Allele-specific methylation across brain and blood: Identification of tissue-specific DMRs

Sarah Marzi

31/05/2014
Allele-specific DNA methylation

- DNA methylation is symmetric on both alleles across most of the human genome
- There are exceptions:
  - X-chromosome inactivation
  - Genomic imprinting (DMRs)
  - Genotype-effects on DNA methylation status (cis, trans)
Previous findings

- ASM is **common** (>35,000 sites)
- It is **quantitative** rather than absolute
- Much of it is **genotype driven** (in cis), not just limited to X-inactivation and imprinting
- Heterogeneity of genome-wide methylation across **tissues** and **individuals**

---

**ARTICLE**

Allelic Skewing of DNA Methylation Is Widespread across the Genome

Matthew N Davies, Manuela Volta, Ruth Pidsley, Katie Lunnion, Abhishek Dixit, Simon Lovestone, Cristian Coarfa, R Alan Harris, Aleksandar Milosavljevic, Claire Troakes, Safa Al-Sarraj, Richard Dobson, Leonard C Schalkwyk, and Jonathan Mill

Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood
Implications of ASM for integrated genetic-epigenetic association studies of common disease

- Genotype-driven ASM as functional mechanism of trait-associated loci?
- ASM quantitative -> dilutes association
- Non-cis ASM might explain missing heritability (hemizygous loci)
- Methylation variability across tissues is known to be high (Davies et al., 2012)
- Studies of tissue-specific ASM haven’t been published yet to our knowledge
Aim of current study

- Study ASM across different **brain regions and blood** obtained from the same individuals
- Identify regions of **tissue-specific ASM**
- Identify regions of **polymorphic ASM**
- Identify contributions of genotype and non-genotype driven ASM
  - Stochastic?
  - Parental origin?
Samples

- 3 elderly control individuals from the MRC London Neurodegenerative Disease Brain Bank
- Post-mortem brain samples
  - BA9 (Dorsolateral PFC)
  - BA10 (Anterior PFC)
  - BA8 (Frontal eye fields)
  - Superior Temporal Gyrus (STG)
  - Entorhinal Cortex (EHC)
  - Visual Cortex
  - Cerebellum
  - Whole blood

- For secondary analyses: 38 control individuals with DNA methylation data from the Illumina 450K human methylation array in multiple brain tissues and blood
Heterozygous AC call

Digestion with methylation-sensitive restriction enzymes

Amplicon not created in Affymetrix protocol

Homozygous A call
Change in Relative Allele Score (RAS)

StyI/NspI StyI/NspI

Heterozygous AC call

A C

Digestion with methylation-sensitive restriction enzymes

Skewed call, A stronger than C

HpaII/HhaI/AciI

Amplicon not created in Affymetrix protocol

HpaII/Hhal/Acil
Allelic skewing of DNA methylation is widespread, a significant amount is tissue-specific

- 220,450 SNPs on the array classified as informative
- 9,311 (4.22%) of these showed allelic skewing of DNA methylation in at least one individual and sample with a change in RAS >0.1
- 57 loci show ASM across all informative amplicons (consistent direction)
- ~50% of ASM is tissue-specific within any one individual, the biggest part of which is found in blood
- Cortex regions show very similar ASM patterns
rs959246

Individual 1
- Cortex
- Cerebellum
- Blood

Individual 46
- Cortex
- Cerebellum
- Blood

Individual 5
- Cortex
- Cerebellum
- Blood

SNP_A-4255628
Clonal bisulfite sequencing verified findings from microarray screen

