 2021;5(1):6.
[15] Wickham H, Averick M, Bryan J, Chang W, McGowan L, François Ret al, Welcome to the Tidyverse. J Open Source Softw [Internet]. 2019;4(43):1686. [Cited 27 Mar 2020]. Available from: https://joss.theoj.org/papers/10.21105/joss.01686
[16] Smith ML, Baggerly KA, Bengtsson H, et al. illuminaio: an open source IDAT parsing tool for Illumina microarrays. *F1000Research*. 2013;2:264.
[17] 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.
[18] Fortin J-P, Triche TJ, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2016;33(4):btw691.
[19] van Iterson M, Tobi EW, Slieker RC, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics. 2014;30(23):3435–3437.
[20] Kling T, Wenger A, Beck S, et al. Validation of the MethylationEPIC BeadChip for fresh-frozen and formalin-fixed paraffin-embedded tumours. Clin Epigenetics. 2017;9(1):33.
[21] Zaimi I, Pei D, Koestler DC, et al. Variation in DNA methylation of human blood over a 1-year period using the Illumina MethylationEPIC array. Epigenetics. 2018;13(10–11):1056–1071.
[22] Bulla A, De Witt B, Ammerlaan W, et al. Blood DNA Yield but not integrity or Methylation Is Impacted after long-term storage. Biopreserv Biobank. 2016;14(1):29–38.
[23] Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. *F1000Research*. 2017;5(3):1281.
[24] Zhou W, Triche TJ, Laird PW, et al. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. Nucleic Acids Res. 2018;46(20):1–15.
[25] Chen YA, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203–209.
[26] Heiss JA, Just AC. Improved filtering of DNA methylation microarray data by detection p values and its impact on downstream analyses. Clin Epigenetics. 2019;11(1):15.
[27] Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:11.
[28] Glinge C, Clauss S, Boddum K, et al. Stability of Circulating Blood-Based MicroRNAs - Pre-Analytic Methodological Considerations. PloS One. 2017;12(2):e0167969.
[29] Jiao C, Zhang C, Dai R, et al. Positional effects revealed in Illumina methylation array and the impact on analysis. Epigenomics. 2018;10(5):643–659.
[30] Verdugo RA, Deschepper CF, Muñoz G, et al. Importance of randomization in microarray experimental designs with illumina platforms. Nucleic Acids Res. 2009;37(17):5610–5618.