

BRIEF REPORT



Correlation of Infinium HumanMethylation450K and MethylationEPIC BeadChip arrays in cartilage

Kathleen Cheung ^{a,b}, Marjolein J. Burgers^a, David A. Young ^a, Simon Cockell^b, and Louise N. Reynard^a

^aSkeletal Research Group, Institute of Genetic Medicine, Newcastle University, Central Parkway, Newcastle upon Tyne, UK; ^bBioinformatics Support Unit, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK

ABSTRACT

DNA methylation of CpG sites is commonly measured using Illumina Infinium BeadChip platforms. The Infinium MethylationEPIC array has replaced the Infinium Methylation450K array. The two arrays use the same technology, with the EPIC array assaying almost double the number of sites than the 450K array. In this study, we compare DNA methylation values of shared CpGs of the same human cartilage samples assayed using both platforms. DNA methylation was measured in 21 human cartilage samples using the both 450K and EPIC arrays. Additional matched 450K and EPIC data in whole tumour and whole blood were downloaded from GEO GSE92580 and GSE86833, respectively. Data were processed using the Bioconductor package Minfi. DNA methylation of six CpG sites was validated for the same 21 cartilage samples by pyrosequencing. In cartilage samples, overall sample correlations of methylation values between arrays were high (Pearson's $r > 0.96$). However, 50.5% of CpG sites showed poor correlation ($r < 0.2$) between arrays. Sites with limited variance and with either very high or very low methylation levels in cartilage exhibited lower correlation values, corroborating prior studies in whole blood. Bisulphite pyrosequencing did not highlight one array as generating more accurate methylation values. For a specific CpG site, the array methylation correlation coefficient differed between cartilage, tumour, and whole blood, reflecting the difference in methylation variance between cell types. Researchers should be cautious when analysing methylation of CpG sites that show low methylation variance within the cell type of interest, regardless of the method used to assay methylation.

ARTICLE HISTORY

Received 19 August 2019
Revised 13 November 2019
Accepted 22 November 2019

KEYWORDS

DNA methylation; 450K vs EPIC; cartilage; 450K; EPIC

Introduction

Illumina methylation BeadChip arrays are a popular method for assaying methylation of CpG sites across the human genome. In 2016, the Illumina MethylationEPIC (EPIC) array replaced the Methylation450K (450K) array, which has since been discontinued. The EPIC array uses the same Infinium probe technology and shares 452,453 common probes, but assays almost double the number of CpG sites compared to the 450K array. The 450K and EPIC arrays utilize Infinium type I and type II probes. Type I probes consist of two probes, each binding to either methylated or unmethylated sites, whereas type II probes use one probe to bind to the target CpG. Type I probes are reported to be more effective at assaying completely unmethylated or completely methylated probes [1]. The combination of type I and type II probes ensures that the whole range of methylation value

can be assayed genome wide. The same CpG is assayed by the same probe type on both arrays; sites do not change probe chemistries between arrays; a comprehensive description of probe types is available [2]. As ongoing studies and meta-analyses may use data from both array platforms, previous studies have assessed the concordance of the two by correlating methylation levels of the same samples assayed on both arrays. Studies comparing the two arrays have been reported in tumour, whole blood, and placental samples [3–6]. Whilst high overall per sample correlations (Pearson's $r > 0.9$) were reported, two of the blood studies found that many individual sites in whole blood showed poor correlations ($r < 0.2$) [2,3]. The studies in whole blood suggested that CpG sites with low variance and that are either completely methylated or un-methylated tend to exhibit a poor correlation between 450K and EPIC arrays. Furthermore,

CONTACT Kathleen Cheung  kat.cheung@ncl.ac.uk

 Supplemental data for this article can be accessed [here](#).

© 2019 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 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

several studies have highlighted sites with low correlation in blood and placental tissue that they have recommended be removed from the analysis when performing a meta-analysis of 450K and EPIC datasets [3–5].

In this study, we investigate the correlation of the 450K and EPIC platforms in human cartilage, a tissue which comprises a single cell type. As variance of DNA methylation in single CpG sites has been found to affect correlation across 450K and EPIC arrays, we hypothesize that tissue types with different inherent DNA methylation variances do not have the same correlation across the two arrays. Tumour tissue is heterogeneous and whole blood consists of many cell types, factors which affect the variance of DNA methylation of CpG sites in a sample. We assayed 21 cartilage samples on both BeadChip arrays and examined per sample and per CpG correlations. We also downloaded publicly available 450K and EPIC data in whole tumour (GEO GSE92580) and whole blood (GSE86833) to compare the variance and correlation across arrays with cartilage samples. We sought to investigate the accuracy of array generated methylation values for six CpG sites using the gold standard of bisulphite pyrosequencing in the 21 samples common to both arrays.

