



## King's Research Portal

DOI:

[10.1080/15592294.2015.1127479](https://doi.org/10.1080/15592294.2015.1127479)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Marzi, S., Meaburn, E. L., Dempster, E. L., Lunnon, K. S., Paya Cano, J. L., Smith, R. G., Volta, M., Troakes, C., Schalkwyk, L. C., & Mill, J. (2016). Tissue-specific patterns of allelically-skewed DNA methylation. *Epigenetics : official journal of the DNA Methylation Society*, 11(1), 24-35. <https://doi.org/10.1080/15592294.2015.1127479>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

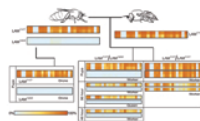
### **General rights**

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

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

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



View Highlights  
Read full article for additional information on this journal article. Please refer to the article text for more information on this journal article.



ISSN: 1559-2294 (Print) 1559-2308 (Online) Journal homepage: <http://www.tandfonline.com/loi/kepi20>

## Tissue-specific patterns of allelically-skewed DNA methylation

Sarah J. Marzi, Emma L. Meaburn, Emma L. Dempster, Katie Lunnon, Jose L. Paya-Cano, Rebecca G. Smith, Manuela Volta, Claire Troakes, Leonard C. Schalkwyk & Jonathan Mill

To cite this article: Sarah J. Marzi, Emma L. Meaburn, Emma L. Dempster, Katie Lunnon, Jose L. Paya-Cano, Rebecca G. Smith, Manuela Volta, Claire Troakes, Leonard C. Schalkwyk & Jonathan Mill (2016) Tissue-specific patterns of allelically-skewed DNA methylation, *Epigenetics*, 11:1, 24-35, DOI: [10.1080/15592294.2015.1127479](https://doi.org/10.1080/15592294.2015.1127479)

To link to this article: <http://dx.doi.org/10.1080/15592294.2015.1127479>



© 2016 The Author(s). Published with license by Taylor & Francis Group, LLC © Sarah J. Marzi, Emma L. Meaburn, Emma L. Dempster, Katie Lunnon, Jose L. Paya-Cano, Rebecca G. Smith, Manuela Volta, Claire Troakes, Leonard C. Schalkwyk, and Jonathan Mill



View supplementary material [↗](#)



Accepted author version posted online: 19 Jan 2016.  
Published online: 19 Jan 2016.



Submit your article to this journal [↗](#)



Article views: 1252



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

RESEARCH PAPER

 OPEN ACCESS

## Tissue-specific patterns of allelically-skewed DNA methylation

Sarah J. Marzi<sup>a</sup>, Emma L. Meaburn<sup>b</sup>, Emma L. Dempster<sup>c</sup>, Katie Lunnon<sup>c</sup>, Jose L. Paya-Cano<sup>a</sup>, Rebecca G. Smith<sup>c</sup>, Manuela Volta<sup>a</sup>, Claire Troakes<sup>a</sup>, Leonard C. Schalkwyk<sup>d</sup>, and Jonathan Mill<sup>a,c</sup>

<sup>a</sup>Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; <sup>b</sup>Department of Psychological Sciences, Birkbeck, University of London, London, UK; <sup>c</sup>University of Exeter Medical School, University of Exeter, Exeter, UK; <sup>d</sup>School of Biological Sciences, University of Essex, Colchester, UK

### ABSTRACT

While DNA methylation is usually thought to be symmetrical across both alleles, there are some notable exceptions. Genomic imprinting and X chromosome inactivation are two well-studied sources of allele-specific methylation (ASM), but recent research has indicated a more complex pattern in which genotypic variation can be associated with allelically-skewed DNA methylation in *cis*. Given the known heterogeneity of DNA methylation across tissues and cell types we explored inter- and intra-individual variation in ASM across several regions of the human brain and whole blood from multiple individuals. Consistent with previous studies, we find widespread ASM with > 4% of the ~220,000 loci interrogated showing evidence of allelically-skewed DNA methylation. We identify ASM flanking known imprinted regions, and show that ASM sites are enriched in DNase I hypersensitivity sites and often located in an extended genomic context of intermediate DNA methylation. We also detect examples of genotype-driven ASM, some of which are tissue-specific. These findings contribute to our understanding of the nature of differential DNA methylation across tissues and have important implications for genetic studies of complex disease. As a resource to the community, ASM patterns across each of the tissues studied are available in a searchable online database: <http://epigenetics.essex.ac.uk/ASMBrainBlood>.

### ARTICLE HISTORY

Received 26 August 2015  
Revised 19 November 2015  
Accepted 25 November 2015

### KEYWORDS

Allele-specific DNA methylation; blood; brain; cerebellum; cortex; epigenetics; genomic imprinting; SNP

### Introduction

DNA methylation is the most widely studied and stable epigenetic mark across the mammalian genome, playing a key role in the developmental regulation of gene expression. DNA methylation is generally symmetrical across both alleles, although exceptions characterized by allelic asymmetry include differentially methylated regions (DMRs) regulating the mono-allelic expression of genes associated with X chromosome inactivation in females and genomic imprinting.<sup>1–5</sup> Recently, it has been shown that the allelic-skewing of DNA methylation can also be driven by DNA sequence variation, with methylation quantitative trait loci (meQTLs) predominantly acting in *cis*.<sup>6–11</sup> ASM can be regarded as a special case of intermediate DNA methylation (IM), which has been found to occur in regions spanning a large portion of the human genome. It has been estimated that ASM contributes up to 18% of IM in the human genome.<sup>12</sup>


DNA methylation patterns are highly dynamic during normal development and cellular differentiation<sup>13–16</sup> and tissue-specific patterns of DNA methylation have been widely studied in humans.<sup>17–20</sup> In complex tissues such as the brain, for example, DNA methylation differentiates between functionally distinct regions<sup>21,22</sup> and cell types.<sup>16,23–26</sup> Patterns of IM can also be tissue-specific,<sup>12</sup> with growing evidence for the widespread prevalence of tissue-specific ASM.<sup>27,28</sup> In mouse, for example, it has been reported that 28% of imprinted genes are monoallelically

expressed in a single tissue type, often the brain or extra-embryonic tissue.<sup>29</sup> Examples of tissue-specifically imprinted genes include *KCNQ1*, which becomes biallelically expressed in embryonic heart development,<sup>30</sup> *GNAS*, which is maternally expressed in a wide-range of tissues including the anterior pituitary, thyroid and ovaries but biallelically expressed in others, such as bone and visceral adipose tissue,<sup>31,32</sup> and *GRB10*, which is maternally expressed in most peripheral tissues but paternally expressed in the brain.<sup>29,33</sup> Genetic influences on DNA methylation can also be tissue-specific, with meQTLs determining allelic patterns of methylation in *cis* in certain tissues or cell types.<sup>11,34</sup>

Increasing evidence supports a role for inter-individual variation in DNA methylation in the etiology and pathogenesis associated with a diverse range of complex disease phenotypes.<sup>35</sup> Allelic differences in DNA methylation may be particularly important in this regard, acting as endophenotypes of genetic variation or additional epi-allelic layers mediating the functional consequences of genotypic variation.<sup>36,37</sup> Teasing apart genetic and non-genetic effects in a tissue- and cell type-specific manner will be a crucial step in understanding the association between non-coding genetic variation, DNA methylation, and complex disease.

To investigate the role of tissue-specific variation of ASM in the human brain and its relation to allelic biases in whole blood, we examined ASM across multiple brain regions and matched blood samples collected from multiple donors. Our data shows

**CONTACT** Jonathan Mill  J.Mill@exeter.ac.uk

 Supplemental data for this article can be accessed on the publishers website.

Published with license by Taylor & Francis Group, LLC © Sarah J. Marzi, Emma L. Meaburn, Emma L. Dempster, Katie Lunnon, Jose L. Paya-Cano, Rebecca G. Smith, Manuela Volta, Claire Troakes, Leonard C. Schalkwyk, and Jonathan Mill

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

that although a large proportion of ASM is conserved across tissues, there are specific differences in the extent and distribution of ASM sites between regions of the brain and whole blood. Genome browser tracks displaying ASM signals as well as an online tool plotting ASM for sites of interest are available for download from a searchable database (<http://epigenetics.essex.ac.uk/ASMBrainBlood>).

## Results

### *DNA methylation is allelically-skewed at specific locations across the genome*

The majority of the genome is not characterized by notable allelic biases in DNA methylation in any of the tissues assessed in this study. The array-wide average of ASM score (in absolute values) is consistently low (mean = 0.025, range = 0.023 to 0.030) (Fig. S1A, Fig. S1B and Table S1). As expected, there is, however, evidence for allelically-biased DNA methylation at a notable number of specific genomic regions; in total 9,311 (4.22%) of the 220,449 informative SNPs in our assay show evidence for allelic-skewing of DNA methylation, defined by an absolute ASM score  $\geq 0.10$  in at least one tissue and individual. The percentage of amplicons characterized by an ASM score  $\geq 0.10$  in each of the 21 profiled samples is given in Table S2. The top-ranked loci showing evidence for allelically-skewed DNA methylation in whole blood, cerebellum, and cortex (BA9), and cerebellum are listed in Tables 1–3. Genome Browser tracks and an online ASM database are available from our laboratory website (<http://epigenetics.essex.ac.uk/ASMBrainBlood>).