CpG sites around rs959246
## Top blood ASM sites

<table>
<thead>
<tr>
<th>Rank</th>
<th>SNP ID</th>
<th>Location</th>
<th>Nearest Gene(s) [bp]</th>
<th>Schalkwyk et al (2010)</th>
<th>Blood ASM Score</th>
<th>Cerebellum ASM Score</th>
<th>BA9 ASM Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SNP_A-2180729</td>
<td>7p15.2</td>
<td>HIBADH, EVX1</td>
<td>0.29</td>
<td>0.25</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>2</td>
<td>SNP_A-8438077</td>
<td>15q21.2</td>
<td>ATP8B4, DTWD1</td>
<td>NA</td>
<td>0.25</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>3</td>
<td>SNP_A-1841543</td>
<td>7p15.3</td>
<td>MGC87042</td>
<td>0.34</td>
<td>0.23</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>4</td>
<td>SNP_A-1946136</td>
<td>20q13.12</td>
<td>STK4</td>
<td>NA</td>
<td>0.23</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>SNP_A-8450837</td>
<td>16q12.2</td>
<td>SLC6A2, LPCAT2</td>
<td>0.28</td>
<td>0.22</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>SNP_A-4279002</td>
<td>1q41</td>
<td>CENPF, KCNK2</td>
<td>NA</td>
<td>0.22</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>7</td>
<td>SNP_A-8450539</td>
<td>1p36.12</td>
<td>CAMK2N1, LOC339505</td>
<td>NA</td>
<td>0.22</td>
<td>0.11</td>
<td>0.15</td>
</tr>
<tr>
<td>8</td>
<td>SNP_A-8633222</td>
<td>5q33.3</td>
<td>CCNJL</td>
<td>NA</td>
<td>0.22</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>9</td>
<td>SNP_A-8472169</td>
<td>10q23.1</td>
<td>NRG3</td>
<td>NA</td>
<td>0.22</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>10</td>
<td>SNP_A-2071005</td>
<td>3q28</td>
<td>UTS2D</td>
<td>NA</td>
<td>0.22</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>11</td>
<td>SNP_A-2185394</td>
<td>7p22.1</td>
<td>RNF216</td>
<td>NA</td>
<td>0.22</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>12</td>
<td>SNP_A-8702215</td>
<td>19p13.11</td>
<td>FCBO1, MAP15</td>
<td>0.19</td>
<td>0.22</td>
<td>0.00</td>
<td>0.07</td>
</tr>
<tr>
<td>13</td>
<td>SNP_A-1807006</td>
<td>8p23.3</td>
<td>CLN8, DLGAP2</td>
<td>NA</td>
<td>0.22</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>14</td>
<td>SNP_A-8326632</td>
<td>2q31.1</td>
<td>HOXD3, HOXD4</td>
<td>NA</td>
<td>0.21</td>
<td>0.24</td>
<td>0.19</td>
</tr>
<tr>
<td>15</td>
<td>SNP_A-2008150</td>
<td>11p15.1</td>
<td>OTOG, MYOD1, USH1C</td>
<td>0.22</td>
<td>0.21</td>
<td>0.16</td>
<td>0.07</td>
</tr>
</tbody>
</table>
### Top cortex (BA9) ASM sites

<table>
<thead>
<tr>
<th>Rank</th>
<th>SNP ID</th>
<th>Location</th>
<th>Nearest Gene(s)</th>
<th>BA9 ASM Score</th>
<th>Cerebellum ASM Score</th>
<th>Blood ASM Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SNP_A-8625237 (rs12493005)</td>
<td>3q26.32</td>
<td>TBL1XR1</td>
<td>0.24</td>
<td>0.24</td>
<td>0.20</td>
</tr>
<tr>
<td>2</td>
<td>SNP_A-8463467 (rs17164474)</td>
<td>7q35</td>
<td>OR2F2, OR2F1</td>
<td>0.23</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>3</td>
<td>SNP_A-1841543 (rs10234308)</td>
<td>7p15.3</td>
<td>MGC87042</td>
<td>0.23</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>4</td>
<td>SNP_A-8652129 (rs2479084)</td>
<td>1p36.21</td>
<td>FHAD1</td>
<td>0.22</td>
<td>0.13</td>
<td>0.18</td>
</tr>
<tr>
<td>5</td>
<td>SNP_A-8301602 (rs398225)</td>
<td>3p25.3</td>
<td>SRGAP3, RAD18</td>
<td>0.21</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>6</td>
<td>SNP_A-2107106 (rs987377)</td>
<td>6q21</td>
<td>AIM1, ATG5</td>
<td>0.20</td>
<td>0.16</td>
<td>0.15</td>
</tr>
<tr>
<td>7</td>
<td>SNP_A-2052542 (rs3098382)</td>
<td>5q13.2</td>
<td>MAP1B</td>
<td>0.19</td>
<td>0.24</td>
<td>0.04</td>
</tr>
<tr>
<td>8</td>
<td>SNP_A-2307481 (rs716591)</td>
<td>15q26.2</td>
<td>LOC400456, MCTP2</td>
<td>0.19</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>9</td>
<td>SNP_A-8643280 (rs3121125)</td>
<td>1q21.1</td>
<td>HFE2, NBPF10</td>
<td>0.19</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>10</td>
<td>SNP_A-8326632 (rs1542180)</td>
<td>2q31.1</td>
<td>HOXD3, HOXD4</td>
<td>0.19</td>
<td>0.24</td>
<td>0.21</td>
</tr>
<tr>
<td>11</td>
<td>SNP_A-1998023 (rs9722212)</td>
<td>9q34.11</td>
<td>TOR1B, PTGES</td>
<td>0.19</td>
<td>0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>12</td>
<td>SNP_A-8640607 (rs12632177)</td>
<td>3q27.1</td>
<td>MCF2L2</td>
<td>0.18</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>13</td>
<td>SNP_A-4291638 (rs12670584)</td>
<td>7p13</td>
<td>YKT6</td>
<td>0.18</td>
<td>0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>14</td>
<td>SNP_A-1892234 (rs17303015)</td>
<td>5p12</td>
<td>MGC42105</td>
<td>0.18</td>
<td>0.11</td>
<td>0.03</td>
</tr>
<tr>
<td>15</td>
<td>SNP_A-8653671 (rs4871852)</td>
<td>8p21.3</td>
<td>TNFRSF10D</td>
<td>0.18</td>
<td>0.09</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Replication on Illumina 450K human methylation array

