Epigenetic variation and complex disease risk

Caroline Relton
Institute of Human Genetics
Newcastle University

ALSPAC Research Symposium
2 & 3 March 2009
Missing heritability

“Even when dozens of genes have been linked to a trait, both the individual and cumulative effects are disappointingly small and nowhere near enough to explain earlier estimates of heritability”

Francis Collins

1. Rare variants
2. Copy number variants
3. Epistasis
4. Epigenetics
5. Something else
Epigenetics and complex disease

- Epigenotype is influenced by environmental factors and stochastic events
- Epigenotype is inherited through the germ-line
- Epigenotype can influence phenotype

- Central & Eastern European Lung Cancer Case Control Study
- North Cumbria Community Genetics Project
- Newcastle Thousand Families Study
- ALSPAC
Causality

- Epigenotype can be considered an intermediate phenotype
- Is epigenotype on the causal pathway to disease?
Epigenotype is influenced by environmental factors and stochastic events
Determinants of DNA methylation

- Nutrition
- Age
- Genetic factors
- Stress
- Smoking
- Infection
- Disease status

Exposure → Δ DNA methylation → Phenotype
Age-related change in methylation

% Methylation (Median IQR)

NCCG Pi (0y)
NCCG Pm (17-40y)
NTFS (50y)

CpG 1
CpG 2
CpG 3
CpG Mean

IGF2
Genetic influences

CHRNB4

<table>
<thead>
<tr>
<th>CpG</th>
<th>% Methylation (median IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-7</td>
<td>rs1562008 P=0.0004</td>
</tr>
<tr>
<td></td>
<td>rs2139444 P=0.0016</td>
</tr>
</tbody>
</table>

IGF2 & DNMT3b

<table>
<thead>
<tr>
<th>CpG</th>
<th>% Methylation (median IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rs437302 P=0.03</td>
</tr>
<tr>
<td>2</td>
<td>P=0.01</td>
</tr>
<tr>
<td>3</td>
<td>P=0.02</td>
</tr>
</tbody>
</table>
Epigenotype can be inherited through the germ-line
Early epigenetic marking of the genome

Methylation reprogramming in the germ line

Methylation reprogramming in pre-implantation embryos
Familial Clustering

Figure 2. Familiality of Global DNA Methylation Change in the Utah Cohort

Change in Hpall methylation as measured by luminometric methylation assay between the 2 time points (average 16-year interval) for the Utah cohort, sorted by family. Each circle represents an individual.
Maternal-child correlation

![Graph showing maternal-child correlation with correlation coefficient 0.489 and p-value less than 0.0001.](image)
Epigenotype can influence phenotype
### CHRNA/B & lung cancer

<table>
<thead>
<tr>
<th></th>
<th>CpG1</th>
<th>CpG2</th>
<th>CpG3</th>
<th>CpG4</th>
<th>CpG5</th>
<th>CpG6</th>
<th>CpG7</th>
<th>CpGmean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHRNA#3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>6.95</td>
<td>7.00</td>
<td>7.10</td>
<td>4.80</td>
<td>5.80</td>
<td>5.85</td>
<td>2.50</td>
<td>5.86</td>
</tr>
<tr>
<td>Control</td>
<td>6.50</td>
<td>7.40</td>
<td>6.80</td>
<td>9.85</td>
<td>5.30</td>
<td>7.80</td>
<td>2.60</td>
<td>6.77</td>
</tr>
<tr>
<td>P-value</td>
<td>0.198</td>
<td>0.317</td>
<td>0.594</td>
<td>&lt;0.0001</td>
<td>0.110</td>
<td>0.0001</td>
<td>0.463</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CpG1</th>
<th>CpG2</th>
<th>CpG3</th>
<th>CpG4</th>
<th>CpG5</th>
<th>CpG6</th>
<th>CpG7</th>
<th>CpGmean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHRNB#4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>6.05</td>
<td>8.65</td>
<td>3.45</td>
<td>11.40</td>
<td>3.65</td>
<td>6.75</td>
<td>6.30</td>
<td>6.60</td>
</tr>
<tr>
<td>Control</td>
<td>9.40</td>
<td>12.60</td>
<td>3.90</td>
<td>11.60</td>
<td>4.60</td>
<td>13.30</td>
<td>21.00</td>
<td>11.20</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.366</td>
<td>0.9992</td>
<td>0.626</td>
<td>0.0006</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Two-sample Wilcoxon rank sum (Mann-Whitney) test for significant differences in CHRN qCpG in PBL DNA from lung cancer cases and controls.
Is methylation of CpG sites correlated?

