

# A comparative analysis of cell-type adjustment methods for epigenome-wide association studies based on simulated and real data sets

Johannes Brägelmann and Justo Lorenzo Bermejo

Corresponding authors: Johannes Brägelmann, Institute of Medical Biometry and Informatics, University Hospital Heidelberg, Heidelberg, Germany. Tel.: +49 221 478 96861; E-mail: johannes.braegelmann@uni-koeln.de; Justo Lorenzo Bermejo, Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany. Tel.: +49 6221 56 4180; E-mail: lorenzo@imbi.uni-heidelberg.de

## Abstract

Technological advances and reduced costs of high-density methylation arrays have led to an increasing number of association studies on the possible relationship between human disease and epigenetic variability. DNA samples from peripheral blood or other tissue types are analyzed in epigenome-wide association studies (EWAS) to detect methylation differences related to a particular phenotype. Since information on the cell-type composition of the sample is generally not available and methylation profiles are cell-type specific, statistical methods have been developed for adjustment of cell-type heterogeneity in EWAS.

In this study we systematically compared five popular adjustment methods: the factored spectrally transformed linear mixed model (FaST-LMM-EWASher), the sparse principal component analysis algorithm ReFACTor, surrogate variable analysis (SVA), independent SVA (ISVA) and an optimized version of SVA (SmartSVA). We used real data and applied a multilayered simulation framework to assess the type I error rate, the statistical power and the quality of estimated methylation differences according to major study characteristics.

While all five adjustment methods improved false-positive rates compared with unadjusted analyses, FaST-LMM-EWASher resulted in the lowest type I error rate at the expense of low statistical power. SVA efficiently corrected for cell-type heterogeneity in EWAS up to 200 cases and 200 controls, but did not control type I error rates in larger studies. Results based on real data sets confirmed simulation findings with the strongest control of type I error rates by FaST-LMM-EWASher and SmartSVA. Overall, ReFACTor, ISVA and SmartSVA showed the best comparable statistical power, quality of estimated methylation differences and runtime.

**Key words:** epigenome-wide association studies; methylation; cell-type adjustment methods; simulation study; epigenetics

## Introduction

Epigenetic modifications such as DNA methylation are major factors governing gene transcription [1, 2]. Since all nucleated

cells in an organism contain identical DNA sequences, regulation of gene expression by additional mechanisms is necessary for the temporal control of highly specialized cell types in multicellular organisms such as humans. Epigenetic

Johannes Brägelmann is a post-doctoral fellow at the University Hospital of Cologne, Germany, with a degree in medical biometry and biostatistics from the University of Heidelberg, Germany, investigating genomic and epigenomic alterations in cancer.

Justo Lorenzo Bermejo is a professor and head of the Statistical Genetics Group at the Institute of Medical Biometry and Informatics, University of Heidelberg, Germany, focusing on the development and application of robust methods in statistical genetics and genetic epidemiology.

Submitted: 4 April 2018; Received (in revised form): 11 June 2018

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

regulation of gene expression, e.g. DNA methylation, is crucial for tissue development and during cellular differentiation. Its alteration may be involved in diseases such as cancer and cardiovascular disease [3–5]. In contrast to heritable genetic variants that alter the DNA sequence, epigenetic marks can be added and removed to reshape cellular gene expression profiles in response to external stimuli such as toxins and pathogens, and during disease progression [2, 6]. DNA methylation represents a dynamic adaptation mechanism governed by an intricate interplay between intrinsic (e.g. genetic) and external (e.g. environmental) factors that may be causal for certain phenotypes and diseases. In addition, changes in DNA methylation can originate during the course of disease development even years before disease onset in a noncausal, associative manner and may thus constitute prognostic and predictive biomarkers [4, 6–8].

The technological development and price reduction of high-density oligonucleotide arrays that simultaneously measure DNA methylation at more than 450 000 positions across the genome have sparked increasing interest in the evaluation of epigenetic changes associated with human disease [9]. Epigenome-wide association studies (EWAS) examine methylation levels across the whole genome according to particular phenotypes (e.g. cancer patients versus unaffected individuals) [5]. One of the major challenges in EWAS based on peripheral blood samples or other tissue types is that biological samples usually contain a mixture of different cell types. Given the role of methylation profiles in cell-lineage differentiation, cell-type heterogeneity among study participants can strongly confound association results, especially when the cell-type composition is associated with the phenotype of interest. Laboratory information on cell-type composition is generally sparse and, in contrast to the investigation of heritable genetic variants, cell-type heterogeneity is a crucial factor that needs to be adjusted for. Nonconsideration of cell-type heterogeneity has probably led to spurious associations in early EWAS, and several adjustment methods have been developed [3, 10].

### Adjustment of cell-type heterogeneity in EWAS

When the cell-type composition of investigated EWAS samples is known, cell-type proportions can easily be integrated into regression models as adjustment covariates. However, information on cell-type composition is usually sparse, and retrieval of high-quality data is technically demanding and expensive. For example, the agreement between pathologists' scores for immune-cell staining tends to be poor [11]. DNA sequencing data may be used for the estimation of tumor purity as well as the number and fractions of tumor cell subpopulations, but still does not allow distinction of functional subgroups among noncancer cells such as fibroblasts and immune cells [12]. In the case of EWAS based on peripheral blood samples, whole blood cell counts could be used for cell-type adjustment, but in practice this information is frequently missing and available counts only apply to a limited number of cell types [13].

To circumvent the limited availability of whole blood cell counts, Houseman *et al.* [14] proposed to infer this information from an external data set that includes methylation data for specific types of blood cells. This adjustment method was quite popular in early EWAS, but it relies heavily on the assumption that all relevant cell types are known and well represented in the external methylation data set [15–17]. However, some cell types needed for adjustment are still unknown and probably depend on investigated phenotypes. In fact, a particular

subtype of lymphocytes not represented in Houseman's data set has been suggested to be responsible for false-positive EWAS results [18].

### Reference-free methods for cell-type heterogeneity adjustment

To overcome the aforementioned limitations several approaches that rely neither on measured cell-type compositions nor on external references have been developed [19–24]. Reference-free methods are expected to correct study results without prespecification of a particular set of cell types, which offers additional flexibility. Popular approaches specifically developed for EWAS are factored spectrally transformed linear mixed model (FaST-LMM-EWASher), ReFACTor and an optimized version of surrogate variable analysis (SVA) (SmartSVA) [21–23]. More general methods like SVA [20], which was originally developed for correction of batch effects in gene expression studies, and independent SVA (ISVA) [24] are also used for EWAS adjustment. In two recent publications SVA has even been suggested to be the best overall reference-free adjustment method [25, 26].

FaST-LMM-EWASher builds on a factored spectrally transformed linear mixed model algorithm initially applied to correct for population stratification in genome-wide association studies [22, 27]. In FaST-LMM a genetic similarity matrix is estimated relying on measured methylation values. This random effect covariance matrix is then combined with fixed-effect factors such as gender and age in a linear mixed model to capture—and adjust for—the possible relatedness/dependence among individuals [28]. To account for cell-type composition in EWAS the approach was further extended, and only the most strongly correlated methylation markers based on a principal component analysis are used for adjustment of cell-type heterogeneity. If genomic inflation is still present, the top principal components (PCs) are consecutively added as covariates [22].

