Genetic and Epigenetic Analysis in DOHaD – past / present / future

Professor Richard Saffery
Murdoch Childrens Research Institute,
Department of Paediatrics, University of Melbourne

Note: for non-commercial purposes only
Environmental Exposure (e.g. diet)

Sub optimal intrauterine environment

Genetic and sex specific effects

Metabolic/endocrine disruption

Modified tissue function/development

In utero

Fetal programming / Maladaption?

Adverse birth outcome including low birth weight

Predisposition to early life and adult onset disease (e.g. T2D)

Where we started
What could mediate these effects?

- Environmentally sensitive
- Capable of biological ‘embedding’ of exposure information
- Compatible with a long latency from exposure to outcome
- Subject to sex-specific and genetic influence

Epigenetic variation?
Fetal programming and DOHaD - mediation

Environmental Exposure (e.g. diet)

Sub optimal intrauterine environment

Molecular (e.g. Epigenetic) disruption

Metabolic/endocrine disruption

Modified tissue function/development

Fetal programming / Maladaption?

Adverse birth outcome including low birth weight

Predisposition to early life and adult onset disease (e.g. T2D)

Where we are…
DOHaD Simple Molecular Mediation Model

Prenatal risk

Birth

Childhood

Adulthood

Epigenetic mediation

Stable and measurable?
Fact -1:

The epigenome is more than just DNA methylation
Mitotic chromosomes are the highest order of DNA packaging.

DNA is wrapped twice around nucleosomes.

DNA exists as a double helix.

DNA strands are linked by hydrogen bonds (C-G and A-T).

nucleosomes are wrapped into chromatin fibres.

Decreasing DNA compaction

1. DNA methylation
2. Histones and Histone variants
3. Nucleosome spacing
4. Non coding RNAs
5. Higher order chromatin modifiers
Chromatin – a combination of epigenetic marks
Complex interplay and interdependence of epigenetic processes
We study DNA methylation because...

- Covalent modification of DNA
  - Highly stable
  - Can be measured in any DNA sample

- Robustly measurable at both genome and locus specific level
  - Enrichment, RE, BeadArray approaches
  - Pyrosequencing, Sequenom, BisSeq

Not because it is any more important than other epigenetic processes.
Limitations of current preferred platform for cohort studies

- Illumina Infinium HumanMethylation450 (450K) arrays
  - Contain many SNP and repetitive probes
  - Assess methylation at <2% of genomic CpG sites
  - Minimal interrogation of non CpG methylation
  - Costs $ several hundred per sample
  - Designed with a gene-centric focus
    - Minimal interrogation of repeat and intergenic sequences
Early-life nutrition modulates the epigenetic state of specific rDNA genetic variants in mice

Michelle L. Holland,¹* Robert Lowe,¹* Paul W. Caton,² Carolina Gemma,¹‡ Guillermo Carbajosa,¹§ Amy F. Danson,¹ Asha A. M. Carpenter,³ Elena Loche,³ Susan E. Ozanne,³ Vardhman K. Rakyan¹†
Consideration 1:

- we are only scratching the tip of the epigenetic iceberg!!

My best FDR corrected p-value is 0.2!!
Fact -2:
The human epigenome is varies with time
Epigenetic variation is dynamic primarily in very early development.
Unsupervised clustering of 27,000 DNA methylation values (HM27) from human placenta across gestation
Early postnatal blood epigenome is dynamic

Unsupervised clustering of 27,000 DNA methylation values

David Martino
Blood of premature infants shows large-scale epigenetic differences

Mark Cruikshank
DNA methylation and CD4+ T-cell development

HM450 array analysis

HT-12 expression analysis

Methylation level (0-100%)

Genome-scale profiling reveals a subset of genes regulated by DNA methylation that program somatic T-cell phenotypes in humans.

D Martino1,2, J Maksimovic1, J-HE Joo1, SL Prescott2,3 and R Saffery1,3
The epigenetic clock and biological ageing

- Several recent studies have demonstrated age-associated changes in DNA methylation that occur independently of sex, tissue type, disease
- Highly variable in early life
- DNAm age, can be used as a biomarker for addressing a host of questions arising in aging research and related fields (Horvath, 2013)
Biological vs Chronological Ageing

- **Biological age**
- **Chronological age**

- **Premature biological ageing**
- **Average biological ageing**
- **Healthy biological ageing**

- **EPIGENETIC DRIFT?**

- Disease ‘threshold’
- Average age of disease onset
Caveat: These findings can potentially be explained by changes in cell composition over time.
In vitro exposure of human blood mononuclear cells to active vitamin D does not induce substantial change to DNA methylation on a genome-scale

Raul A. Chavez Valencia\textsuperscript{a,b,1}, David J. Martino\textsuperscript{b,c,1}, Richard Saffery\textsuperscript{b,c}, Justine A. Ellis\textsuperscript{a,b,1}