Results

Overall sample correlation between 450K and EPIC arrays

DNA methylation in human cartilage samples, 9 from patients undergoing total joint replacement with osteoarthritis (OA) of the knee and 12 with neck of femur fracture (NOF), was assessed using the 450K and EPIC arrays. The correlation between the same samples ran on both arrays was examined, with probes common to both arrays included in the analysis. In total, 450K and EPIC share 452,453 probes. Quality control measures and probe filters including removal of sex chromosome and cross-reacting probes were applied to ensure only probes which could be measured with confidence were used for further analysis (see Methods section). After quality control of probes, 370,669 remained. Additionally, all cartilage samples had >95% probes with a detection p value <0.01 on both 450K and EPIC arrays. Both arrays were normalized using the

functional normalization method [7], which uses control probes to account for unwanted variation in the data.

To investigate the correlation between the arrays, the Pearson correlation coefficient was calculated per sample for the 370,699 probes common to both arrays and passing quality filters. Corroborating previous studies [3,4,6,8], we found that the per sample correlation of methylation β -value was high. A 450K vs EPIC scatter plot of methylation β -value for one cartilage sample illustrates that the overall correlation was high (Pearson's $r = 0.99$) although many CpGs show substantial differences between arrays (Figure 1a). The overall per sample correlation between the same sample assayed on the EPIC array twice was also high at $r = 0.99$ (Figure 1b). However, similar to Logue et al. [3], we also observed that pairwise correlations between any two samples from different individuals and arrays were high. A correlation heatmap showed that samples were more likely to be similar to other samples assayed on the same array, rather than to itself run on the other array (Figure 1c). The separation between arrays is also seen and is likely a batch effect; nevertheless, we still observe high Pearson's r across all samples which implies that a high overall correlation may not necessarily reflect biological similarity.

Variance and correlation of CpG sites by probe type

We next examined the correlation between 450K- and EPIC-generated β values for individual CpG sites common to the 450K and EPIC arrays. Similar to prior studies, we found that over 50% of the sites showed poor correlation (Pearson's $r < 0.2$), with a median per CpG correlation of 0.195 (mean 0.218) (Supplemental Table 1A). Only 3.56% of the probes showed high ($r \geq 0.8$) correlation values. We also observed that CpG site methylation variance was generally low across our cartilage samples (Figure 2a). Poor correlation CpG sites ($r < 0.2$) had a significantly lower β value variance than high correlation ($r > 0.8$) sites (median variance of 0.0018 and 0.0064, respectively, $p < 2.2 \times 10^{-16}$, Fligner–Killeen non-parametric variance test). This confirms that the pattern of low variance sites and poor correlation observed in whole blood [3,4] and placenta [5] is also seen in cartilage.

