A second DNA sample was extracted from peripheral blood drawn at the 7-year clinic visit, and a third at a clinic held either at age 15 or at age 17 years. Maternal samples were provided during the antenatal period [mean 25.7 (SD 9.5) weeks of gestation] and approximately 15–17 years later when mothers accompanied their child to a clinic visit or when they attended a clinic themselves.

T e-_ ec f c DNA e a _ f g

DNA methylation patterns differ between tissues. To generate a reference resource to facilitate comparison of peripheral blood DNA methylation with that of other tissues, we undertook profiling of DNA methylation from a range of other sources, where possible with paired peripheral

annotated with version numbers, with all previous versions stored and available on request. Notification of version updates will be posted on the study website.

Additional data generation is under way, with further work planned. This includes Illumina 450 K array analysis of more ALSPAC participants including fathers ($n = 312$ ARIES

were flagged if there was a sex-mismatch based on X-chromosome methylation.

Tissue samples from autopsy specimens, post-mortem brain and fetal samples were processed in exactly the same way as blood samples described above.

The Illumina 450 K BeadChip assay detects the proportion of molecules methylated at each CpG site on the array. For each sample, the estimated methylation level at each CpG site was expressed as a beta value (β), which is the ratio of the methylated probe intensity and the overall intensity and ranges from 0 (no cytosine methylation) to 1 (complete cytosine methylation). Methylation data were pre-processed using R (version 3.0.1), with background correction and subset quantile normalization performed within each time point using the pipeline described by Touleimat and Tost.[4] Data are available both pre- and post-normalization.

It is widely recognized that white blood cell heterogeneity can confound DNA methylation measurement.[5] The overwhelming majority of ALSPAC biosamples did not have whole blood cell counts taken prior to DNA extraction and therefore the only viable option is to apply a post hoc correction for cellular heterogeneity to the DNA methylation data generated. In order to ensure that the results of any downstream analysis are not influenced by variation in cell type fraction between samples, we estimated the fraction of CD8 T cells, CD4 T cells, NK cells, B cells, monocytes and granulocytes in the samples using the estimateCellCounts function in the minfi Bioconductor package implemented in R.[6] Both cell-type corrected and uncorrected data are available to facilitate use of either data set in a main analysis or the corrected data set as a sensitivity analysis.

W e-ge e b e e e c g

For BS-seq library preparation, DNA samples were fragmented by sonication (Covaris) and ad3ftion,(iga11 (by)-d

evidence base for future epigenetic studies, a Pregnancy and Childhood Epigenetics Consortium has been established (PACE), to facilitate meta-analyses across several exposures and outcomes pertinent to this sector of the population. EWAS of many other exposures and phenotypes have been conducted (e.g. alcohol intake, trace elements, heavy metals, stress, autism spectrum disorders, children born following assisted reproductive technologies, etc.).

ii. Replication studies. Frequent requests are made by other investigators to access ARIES data relating to one or a few loci that have been identified as being associated with a specific trait or exposure in an independent cohort.[13,14] Such replication is an essential step in a study design, and ARIES has developed reciprocal arrangements for replication with several other cohort studies. Applications to access and use ARIES data for replication can be made at any time with no requirement to involve ALSPAC investigators.

iii. Genome-wide association studies have been undertaken to identify common genetic variation associated with DNA methylation variation. A large-scale GWAS has been completed, indicating a stable genetic influence of genotype on DNA methylation across the life course. Findings are currently being prepared for