\textsuperscript{a} Genes, Environment and Complex Disease, Murdoch Childrens Research Institute, 50 Flemington Road, Parkville, Victoria 3052, Australia
\textsuperscript{b} Department of Paediatrics, The University of Melbourne, Victoria 3010, Australia
\textsuperscript{c} Cancer and Disease Epigenetics, Murdoch Childrens Research Institute, 50 Flemington Road, Parkville, Victoria 3052, Australia

**A**

MDS; 16 samples* 417659 CpGs

**B**

CYP24A1

<table>
<thead>
<tr>
<th>Dose (nM)</th>
<th>0</th>
<th>0.1</th>
<th>1</th>
<th>10</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative gene expression</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

- 24hr
- 96hr
- 120hr

**MDS; 5000 most variable CpGs**

*Note: *MDS stands for Multi-Dimensional Scaling.
A substantial proportion of DMRs are VDR binding elements.
Considerations:

2. Epigenetic variation observed over time reflects cell proportion changes in addition to _bona fide_ time dependent effects in a specific cell type.

3. Specific environments may only influence a subset of cells in blood.

4. Epigenetic variation in minor cell populations within a tissue can potentially have major phenotypic consequences.
Fact 3:

Epigenetic profile is sensitive to environmental influence
Twin studies

• Study plan: **Twin Model**
  – Monozygotic (MZ) twins share 100% of genetic variation
  – Dizygotic (DZ) twins share ~50%

• Within pair: MZ (1) vs MZ (2)
  – Within pair MZ differences
    (non-shared environmental/ stochastic influence)

• DZ (group) vs MZ (group): relative role of genes and environment to phenotypic trait
  – MZ correlation > DZ correlation (**genetic influence**)
PETS timeline

- **18-20 weeks** – recruitment, diet, stress, lifestyle, conception questionnaires
- **28 weeks** - maternal questionnaires (as above if needed)
  - maternal blood → serum/plasma storage
- **Birth** - baby measurements, birth data, (questionnaire data)
  - cord blood → serum/plasma (-70C), WBC/ CBMC (LN₂)
  - placenta → multiple biopsies in RNA later (-70C)
  - cord tissue → biopsy (-70C), HUVECs (LN₂)
  - buccal swabs → DNA (-70C),
- **18mth follow up** - questionnaires (diet, lifestyle, general), baby measurements
  - peripheral blood → serum/plasma (-70C), WBC/ PBMC (LN₂)
  - buccal swabs → DNA (-70C),
- **6 year follow up complete** - questionnaires, baby measurements, dental exam
  - peripheral blood → serum/plasma
  - buccal swabs → DNA (-70C)
  - Stool → gut microbiome
  - Dental swabs → oral microbiome

---

*Pregnancy*

- 20wk
- 28wk
- birth
- 18 month
- 6 years
- 8 years (planned)
Genome-wide methylation in newborn twins