CHRN4

<table>
<thead>
<tr>
<th></th>
<th>CpG1</th>
<th>CpG2</th>
<th>CpG3</th>
<th>CpG4</th>
<th>CpG5</th>
<th>CpG6</th>
<th>CpG7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpG1</td>
<td>0.48</td>
<td>0.52</td>
<td>0.15</td>
<td><strong>0.60</strong></td>
<td>0.81</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>CpG2</td>
<td>0.03</td>
<td>0.18</td>
<td>0.24</td>
<td>0.41</td>
<td><strong>0.72</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG3</td>
<td>0.30</td>
<td>0.44</td>
<td>0.17</td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG4</td>
<td>0.45</td>
<td>0.59</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG5</td>
<td><strong>0.65</strong></td>
<td><strong>0.62</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CpG6</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Is methylation of CpG sites correlated?

CHRN3A

<table>
<thead>
<tr>
<th>CpG1</th>
<th>CpG2</th>
<th>CpG3</th>
<th>CpG4</th>
<th>CpG5</th>
<th>CpG6</th>
<th>CpG7</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.72</td>
<td>0.62</td>
<td>0.15</td>
<td>0.07</td>
<td>0.07</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>0.61</td>
<td>0.22</td>
<td>0.55</td>
<td>0.22</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.18</td>
<td>0.27</td>
<td>-0.06</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.23</td>
<td>-0.03</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.06</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cross-sectional study of DNA methylation and phenotype at age 50

- Bone health
- Respiratory health
- Cardiovascular health
- Metabolic health
- Body composition
Newcastle Thousand Families Study

Bone mineral density
Cross-sectional study, age 50y

Regression co-efficient
0.84 (0.76, 0.93) p=0.001
Epigenetics and developmental programming
Early life programming mediated by DNA methylation

Persistent epigenetic differences associated with prenatal exposure to famine in humans

Bastiaan T. Heijmans¹,², Elmar W. Tobi¹,², Aryeh D. Stein¹, Hein Putter³, Gerard J. Blauw⁴, Ezra S. Susser⁵,⁶, P. Eline Slagboom⁷, and L. H. Lumey⁸,¹

- 5 CpG sites in the $IGF2$ gene measured in exposed (n=60) and their same sex siblings (n=60)
- Prenatal early gestation exposure to famine is associated with changes in methylation 6 decades later; ↓5.2% in exposed vs unexposed
- No association observed with those exposed in late gestation
- 122 additional controls analysed to assess influence of age; ↓3.6% with each 10 year increase in age

Proc Natl Acad Sci 2008; 105(44): 17046-49
**In utero influences**

The figure shows bar charts comparing maternal and child methylation levels. The x-axis represents different genotypes for the *MTHFR* 677C>T polymorphism (CC, CT, TT) and maternal methylation status (Yes, No). The y-axis represents HpaI % Methylation (mean, SD). The chart indicates that maternal methylation levels are generally higher than child methylation levels for the same genotypes.
DNA methylation in the ACSL3 gene, PAH exposure and childhood asthma

- Maternal airborne PAH exposure
- Methylation of the ACSL3 gene in UCWBC DNA
- Asthma symptoms in children <5yr

- Methylation profiling n=10 exposed & n=10 unexposed
- 31 candidate genes demonstrated differential methylation
- 6/31 genes had CGI in promoter region
- Validated association in 6 promoters in n=20
- Established association between ACSL3 CGI promoter methylation, PAH exposure and childhood asthma in n=56

Illumina GoldenGate

**FIGURE 1: THE INFINIUM ASSAY FOR METHYLATION**

The Infinium Assay for Methylation detects methylation status at individual CpG loci by typing bisulfite-converted DNA. Methylation protects C from conversion (left), whereas unmethylated C is converted to T (right). A pair of bead-bound probes is used to detect the presence of T or C by hybridization followed by single-base extension with a labeled nucleotide.
Relationship between DNA methylation and body composition