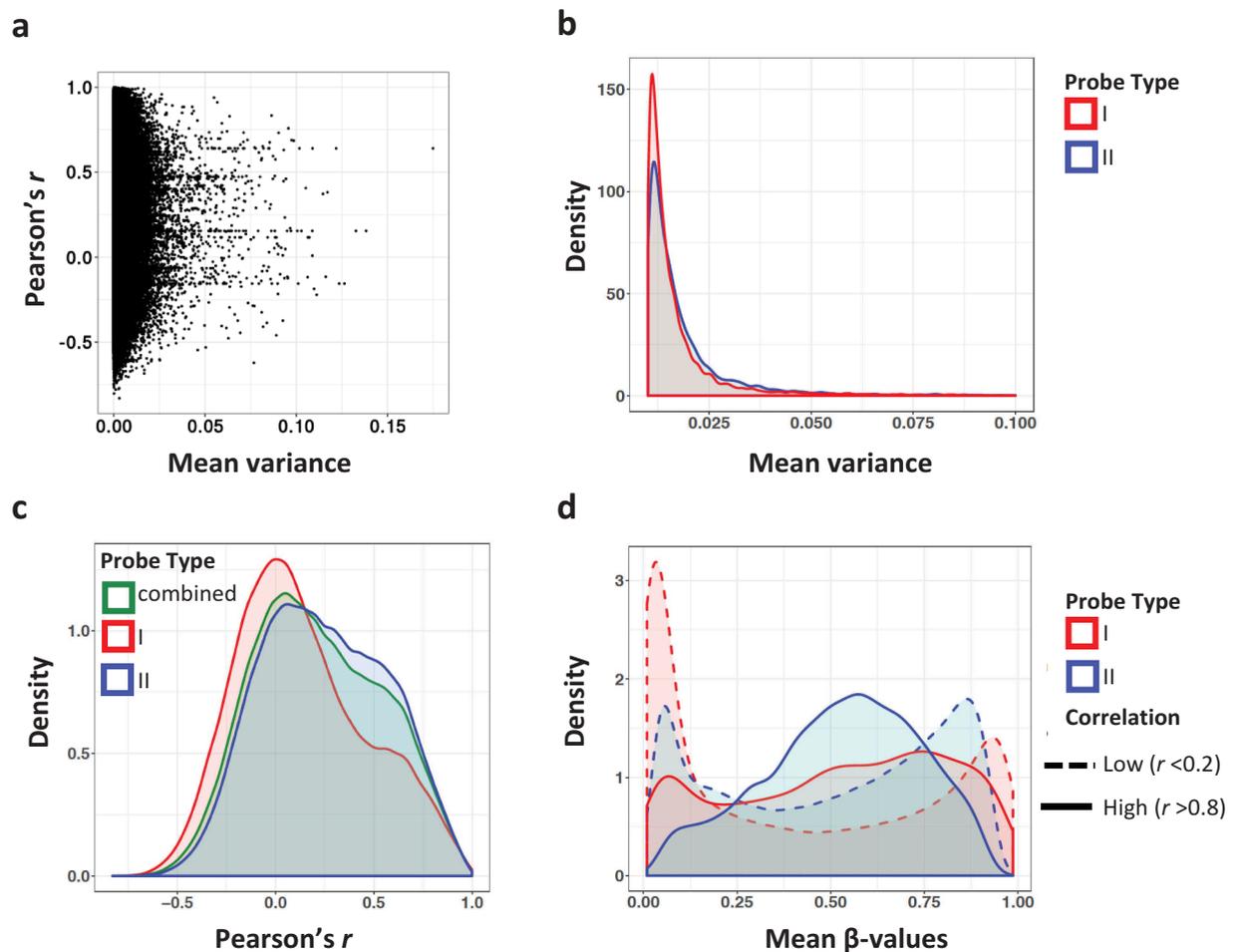


Figure 2. Correlation and variance of methylation β values by probe type. (a) Methylation (β value) variance plotted against the Pearson correlation coefficient across common CpG sites in 450K and EPIC arrays. Variance is the mean of the 450K and EPIC variances across the 21 samples. (b) Density of β value variance by probe type. Type I probes are shown in red and type II probes are in blue. (c) Density of Pearson correlation coefficient r by probe type. Type I probes are red, type II probes are in blue, and the combined Pearson coefficient of both probe types are in green. (d) Mean methylation β values by probe type and correlation group. CpG sites with low correlation ($r < 0.2$) are denoted by a dashed line, whereas high correlation ($r > 0.8$) is marked with solid lines.

We sought to determine whether there were differences in CpG variance and correlation assayed by probe types in cartilage, as has been reported in blood [3]. Solomon et al. [4] found that type II probes had a larger proportion of high correlation probes in comparison to type I. The 450K array consisted of 28% Infinium type I probes, whereas the EPIC array consists of 16% type I probes. We observed that CpG sites assayed using Infinium type I probes had significantly lower variance compared to sites measured with type II probes (Figure 2b; $p < 2.2 \times 10^{-16}$, Fligner–Killeen test). Furthermore, there was a trend for type I probes to show lower CpG correlations compared to type II (Figure 2c).

It has previously been observed that CpG sites with methylation levels at or close to extremes (either 0% or 100% methylation) were more prone to low variance and poor reproducibility [2,3]. We determined the mean methylation β -values per CpG in four groups: probe type I and high correlation, probe type I and low correlation, probe type II and high correlation, and probe type II and low correlation (Figure 2d). In the low correlation groups, regardless of probe chemistry, methylation β levels were indeed skewed towards 0% or 100%. This skew is more pronounced in type I probes, although the same trend is also seen in type II probes.