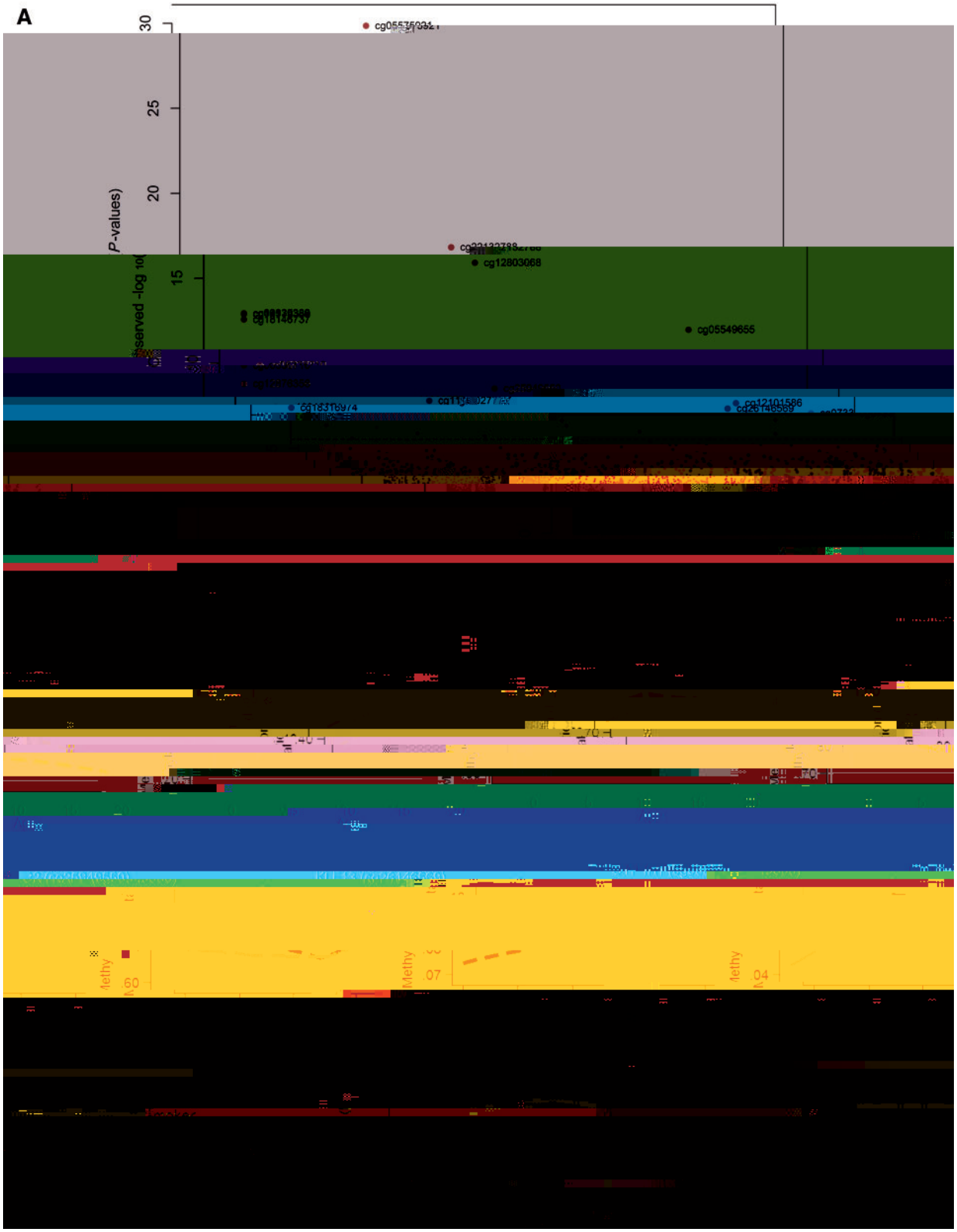

Figure 2.

upon the inferences that can be made regarding tissue-specific DNA methylation patterns. The creation of a reference data series of numerous tissue types mitigates this limitation to some extent but is an imperfect solution. Other limitations include reliance on the Illumina Infinium HumanMethylation450K BeadChip which has limited coverage (1.7% of all CpG sites in the genome) and recognized technical biases.[23] Complementary methods can be (and have been) used for validation and verification of differential DNA methylation, such as Pyrosequencing or targeted BS-Seq. Applications can be made to access and use biological samples from the ALSPAC bio-repository to undertake additional epigen-

Funding

ARIES was funded by the BBSRC (BBI025751/1 and BB/I025263/1). Core programme support for ALSPAC is provided by the Medical Research Council (MRC) and the Wellcome Trust (Grant

24. Pidsley R, Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data driven approach to preprocessing Illumina 450 K methylation array data. BMC Genomics 2013;14:293.
25. 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; 30:1363–69.
26. Barfield RT, Kilaru V, Smith AK, Coneely KN. CpGassoc: an R function for analysis of DNA methylation microarray data. Bioinformatics