ReFACTor was specifically developed by Rahmani *et al.* [21] to adjust for cell-type confounding in EWAS. As in FaST-LMM-EWASher, the correction of cell-type heterogeneity is unsupervised, i.e. the phenotype of interest is not considered in the adjustment. Basic ReFACTor assumptions are that only a small subset of  $t$  methylation markers is associated with cell-type composition and that measured  $\beta$ -methylation values in an individual are the result of a weighted sum of average methylation in  $k$  different cell types with weights equaling the individual cell-type proportions. Using the matrix  $O$  of observed/measured  $\beta$ -methylation values across individuals, ReFACTor calculates a  $k$ -rank approximation matrix  $O'$ , which retains the  $t$  most informative markers. The top  $k$  PCs from  $O'$  are then used as covariates in association tests to adjust for cell-type heterogeneity.

Unlike FaST-LMM-EWASher and ReFACTor, SVA was originally designed to deal with confounding by unmeasured factors such as environmental and batch effects in gene-expression data sets [20]. SVA is a supervised adjustment method. First, a linear model with the phenotype is fitted to extract the 'signal of interest' from the measured methylation matrix. Subsequently, a singular value decomposition of the residual matrix is used to construct surrogate variables in an iterative process that captures methylation variability attributable to unmeasured confounders. The surrogate variables are then used to extend the initial linear model and calculate adjusted effects and corresponding probability values.

ISVA and SmartSVA build upon the original SVA algorithm. ISVA relies on independent component analysis instead of

singular value decomposition of the residual matrix to derive the surrogate variables. Compared with SVA, ISVA identifies latent variables that are neither linearly nor nonlinearly correlated and thus statistically independent [24]. SmartSVA aims to improve SVA in the common situation where the primary variable of interest and the confounders (in the present context cell-type composition) are correlated and SVA does not converge to a reliable solution during iterative construction of surrogate variables. SmartSVA imposes an explicit convergence criterion and predetermines the number of surrogate variables based on random matrix theory. In order to mitigate the effect of potential correlations between primary and latent variables, SmartSVA reweights the probability that methylation markers are related to the phenotype of interest conditional on unmodeled confounders [23].

Given the large impact of cell-type heterogeneity on the design and statistical analysis of EWAS, the present comparison of adjustment methods aims at providing practical recommendations on the main relevant issues in the field. We investigate first the parameters that influence the nature and magnitude of cell-type methylation confounding. To reach this objective, we develop and apply a multilayered simulation framework based on real methylation data, which allows controlled modifications of critical parameters. Then, we use simulated and real data sets to compare the reference-free adjustment methods FaST-LMM-EWASher, ReFACTor, SVA, ISVA and SmartSVA. We identify method-specific strengths and limitations and guide researchers in their choice of the most appropriate adjustment method.

## Materials and methods

### Simulation framework

Synthetic data sets offer the advantage that the ground truth is known and that relevant parameters can be adjusted deliberately to assess their impact on obtained results. We therefore developed a multilayered simulation design based on real blood cell-type-specific methylation levels with a generative model to create ‘measured/observed’ methylation  $\beta$  values ( $\beta_j \in [0, 1]$ ) in a mixture of  $K$  different blood cell types (Figure 1). Methylation was modeled as cell-type and marker-specific  $\beta_{jk} = \tau_{jk} + \epsilon_{jk}$  at methylation marker  $j$  in cell-type  $k = 1, \dots, K$  based on cell-type-specific average methylation  $\tau_{jk}$  and  $\epsilon_{jk} \sim N(0; \sigma_{jk}^2)$  with cell-type and marker-specific  $\sigma_{jk}^2$ . Observed methylation at  $j$  is the average cell-type methylation weighted by cell-type proportions:  $\beta_j = \frac{\sum_k p_k \beta_{jk}}{\sum_k p_k}$  (note that  $\sum_k p_k = 1$ ). Cell-type proportions  $p$  may differ systematically between cases and controls. A set  $S$  of markers was differentially methylated in cases and controls with cell-type  $c$ .

Publicly available methylation data from flow cytometry-sorted blood-cell types (CD14+ monocytes, CD19+ B cells, CD4+ helper T cells, CD56+ NK cells, CD8+ cytotoxic T cells, eosinophils and neutrophils) were used as a basis for simulation [29]. Raw data per cell type were background corrected, followed by functional normalization before  $\beta$  values were calculated; markers with single nucleotide polymorphisms (SNPs) or cross-reactivity were excluded [30, 31]. For computational reasons simulations were restricted to markers on chromosome 1 ( $j = 1, \dots, 42772$  methylation markers). Average methylation  $\tau_{jk}$  and its variation  $\sigma_{jk}$  were calculated separately for marker  $j$  in cell-type  $k$  across the reference subjects.

For simulation of phenotype-dependent cell-type proportions per individual, blood cell count data on neutrophils, monocytes, eosinophils and total lymphocytes from a cohort study involving colon cancer patients and healthy controls

were used [32]. Total lymphocytes were split into B-cell, CD4+ and CD8+ T-lymphocyte subpopulations according to [29], and average cell-type proportions for cases and controls were summarized in  $k \times 1$  vectors  $\alpha_{case}$  and  $\alpha_{control}$  with  $k=6$ . For  $n=1000$  individuals  $i$  (500 cases, 500 controls), cell-type composition  $p_i$  was generated using a Dirichlet distribution as  $\text{Dirichlet}(\rho\alpha_{case})$  or  $\text{Dirichlet}(\rho\alpha_{control})$  with  $\rho = 20$ .

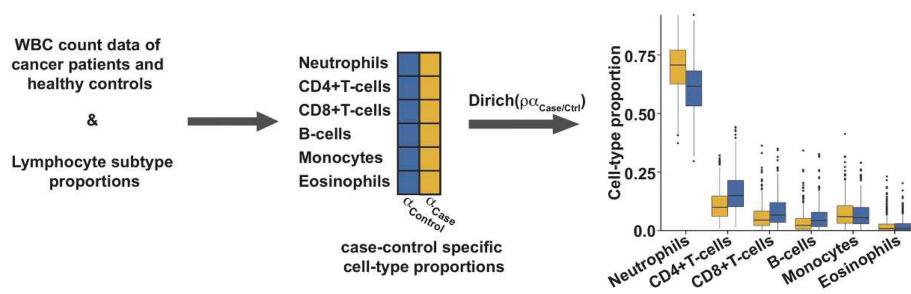
To test a range of scenarios, varying numbers of methylation markers  $S \in \{0, 5, 50, 500, 5000\}$  were simulated as differentially methylated in  $c = \text{neutrophils}$ . The methylation difference  $\delta$  between cases and controls was chosen to achieve 80% power in an unadjusted analysis of 200 cases and 200 controls ( $\delta = 0.0045$ ) and to mimic a strong biological signal ( $\delta = 0.1$ ). In addition, 25 scenarios ( $S=50$  markers with a differential methylation  $\delta \sim \text{Unif}(0.001, 0.01)$  between  $n=200$  cases and 200 controls) were simulated.  $\beta_{ijk}$  per cell type  $k$  and marker  $j$  were generated for each individual  $i \in 1, \dots, n$  as described above and combined as a cell-type-specific matrix  $B_k$ . For case subjects and  $k = \text{neutrophils}$ , the markers  $j$  among the  $S$  differentially methylated markers were altered as  $\beta'_{ijk} = \beta_{ij} - \delta_j$  with  $\delta_j = \delta$  if  $\beta_{ijk} > 0.5$  and  $\delta_j = -\delta$ , otherwise, to ensure that  $\beta'_{ijk} \in [0, 1]$ . Cell-type-specific  $\beta$  values were subsequently combined per individual as the weighted average according to cell-type proportions  $p_i$  before combining all individuals into one final ‘observed’  $\beta$ -value matrix for further processing.