Differences in variance and correlation between tissue types

To examine if there are differences in DNA methylation variance and array β value correlations between tissue types, we compared our cartilage correlations with those of previously published 450K vs EPIC datasets. The raw data from matched 450K and EPIC samples collected from fresh-frozen paediatric brain tumours and whole blood from Guthrie cards were accessed from GEO GSE92580 and GSE86833, respectively. Data were filtered and processed in the same way as our cartilage datasets.

The Guthrie card whole blood dataset showed similar median ($r = 0.285$) and mean ($r = 0.221$) individual CpG correlations to cartilage (Supplemental Table 1A) and other whole blood comparisons [3,4], with only 16.68% of probes having Pearson correlations above 0.8 (Figure 3a, Supplemental Table 1A). Interestingly, in contrast to whole blood and cartilage, the majority (58.6%) of individual CpG sites within the tumour dataset exhibited high correlation ($r \geq 0.8$) between the two arrays, with a median Pearson's r of 0.888 and a mean of 0.695 (Figure 3a-b, Supplemental Table 1A). We further observed that the variance of individual CpG sites in tumour was higher than in whole blood and cartilage (Figure 3c). Together, our data support the hypothesis that low methylation variance sites tend to have poorly correlating β values between the two arrays, a pattern we observe across both probe types and across different tissue types.

Given the difference in Pearson r distribution and mean β variance between cartilage, tumour, and whole blood, we investigated the degree of overlap of high ($r \geq 0.8$) and poorly correlating probes ($r < 0.2$) in the three tissues (Figure 3d, Supplemental Table 1B). For a specific CpG site, DNA methylation variance is different between the three tissue types analysed, such that CpG sites with low 450K–EPIC correlations are different between the tissue types (Supplemental Table 1B). Of 261,995 CpG sites that show low ($r < 0.2$) β value correlations between the two arrays in at least one tissue, only 8.19% were common between them; this rises to 61% when tumour data are excluded (Figure 3d). As low correlation

sites are largely distinct between these tissue types, researchers should not extrapolate data on 450K–EPIC probe correlations from one tissue to inform another study using different tissue samples. This is further confirmed by examining the Pearson's r distribution in cartilage, tumour, and blood of 2668 probes recently reported to show poor correlations ($r \leq 0.2$) between arrays in 108 placental samples [5]. Although more than half of these low reproducibility placental probes also poorly correlate across the 450K and EPIC arrays in cartilage and blood samples, the majority (56.07%) show Pearson correlations of ≥ 0.8 in tumour (Figure 3e, Supplemental Table 1C).

Comparison with pyrosequencing

To investigate whether the 450K or EPIC array gave more accurate measurements, we assayed DNA methylation of six CpG sites that are present on both arrays using bisulphite pyrosequencing. Pyrosequencing is a technique that uses sequencing-by-synthesis to determine nucleotide polymorphisms in DNA and can be used to quantify DNA methylation after bisulphite conversion of DNA [7]. The six CpGs were selected to reflect the range of β value correlations observed in cartilage; two with negative correlations and very different β values between the two arrays (Supplemental Fig. S1), three with intermediate r values ($r = 0.420$ to 0.711), and one with high correlation (cg13782176, $r = 0.976$). After confirming the accuracy of pyrosequencing DNA methylation values using DNA of known methylation levels, we measured the same 21 cartilage DNA samples that we had matched 450K and EPIC data for (Supplemental Fig. S1 and S2). Our limited pyrosequencing validation was inconclusive for deducing which array platform generated more accurate methylation values; for the two CpGs that were non concordant between 450K and EPIC arrays, the pyrosequencing methylation values for one (cg24555089; Fig. S1A) was more similar to the 450K values and the other (cg20806296; Fig. S1B) was more similar to the EPIC values. CpG sites with similar methylation β values between 450K and EPIC also yielded similar methylation values measured by pyrosequencing (Fig. S1C). Prior studies examining the concordance between the 450K