### *Patterns of ASM in whole blood overlap with those identified in a previous study*

In a previous study we characterized allelically-skewed DNA methylation in whole blood derived from 5 monozygotic twin pairs.<sup>7</sup> There is a highly significant correlation between absolute ASM scores across all probes informative in both data sets ( $n = 129,559$ ,  $r = 0.21$ ,  $P < 1.0 \times 10^{-50}$ , Fig. S2), even though the majority of the probes do not exhibit ASM. Of the 2,704

ASM loci identified in Schalkwyk et al., 1,717 (63.50%) are informative in the current study, with a highly significant cross-study correlation of ASM scores at these probes ( $r = 0.52$ ,  $P < 1.0 \times 10^{-50}$ ). Likewise, there is a highly significant correlation between ASM scores across the two studies at sites showing allelically-skewed DNA methylation in blood in the current study which were also informative in our previous study ( $r = 0.38$ ,  $P = 3.0 \times 10^{-28}$ ). Of the 15 top-ranked blood ASM sites identified in our current study (Table 1), 7 of the 9 sites (78%) also informative in our previous study of ASM in blood<sup>7</sup> were characterized by an absolute ASM score  $\geq 0.10$  in both analyses. These data confirm the validity of the MSNP approach for identifying allelically-skewed DNA methylation, reinforcing our previous conclusions about the extent of ASM in whole blood.<sup>7</sup>

### *The extent and distribution of ASM differs across tissues*

The average proportion of informative sites characterized by allelically-skewed DNA methylation (absolute ASM score  $\geq 0.10$ ) in each of the 8 tissues profiled was examined (Fig. 1A). Table 4 lists the top-ranked consistently allelically-skewed probes across cerebellum, whole blood, and cortex (BA9), with specific examples shown in Fig. 2A and Fig. 2B. Allelically-skewed DNA methylation appears to be consistently less prevalent in cortical regions (informative probes with ASM score  $\geq 0.10 = 0.54\%$ ) compared to the cerebellum (1.14%) and whole blood (0.84%). The elevated level of allelically-skewed DNA methylation in the cerebellum and whole blood relative to cortex is more pronounced at more extreme ASM score thresholds (i.e., ASM score  $\geq 0.20$ , cortex = 0.003%, cerebellum = 0.019%, whole blood = 0.013%) (Fig. S1C, Fig. S1D and Fig. S3). Of note, there is little variation in the prevalence and distribution of ASM scores between different regions of the cortex (average correlation between 2 cortical areas = 0.52, Fig. 1B and Fig. S4). We therefore selected one representative cortical region (BA9) for inclusion in subsequent analyses. In contrast, we find more striking differences between cortex, cerebellum and whole blood samples with inter-tissue correlations ranging from  $r = 0.42$  to 0.48 (Figs. 1C–E). Table 5 lists the probes showing the highest level of variation in ASM scores across tissues with specific

**Table 1.** Top 15 ASM sites in whole blood, ASM score averaged across individuals.

Rank	SNP ID	Location	Associated gene(s)	Schalkwyk et al. (2010)	Blood ASM score	Cerebellum ASM score	BA9 ASM score
1	SNP_A-2180729 (rs10276966)	7p15.2	<i>HIBADH, EVX1<sup>b</sup></i>	0.29	0.25	0.01	0.03
2	SNP_A-8438077 (rs585451)	15q21.2	<i>ATP8B4, DTWD1</i>	NA	0.25	0.01	0.03
3	SNP_A-1841543 (rs10234308)	7p15.3	<i>MGC87042</i>	0.34	0.23	0.25	0.23
4	SNP_A-1946136 (rs927000)	20q13.12	<i>STK4</i>	NA	0.23	0.04	0.01
5	SNP_A-8450837 (rs4404067)	16q12.2	<i>SLC6A2, LPCAT2</i>	0.28	0.22	0.05	0.04
6	SNP_A-4279002 (rs335554)	1q41	<i>CENPF, KCNK2</i>	NA	0.22	0.08	0.09
7	SNP_A-8450539 (rs10916799)	1p36.12	<i>CAMK2N1, LOC339505</i>	0.18	0.22	0.11	0.15
8	SNP_A-8633222 (rs13165930)	5q33.3	<i>CCNJL</i>	NA	0.22	0.02	0.12
9	SNP_A-8472169 (rs11193683)	10q23.1	<i>NRG3</i>	0.01	0.22	0.09	0.07
10	SNP_A-2071005 (rs4687210)	3q28	<i>UTS2D</i>	NA	0.22	0.16	0.17
11	SNP_A-2185394 (rs852454)	7p22.1	<i>RNF216</i>	0.06	0.22	0.11	0.11
12	SNP_A-8702215 (rs12978286)	19p13.11	<i>FCHO1, MAP1S</i>	0.19	0.22	0.00	0.07
13	SNP_A-1807006 (rs10481354)	8p23.3	<i>CLN8, DLGAP2<sup>a</sup></i>	0.22	0.22	0.00	0.05
14	SNP_A-8326632 (rs1542180)	2q31.1	<i>HOXD3, HOXD4</i>	NA	0.21	0.24	0.19
15	SNP_A-2008150 (rs869108)	11p15.1	<i>OTOG, MYOD1, USH1C</i>	0.22	0.21	0.16	0.07

<sup>a</sup> Known imprinted gene

<sup>b</sup> Suspected imprinted gene

**Table 2.** Top 15 ASM sites in cerebellum, ASM score averaged across individuals.

Rank	SNP ID	Location	Associated gene(s)	Cerebellum ASM score	Blood ASM score	BA9 ASM score
1	SNP_A-4255628 (rs959246)	18q12.3	<i>SLC14A2, SETBP1</i>	0.30	0.03	0.05
2	SNP_A-8696273 (rs1003533)	5q31.1	<i>C5orf56</i>	0.29	0.16	0.07
3	SNP_A-1820553 (rs7959070)	12q22	<i>CLU10S, BTG1</i>	0.26	0.03	0.05
4	SNP_A-1841543 (rs10234308)	7p15.3	<i>MGC87042</i>	0.25	0.23	0.23
5	SNP_A-2002282 (rs12246813)	10q22.1	<i>COL13A1, C10orf35</i>	0.24	0.10	0.10
6	SNP_A-8625237 (rs12493005)	3q26.32	<i>TBL1XR1</i>	0.24	0.20	0.24
7	SNP_A-8326632 (rs1542180)	2q31.1	<i>HOXD3, HOXD4</i>	0.24	0.21	0.19
8	SNP_A-2160121 (rs7205794)	16q23.3	<i>CDH13, MPHOSPH6</i>	0.24	0.01	0.01
9	SNP_A-2052542 (rs3098382)	5q13.2	<i>MAP1B</i>	0.24	0.04	0.19
10	SNP_A-2040586 (rs17097827)	14q32.2	<i>BCL11B, C14orf177</i>	0.23	0.10	0.18
11	SNP_A-8420373 (rs10186346)	2p22.1	<i>TMEM178</i>	0.23	0.01	0.07
12	SNP_A-1788157 (rs7158663)	14q32.2	<i>MEG3<sup>a</sup></i>	0.23	0.01	0.03
13	SNP_A-4235630 (rs9641549)	7q31.2	<i>TFEC, MDFIC</i>	0.23	0.05	0.13
14	SNP_A-4264458 (rs1358229)	4q27	<i>TRPC3, KIAA1109</i>	0.23	0.01	0.00
15	SNP_A-1931666 (rs6707698)	2q34	<i>IKZF2, ERBB4</i>	0.22	0.03	0.02

<sup>a</sup> Known imprinted gene

examples shown in Fig. 3A and Fig. 3B. We used clonal bisulfite sequencing to validate tissue-specific ASM identified by the MSNP method in these 2 regions (Fig. 3C and Fig. 3D), confirming the patterns observed in our array data for both loci.

### Informative MSNP probes within DNase I hypersensitive regions are characterized by elevated ASM scores

Enrichment analyses were performed using a Kruskal-Wallis rank-sum test for ASM rank differences between the annotated genic regions (see **Materials and Methods**). We observed a differential distribution of ASM scores across annotated genic regions (i.e., coding, 5'UTR, intergenic, intron, promoter, 3'UTR) in cortex (BA9) ( $P = 1.29 \times 10^{-15}$ ), cerebellum ( $P = 3.98 \times 10^{-14}$ ), whole blood ( $P = 2.06 \times 10^{-12}$ ), and the cross-tissue analysis ( $P = 2.12 \times 10^{-26}$ ). Post-hoc tests identified these differences to be primarily driven by an enrichment of high ASM scores in promoter regions (Fig. S5). We used data from ENCODE<sup>38</sup> to assess whether ASM is enriched in regions associated with DNase I hypersensitive (DHS) sites identified in multiple tissues including frontal cortex and cerebellum, as well as CD14+ monocytes and naïve B cells (as a proxy for blood). We compared the ASM score ranks for informative probes between regions defined by the presence or absence of DHS sites using a Wilcoxon rank-sum test (see

**Materials and Methods**). DHS peaks across all tissues are enriched for higher ASM scores identified in cortex (BA9), cerebellum, and whole blood (Table S3 and Fig. S6). Of note, the most striking enrichment is found for cerebellum ASM scores in regions characterized by cerebellum DHS peaks in ENCODE ( $P = 3.51 \times 10^{-220}$ ).