### Real methylation data sets

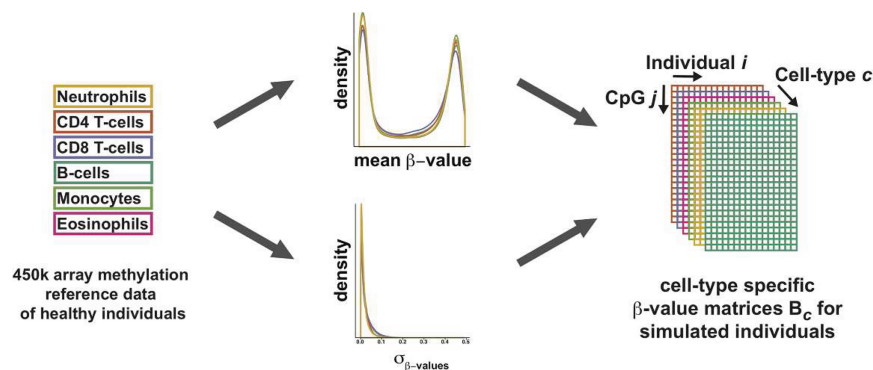
In addition to simulated data, nine real methylation data sets were used to benchmark the adjustment methods. They comprised the exemplary data with methylation measurements in tumor and normal breast tissue samples provided by Zou et al. [22] with FaST-LMM-EWASher ([www.microsoft.com/en-us/download/details.aspx?id=52501](http://www.microsoft.com/en-us/download/details.aspx?id=52501)) and eight studies retrieved from Gene Expression Omnibus (GEO) ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)). Seven data sets contained blood methylation measurements: Liu et al. [17] investigated patients with rheumatoid arthritis (RA) and healthy controls (GSE42861); Tsaprouni et al. [33] selected current, former and never smokers, here grouped as current smokers versus nonsmokers (GSE50660); and Veldhoven et al. [34] prospectively investigated women who were disease-free at the time of blood collection and compared methylation profiles of women who subsequently developed breast cancer compared with those who remained healthy (GSE51057). Heyn et al. [35] examined the methylation profiles in cord blood from newborns and peripheral blood from adults older than 90 years (GSE30870). Two samples with sorted blood cell subtypes were included to increase overall variability. Hannon et al. [36] investigated schizophrenic patients versus non-psychiatric controls (GSE80417, phase I cohort), while Chen et al. [23] examined peripheral blood of patients with congenital hypopituitarism and age-matched controls (GSE107737). Philibert and colleagues compared the methylation profiles of drinkers versus non-drinkers (GSE110043). To assess the performance of adjustment methods in tissue methylation data, we included an additional data set with breast cancer and normal samples generated by Stefansson et al. [37] (GSE52865).

To allow comparison with simulated scenarios and to reduce the computing time and memory requirements, the methylation data sets were restricted to markers on chromosome 1, with the exception of the FaST-LMM-EWASher methylation data set, which contains only 22 690 markers. Cross-reactive CpG markers, markers that contained SNPs and markers with missing methylation data were excluded prior to downstream analysis [31].

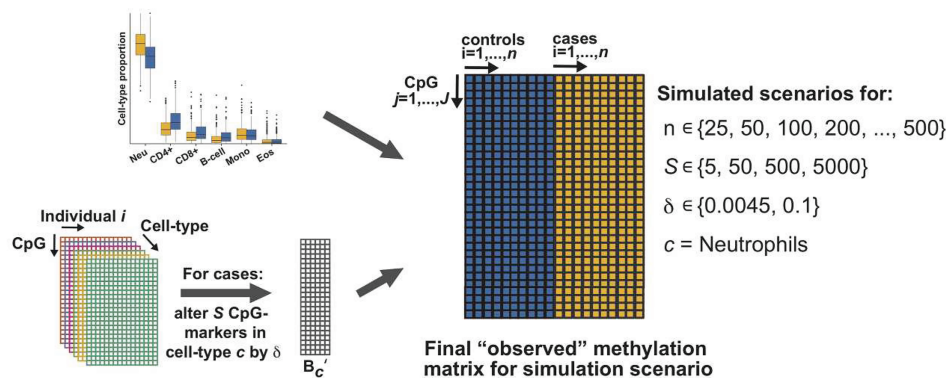
## 1) Simulation of cell-type proportions



## 2) Location and dispersion of $\beta$ -values



## 3) Simulation of study scenarios



**Figure 1.** Simulation of EWAS data sets. (1) Blood cell-type proportions of cases and controls for six cell type were simulated using Dirichlet distributions based on observed cell-type compositions of cancer patients and healthy controls. (2) Public flow-sorted blood cell-type methylation data of healthy individuals were used to simulate matrices of cell-type-specific  $\beta$ -values. (3) The simulated cell-type proportions and  $\beta$ -values per individual were combined to generate a matrix of 'observed' methylation values. In case subjects,  $S$  CpG markers in neutrophils were altered by  $\delta$  before generating the final methylation matrix. Different scenarios were created using varying values for individuals per group  $n$ , number of differentially methylated CpG markers  $S$  and the  $\delta$  between cases and controls.

### Association analysis and cell-type adjustment

To assess the impact of simulated cell-type confounding and to compare the performance of the cell-type adjustment methods, unadjusted results were compared with results after adjustment. The linear model  $\beta_j = a_j + Xb_j + e_j$  was fitted for unadjusted analyses, where the methylation  $\beta$  at marker  $j$  depended on the intercept  $a_j$ ,  $X$  was equal to 1 for cases and 0 for controls,  $b_j$  was used to estimate the methylation difference in cases versus controls and  $e_j$  was a random error term. No additional clinical covariates were considered.

The adjustment of cell-type heterogeneity with FaST-LMM-EWASher was conducted using the R version after applying the

patch described by McGregor et al. [25]. Per default, FaST-LMM-EWASher excludes markers with average methylation values outside the interval [0.2, 0.8]. This filter excludes the majority of markers in an average methylation data set and may be one of the causes of previously reported conservative results after FaST-LMM-EWASher adjustment [25]. This filter was therefore not applied in the present comparison. Apart from this exception, FaST-LMM-EWASher was run using default parameters including a maximum of 10 PCs for adjustment. ISVA was run with default parameters, with the exception of 'fastICA' for independent component analysis. ReFACTOR, SVA and SmartSVA were run using default parameters.