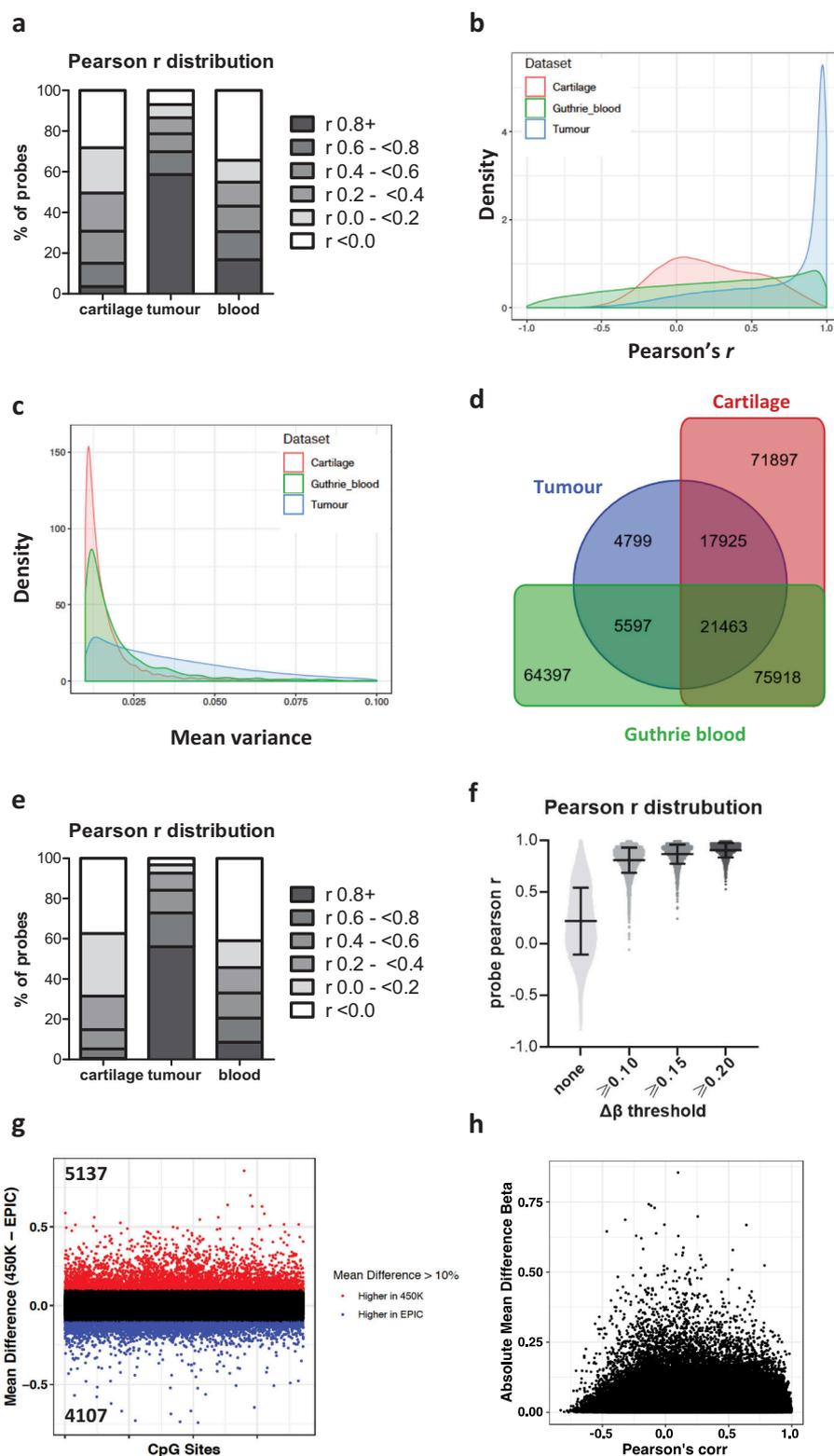


Figure 3. Pearson correlation and β value variance of CpG sites across 450K and EPIC arrays in three tissue types. Common samples assayed on both 450K and EPIC arrays were investigated in whole blood (GEO GSE86833) and paediatric brain tumour (GEO GSE92580). (a) Percentage of probes grouped by Pearson's r values. (b) Density plot of Pearson's r of CpG sites in cartilage (red), whole blood (green), and whole tumour (blue). (c) Density plot of mean variance (450K and EPIC) in cartilage, whole blood, and whole tumour. (d) Venn diagram of low correlation ($r < 0.2$) sites in cartilage, tumour, and whole blood. (e) The grouped distribution of Pearson's r values in cartilage, tumour, and whole blood of low correlating probes ($r \leq 0.2$) identified in 108 placental samples by Fernandez-Jimenez and colleagues (ref [5]). (f) Pearson r distribution of differentially methylated CpGs between OA knee and NOF cartilage samples plotted against $\Delta\beta$ thresholds. The cartilage r distribution for all probes is shown on the left for comparison. (g) Mean β differences per CpG between 450K and EPIC arrays. Sites with higher mean β values in the 450K array are highlighted in red and those higher in the EPIC array are highlighted in blue. (h) Per CpG Pearson's r against absolute mean difference β values between 450K and EPIC arrays.

array and pyrosequencing have found that, whilst showing a good overall correlation, measurements of some CpG sites do not replicate between the two methods [9].