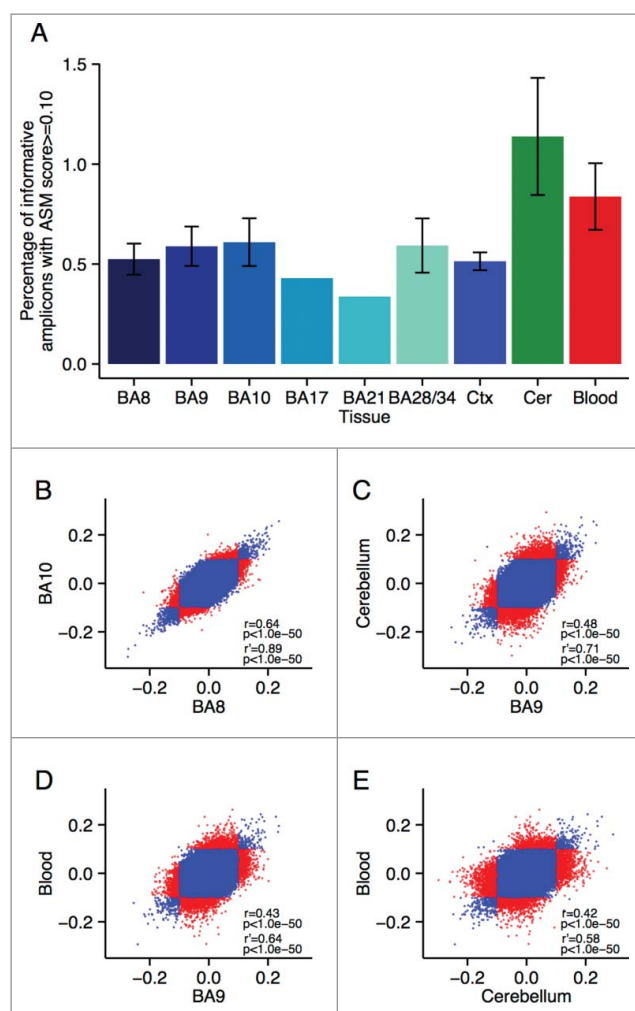
### Inter-individual variation in ASM

We next examined inter-individual differences in ASM score at specific loci, defining probes with a large range of ASM scores across the 3 individuals as being characterized by “variable ASM.” Differentially methylated regions (DMRs) associated with genomic imprinting, for example, are characterized by parental-origin-specific ASM and are expected to show consistently large ASM scores that exhibit allelic-flipping, resulting from genotype-independent ASM. Genotype-driven ASM, in contrast, is likely to be exemplified by consistent allelic biases in DNA methylation across individuals, and is generally not variable between individuals of the same genotype. Fig. S7 shows the correlation in ASM scores across the 3 individuals profiled by MSNP, with tissue-specific correlations given in Table S4. As expected, the individuals were more highly correlated for loci characterized by high ASM scores. For probes informative in at least 2 individuals we examined the range of

**Table 3.** Top 15 ASM sites in cortex (BA9), ASM score averaged across individuals.

Rank	SNP ID	Location	Associated gene(s)	BA9 ASM score	Cerebellum ASM score	Blood ASM score
1	SNP_A-8625237 (rs12493005)	3q26.32	<i>TBL1XR1</i>	0.24	0.24	0.20
2	SNP_A-8463467 (rs17164474)	7q35	<i>OR2F2, OR2F1</i>	0.23	0.17	0.17
3	SNP_A-1841543 (rs10234308)	7p15.3	<i>MGC87042</i>	0.23	0.25	0.23
4	SNP_A-8652129 (rs2479084)	1p36.21	<i>FHAD1</i>	0.22	0.13	0.18
5	SNP_A-8301602 (rs398225)	3p25.3	<i>SRGAP3, RAD18</i>	0.21	0.18	0.21
6	SNP_A-2107106 (rs987377)	6q21	<i>AIM1<sup>a</sup>, ATG5</i>	0.20	0.16	0.15
7	SNP_A-2052542 (rs3098382)	5q13.2	<i>MAP1B</i>	0.19	0.24	0.04
8	SNP_A-2307481 (rs716591)	15q26.2	<i>LOC400456, MCTP2</i>	0.19	0.21	0.17
9	SNP_A-8643280 (rs3121125)	1q21.1	<i>HFE2, NBP10</i>	0.19	0.17	0.15
10	SNP_A-8326632 (rs1542180)	2q31.1	<i>HOXD3, HOXD4</i>	0.19	0.24	0.21
11	SNP_A-1998023 (rs9722212)	9q34.11	<i>TOR1B, PTGES</i>	0.19	0.15	0.19
12	SNP_A-8640607 (rs12632177)	3q27.1	<i>MCF2L2</i>	0.18	0.11	0.07
13	SNP_A-4291638 (rs12670584)	7p13	<i>YKT6</i>	0.18	0.19	0.01
14	SNP_A-1892234 (rs17303015)	5p12	<i>MGC42105</i>	0.18	0.11	0.03
15	SNP_A-8653671 (rs4871852)	8p21.3	<i>TNFRSF10D</i>	0.18	0.09	0.11

<sup>a</sup> Known imprinted gene



**Figure 1.** Allelic-skewing is less prevalent and less variable across cortical regions compared to cerebellum and whole blood. (A) The average proportion of informative amplicons showing an ASM score  $\geq 0.10$  for the 6 cortical regions profiled as well as averaged across the cortical areas (Ctx) is consistently lower (ASM score range = 0.34–0.61%) than in cerebellum (Cer) (ASM prevalence = 1.14%) or blood (ASM prevalence = 0.84%). Standard errors are shown for tissues for which samples were available from all 3 individuals. (B) – (E) Correlations of ASM scores are shown with each point representing one probe in one individual. Probes classified as allelically-skewed at an ASM score  $\geq 0.10$  in only one of the 2 compared tissues are highlighted in red. A higher degree of between-tissue variability is observed between cerebellum, cortex, and whole blood than between different cortical regions (shown as an example is BA8 vs. BA10). This difference becomes even more pronounced when restricting the set of probes to those that show allelic-skewing at an ASM score  $\geq 0.10$  in at least one of the 2 compared tissues (see subset correlation  $r$ ).

ASM scores across individuals and identified the top ranked variable ASM probes in each tissue (Table S5–S7) as well as cross-tissue variable sites, which show consistent inter-individual variation across all tissues (Table 6). Some sites show evidence of allelic-flipping in ASM score between individuals, indicative of genomic imprinting. These included several probes in the vicinity of the imprinted gene cluster on chromosome 15q11.2 (Figs. 4A–C). High ASM scores were also observed in the vicinity of other known imprinted loci, for example, *SNRPN* (Fig. S8A), *DLGAP2* (Fig. S8B), *AIM1* (Fig. S8C), *MEG3* (Fig. S8D), *BLCAP* (Fig. S8E), and *GRB10* (Fig. S8F), in addition to loci suspected to be imprinted, e.g., *TRAPPC9* (Fig. S8G), *EVX1* (Fig. S8H), and *TGFBI/VTRNA2* (Fig. S8I); however, we were unable to examine variable ASM

in many of these regions because they were only informative (i.e., heterozygous) in a single individual. Notably, we also identified allelic-flipping in the vicinity of loci not previously characterized as being imprinted, for example *WRB* (Fig. 4D) and *ITPK1* (Fig. S9). Other variable ASM sites are marked by both high and low ASM scores in different individuals, rather than allelic-flipping between them, for example, *MGST3/LOC400794* (Table S5). Interestingly, we identified a number of sites characterized by tissue-specific variable ASM. A notable example is the imprinted gene *GRB10*, which has been previously shown to be differentially maternally- and paternally-expressed in a tissue-specific manner<sup>29,33</sup> (Fig. S8F).

### Variable ASM sites are flanked by extended regions of intermediate DNA methylation

We next quantified genome-wide patterns of DNA methylation in a larger sample ( $n=39$ ) of matched whole blood, cortex (BA9), and cerebellum samples using the Illumina Infinium HumanMethylation450 BeadChip (450K array). For the 100 top-ranked ASM sites in each of the 3 tissues, plus the 100 top-ranked cross-tissue, tissue-specific, and variable ASM sites we identified probes on the array located within 1 kb of the ASM marker SNPs (Table S8; detailed in Table S9–S17) to investigate patterns of DNA methylation across an extended region. As expected, regions around known imprinted loci identified by our ASM analysis are flanked by extended regions of intermediate DNA methylation (i.e., average levels of DNA methylation between 0.4 and 0.6) (Fig. 5A and Fig. 5B). We observe a highly significant enrichment ( $P$  range =  $6.82 \times 10^{-11}$  – 0.005) of intermediate DNA methylation relative to overall levels identified on the 450K array in regions flanking variable ASM sites in all 3 tissues (Table 7 and Fig. 6). For example, intermediate DNA methylation was observed around the variable ASM site overlapping *WRB* (Fig. 5C), that showed evidence of allelic-flipping (Fig. 4D). Another probe exhibiting variable ASM annotated to *TGFBI/VTRNA2-1* on chromosome 5 also shows a similar pattern of intermediate DNA methylation (Fig. 5D). Interestingly, 4 individuals are distinguished by consistent hypomethylation in whole blood and cortex across 16 of the 19 sites, consistent with previous reports describing polymorphic imprinting of this locus.<sup>39,40</sup>

### Identification of tissue-specific genotype-driven ASM

In contrast to the intermediate DNA methylation patterns enriched in the vicinity of variable ASM sites, non-variable ASM (i.e., characterized by consistent ASM scores across individuals) is not significantly enriched for intermediate DNA methylation and in some cases is exemplified by trimodal patterns of DNA methylation, which are highly suggestive of genotype-driven ASM acting in *cis*. Of note, we also observe examples of tissue-specific genotype-driven ASM. For example, a tissue-specific ASM site identified as showing allelically-skewed DNA methylation in cerebellum (ASM score = 0.16) but not whole blood (ASM score = 0.03) or cortex (ASM score = 0.04) (Fig. 7A) located in an intron of the gene *SYNJ2* is