- CpG sites on the array were identified within 1kb distance of the top ASM SNP hits for
  - Top 100 **cross tissue** ASM sites (defined by mean)
    - 58 probes, 50 differentially methylated (86%)
  - Top 100 **tissue-specific** ASM sites (defined by SD)
    - 58 probes, 53 differentially methylated (91%)
  - Top 100 **polymorphic** ASM sites (defined by range)
    - 69 probes, 67 differentially methylated (97%)
- Expected tissue-specific DMR rate: 0.56
- P<0.001 by random sampling
Enrichment in tissue-specific DMRs in vicinity of tissue-specific ASM

cg03557916
Evidence of differential methylation between cerebellum and STG, 711 bp from rs14067
DCUN1D2
Enrichment in hemimethylated probes

- DNA methylation density of 68 CpG sites in vicinity of top 100 polymorphic ASM sites
- All 5 tissues show trimodal methylation distribution
  - A = BA9
  - E = EHC
  - F = STG
  - H = Cerebellum
Enrichment in hemimethylated probes

cg19131313
Hemimethylation in both cerebellum and STG, 17 bp from rs10481354
CLN8, DLGAP2
Identification of CpG sites with indication of genotype-driven ASM

cg02380521
Indication of allele-specific methylation in both cerebellum and STG,
262 bp from rs12978286
FCHO1, MAP1S
Identification of CpG sites with tissue-specific genotype-driven ASM

cg18559896
Indication of allele-specific methylation in cerebellum but not STG,
398 bp from rs1009014
SYNJ2
Conclusion

- ASM is widespread across the genome
- A significant amount of it seems to be tissue-specific
- ASM regions are enriched in tissue-specific DMRs
- Abundance of hemimethylated probes
  - Parental-origin effects?
  - Stochastic ASM?
- Identification of what appears to be genotype-driven ASM across tissues and tissue-specifically
Discussion

- What proportion of ASM occurrences are tissue-general (more flexible definition?)
- Are they particularly interesting? Connection to lncRNAs?
- Systematic overlap with pathways and known imprinted genes?
- Consequences for GWAS
  - Quantitative association
  - Functional consequence of non-coding SNPs
  - Hemizygosity
  - Integrated genetic-epiallelic approach needed
- Selection of disease relevant tissues
- Sample size, more tissues
Acknowledgement

Leonard Schalkwyk

Jonathan Mill

Emma Meaburn

Psychiatric Epigenetics Group

university of
EXETER
MEDICAL SCHOOL

Institute of Psychiatry
at The Maudsley

KING'S COLLEGE
LONDON

ROADMAP epigenomics project

EpiTrain
Training in Epigenetics of Common Disease