For epigenome-wide association studies (EWASs), the ability to combined 450K- and EPIC-generated DNA datasets to perform meta-analyses is of great interest. The majority of EWASs use both a $\Delta\beta$ threshold and p value threshold for identifying differentially methylated sites (DMSs), although the $\Delta\beta$ thresholds vary widely between studies. We performed differential methylation analysis of 12 NOF cartilage samples and 9 OA knee samples in order to examine if changes in $\Delta\beta$ threshold altered the probe Pearson's r distribution between the 450K and EPIC arrays (figure 3f, Supplemental Table 1D). For DMSs identified by both arrays, the mean and median Pearson's r increases with increasing $\Delta\beta$ threshold (figure 3f), presumably due to the selection of probes increasingly variable methylation levels. Furthermore, the percentage of highly correlating probes ($r \geq 0.8$) rising from 61.90% to 90.78% when the $\Delta\beta$ threshold is raised from 0.1 to 0.2. Thus, when performing a meta-analysis of data generated by the two arrays, the correlation between the datasets will increase with increasing $\Delta\beta$ threshold used for identifying DMSs.

Absolute beta differences

We assessed the absolute mean difference in β values per CpG between the 450K and EPIC arrays and found that 9244 CpG sites (Supplementary Table 3) showed greater than 10% absolute mean β difference between arrays. Of these 9244 sites, there were 5137 CpG sites showing higher values in the 450K array and 4107 probes with higher values on the EPIC array (Figure 3g). There was a trend towards sites exhibiting poor correlation to show larger absolute mean β differences (Figure 3h). However, similar to other findings, sites with large differences were not necessarily those exhibiting low variance [4,5]. There were 4321 probes with poor correlation ($r < 0.2$) and an absolute mean β difference $>10\%$ compared to 327 probes with high correlation ($r > 0.8$) and a mean β difference $>10\%$. The 4321 probes with poor correlation and $>10\%$ beta difference between arrays had a variance of 0.0082, whereas the 327 CpG sites with high correlation and $>10\%$ beta difference had a higher variance of 0.0136. This suggests that

variance does not explain all of the differences seen between CpG sites measured on Illumina arrays. We advise excluding the 9244 CpG sites with an absolute mean β difference $>10\%$ from meta-analyses of DNA methylation in cartilage tissue measured using the 450K and EPIC arrays.

Discussion

We have shown that whilst overall correlation is high between the same cartilage samples assayed on 450K and EPIC arrays, some individual CpG sites show poor correlation. We found that type I probes are more prone to low correlation than type II. Type I probes are more effective than type II probes at measuring the extremes of methylation (0% or 100%), but it is at these extremes that CpG sites show poor correlation. Studies in whole blood also reported poor individual CpG site correlations [3,4]. These blood studies used different normalization procedures, suggesting that this artefact is present regardless of pre-processing pipeline. Furthermore, this study encompasses fewer samples compared to previously published comparisons and the conclusions remain the same.

Comparisons across tissues reveal that this trend of poor correlation in low variance sites is likely due to the inherent variability of DNA methylation in tissue or cell types rather than the analysis method. Cartilage tissue has a lower DNA methylation variance in comparison to whole tumour and whole blood. This could be in part due to the composition of cell types within these tissues; whole blood comprises numerous cell types and tumour tissue consists of normal and abnormal cells, factors which contribute to heterogeneity in DNA methylation [10–12]. A study comparing 450K and EPIC in placental tissue, a tissue comprising multiple cell types, found that per CpG correlations were high [5]. CpG sites with poor correlation vary between tissue types, and therefore, the specification of a common list of CpGs to exclude would not be applicable. Researchers studying DNA methylation in tissues or cell types exhibiting low variance or extreme methylation values should be aware that some CpG sites may not be reproducible.