**Table 4.** Top 15 loci characterized by consistent ASM across cerebellum, whole blood, and cortex (BA9).

Rank	SNP ID	Location	Associated gene(s)	Cerebellum ASM score	Blood ASM score	BA9 ASM score	Tissue average
1	SNP_A-1841543 (rs10234308)	7p15.3	<i>MGC87042</i>	0.25	0.23	0.23	0.24
2	SNP_A-8625237 (rs12493005)	3q26.32	<i>TBL1XR1</i>	0.24	0.20	0.24	0.22
3	SNP_A-8326632 (rs1542180)	2q31.1	<i>HOXD3, HOXD4</i>	0.24	0.21	0.19	0.21
4	SNP_A-8301602 (rs398225)	3p25.3	<i>SRGAP3, RAD18</i>	0.18	0.21	0.21	0.20
5	SNP_A-8463467 (rs17164474)	7q35	<i>OR2F2, OR2F1</i>	0.17	0.17	0.23	0.19
6	SNP_A-2307481 (rs716591)	15q26.2	<i>MCTP2, LOC400456</i>	0.21	0.17	0.19	0.19
7	SNP_A-1894705 (rs986324)	Xp22.11	<i>DDX53, ZNF645</i>	0.18	0.20	0.17	0.18
8	SNP_A-2071005 (rs4687210)	3q28	<i>UTS2D</i>	0.16	0.22	0.17	0.18
9	SNP_A-1998023 (rs9722212)	9q34.11	<i>PTGES, TOR1B</i>	0.15	0.19	0.19	0.18
10	SNP_A-8652129 (rs2479084)	1p36.21	<i>FHAD1</i>	0.13	0.18	0.22	0.18
11	SNP_A-8696273 (rs1003533)	5q31.1	<i>C5orf56</i>	0.29	0.16	0.07	0.17
12	SNP_A-1862496 (rs17578280)	1p31.1	<i>LRRIQ3, NEGR1</i>	0.18	0.17	0.17	0.17
13	SNP_A-2040586 (rs17097827)	14q32.2	<i>C14orf177, BCL11B</i>	0.23	0.10	0.18	0.17
14	SNP_A-1932077 (rs220030)	15q11.2	<i>SNRPN<sup>a</sup></i>	0.18	0.18	0.15	0.17
15	SNP_A-2107106 (rs987377)	6q21	<i>AIM1<sup>a</sup>, ATG5</i>	0.16	0.15	0.20	0.17

<sup>a</sup> Known imprinted gene

flanked by trimodal levels of DNA methylation in cerebellum but not blood or cortex (Fig. 7B).

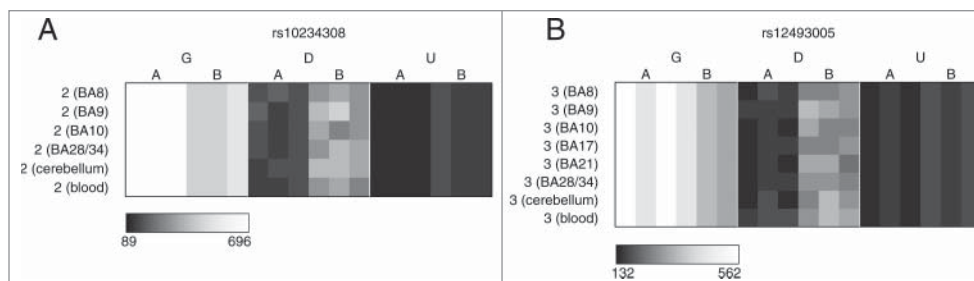
## Discussion

This study confirms the relatively widespread distribution of allelically-skewed DNA methylation in the human genome, corroborating our previous data generated in whole blood.<sup>7</sup> We also present evidence for tissue-specific differences in the quantity and distribution of ASM between different regions of the human brain, and between brain and whole blood. Our findings are in line with previous reports, confirming the importance of tissue-specific DNA methylation profiles across the brain.<sup>22,41</sup>

Although our data confirm previous studies, identifying more between-tissue variation than inter-individual variation,<sup>28</sup> we find clear examples where ASM is variable between individuals. While the number of samples profiled in this study is too small to accurately determine how much of the observed inter-individual variation in ASM results from genetic and non-genetic effects, previous studies suggest that the majority of such variation is likely to be genetically driven.<sup>7,8,11</sup> Interestingly, we identify instances of tissue-specific allelically-skewed DNA methylation resulting from both genomic imprinting and genotypic effects. For example, we observe tissue-specific variable ASM around the imprinted growth factor receptor-bound

protein 10 gene (*GRB10*), which encodes a protein that interacts with insulin-like growth factors<sup>42,43</sup> and we observe genotype-driven ASM exclusively in cerebellum for several probes within the synaptojanin 2 gene (*SYNJ2*), which encodes a protein involved in the uncoating of vesicles.<sup>44,45</sup> Such tissue-specific ASM has important implications for epigenetic epidemiology, and provides a mechanism by which genotype may exert an effect on gene function and regulation in a tissue-specific manner.

This study has a number of important limitations. First, although we used a unique set of samples comprising of matched tissues obtained from the same donors, the number of individuals profiled in our analysis was small, meaning we cannot definitively distinguish between genetic and non-genetic effects, or make broad statements about general patterns of inter-individual variation of ASM. Our Illumina 450K array validation studies were undertaken in a larger set of individuals, but could only confirm intermediate levels of DNA methylation and not detect allele-specific patterns. Second, given the limited availability of RNA from the same samples, we were unable to relate our ASM findings to allelic patterns of gene expression in the same individuals. Previous studies, however, have shown that ASM is linked to allele-specific expression of nearby genes.<sup>28</sup> Third, our analyses were undertaken on whole tissue, and represent aggregate values across a number of individual cell types. Fourth, the MSNP approach utilizes SNP



**Figure 2.** Multiple loci are characterized by consistent allelic-skewing of DNA methylation across all tissues. Heatmaps display allele signal intensities for genomic DNA (G), MSRE-digested DNA (D) and fully unmethylated, MSRE-digested DNA (U) in all tissues. A and B denote the 2 alleles of the SNP and brightness represents the quantile normalized signal intensity, with the scale displayed below the heatmap. Shown are the 2 top-ranked probes characterized by consistent ASM across tissues. These two probes were informative in (A) individual 2 for rs10234308 and (B) individual 3 for rs12493005.

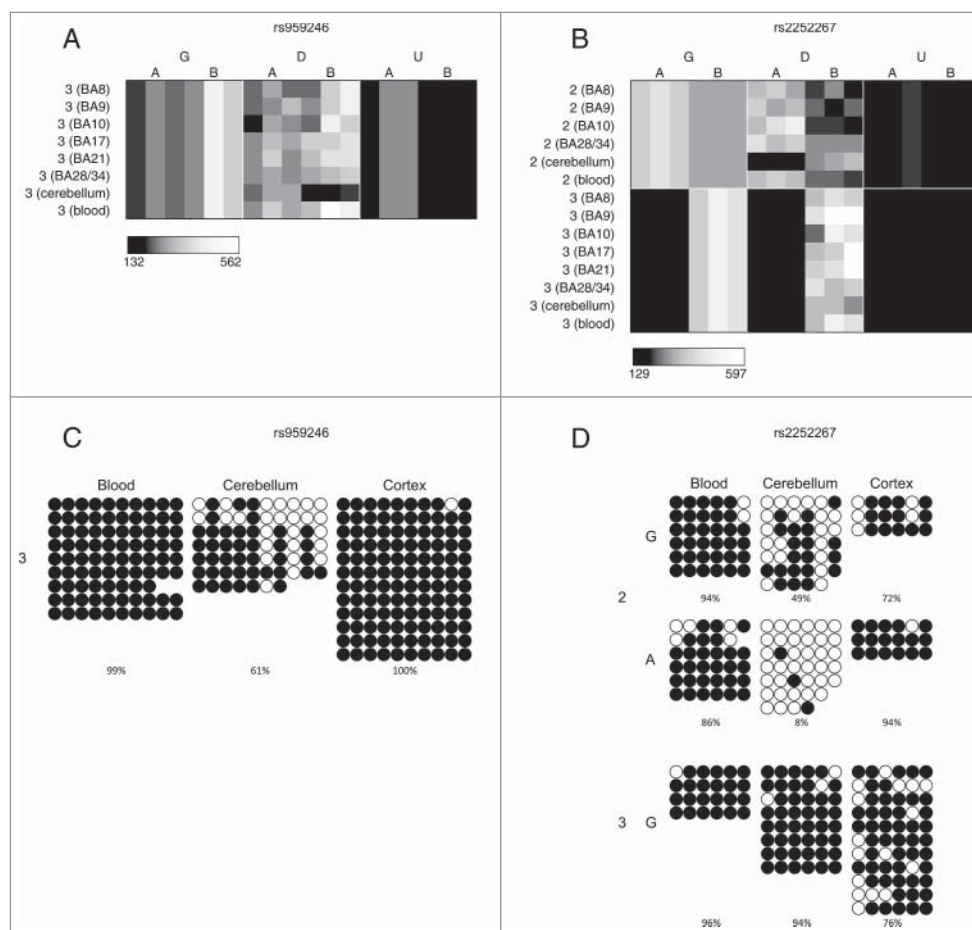
**Table 5.** Top 15 tissue-specific ASM sites, defined by highly variable ASM scores across cerebellum, whole blood, and cortex (BA9).