- 1,505 CpG sites in 807 genes
- 234 CpG sites in 100 genes associated with body composition [height, BMI] based on Medline evidence and GWAS data
- Informative n=?
- Is methylation status in cord blood DNA associated with:
  - BMI at age 7w, 40w, 1.5y, 3.5y, 7y, 9y, 10y, 11y
  - Height at age 7w, 40w, 1.5y, 3.5y, 7y, 9y, 10y, 11y
DNA methylation in cord blood
DNA predicts BMI at age 7y

BMI at age 7y, IGF1 methylation in cord blood DNA

Regression co-efficient
-5.47 (-9.96, -1.53) p=0.008
-0.66 (-1.19, -0.18)

Mean (SD) = 0.676 (0.12)

BMI at age 7y, IGFBP3 methylation in cord blood DNA

Regression co-efficient
37.97 (12.69, 63.25) p=0.004
0.68 (0.23, 1.14)

Mean (SD) = 0.022 (0.018)
Tracking of DNA methylation effects through childhood

Regression coefficient (95% CI)

Unit change in BMI per quintile change in promoter methylation

POMC
TCF7L2
Tracking of DNA methylation effects through childhood

Regression coefficient (95% CI)

Unit change in height per quintile change in promoter methylation

CDKN2A
CDKN2B
Challenges

- Which tissue?
- Which loci?
- Which methods?
- How many CpG sites?
- Are CpG sites correlated (epitypes)?
- How do genetic and epigenetic variation interact (hepitypes)?
- CGIs and CGI-shores?
- Skewed or bimodal distribution
- Subtle exposures – subtle effects
- Establishing causality
Epigenetic biomarker discovery

- Illumina GoldenGate/Infinium arrays
  - Cancer Panel (1,505 CpG sites in 807 genes)
  - 27K array (all annotated promoters)
  - 50K array

- MeDIP-chip
  - Requires 20-30 cases/controls

- MeDIP-seq
  - 12 samples pooled using unique molecular tags. Full methylome sequenced. 1.2Gb sequence = 1 Solexa 1G run
CpG coverage

CDKN2A 21984138 - 21984490
CGCTCAGGGGAAGGCAGGTGGTGCAGCCGTCTGCGGGGCGGAGATGGGCA
GGGGGGCGGTGGTGTGGTGGTCCCACTGCTGCAAGTTAAGGGGGGCAGGAG
TGGCGCTGCTCACCTCTGGTGCCAAAGGGGGGGGGGCGCAGGGCTGGCC
GAGCTCGGGGCTGAGGGCCGCGGAGACATGGTGGCGCAGGTTCCTTG
GTGACCCCTCCGGATTCGGCGCAGGTGCCGGGCCGCGCCGCGCCGCGGCAGGAG
GGTTTTTCGTCGTTTACATCCCCGCGGCTCAGGGGGAGATGGGGCGACG
GCCAGGGCGCCCGCGCGCTGTGGCCCTCTGCTGTGATGCTACTGAG
GAGCCAGCCTCTAGGGCAGCAGCCCTGTCCTTAGAAGACCAGgtagg
aaaggccct

TGGT Sequenom (29) £12 / sample
TGGT Illumina GoldenGate (1) £0.10 / sample
TGGT Pyrosequencer (7) £3.50 / sample
## Quantitative DNA methylation analysis in epidemiological studies

<table>
<thead>
<tr>
<th>Standard Deviation</th>
<th>Independent groups</th>
<th>Paired groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.05$</td>
<td>$\alpha = 0.001$</td>
</tr>
<tr>
<td>0.10</td>
<td>1238</td>
<td>3095</td>
</tr>
<tr>
<td>0.25</td>
<td>199</td>
<td>498</td>
</tr>
<tr>
<td>0.50</td>
<td>64</td>
<td>140</td>
</tr>
<tr>
<td>1.00</td>
<td>17</td>
<td>37</td>
</tr>
</tbody>
</table>
CGIs and CGI shores

Irizarry et al. *Nature Genetics* 2009
Epigenetic mechanisms transduce stimuli to influence body composition.
Acknowledgements

- Newcastle University
  - Alix Groom
  - Hannah Elliott
  - Jill McKay
  - James McConnell
  - Mark Pearce
  - Heather Cordell
  - John Mathers

- Bristol University
  - George
  - Beate
  - Sue

- IARC
  - Paul Brennan
  - Zdenko Herceg