**Table 1.** Average type I error rates with the corresponding 95% intervals for simulated null scenarios after different cell-type adjustment methods

Individuals per group	Cell-type adjustment method													
	Unadjusted		FaST-LMM-EWASher		ReFACTor		SVA		ISVA		SmartSVA			
	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI		
25	0.182	0.170 0.194	0.022	0.015 0.033	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	0.053	0.050 0.055	0.051	0.050 0.051
50	0.222	0.188 0.257	0.019	0.011 0.026	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	<b>0.050</b>	<b>0.049</b>	<b>0.051</b>	0.051	0.050 0.051	<b>0.050</b>	<b>0.049</b> <b>0.051</b>
100	0.380	0.360 0.398	0.018	0.010 0.025	0.049	0.048	0.049	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	<b>0.05</b>	<b>0.049</b>	<b>0.050</b>	<b>0.050</b> <b>0.049</b> <b>0.051</b>
200	0.484	0.481 0.487	<b>0.045</b>	<b>0.041</b> <b>0.050</b>	0.049	0.049	0.049	0.053	0.051	0.053	<b>0.05</b>	<b>0.049</b>	<b>0.051</b>	<b>0.050</b> <b>0.049</b> <b>0.050</b>
300	0.545	0.541 0.549	0.048	0.048 0.049	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	0.105	0.073	0.136	0.049	0.049	0.050	0.049 0.049 0.050
400	0.582	0.580 0.584	0.048	0.047 0.048	0.049	0.048	0.049	0.224	0.184	0.265	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	0.049 0.049 0.049
500	0.616	0.615 0.616	0.048	0.048 0.049	<b>0.049</b>	<b>0.049</b>	<b>0.050</b>	0.518	0.453	0.583	<b>0.05</b>	<b>0.049</b>	<b>0.052</b>	<b>0.049</b> <b>0.049</b> <b>0.050</b>

Bold type denotes 95% CI that includes 0.05

## Performance metrics

The impact of simulated cell confounding and the improvements in detection and estimation achieved by the adjustment methods were evaluated by the type I error rate, the statistical power and the genomic inflation factor (GIF), as well as the bias, variance and mean squared error (MSE) of estimated methylation differences. The type I error rate was calculated as

$$\text{Type I error rate} = \frac{\#p\text{-values} < \alpha}{\# \text{ of all tests}}$$

for a significance level of  $\alpha = 0.05$ . A type I error rate above 5% generally indicates too many false-positive findings due to confounding, whereas error rates under 5% are usually indicative of overcorrection. No correction for multiplicity was applied.

The statistical power was calculated as

$$\text{Power} = \frac{\# \text{ of diff. methylated markers with } p \text{ values} < \alpha}{\# \text{ of diff. methylated markers}}$$

The methylation difference  $\delta = 0.0045$  between cases and controls in the majority of simulated scenarios was chosen to reach 80% power in the analyses of neutrophils alone, thus resulting in an expected power of 60–80% in mixed-cells analyses.

The GIF was initially developed to quantify the inflation of genetic association tests due to population stratification [38–40]. It is defined as the ratio of empirically observed test statistic to its expected median [41] and was calculated here as  $\text{GIF} = \frac{\text{median observed test statistics}}{\text{median expected test statistic}}$ , with the median of the expected test statistic being the median of a  $\chi^2_{1df}$ -distribution with one degree of freedom equal to 0.455. In the absence of confounding the GIF approaches 1.0, whereas a  $\text{GIF} > 1.0$  indicates a systematic deviation from the null distribution that can be attributed to confounding.

## Results

### Simulated data sets without differential methylation

To assess the overall influence of cell-type heterogeneity on EWAS results we first investigated ‘null’ scenarios, i.e. scenarios without simulated differentially methylated markers in cases and controls. The type I error rate equals the false-positive rate under the null scenarios, and it should be close to the chosen  $\alpha$ -level when cell-type confounding has been perfectly adjusted for. The GIF due to stratification has been found to scale up

with increasing population size in GWAS [40, 41]. We therefore simulated studies of different sizes by sampling patients and controls with a 1:1 ratio with  $n \in \{25, 50, 100, 200, 300, 400, 500\}$  individuals per group. In unadjusted analyses, the type I error rate was above the nominal  $\alpha$  level for  $n = 25$  individuals and increased with increasing study size (Table 1). Across all study sizes, FaST-LMM-EWASher, ReFACTor, ISVA and SmartSVA corrected the type I error inflation close to the nominal level, with a tendency toward overcorrection by FaST-LMM-EWASher. The correction performance of SVA broke down for studies with more than 200 individuals per group. Similar results were found examining the GIF with increasing study sizes (Figure 2).

The limited adjustment ability of SVA for large studies could be related to the first step of this method, where the signal of interest is extracted from the matrix of methylation data taking into account the case–control phenotype. In the case of large studies with marked cell-type differences between cases and controls, the initially fitted linear model may remove an excess of methylation variability, resulting in insufficient posterior adjustment. To evaluate this hypothesis, we repeated the SVA adjustment for  $n = 500$  individuals per group using randomly permuted case–control information for the first step of the SVA correction. Interestingly, permuted phenotypes led to a controlled type I error rate and GIF (Supplementary Figure S1). This result indicates that in situations with a strong correlation between phenotype and cell-type composition, especially in the case of weak additional phenotype-dependent methylation signals, SVA may be of limited use.

We found considerable differences among the investigated adjustment methods regarding computation time. While runtime always increased with higher  $n$ , this effect was much stronger for SVA than for FaST-LMM-EWASher and ISVA and was hardly present for ReFACTor and SmartSVA, which showed a similar runtime to unadjusted analyses (Supplementary Figure S2).

### Simulated data sets with differential methylation, low $\delta$

In the first set of simulated alternative scenarios, the case–control difference in  $\beta$ -methylation values at the differentially methylated markers was fixed to  $\delta = 0.0045$  in neutrophils to achieve 80% power in studies with  $n = 200$  individuals per group. Different numbers of differentially methylated markers were considered ( $S \in \{5, 50, 500, 5000\}$ ). Overall, the statistical power to detect differential methylation was highest, close to the nominal