Rank	SNP ID	Location	Associated gene (s)	Cerebellum ASM score	Blood ASM score	BA9 ASM score	SD
1	SNP_A-4255628 (rs959246)	18q12.3	<i>SLC14A2, SETBP1</i>	-0.30	-0.03	-0.05	0.15
2	SNP_A-1946136 (rs927000)	20q13.12	<i>STK4</i>	-0.04	0.23	0.01	0.15
3	SNP_A-8438077 (rs585451)	15q21.2	<i>ATP8B4, DTWD1</i>	0.01	-0.25	0.00	0.15
4	SNP_A-8397727 (rs1123514)	5q13.3	<i>ENC1, RGENF</i>	-0.16	0.11	-0.01	0.13
5	SNP_A-2273834 (rs2252267)	14q23.1	<i>PRKCH</i>	0.19	-0.03	-0.05	0.13
6	SNP_A-4264458 (rs1358229)	4q27	<i>TRPC3, KIAA1109</i>	-0.23	0.01	0.00	0.13
7	SNP_A-8643616 (rs11700515)	21q22.3	<i>COL6A1, PCBP3</i>	0.20	-0.01	-0.03	0.13
8	SNP_A-2180729 (rs10276966)	7p15.2	<i>HIBADH, EVX1<sup>b</sup></i>	-0.01	-0.25	-0.03	0.13
9	SNP_A-4210659 (rs10517764)	4q32.2	<i>NAF1, FSTL5</i>	0.07	-0.14	-0.15	0.13
10	SNP_A-1997061 (rs10512149)	9q21.33	<i>SLC28A3, NTRK2</i>	-0.19	0.02	0.05	0.13
11	SNP_A-2160121 (rs7205794)	16q23.3	<i>CDH13, MPHOSPH6</i>	-0.24	-0.01	-0.01	0.13
12	SNP_A-8667432 (rs19782)	1p36.11	<i>C1orf201</i>	0.19	-0.04	-0.01	0.13
13	SNP_A-1787058 (rs10951911)	7p12.3	<i>TNS3, C7orf65</i>	-0.20	-0.01	0.04	0.13
14	SNP_A-1931666 (rs6707698)	2q34	<i>IKZF2, ERBB4</i>	-0.22	-0.03	0.02	0.13
15	SNP_A-8502638 (rs10989120)	9q31.1	<i>TMEFF1, C9orf30</i>	0.04	-0.20	-0.15	0.13

<sup>b</sup> Suspected imprinted gene

microarrays—these do not interrogate the whole genome, and can only assess pools of DNA molecules. Allelic patterns of DNA methylation across individual DNA molecules cannot be

directly assessed using this approach, as would be possible using bisulfite-sequencing methods. Furthermore, our threshold for calling allelic imbalances in DNA methylation is



**Figure 3.** A number of loci are characterized by allelic-skewing of DNA methylation in only one tissue with minimal ASM present in any of the other tissues examined. Heatmaps display allele signal intensities for genomic DNA (G), MSRE-digested DNA (D) and fully unmethylated, MSRE-digested DNA (U) in all tissues. A and B denote the 2 alleles of the SNP and brightness represents the quantile normalized signal intensity, with the scale displayed below the heatmap. Two of the top-ranked cerebellum-specific ASM signals are (A) rs959246 (informative for individual 3) and (B) rs2252267 (informative for individual 2). The tissue-specific patterns of DNA methylation in these 2 loci were validated by clonal bisulfite sequencing, confirming the findings from the MSNP assay (C,D). Each row represents a single DNA molecule, with black dots depicting methylated cytosines and white dots depicting unmethylated cytosines. The percentage of methylated cytosines for each sample is displayed below the plots. The amplicon spanning the DMR associated with rs959246 in (C) did not encompass a SNP enabling us to distinguish between the 2 alleles, however the methylation pattern shows evidence for tissue-specific intermediate methylation (IM) in the cerebellum. G and A in (D) denote the 2 alleles determined by SNP variation within the amplicon in heterozygous individual 2. Cerebellum-specific hypomethylation of the A allele is observed surrounding rs2252267, with individual 3 (homozygous for the G allele) being highly methylated in all 3 tissues.

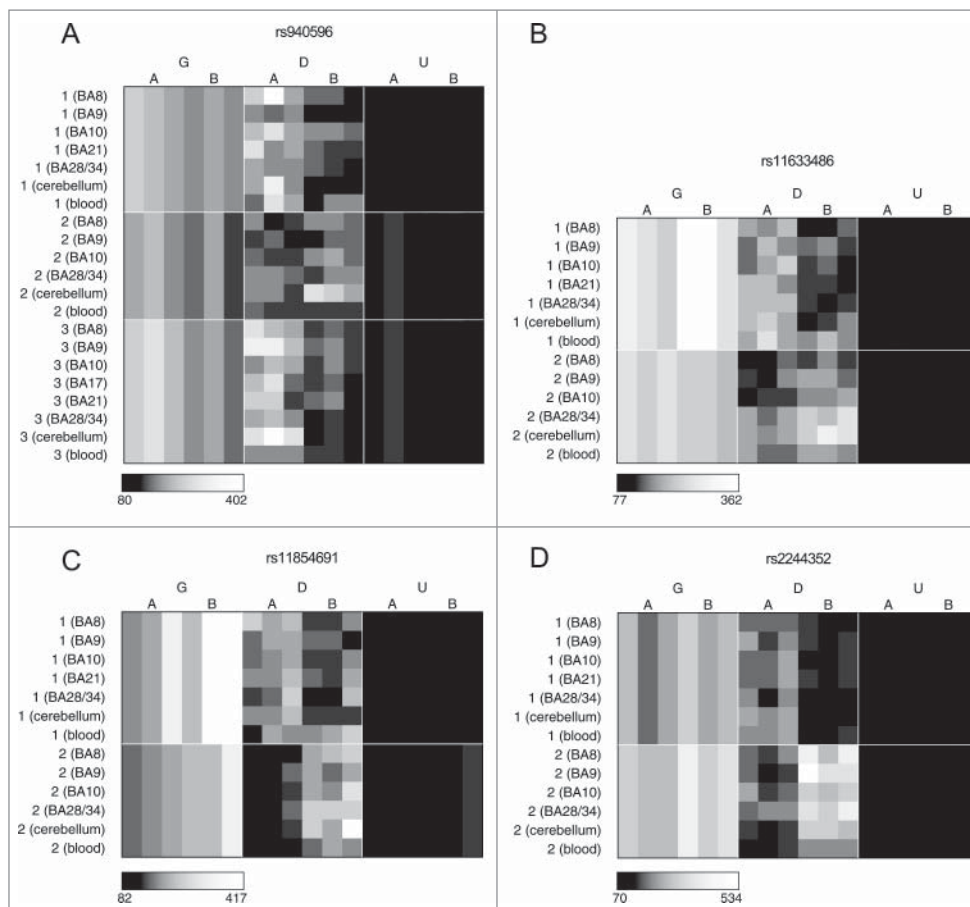
**Table 6.** Top 15 variable ASM sites, defined by average range of ASM scores between individuals across cerebellum, whole blood, and cortex (BA9).

Rank	SNP ID	Location	Associated gene(s)	Cerebellum range	Blood range	BA9 range	Average range
1	SNP_A-8579417 (rs1538116)	1q24.1	<i>MGST3, LOC400794</i>	0.24	0.35	0.32	0.30
2	SNP_A-2019421 (rs2244352)	21q22.2	<i>WRB</i>	0.36	0.24	0.29	0.30
3	SNP_A-4219174 (rs2346019)	5q31.1	<i>TGFBI, VTRNA2<sup>b</sup></i>	0.23	0.23	0.30	0.25
4	SNP_A-4208914 (rs927651)	20q13.2	<i>CYP24A1</i>	0.25	0.23	0.17	0.22
5	SNP_A-8717059 (rs12713666)	2p13.3	<i>ARHGAP25</i>	0.26	0.15	0.21	0.21
6	SNP_A-8692937 (rs4605656)	4q24	<i>CXXC4, TACR3</i>	0.16	0.18	0.25	0.20
7	SNP_A-8634251 (rs1209228)	14q24.2	<i>RGS6, SIPA1L1</i>	0.19	0.15	0.24	0.19
8	SNP_A-8329713 (rs16825906)	3q13.31	<i>LSAMP, IGSF11, LOC285194</i>	0.21	0.13	0.22	0.19
9	SNP_A-8713358 (rs16890883)	4p15.33	<i>CPEB2, LOC152742</i>	0.24	0.14	0.17	0.19
10	SNP_A-8638348 (rs4525744)	2p21	<i>SRBD1</i>	0.20	0.18	0.17	0.18
11	SNP_A-1855770 (rs3764124)	13q34	<i>CUL4A</i>	0.16	0.23	0.15	0.18
12	SNP_A-4259064 (rs7766133)	6p22.3	<i>MBOAT1</i>	0.17	0.17	0.20	0.18
13	SNP_A-2118217 (rs1695824)	1p36.33	<i>VWA1, TMEM88B</i>	0.25	0.16	0.13	0.18
14	SNP_A-1880775 (rs6116750)	20p12.3	<i>PROKR2</i>	0.23	0.13	0.18	0.18
15	SNP_A-8424056 (rs3922835)	18q12.1	<i>CDH2<sup>b</sup>, CHST9</i>	0.18	0.04	0.32	0.18