The 450K array is no longer available; researchers wishing to assay methylation using BeadChip technology only have the option of the EPIC array. Although the 450K array has been discontinued by Illumina, the array may still have relevance in ongoing studies and meta-analyses. The comparisons between the 450K and EPIC have revealed an interesting technical artefact in measuring the methylation of low variance sites with either 0% or 100% methylation levels. However, it is not the first time that poorly correlated CpG sites have been described. Comparisons between the 450K and its predecessor, the Infinium HumanMethylation27 BeadChip, showed strong overall correlations but weak correlation of some CpG sites [1,13]. Past studies examining the similarity of measurements in 450K and pyrosequencing in cell lines and brain tissue [9,14] have also observed that some CpG sites do not reproduce. Moreover, comparisons of 450K to whole-genome bisulphite sequencing showed the same conclusions [1,15,16] as did reduced representation bisulphite sequencing [17,18]. Multiple published manuscripts have described the correlation of CpG sites measured using different methods. These studies were performed in various cell lines and tissue types. All report a high overall correlation between DNA methylation of the same samples measured using different methods. Not all of these studies commented on individual CpG site correlation; however, manual inspection of the correlation scatter plots reveals that a subset of CpGs shows large differences in methylation values between the platforms being compared [1,15–18]. Altogether, these studies demonstrate that some CpGs reproduce poorly, regardless of sample type or method used.

It is not often feasible for researchers to validate CpGs genome wide using a different method and compare between methods in order to find sites with poor correlation. However, we and others have shown that CpG sites with poor correlation also tend, but not always, to be those with low variance and very high or very low methylation. Therefore, it may be beneficial to apply a filter based on variance and/or percentage methylation, especially where the methylation level of individual CpGs is important for downstream analysis or if DNA methylation in the tissue type being studied is relatively homogeneous. For example,

studies investigating the effect of single CpG sites such as methylation quantitative trait loci (mQTL) or EWASs may wish to apply additional filters. Whilst researchers may wish to validate potentially important CpG sites using a different assay, some sites may not reliably reproduce regardless of method. Additionally, it can be difficult to reliably deduce which method shows the correct methylation level.

To conclude, we have corroborated past studies comparing the reproducibility between the 450K and EPIC platforms. We found that whilst per sample correlation was high, correlation between individual CpGs tended to be low in cartilage. We also report that CpG sites with poor correlation differ between tumour, whole blood, and cartilage, and this could be due to differences in DNA methylation variance between tissues.

Methods

Collection of cartilage samples and DNA methylation analysis

Cartilage samples were obtained from patients with knee OA ($n = 9$) or NOF ($n = 12$) undergoing total joint replacement surgery at the Newcastle upon Tyne NHS Foundation Trust hospitals. The Newcastle and North Tyneside Research Ethics Committee granted ethical approval for the collection of the samples (REC reference number 14/NE/1212), with each donor providing informed consent. Cartilage was ground as previously described (Raine et al. 2012 PMID 22586168) and DNA extracted using the E.Z.N.A Tissue DNA kit (omega BIO-TEK). One microgram of DNA was bisulphite converted using the EZ DNA methylation kit (Zymo Research) and quantified using the 450K and EPIC arrays by Edinburgh Clinical Research Facility.

Analysis of 450K and EPIC data

Data were processed using the Bioconductor package *minfi* [19]. Samples were normalized using the functional normalization method implemented in the *minfi* package. All samples had >95% probes passing a detection p value of 0.01. Probes with a detection p value <0.01 were removed from the analysis. Probes within sex chromosomes were

removed from the analysis, as were probes previously reported to be cross-reacting [20] and probes containing single nucleotide polymorphisms (SNPs) with a minimum allele frequency >0.05 in European populations. Only probes common to the 450K and EPIC array were selected, this left 372,278 probes (96,514 type I and 275,764 type II) for analysis. Age and sex covariates of cartilage samples were corrected for using the `removeBatchEffect` function in *limma* [21]. Correction was performed using methylation M-values before conversion to methylation β values. Pearson correlation coefficients were calculated per sample and per CpG site across both 450K and EPIC arrays. Datasets GEO GSE92580 (paediatric brain tumour, $n = 12$) and GSE86833 (infant whole blood, $n = 5$) were analysed using exactly the same method. Correlation and density plots were generated using the `ggplot2` package in R. An online webtool was used to create the Venn diagram (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Targeted pyrosequencing