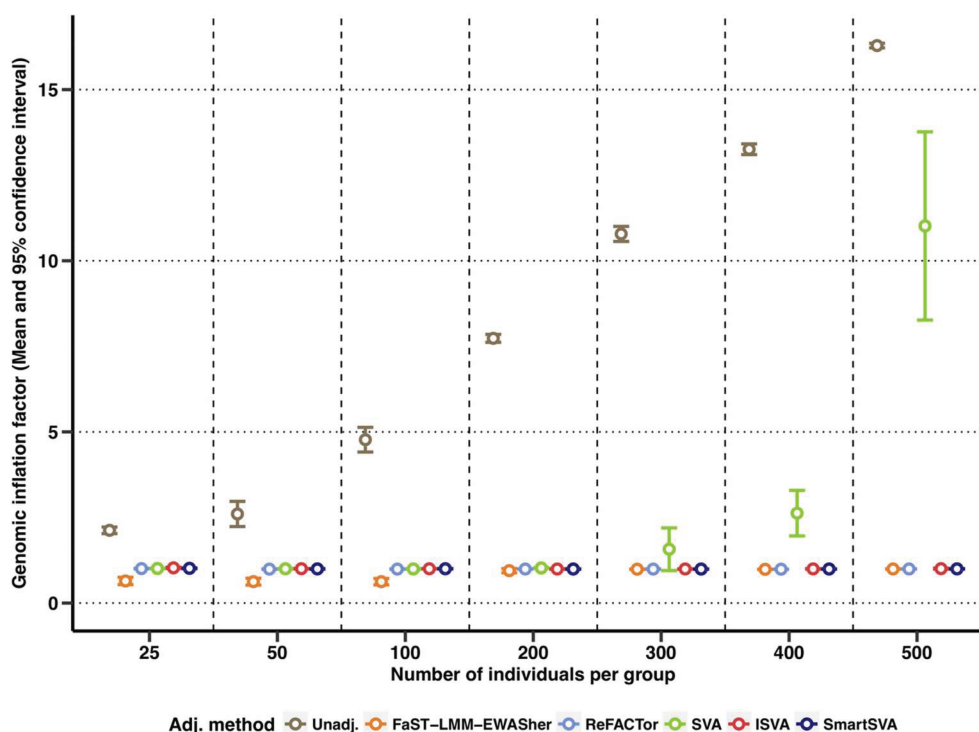


Figure 2. Distribution of the GIF in simulated null scenarios according to the sample size of the study.

power of 0.8, for unadjusted analyses, with average powers of 0.83 ( $S = 5$ ) and 0.79–0.80 ( $S = 50, 500$  or 5000) (Figure 3). FaST-LMM-EWASher resulted in decreasing power with increasing  $S$ , with only 13% power for  $S = 500$  and 7% for  $S = 5000$ . In contrast, adjustment with ReFACTor, SVA, ISVA and SmartSVA resulted in a  $\sim 10\%$  power reduction compared with unadjusted analyses independently of the number of differentially methylated markers  $S$  (Figure 3). In simulation scenarios with  $\delta$  ranging between 0.001 and 0.01 for differentially methylated markers, the statistical power increased with increasing  $\delta$  (Supplementary Figure S3). In agreement with Figure 3, the statistical power was similarly high for SVA, ISVA and ReFACTor, and markedly lower for FaST-LMM-EWASher (Supplementary Figure S3a).

In addition to the detection power, the availability and quality of estimated methylation differences for differentially methylated markers is relevant in the analysis of EWAS. FaST-LMM-EWASher provides adjusted  $P$ -values but does not provide estimated methylation differences, representing an important limitation. The average difference of  $\beta$  values between cases and controls was estimated using unadjusted linear models, ReFACTor, SVA, ISVA and SmartSVA. The quality of estimated methylation differences was examined by the bias, the variance and the MSE compared to the expected  $\delta = 0.0031$  in the simulated cell mixture [42, 43]. Comparison of unadjusted and adjusted analyses revealed that ReFACTor, SVA, ISVA and SmartSVA are clearly superior to unadjusted analyses with very similar estimation quality for the four methods (Figure 4, Supplementary Table S1).

In scenarios with variable  $\delta$ , cell-type adjustment also improved the quality of estimated methylation differences. Interestingly, low ( $\delta < 0.0033$ ) methylation differences were overestimated, while high ( $\delta > 0.0066$ ) methylation differences were underestimated after adjustment with by all four methods (Supplementary Figure S3b). Analysis of the bias, the variance

and the MSE revealed good overall performances for the four adjustment methods, clearly superior to unadjusted estimates (Supplementary Table S2).

### Simulated data sets with differential methylation, high $\delta$

Adjustment methods that do not consider case–control status (unsupervised methods such as FaST-LMM-EWASher and ReFACTor) rely on the markers that show the strongest methylation differences among all samples to estimate adjustment components. As demonstrated above, this may be advantageous compared with supervised methods such as SVA when cell-type heterogeneity is the strongest source of methylation variability and correlates with the variable of interest. However, when methylation differences related to the case–control phenotype are stronger than cell-type heterogeneity effects, supervised adjustment considering case–control status may lead to overcorrection and loss of power. To test this hypothesis, methylation data were simulated assuming  $n = 200$ ,  $S \in \{50, 500, 5000\}$  and  $\delta = 0.1$  for differentially methylated markers. This larger difference in methylation levels compared with the previous analyses was chosen to achieve a relevant influence of the differentially methylated markers on the overall methylation patterns on top of variability due to cell-type composition. Biologically, this corresponds to a 10% difference in average methylation at a given locus and is within the range of methylation differences reported in EWAS [44–46].

While all adjustment methods adequately corrected the type I error rate inflation as seen before, differences in statistical power became evident for higher  $S$  (Supplementary Table S2). In particular, the very high power reached in unadjusted analyses was retained by SVA, ISVA and SmartSVA. In contrast,

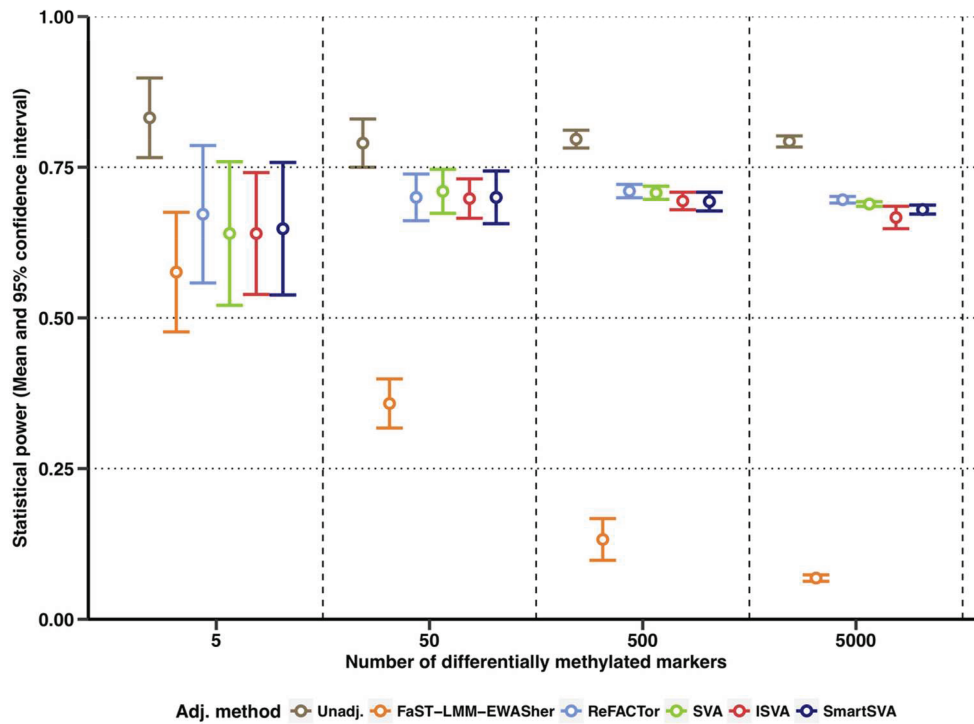


Figure 3. Statistical power to detect differences in methylation across simulated scenarios according to the number of differentially methylated markers ( $n = 200$  per group,  $\delta = 0.0045$ ).