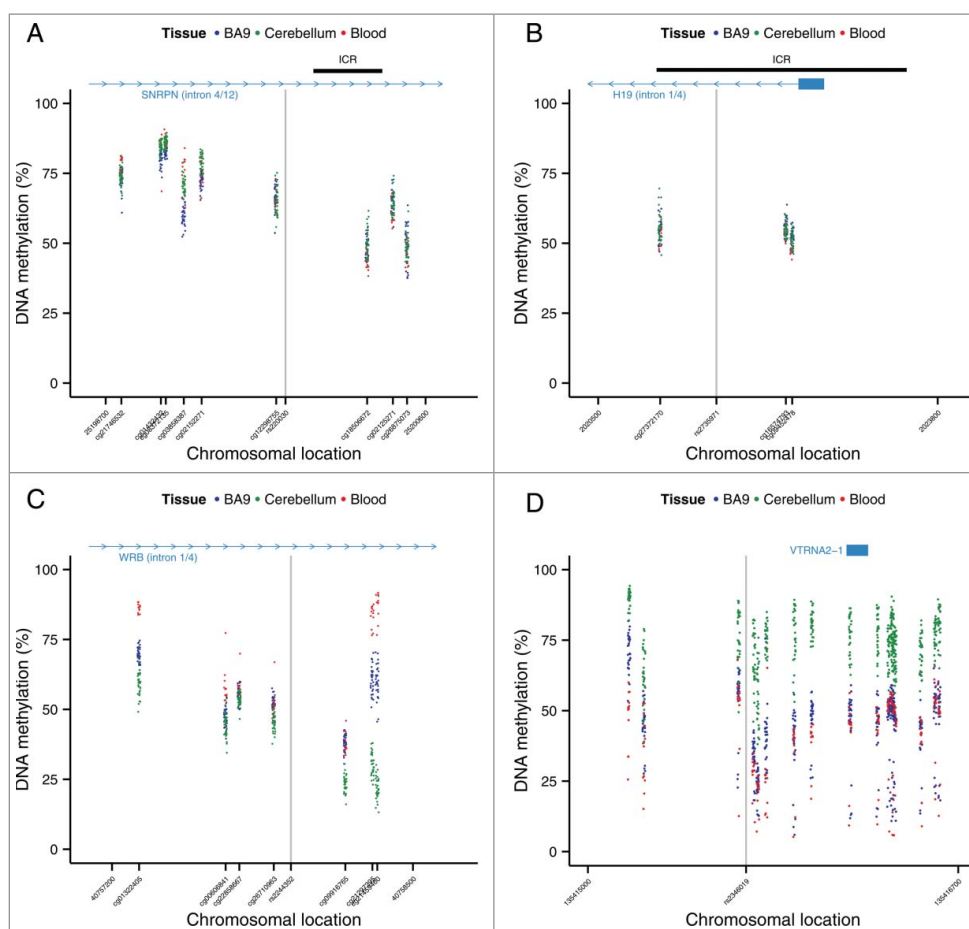
<sup>b</sup> Suspected imprinted gene

somewhat arbitrary; it is likely that our data is confounded by both false positives and negatives. We did, however, find very consistent overlap in whole blood ASM data with that reported in our previous study using the same laboratory and analysis methods,<sup>7</sup> confirming the validity of the MSNP approach. Furthermore, we validated our findings using 2 independent platforms: clonal bisulfite sequencing and the Illumina 450K

Human methylation array. Using the latter, we were able to show that variable ASM sites are located in an extended context of intermediate DNA methylation, supporting a regional regulatory role of DNA methylation in these domains, which is potentially driving intermediate expression levels in a quantitative manner across gene regulation clusters.<sup>12</sup> In addition, we observed a significant enrichment of ASM in regions



**Figure 4.** Cases of allelic-flipping of DNA methylation between individuals are found both in known imprinted gene clusters as well as regions not previously confirmed as being imprinted. The heatmaps show allele signal intensities for genomic DNA (G), MSRE-digested DNA (D) and fully unmethylated, MSRE-digested DNA (U) in all tissues. A and B denote the 2 alleles of the SNP and brightness represents the quantile normalized signal intensity, with the scale displayed below the heatmap. (A)–(C) Three probes in the vicinity of the known imprinted cluster on chromosome 15q11.2 show variable ASM and allelic-flipping in cerebellum ((A) rs940596, (B) rs11633486, (C) rs11854691). (D) Allelic-flipping of DNA methylation across all tissues is observed in the vicinity of *WRB* (rs2244352), which has not been reported as imprinted previously.



**Figure 5.** Regions around variable ASM sites are enriched for genomic domains characterized by intermediate DNA methylation. Flanking regions (1 kb) of variable ASM sites show a significant enrichment in intermediately methylated probes on the Illumina 450K Human Methylation Array ( $P$  range =  $6.82 \times 10^{-11}$  – 0.005). The scatter plots (A)–(D) show the location of genes and imprinting control regions (ICR) if overlapping the plotting window, as well as the location of the SNP from the MSNP assay (gray vertical line). These intermediate DNA methylation patterns span several known imprinted regions, for example (A) a flanking region of rs220030 in the *SNRPN* imprinted DMR, and (B) a region around rs2735971 in the imprinted gene *H19*. (C) Other sites showing allelic-flipping and intermediately methylated flanking regions lie in areas not previously known to be imprinted, for example rs2244352, which lies in an intron of *WRB*. (D) A polymorphic ASM pattern is observed in a flanking region around rs2346019, a downstream gene variant of *VTRNA2-1*, in which the majority of samples display intermediate DNA methylation in cortex (BA9) and whole blood. Of note, 4 samples show consistent hypomethylation across this region.

characterized by DHS peaks across several tissues. This enrichment of ASM in the vicinity of markers of open chromatin supports the involvement of ASM in transcriptional activity.

To conclude, we explored inter- and intra-individual variation in ASM across several regions of the human brain and whole blood from multiple individuals. Consistent with previous

studies, we find relatively widespread ASM, observing allelically-skewed DNA methylation flanking known imprinted regions, and show that ASM sites are often located in an extended genomic context of intermediate DNA methylation. Interestingly, we detect cases of genotype-driven ASM, which are also tissue-specific. These findings contribute to our understanding about the nature of differential DNA methylation across tissues and have important implications for genetic studies of complex disease. As a resource to the community, ASM patterns across each of the tissues studied are available in a searchable online database: <http://epigenetics.essex.ac.uk/ASMBrainBlood>.

## Materials and methods

### Genome-wide analysis of allelically-skewed DNA methylation

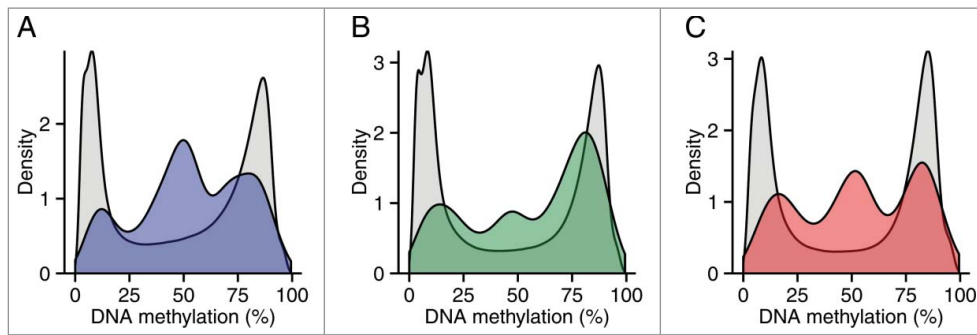
Post-mortem brain and pre-mortem whole blood samples from 2 female and one male donors were provided by the MRC London Neurodegenerative Disease Brain Bank (<http://www.kcl.ac.uk/ioppn/depts/bcn/Our-research/Neurodegeneration/brain-bank.aspx>). Subjects were approached in life for written consent for

**Table 7.** Enrichment in intermediate methylation (IM) near ASM regions.

	IM rate	O/E	$P$
Blood <sup>d</sup>	5.88%	0.98	0.576
Cerebellum <sup>d</sup>	11.49%	1.75	0.059
BA9 <sup>d</sup>	13.29%	1.45	0.065
Cross-tissue	8.11%	0.87	0.741
Tissue-specific	8.00%	0.85	0.701
Variable	38.46%	4.10	$9.56 \times 10^{-11}$
Variable (BA9) <sup>d</sup>	37.50%	4.09	$8.86 \times 10^{-9}$
Variable (cerebellum) <sup>d</sup>	16.95%	2.58	0.005
Variable (blood) <sup>d</sup>	34.48%	5.77	$6.82 \times 10^{-11}$

<sup>c</sup> IM is defined as an average methylation between 0.4 and 0.6, enrichment  $P$  values are based on a hypergeometric test based on the background distribution of IM.

<sup>d</sup> We used tissue-specific background distributions for ASM types based on single tissues.



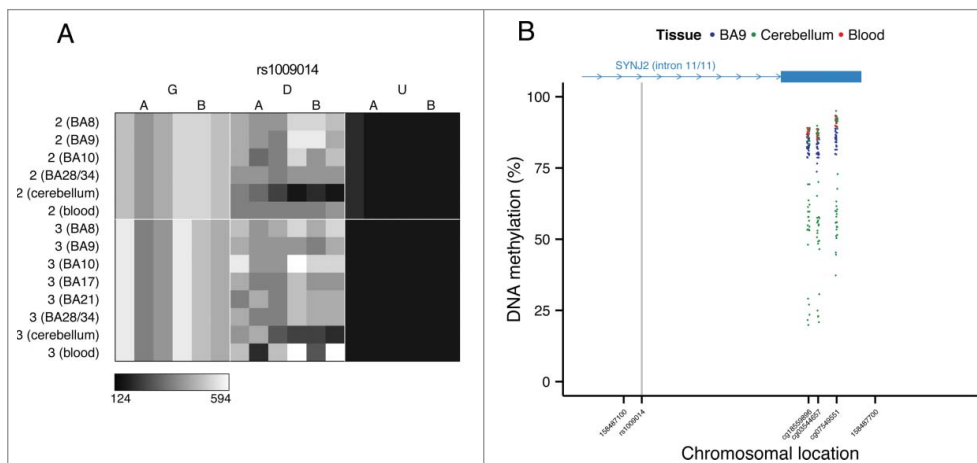
**Figure 6.** Regions around variable ASM are enriched in intermediate DNA methylation. The distributions of DNA methylation at the 65 Illumina 450K Human Methylation Array probes within 1 kb of the top 100 variable ASM sites (Table S14) show an enrichment in intermediately methylated probes compared to DNA methylation levels across the whole array (shown in gray) in (A) cortex (BA9), (B) cerebellum, and (C) whole blood.