Genomic DNA (750 ng) of the same 21 cartilage samples was bisulphate converted as above and eluted in 40 μ l elution buffer. Pyrosequencing assays for the six selected CpGs were designed using PyroMark assay design software 2.0 (Qiagen); the primers used are listed in Supplementary Table S2. Pyrosequencing was performed using PyroMark Q24 (Qiagen). A calibration curve was made to confirm the accuracy of the pyrosequencer. For each CpG, two gBlocks (IDT) of the 300 bp region surrounding the CpG site of interest were designed to the bisulphite sequence with either the CpG site methylated (C) or unmethylated (T); then, the pyrosequencing calibration curve was performed using input DNA of 0–100% methylation with 10% increments. The calibration curve for cg1925474, cg13979708, and cg13782176 deviated from the expected value, and therefore, the methylation values generated by pyrosequencing were corrected using the equation $x = (y-b)/a$, where x is the corrected level of methylation, y is the level of methylation measured by pyrosequencing, and a is the slope and b is the intercept of the calibration curve.

Authors' contributions

LNR, DAY, SC, and KC designed the study. LNR and MJB collected the data. KC performed the data analysis. All authors read and approved the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

KC is funded by the Medical Research Council (MRC) and Versus Arthritis as part of the MRC-Arthritis Research UK Centre for Integrated Research into Musculoskeletal Ageing (CIMA, grant references JXR [10641 and MR/P020941/1]) and MB by a PhD studentship funded by the Community Foundation.

ORCID

Kathleen Cheung  <http://orcid.org/0000-0002-6449-7517>

David A. Young  <http://orcid.org/0000-0002-7078-6745>

References

- [1] Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98:288–295.
- [2] Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res*. 2016;45:gw967.
- [3] Logue MW, Smith AK, Wolf EJ, et al. The correlation of methylation levels measured using Illumina 450K and EPIC BeadChips in blood samples. *Epigenomics*. 2017;9:1363–1371.
- [4] Solomon O, MacIsaac J, Quach H, et al. Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. *Epigenetics*. 2018;13:655–664.
- [5] Fernandez-Jimenez N, Allard C, Bouchard L, et al. Comparison of Illumina 450K and EPIC arrays in placental DNA methylation. *Epigenetics*. 2019. 14:1177–1182.
- [6] 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;17:208.
- [7] Fortin J-P, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol*. 2014;15:503.
- [8] 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;9:33.
- [9] Roessler J, Ammerpohl O, Gutwein J, et al. Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. *BMC Res Notes*. 2012;5:210.
- [10] Adalsteinsson BT, Gudnason H, Aspelund T, et al. Heterogeneity in white blood cells has potential to confound DNA methylation measurements. *PLoS One*. 2012;7:e46705.
- [11] Quek K, Li J, Estecio M, et al. DNA methylation intratumor heterogeneity in localized lung adenocarcinomas. *Oncotarget*. 2017;8:21994–22002.
- [12] Mikeska T, Candiloro IL, Dobrovic A. The implications of heterogeneous DNA methylation for the accurate quantification of methylation. *Epigenomics*. 2010;2:561–573.
- [13] Sandoval J, Heyn H, Moran S, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*. 2011;6:692–702.
- [14] Bock C, Halbritter F, Carmona FJ, et al. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. *Nat Biotechnol*. 2016;34:726–737.
- [15] Teh AL, Pan H, Lin X, et al. Comparison of Methyl-capture sequencing vs. Infinium 450K methylation array for methylome analysis in clinical samples. *Epigenetics*. 2016;11:36–48.
- [16] Thurner M, van de Bunt M, Torres JM, et al. Integration of human pancreatic islet genomic data refines regulatory mechanisms at type 2 diabetes susceptibility loci. *Elife*. 2018;7:7.
- [17] Tang B, Zhou Y, Wang C-M, et al. Integration of DNA methylation and gene transcription across nineteen cell types reveals cell type-specific and genomic region-dependent regulatory patterns. *Sci Rep*. 2017;7:3626.
- [18] Pan H, Chen L, Dogra S, et al. Measuring the methylome in clinical samples: improved processing of the Infinium Human Methylation450 BeadChip array. *Epigenetics*. 2012;7:1173–1187.
- [19] 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;30:1363–1369.
- [20] Chen Y, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8:203–209.
- [21] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47–e47.