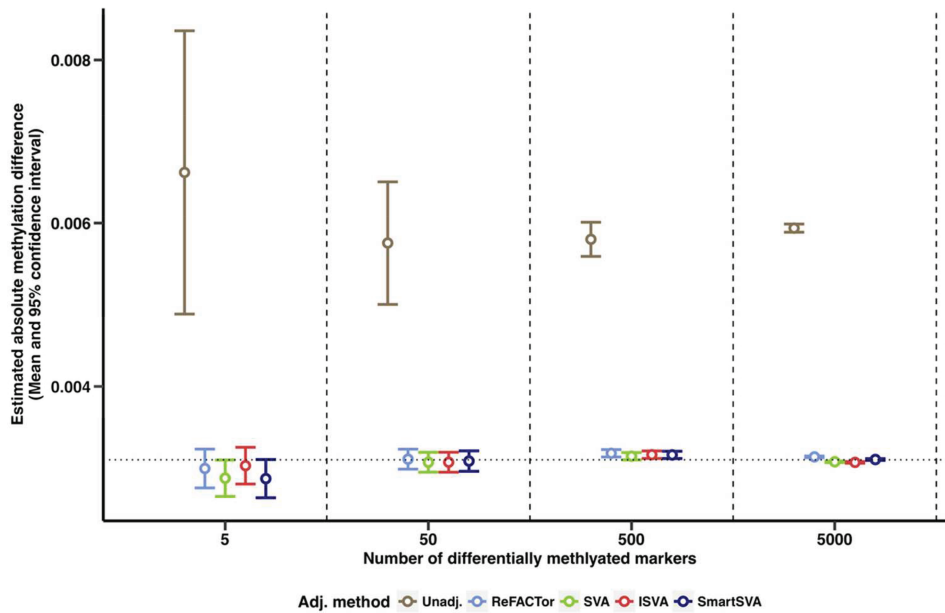


Figure 4. Estimated absolute methylation differences in simulation scenarios with differential methylation. (The dotted line indicates the expected methylation difference  $\delta = 0.0031$  in a simulated cell mixture,  $n = 200$  per group; FaST-LMM-EWASher does not provide estimated methylation differences).

FaST-LMM-EWASher and ReFACTor adjustment resulted in considerable power loss with increasing number of differentially methylated markers. This result indicates that the magnitude of true biological effects relative to cell-type confounding is a key factor that should be considered in the choice of cell-type adjustment method.

#### Investigation of real data sets

Simulations were complemented by real data from methylation studies covering a wide range of study designs (Table 2). Overall, potential confounding as indicated by increased GIFs was present in unadjusted analyses for all data sets (Table 2). The

Table 2. GIF and proportion of probability values higher than 0.05 in four real data sets after different cell-type adjustment methods

Dataset	Investigated phenotype	Tissue type	Cases Controls	No. of markers	GIF	% P-value < 0.05										
						Cell-type adjustment method					Cell-type adjustment method					
						Unadj.	FaST-LMM-EWASher	RefACTor	SVA	ISVA	Smart-SVA	Unadj.	FaST-LMM-EWASher	RefACTor	SVA	ISVA
Zou et al.[22]	Breast cancer	Breast tissue	16737	22,698	29.6	1.00	2.85	42.9	3.25	1.22	0.756	0.048	0.278	0.773	0.274	0.086
Liu et al. (GSE42861)	Rheumatoid arthritis	Blood	354335	42,772(Chr1)	24.7	1.33	2.13	11.0	1.26	1.18	0.707	0.088	0.180	0.556	0.083	0.073
Tsprouni et al. (GSE50660)	Smoking	Blood	285179	42,495(Chr1)	4.83	0.97	1.67	1.49	1.09	1.00	0.352	0.048	0.138	0.116	0.061	0.050
Veldhoven et al. (GSE51057)	Breast cancer	Blood	152177	26,080(Chr1)	3.22	1.03	2.02	1.14	1.20	1.07	0.309	0.054	0.181	0.068	0.071	0.057
Heyn et al. (GSE30870)	Newborn vs. old age	Blood	2020	42,072(Chr1)	4.29	0.72	1.65	6.31	4.44	1.08	0.364	0.035	0.113	0.445	0.377	0.076
Hannon et al. (GSE80417)	Schizophrenia	Blood	353322	42,681(Chr1)	11.4	1.13	1.77	3.17	1.30	1.21	0.559	0.065	0.161	0.287	0.089	0.078
Chen et al. (GSE107737)	Hypopituitarism	Blood	1212	42,772(Chr1)	1.58	0.96	1.03	1.39	1.86	0.98	0.130	0.057	0.054	0.099	0.154	0.049
Philibert et al. (GSE110043)	Alcohol consumption	Blood	4747	38,961(Chr1)	1.56	0.93	1.17	1.15	1.24	1.10	0.129	0.049	0.076	0.070	0.082	0.060
Stefansson et al. (GSE52865)	Breast cancer	Breast tissue	4017	42,772(Chr1)	8.08	0.84	1.25	4.88	3.66	1.20	0.490	0.037	0.090	0.401	0.328	0.085



**Table 3.** Overall performance of the five investigated cell-type adjustment methods

Adjustment method	Type I error rate control <sup>1</sup>	Statistical power	Estimated methylation differences	Runtime <sup>2</sup>
FaST-LMM-EWASher	+++	-	Not available	+
ReFACTor	+	+	+	++
SVA	-	+	+	-
ISVA	+	+	+	+
SmartSVA	++	+	+	++

<sup>1</sup>+++; in simulations and 6 out of 9 investigated real datasets; ++; in simulations and 2 out of 9 investigated real datasets; +; in simulations, -: maximum 200 cases and 200 controls (Table 1 and Table 2).

<sup>2</sup>See Supplementary Figure S2.

Zou *et al.* [22] breast cancer data set and the Liu *et al.* [17] RA data showed the strongest signs of confounding with GIFs of 29.6 and 24.7, respectively. This was an even larger inflation than in the simulated null scenarios with 500 subjects per group (GIF 16.3, Figure 2). SVA was not able to control the type I error rate inflation in the Liu *et al.* [17] RA data (adjusted GIF 11.0, Table 2). This result is in line with our simulations and could be due to a strong correlation between cell-type composition and phenotype together with weak methylation differences between cases and controls. ReFACTor, ISVA, SmartSVA and in particular FaST-LMM-EWASher strongly reduced the GIF and the proportion of significant markers for the Zou *et al.* [22] data set, but the GIF actually increased after SVA correction (Table 2).

In order to comprehensively compare the different adjustment methods, we examined six additional real blood-cell methylation data sets covering a spectrum of investigated phenotypes, plus one breast tissue methylation data set. Unadjusted analyses showed varying degrees of genomic inflation that may have been attributable to real differences, cell-type composition and other confounding factors (Table 2). The adjustment methods improved the control of genomic inflation and the false-positive rates for all datasets. The tightest control of genomic inflation was achieved by FaST-LMM-EWASher, with a GIF equal or lower than 1.00 in six of nine datasets. ReFACTor and ISVA decreased the GIF and the type I error rates, but nominal levels were not reached. SVA did not sufficiently control the genomic inflation, and the GIF even increased after SVA adjustment in the data sets by Liu *et al.* [17] and Heyn *et al.* [35]. SmartSVA resulted in a GIF equal to or lower than 1.00 in two of nine datasets.