brain banking, and all tissue donations were collected and stored following legal and ethical guidelines (NHS reference number 08/MRE09/38; the HTA license number for the LBBND brain bank is 12293). All samples were free from neuropathological and neuropsychiatric disease. A detailed list of brain regions obtained for each individual is provided in Table S18. Genomic DNA was isolated from all tissue samples using a standard phenol-chloroform protocol and assessed for purity and degradation prior to analysis (see Davies et al.<sup>22</sup> for additional information about the samples used in this study). The MSNP method, described previously,<sup>6,7</sup> was used to quantitatively assess allelic-skewing of DNA methylation across the genome. Briefly, Affymetrix Genome-wide Human SNP 6.0 Arrays were used to genotype a) DNA from each tissue sample digested with a cocktail of MSREs (*Hpa*II: 5' -C<sup>^</sup>C G G-3', *Hha*I: 5' -G C G<sup>^</sup>C-3', and *Aci*I: 5' -C<sup>^</sup>C G C-3') (D arrays), b) unmethylated whole-genome-amplified DNA for each individual digested with the same cocktail of MSREs to control for possible confounding effects of DNA sequence polymorphisms located at MSRE cut-sites (U arrays), and c) genomic DNA from each of the 3 individuals to identify heterozygous (informative) SNPs (G arrays). Unmethylated DNA was produced by whole-genome amplifying 100 ng cerebellum DNA using the Qiagen RepliG kit (Qiagen, Crawley, UK) using the manufacturer's protocol. In total 28 genotyping arrays were

processed: 22 D arrays (DNA from between 6 and 7 brain regions plus whole blood, for each individual), 3 U arrays (one for each individual), and 3 G arrays (one for each individual). Additional methodological details are available in Schalkwyk et al.<sup>7</sup>

### Selection of informative SNPs and quantification of ASM

To be informative in the ASM assay, SNPs must be heterozygous, and the amplicon must contain an MSRE cut site.<sup>6,7</sup> To guard against poorly performing SNP probes we also removed consistently low signal intensity SNPs across the 22 G arrays, and those yielding a highly variable U/G signal ratio ( $SD > 0.077$ ) across all samples. A total of 220,449 SNPs passed our stringent filtering criteria and were classified as informative and heterozygous in at least one individual. The number of informative SNPs in each of the individual samples profiled by the MSNP method is shown in Table S19. Quantitative measures of ASM were derived by comparing signal intensities between the D (MSRE digested) and G (genomic DNA) arrays using the SNPMAp package (v1.02) in R that was developed for the estimation of allele frequencies in DNA pools genotyped on SNP arrays.<sup>46</sup> Briefly, relative allele score (RAS) values were generated for all SNPs on the array, which are defined as  $A/(A + B)$ , where A and B are the intensities of the probes for the 2 alleles of



**Figure 7.** Genotype-driven ASM can be tissue-specific. (A) The tissue-specific ASM site rs1009014, located in an intron of *SYNJ2*, shows allelic-skewing of DNA methylation in cerebellum (ASM score = 0.16) but not in cortex (ASM score = 0.04) or whole blood (ASM score = 0.03). (B) DNA methylation levels for Illumina 450K Human Methylation Array probes in cerebellum across a flanking region of this locus exhibit a genotype-driven tissue-specific DNA methylation pattern. The scatter plot shows the location of *SYNJ2* (transcript variant 1), as well as the location of the informative SNP from the MSNP assay (gray vertical line).

a given SNP. For a given SNP in a heterozygous individual, ASM (or allelic-skewing of DNA methylation) is detected as a difference in RAS between the G and D arrays. We call this difference in RAS “ASM score” and define probes showing an absolute ASM score  $\geq 0.10$  as “allelically-skewed.” A UCSC custom annotation track showing the location of all 220,449 loci and the degree of allelic-skewing in DNA methylation across each tissue and individual is available for download from our website (<http://epigenetics.essex.ac.uk/ASMBrainBlood>). Enrichment analyses were performed using a Kruskal-Wallis rank-sum test for ASM rank differences between the annotated genic regions. This non-parametric method tests whether multiple samples were drawn from the same distribution and is the multivariate extension of the better-known Wilcoxon rank-sum test. This test allowed us to avoid selecting a specific threshold for ASM scores and does not assume a normal distribution of residuals. Of the 220,449 informative probes, 219,921 could be annotated to specific defined genic regions. Annotations were based on the *Homo sapiens* hg19 build from UCSC using the AnnotationHub Bioconductor package<sup>47</sup> classifying probes as residing in introns ( $n = 100,254$ ), 5'UTRs ( $n = 347$ ), 3'UTRs ( $n = 2,341$ ), coding regions ( $n = 2,569$ ), intergenic regions ( $n = 110,186$ ) and promoters ( $n = 4,224$ ). A Nemenyi test for pairwise multiple comparisons of mean rank sums as implemented in the PMCMR R package<sup>48</sup> was used for post-hoc comparisons. ENCODE tracks for DHS peaks in frontal cortex, frontal cerebrum, cerebellum, CD14+ monocytes, naïve B cells, H1 human embryonic stem cells (H1-hesc), heart, and fibroblasts were obtained from the UCSC genome browser (<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeOpenChromDnase/>). For the DHS enrichment analyses we tested whether the rank-sums for ASM scores differed significantly in informative probes defined by the presence or absence of DHS peaks using a Wilcoxon rank-sum test. Informative probes were ranked according to ASM scores with higher absolute ASM scores corresponding to lower ranks.

### Clonal bisulfite sequencing

Two regions were subsequently selected for clonal bisulfite sequencing analysis to further verify our findings and determine the precise allele-specific patterns of DNA methylation. Following sodium bisulfite treatment and bisulfite-PCR amplification, amplicons were cloned using the TOPO TA cloning method (Invitrogen, Paisley, UK) and sequenced with BigDye v1.1 sequencing chemistry (Applied Biosystems) (Table S20). Sequencing traces were visualized, quality controlled, and aligned using BiQ Analyzer.<sup>49</sup> All data were tested for complete sodium bisulfite conversion, with an overall conversion rate  $> 99.9\%$  estimated by BiQ Analyzer.

### Validation of ASM on the Illumina 450K HumanMethylation microarray

Further analysis of ASM sites was undertaken on a larger collection of post-mortem brain samples ( $n = 34$ ), comprising BA9, BA21, BA28/34, and cerebellum, which were also free of any neuropathology and neuropsychiatric disease. Additionally we analyzed matched pre-mortem whole blood samples, which

were available for a subset ( $n = 8$ ), as well as 5 unmatched blood samples (see Table S18). The 3 individuals profiled by MSNP were included in this analysis. DNA (500 ng) from each sample was treated with sodium bisulfite in duplicate, using the EZ-96 DNA methylation kit (Zymo Research, CA, USA). DNA methylation was quantified using the Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., CA, USA) run on an Illumina HiScan System (Illumina) using the manufacturers' standard protocol, with pre-processing and stringent quality control performed as previously described.<sup>50</sup> We used the GenomicRanges package<sup>51</sup> to extract data for all CpG sites within 1 kb of candidate ASM SNPs and examined patterns of DNA methylation across the 3 tissues. Intermediate DNA methylation was defined as an average methylation value between 0.4 and 0.6 across all individuals. To test for statistical significance of enrichment in intermediately methylated probes we used a hypergeometric distribution based on the number of probes tested and the background of intermediately methylated probes across the whole array. Annotation of genes in the methylation plots (Fig. 5 and Fig. 7B) was obtained from the UCSC Genome Browser hg19 assembly. Imprinting control region (ICR) annotation was obtained from the web resource on human DMRs provided by the Department of Medical and Molecular Genetics, Kings College London (<https://atlas.genetics.kcl.ac.uk>) and lifted over from hg18 to hg19.

### Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

### Acknowledgments

This work was supported by grants from the UK Medical Research Council (MRC) (grant number MR/K013807/1) and US National Institutes of Health (grant number AG036039) to JM. SJM is funded by the EU-FP7 project EpiTrain (REA grant agreement n° 316758).

### References

1. Ferguson-Smith AC, Sasaki H, Cattanach BM, Surani MA. Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature* 1993; 362:751-5; PMID:8469285; <http://dx.doi.org/10.1038/362751a0>
2. Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature* 1997; 389:745-9; PMID:9338788; <http://dx.doi.org/10.1038/39631>
3. Lee JT, Bartolomei MS. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* 2013; 152:1308-23; PMID:23498939; <http://dx.doi.org/10.1016/j.cell.2013.02.016>
4. Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genet* 2011; 12:565-75; PMID:21765458; <http://dx.doi.org/10.1038/nrg3032>
5. Barlow DP, Bartolomei MS. Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol* 2014; 6:pii: a018382; PMID:24492710; <http://dx.doi.org/10.1101/cshperspect.a018382>
6. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, Li K, Murty VV, Schupf N, Vilain E, et al. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat Genet* 2008; 40:904-8; PMID:18568024; <http://dx.doi.org/10.1038/ng.174>
7. Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN, Plomin R, Mill J. Allelic skewing of DNA methylation is