Taken together, the comparison of cell-type adjustment methods based on these nine data sets clearly demonstrates that cell-type heterogeneity is present at varying degrees in real-world data sets and requires correction by appropriate adjustment methods. It also shows that study characteristics such as anticipated cell-type differences between cases and controls influence not only confounding, but also the performance of the adjustment methods.

## Discussion

In the present study we investigated two related issues that are currently of central importance for the analysis of EWAS data: (1) what study parameters influence the type and impact of cell-type confounding and (2) what adjustment methods should be used to mitigate false findings attributable to cell-type heterogeneity. We developed and applied a simulation framework that allows fine tuning of critical study parameters, for example the study size or changes in the number and associated effects of differentially methylated markers, facilitating a systematic assessment. We complemented simulation results

with analyses of real data sets, which allowed us to examine a broad range of cell-type compositions and hidden confounders such as age, medication or batch effects that may possibly have different distributions in cases and controls.

Using our simulation framework, we demonstrated that the impact of cell-type confounding increases with increasing study size. A similar observation with respect to genomic inflation was made for GWAS [38–40], but it has not been reported in the context of EWAS, probably because previous simulations considered smaller study sizes [25, 26]. The impact of cell-type confounding is more pronounced in the case of large differences in cell-type compositions between cases and controls. Accordingly, the genomic inflation was highest in the breast cancer data set of Zou *et al.*, [22] in the data set of Stefansson *et al.*, [37] and in Liu *et al.*'s [17] RA data set.

Across all scenarios and data sets, FaST-LMM-EWASher achieved the tightest control of false-positive rates. In fact, FaST-LMM-EWASher was overly conservative in the simulated null scenarios where the GIF and false-positive rate were below the nominal level. Moreover, FaST-LMM-EWASher suffered from a dramatic loss of power in scenarios with simulated differential methylation. Low power has been previously observed [25, 26], which may be attributable to the filtering of markers with  $\beta$ -methylation values over 0.8 and under 0.2, i.e. the majority of markers including differentially methylated ones. In the present analysis this filter was therefore removed. Nevertheless, very few of the simulated markers with differential methylation were recovered after FaST-LMM-EWASher adjustment. This renders FaST-LMM-EWASher primarily suitable for situations in which avoidance of false positives is a major concern, even if true positives are missed in the process.

ReFACTor was specifically designed to correct for cell-type heterogeneity in EWAS [21] and exhibited reliable and robust adjustment of cell-type confounding across the largest proportion of simulated and real data sets. In contrast to SVA-related adjustment methods, ReFACTor uses an approach that is unsupervised with respect to the phenotype and relies for the cell-type adjustment on the most variable markers across the complete data set. For this reason, the adjustment performance of ReFACTor does not depend on the study size. For the same reason, however, adjustment with ReFACTor may translate into decreased statistical power when large methylation differences ( $\delta = 0.1$ ) are present. The methylation differences identified in most blood-based studies are few and slight, and one potential way to improve the performance of ReFACTor when many and large methylation differences are expected is to select the markers for adjustment based on the controls only [21].

SVA was proposed as the method of choice in two recent comparative studies by McGregor *et al.* [25] and Kaushal *et al.* [26]. Overall, SVA efficiently minimized the MSE and variance in simulated scenarios with varying and fixed  $\delta$ , but SVA did not control

the type I error rate inflation in studies with more than 200 cases and 200 controls. Similarly, SVA incompletely controlled the GIF in the real-world EWAS data sets. In advance to the construction of surrogate variables, SVA removes the phenotype-related variation by fitting a linear model to the methylation data. This first step may result in a limited adjustment if the cell-type composition is the dominating driver of methylation differences between cases and controls. Given its design, SVA may thus be of limited utility in scenarios with strong phenotype-dependent differences in cell-type compositions, especially in large studies. In contrast, SVA may be well suited if strong biological differences are present, as in our simulations with large  $\delta$  of differentially methylated markers between cases and controls. In those scenarios, SVA controlled the type I error rate and retained very high power independent of the number of truly differentially methylated markers.

Given that previous studies suggested SVA as the method of choice [25, 26], we evaluated ISVA [24] and SmartSVA [23], which extend and aim to improve on the original SVA algorithm. Across the simulated data sets, ISVA showed a better performance than SVA regarding genomic inflation and type I error rates, while retaining high power. In some of the real data sets, ISVA incompletely controlled the genomic inflation. However, since the true number of methylation differences in real data sets is unknown, it is difficult to quantify the extent of confounding and the comparative performance of adjustment methods in real data sets.

The recently proposed SmartSVA algorithm [23] specifically addresses the limitations of SVA in situations where the primary variable of interest (usually a particular phenotype) is correlated with potential confounders (usually the cell-type composition). In agreement with the present results, this situation often leads to a lack of convergence and insufficient adjustment of genomic inflation by SVA. By imposing an explicit convergence criterion, SmartSVA aims to control the inflation of type I error rates even in such circumstances. SmartSVA indeed showed excellent control of type I error rate while retaining high power across all simulation scenarios. In the real data sets, SmartSVA showed the second most stringent control of type I error inflation after FaST-LMM-EWASher, but again, ranking the performance of adjustment methods based on real data is difficult.

A novel aspect of this study was the assessment of the quality of estimated methylation differences (bias, variance and MSE) according to the adjustment method and major study characteristics, which is fundamental for the design and implementation of subsequent validation assays. FaST-LMM-EWASher does not provide adjusted estimated methylation differences between cases and controls. The quality of estimated methylation differences after adjustment with ReFACTor, SVA, ISVA and SmartSVA was clearly superior to unadjusted estimates, low ( $\delta < 0.0033$ ) methylation differences were overestimated and high ( $\delta > 0.0066$ ) methylation differences were underestimated. The comparison of cell-type adjustment methods by users instead of method developers is another strength of the present study. Because of publication bias and authors' better knowledge of their own algorithms than previous software, the real advantage of new methods over established ones is often overestimated, and unbiased comparisons by independent researchers are beneficial [47]. Table 3 summarizes our findings based on simulated and real data sets. FaST-LMM-EWASher resulted in the lowest type I error rate at the expense of low statistical power. SVA did not control type I error rates in EWAS with more than 200 cases and 200 controls. The statistical power, the quality of estimated methylation differences and

the runtime were best, and very similar with ReFACTor, ISVA and SmartSVA.

### Key Points

- Cell-type heterogeneity is an important issue in EWAS based on peripheral blood samples.
- Laboratory information on the cell-type composition of investigated samples is usually sparse, and several techniques have been developed to adjust EWAS results for cell-type heterogeneity.
- Here we describe the development and application of a flexible, multilayered simulation framework and use real data sets to compare five popular cell-type adjustment methods: FaST-LMM-EWASher, ReFACTor, SVA, ISVA and SmartSVA.
- In most investigated scenarios, FaST-LMM-EWASher tends to be conservative. ReFACTor, ISVA and SmartSVA control false-positive rates and achieve a high statistical power, while SVA minimizes the mean squared error of estimated methylation differences, but does not control false-positive rates in large EWAS.

### Supplementary data

Supplementary data are available online at <https://academic.oup.com/bib>.

### Funding

This study was financially supported by the German Federal Ministry of Education and Research (BMBF, grant 01DN15021), the Deutsche Forschungsgemeinschaft and the Ruprecht-Karls-Universität Heidelberg within the funding program Open Access Publishing.

### References

1. Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell* 2013;**153**(1):38–55.
2. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell* 2007;**128**(4):669–81.
3. Birney E, Smith GD, Gready JM. Epigenome-wide association studies and the interpretation of disease-omics. *PLoS Genet* 2016;**12**(6):1–9.
4. Egger G, Liang G, Aparicio A, et al. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004;**429**(6990):457–63.
5. Rakyan VK, Down TA, Balding DJ, et al. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 2011;**12**(8):529–41.
6. Heyn H, Carmona FJ, Gomez A, et al. DNA methylation profiling in breast cancer discordant identical twins identifies DOK7 as novel epigenetic biomarker. *Carcinogenesis* 2013;**34**(1):102–8.
7. Baglietto L, Ponzi E, Haycock P, et al. DNA methylation changes measured in pre-diagnostic peripheral blood samples are associated with smoking and lung cancer risk. *Int J Cancer* 2016;**140**(1):50–61.
8. Mikeska T, Craig J. DNA methylation biomarkers: cancer and beyond. *Genes* 2014;**5**(3):821–64.
9. Sandoval J, Heyn H, Moran S, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 2014;**6**(6):692–702.

10. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol* 2014;**15**(2):R31.
11. Rimm DL, Han G, Taube JM, et al. A prospective, multi-institutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-L1 expression in non-small cell lung cancer. *JAMA Oncol* 2017;**3**(8):1051–8.
12. Roth A, Khattra J, Yap D, et al. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods* 2014;**11**(4):396–8.
13. Chen W, Wang T, Pino-Yanes M, et al. An epigenome-wide association study of total serum IgE in Hispanic children. *J Allergy Clin Immunol* 2017;**140**(2):571–7.
14. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;**13**:86.
15. Chuang Y-H, Quach A, Absher D, et al. Coffee consumption is associated with DNA methylation levels of human blood. *Eur J Hum Genet* 2017;**25**(5):1–9.
16. Langevin SM, Houseman EA, Accomando WP, et al. Leukocyte-adjusted epigenome-wide association studies of blood from solid tumor patients. *Epigenetics* 2014;**9**(6):884–95.
17. Liu Y, Aryee MJ, Padyukov L, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol* 2013;**31**(2):142–7.
18. Bauer M, Linsel G, Fink B, et al. A varying T cell subtype explains apparent tobacco smoking induced single CpG hypomethylation in whole blood. *Clin Epigenet* 2015;**7**:81.
19. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics* 2014;**30**(10):1431–9.
20. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* 2007;**3**(9):1724–35.
21. Rahmani E, Zaitlen N, Baran Y, et al. Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies. *Nat Methods* 2016;**13**(9):443–5.
22. Zou J, Lippert C, Heckerman D, et al. Epigenome-wide association studies without the need for cell-type composition. *Nat Methods* 2014;**11**(3):309–11.
23. Chen J, Behnam E, Huang J, et al. Fast and robust adjustment of cell mixtures in epigenome-wide association studies with SmartSVA. *BMC Genomics* 2017;**18**(1):413.
24. Teschendorff AE, Zhuang J, Widschwendter M. Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. *Bioinformatics* 2011;**27**(11):1496–505.
25. McGregor K, Bernatsky S, Colmegna I, et al. An evaluation of methods correcting for cell-type heterogeneity in DNA methylation studies. *Genome Biol* 2016;**17**:84.
26. Kaushal A, Zhang H, Karmaus WJJ, et al. Comparison of different cell type correction methods for genome-scale epigenetics studies. *BMC Bioinformatics* 2017;**18**(1):216.
27. Lippert C, Listgarten J, Liu Y, et al. FaST linear mixed models for genome-wide association studies. *Nat Methods* 2011;**8**(10):833–5.
28. Listgarten J, Lippert C, Kadie CM, et al. Improved linear mixed models for genome-wide association studies. *Nat Methods* 2012;**9**(6):525–6.
29. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One* 2012;**7**(7):e41361.
30. 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**(12):503.
31. Chen Y-A, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2014;**8**(2):203–9.
32. Kinar Y, Kalkstein N, Akiva P, et al. Development and validation of a predictive model for detection of colorectal cancer in primary care by analysis of complete blood counts: a binational retrospective study. *J Am Med Inform Assoc* 2016;**23**(5):879–90.
33. Tsaprouni LG, Yang T-P, Bell J, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* 2014;**9**(10):1382–96.
34. van Veldhoven K, Polidoro S, Baglietto L, et al. Epigenome-wide association study reveals decreased average methylation levels years before breast cancer diagnosis. *Clin Epigenet* 2015;**7**:67.
35. Heyn H, Li N, Ferreira HJ, et al. Distinct DNA methylomes of newborns and centenarians. *Proc Natl Acad Sci* 2012;**109**(26):10522–7.
36. Hannon E, Dempster E, Viana J, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biol* 2016;**17**(1):176.
37. Stefansson OA, Moran S, Gomez A, et al. A DNA methylation-based definition of biologically distinct breast cancer subtypes. *Mol Oncol* 2014;**9**(3):555–68.
38. Bacanu SA, Devlin B, Roeder K. The power of genomic control. *Am J Hum Genet* 2000;**66**(6):1933–44.
39. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* 2001;**60**(3):155–66.
40. Freedman ML, Reich D, Penney KL, et al. Assessing the impact of population stratification on genetic association studies. *Nat Genet* 2004;**36**(4):388–93.
41. De Bakker P, Ferreira M, Jia X. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet* 2008;**17**(R2):R122–8.
42. Kesselmeier M, Lorenzo Bermejo J. Robust logistic regression to narrow down the winner's curse for rare and recessive susceptibility variants. *Brief Bioinform* 2017;**18**(6):962–72.
43. Walther BA, Moore JL. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography* 2005;**28**:815–29.
44. Gentilini D, Scala S, Gaudenzi G, et al. Epigenome-wide association study in hepatocellular carcinoma: Identification of stochastic epigenetic mutations through an innovative statistical approach. *Oncotarget* 2017;**8**(26):41890–902.
45. Tang Q, Holland-Letz T, Slynko A, et al. DNA methylation array analysis identifies breast cancer associated RPTOR, MGRN1 and RAPSN hypomethylation in peripheral blood DNA. *Oncotarget* 2016;**7**(39):64191–202.
46. Zhang J, Liu Z, Umukoro PE, et al. An epigenome-wide association analysis of cardiac autonomic responses among a population of welders. *Epigenetics* 2017;**12**(2):71–6.
47. Boulesteix A-L, Stierle V, Hapfelmeier A. Publication bias in methodological computational research. *Cancer Inform* 2015;**14**(suppl 5):11–9.