- widespread across the genome. *Am J Hum Genet* 2010; 86:196-212; PMID:20159110; <http://dx.doi.org/10.1016/j.ajhg.2010.01.014>
8. Shoemaker R, Deng J, Wang W, Zhang K. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res* 2010; 20:883-9; PMID:20418490; <http://dx.doi.org/10.1101/gr.104695.109>
  9. Tycko B. Allele-specific DNA methylation: beyond imprinting. *Hum Mol Genet* 2010; 19:R210-20; PMID:20855472; <http://dx.doi.org/10.1093/hmg/ddq376>
  10. Gertz J, Varley KE, Reddy TE, Bowling KM, Pauli F, Parker SL, Kucera KS, Willard HF, Myers RM. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. *PLoS genetics* 2011; 7:e1002228; PMID:21852959; <http://dx.doi.org/10.1371/journal.pgen.1002228>
  11. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, Gilad Y, Pritchard JK. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome biology* 2011; 12:R10; PMID:21251332; <http://dx.doi.org/10.1186/gb-2011-12-1-r10>
  12. Elliott G, Hong C, Xing X, Zhou X, Li D, Coarfa C, Bell RJ, Maire CL, Ligon KL, Sigaroudinia M, et al. Intermediate DNA methylation is a conserved signature of genome regulation. *Nat Commun* 2015; 6:6363; PMID:25691127; <http://dx.doi.org/10.1038/ncomms7363>
  13. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 2007; 447:425-32; PMID:17522676; <http://dx.doi.org/10.1038/nature05918>
  14. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 2008; 454:766-70; PMID:18600261; <http://dx.doi.org/10.1038/nature07107>
  15. Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol Cell* 2012; 48:849-62; PMID:23219530; <http://dx.doi.org/10.1016/j.molcel.2012.11.001>
  16. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, et al. Global epigenomic reconfiguration during mammalian brain development. *Science* 2013; 341:1237905; PMID:23828890; <http://dx.doi.org/10.1126/science.1237905>
  17. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 2009; 462:315-22; PMID:19829295; <http://dx.doi.org/10.1038/nature08514>
  18. Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, De Jager PL, Rosen ED, Bennett DA, Bernstein BE, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* 2013; 500:477-81; PMID:23925113; <http://dx.doi.org/10.1038/nature12433>
  19. Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, Cross MK, Williams BA, Stamatoyannopoulos JA, Crawford GE, et al. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res* 2013; 23:555-67; PMID:23325432; <http://dx.doi.org/10.1101/gr.147942.112>
  20. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015; 518:317-30; PMID:25693563; <http://dx.doi.org/10.1038/nature14248>
  21. Ladd-Acosta C, Pevsner J, Sabuncian S, Yolken RH, Webster MJ, Dinkins T, Callinan PA, Fan JB, Potash JB, Feinberg AP. DNA methylation signatures within the human brain. *American journal of human genetics* 2007; 81:1304-15; PMID:17999367; <http://dx.doi.org/10.1086/524110>
  22. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, Coarfa C, Harris RA, Milosavljevic A, Troakes C, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology* 2012; 13:R43; PMID:22703893; <http://dx.doi.org/10.1186/gb-2012-13-6-r43>
  23. Takizawa T, Nakashima K, Namihira M, Ochiai W, Uemura A, Yanagisawa M, Fujita N, Nakao M, Taga T. DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain. *Developmental cell* 2001; 1:749-58; PMID:11740937; [http://dx.doi.org/10.1016/S1534-5807\(01\)00101-0](http://dx.doi.org/10.1016/S1534-5807(01)00101-0)
  24. Kozlenkov A, Roussos P, Timashpolsky A, Barbu M, Rudchenko S, Bibikova M, Klotzle B, Byne W, Lyddon R, Di Narzo AF, et al. Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. *Nucleic Acids Res* 2014; 42:109-27; PMID:24057217; <http://dx.doi.org/10.1093/nar/gkt838>
  25. Ziller MJ, Edri R, Yaffe Y, Donaghey J, Pop R, Mallard W, Issner R, Gifford CA, Goren A, Xing J, et al. Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* 2015; 518:355-9; PMID:25533951; <http://dx.doi.org/10.1038/nature13990>
  26. Mo A, Mukamel EA, Davis FP, Luo C, Henry GL, Picard S, Urich MA, Nery JR, Sejnowski TJ, Lister R, et al. Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron* 2015; 86:1369-84; PMID:26087164; <http://dx.doi.org/10.1016/j.neuron.2015.05.018>
  27. Paliwal A, Temkin AM, Kerker K, Yale A, Yotova I, Drost N, Lax S, Nhan-Chang CL, Powell C, Borczuk A, et al. Comparative anatomy of chromosomal domains with imprinted and non-imprinted allele-specific DNA methylation. *PLoS genetics* 2013; 9:e1003622; PMID:24009515; <http://dx.doi.org/10.1371/journal.pgen.1003622>
  28. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* 2015; 523(7559):212-6; PMID:26605523; <http://dx.doi.org/10.1038/nature16179>
  29. Prickett AR, Oakey RJ. A survey of tissue-specific genomic imprinting in mammals. *Mol Genet Genomics* 2012; 287:621-30; PMID:22821278; <http://dx.doi.org/10.1007/s00438-012-0708-6>
  30. Mancini-Dinardo D, Steele SJ, Levorse JM, Ingram RS, Tilghman SM. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* 2006; 20:1268-82; PMID:16702402; <http://dx.doi.org/10.1101/gad.1416906>
  31. Kelsey G. Imprinting on chromosome 20: tissue-specific imprinting and imprinting mutations in the GNAS locus. *Am J Med Genet C Semin Med Genet* 2010; 154C:377-86; PMID:20803660; <http://dx.doi.org/10.1002/ajmg.c.30271>
  32. Williamson CM, Ball ST, Dawson C, Mehta S, Beechey CV, Fray M, Teboul L, Dear TN, Kelsey G, Peters J. Uncoupling antisense-mediated silencing and DNA methylation in the imprinted *Gnas* cluster. *PLoS genetics* 2011; 7:e1001347; PMID:21455290; <http://dx.doi.org/10.1371/journal.pgen.1001347>
  33. Arnaud P, Monk D, Hitchins M, Gordon E, Dean W, Beechey CV, Peters J, Craigen W, Preece M, Stanier P, et al. Conserved methylation imprints in the human and mouse *GRB10* genes with divergent allelic expression suggests differential reading of the same mark. *Hum Mol Genet* 2003; 12:1005-19; PMID:12700169; <http://dx.doi.org/10.1093/hmg/ddg110>
  34. Grundberg E, Meduri E, Sandling JK, Hedman AK, Keildson S, Buil A, Busche S, Yuan W, Nisbet J, Sekowska M, et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am J Hum Genet* 2013; 93:876-90; PMID:24183450; <http://dx.doi.org/10.1016/j.ajhg.2013.10.004>
  35. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet* 2013; 14:585-94; PMID:23817309; <http://dx.doi.org/10.1038/nrg3405>
  36. Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics* 2010; 5:578-82; PMID:20716955; <http://dx.doi.org/10.4161/epi.5.7.12960>
  37. Hutchinson JN, Raj T, Fagerness J, Stahl E, Vilorio FT, Gimelbrant A, Seddon J, Daly M, Chess A, Plenge R. Allele-specific methylation occurs at genetic variants associated with complex disease. *PLoS one*

- 2014; 9:e98464; PMID:24911414; <http://dx.doi.org/10.1371/journal.pone.0098464>
38. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; 489:57-74; PMID:22955616; <http://dx.doi.org/10.1038/nature11247>
39. Treppendahl MB, Qiu X, Sogaard A, Yang X, Nandrup-Bus C, Hother C, Andersen MK, Kjeldsen L, Mollgard L, Hellstrom-Lindberg E, et al. Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML. *Blood* 2012; 119:206-16; PMID:22058117; <http://dx.doi.org/10.1182/blood-2011-06-362541>
40. Romanelli V, Nakabayashi K, Vizoso M, Moran S, Iglesias-Platas I, Sugahara N, Simon C, Hata K, Esteller M, Court F, et al. Variable maternal methylation overlapping the nc886/vtRNA2-1 locus is locked between hypermethylated repeats and is frequently altered in cancer. *Epigenetics* 2014; 9:783-90; PMID:24589629; <http://dx.doi.org/10.4161/epi.28323>
41. Illingworth RS, Gruenewald-Schneider U, De Sousa D, Webb S, Merusi C, Kerr AR, James KD, Smith C, Walker R, Andrews R, et al. Inter-individual variability contrasts with regional homogeneity in the human brain DNA methylome. *Nucleic Acids Res* 2015; 43:732-44; PMID:25572316; <http://dx.doi.org/10.1093/nar/gku1305>
42. Holt LJ, Siddle K. Grb10 and Grb14: enigmatic regulators of insulin action—and more? *Biochem J* 2005; 388:393-406; PMID:15901248; <http://dx.doi.org/10.1042/BJ20050216>
43. Garfield AS, Cowley M, Smith FM, Moorwood K, Stewart-Cox JE, Gilroy K, Baker S, Xia J, Dalley JW, Hurst LD, et al. Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. *Nature* 2011; 469:534-8; PMID:21270893; <http://dx.doi.org/10.1038/nature09651>
44. Rusk N, Le PU, Mariggio S, Guay G, Lurisci C, Nabi IR, Corda D, Symons M. Synaptojanin 2 functions at an early step of clathrin-mediated endocytosis. *Curr Biol* 2003; 13:659-63; PMID:12699622; [http://dx.doi.org/10.1016/S0960-9822\(03\)00241-0](http://dx.doi.org/10.1016/S0960-9822(03)00241-0)
45. Verstreken P, Koh TW, Schulze KL, Zhai RG, Hiesinger PR, Zhou Y, Mehta SQ, Cao Y, Roos J, Bellen HJ. Synaptojanin is recruited by endophilin to promote synaptic vesicle uncoating. *Neuron* 2003; 40:733-48; PMID:14622578; [http://dx.doi.org/10.1016/S0896-6273\(03\)00644-5](http://dx.doi.org/10.1016/S0896-6273(03)00644-5)
46. Davis OS, Plomin R, Schalkwyk LC. The SNPMap package for R: a framework for genome-wide association using DNA pooling on microarrays. *Bioinformatics* 2009; 25:281-3; PMID:19008252; <http://dx.doi.org/10.1093/bioinformatics/btn587>
47. Morgan M, Carlson M, Tenenbaum D, Arora S. *AnnotationHub: Client to access AnnotationHub resources*. R package version 2.2.2.
48. Pohlert T (2014). *The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR)*. R package, <http://CRAN.R-project.org/package=PMCMR>.
49. Bock C, Reither S, Mikeska T, Paulsen M, Walter J, Lengauer T. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* 2005; 21:4067-8; PMID:16141249; <http://dx.doi.org/10.1093/bioinformatics/bti652>
50. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC genomics* 2013; 14:293; PMID:23631413; <http://dx.doi.org/10.1186/1471-2164-14-293>
51. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. Software for computing and annotating genomic ranges. *PLoS Comput Biol* 2013; 9:e1003118; PMID:23950696; <http://dx.doi.org/10.1371/journal.pcbi.1003118>