<table>
<thead>
<tr>
<th>Twin pair ID no.</th>
<th>Zygosity</th>
<th>Chorionicity</th>
<th>Twin 1 sex</th>
<th>Twin 2 sex</th>
<th>Twin 1 birth weight (g)</th>
<th>Twin 2 birth weight (g)</th>
<th>Birth weight discordance (%)</th>
<th>Gestational age</th>
<th>Cell types studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>F</td>
<td>2530</td>
<td>1785</td>
<td>29.4%</td>
<td>34</td>
<td>H, C</td>
</tr>
<tr>
<td>1002</td>
<td>MZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>3020</td>
<td>3175</td>
<td>4.9%</td>
<td>37</td>
<td>H, C</td>
</tr>
<tr>
<td>1012</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>2155</td>
<td>2635</td>
<td>18.2%</td>
<td>37</td>
<td>H, C</td>
</tr>
<tr>
<td>1016</td>
<td>MZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2902</td>
<td>2746</td>
<td>5.4%</td>
<td>37</td>
<td>H, C</td>
</tr>
<tr>
<td>1019</td>
<td>MZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2430</td>
<td>2330</td>
<td>4.1%</td>
<td>37</td>
<td>H, C</td>
</tr>
<tr>
<td>1021</td>
<td>DZ</td>
<td>DC</td>
<td>F</td>
<td>M</td>
<td>3038</td>
<td>3096</td>
<td>1.9%</td>
<td>37</td>
<td>C</td>
</tr>
<tr>
<td>1022</td>
<td>MZ</td>
<td>MC</td>
<td>M</td>
<td>M</td>
<td>2622</td>
<td>2298</td>
<td>12.4%</td>
<td>37</td>
<td>H, C</td>
</tr>
<tr>
<td>1048</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2030</td>
<td>2440</td>
<td>16.8%</td>
<td>38</td>
<td>H, C</td>
</tr>
<tr>
<td>1049</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>3155</td>
<td>3065</td>
<td>2.9%</td>
<td>34</td>
<td>C</td>
</tr>
<tr>
<td>1065</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2455</td>
<td>2330</td>
<td>5.1%</td>
<td>38</td>
<td>H, C</td>
</tr>
<tr>
<td>1072</td>
<td>DZ</td>
<td>DC</td>
<td>F</td>
<td>F</td>
<td>2664</td>
<td>3114</td>
<td>14.5%</td>
<td>36</td>
<td>P</td>
</tr>
<tr>
<td>1074</td>
<td>DZ</td>
<td>DC</td>
<td>F</td>
<td>M</td>
<td>2625</td>
<td>2455</td>
<td>6.5%</td>
<td>37</td>
<td>P, H, C</td>
</tr>
<tr>
<td>1082</td>
<td>MZ</td>
<td>DC</td>
<td>F</td>
<td>F</td>
<td>1710</td>
<td>1955</td>
<td>12.5%</td>
<td>34</td>
<td>C</td>
</tr>
<tr>
<td>1104</td>
<td>MZ</td>
<td>MC</td>
<td>M</td>
<td>M</td>
<td>2130</td>
<td>1785</td>
<td>16.2%</td>
<td>38</td>
<td>P, H</td>
</tr>
<tr>
<td>1108</td>
<td>MZ</td>
<td>MC</td>
<td>M</td>
<td>M</td>
<td>2655</td>
<td>2525</td>
<td>4.9%</td>
<td>32</td>
<td>C</td>
</tr>
<tr>
<td>1114</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>2485</td>
<td>2610</td>
<td>4.8%</td>
<td>36</td>
<td>P, H, C</td>
</tr>
<tr>
<td>1126</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>1746</td>
<td>1270</td>
<td>27.3%</td>
<td>37</td>
<td>P</td>
</tr>
<tr>
<td>1130</td>
<td>MZ</td>
<td>DC</td>
<td>F</td>
<td>F</td>
<td>2140</td>
<td>2755</td>
<td>22.3%</td>
<td>38</td>
<td>P, H, C</td>
</tr>
<tr>
<td>1138</td>
<td>MZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2788</td>
<td>2612</td>
<td>6.3%</td>
<td>38</td>
<td>C</td>
</tr>
<tr>
<td>2010</td>
<td>MZ</td>
<td>MC</td>
<td>M</td>
<td>M</td>
<td>2533</td>
<td>2407</td>
<td>5.0%</td>
<td>35</td>
<td>H, C</td>
</tr>
<tr>
<td>2027</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>F</td>
<td>3510</td>
<td>2810</td>
<td>19.9%</td>
<td>38</td>
<td>P, H</td>
</tr>
<tr>
<td>2028</td>
<td>MZ</td>
<td>DC</td>
<td>F</td>
<td>F</td>
<td>2630</td>
<td>2210</td>
<td>16.0%</td>
<td>36</td>
<td>C</td>
</tr>
<tr>
<td>2029</td>
<td>MZ</td>
<td>DC</td>
<td>F</td>
<td>F</td>
<td>2230</td>
<td>1470</td>
<td>34.1%</td>
<td>34</td>
<td>H, C</td>
</tr>
<tr>
<td>2036</td>
<td>MZ</td>
<td>MC</td>
<td>M</td>
<td>M</td>
<td>2349</td>
<td>2295</td>
<td>2.3%</td>
<td>38</td>
<td>P, H</td>
</tr>
<tr>
<td>2051</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>F</td>
<td>2660</td>
<td>1890</td>
<td>28.9%</td>
<td>36</td>
<td>P, H, C</td>
</tr>
<tr>
<td>2064</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>2120</td>
<td>1920</td>
<td>9.4%</td>
<td>37</td>
<td>P, H</td>
</tr>
<tr>
<td>2071</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>2850</td>
<td>2650</td>
<td>7.0%</td>
<td>38</td>
<td>P, H</td>
</tr>
<tr>
<td>2075</td>
<td>DZ</td>
<td>DC</td>
<td>F</td>
<td>M</td>
<td>3050</td>
<td>2980</td>
<td>2.3%</td>
<td>37</td>
<td>P, H, C</td>
</tr>
<tr>
<td>3001</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>2480</td>
<td>2680</td>
<td>7.5%</td>
<td>34</td>
<td>H, C</td>
</tr>
<tr>
<td>3006</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2765</td>
<td>2665</td>
<td>3.6%</td>
<td>34</td>
<td>C</td>
</tr>
<tr>
<td>3027</td>
<td>DZ</td>
<td>DC</td>
<td>M</td>
<td>M</td>
<td>2500</td>
<td>2280</td>
<td>8.8%</td>
<td>36</td>
<td>P, H, C</td>
</tr>
<tr>
<td>3034</td>
<td>MZ</td>
<td>DC</td>
<td>F</td>
<td>F</td>
<td>2950</td>
<td>2550</td>
<td>13.6%</td>
<td>38</td>
<td>C</td>
</tr>
<tr>
<td>3051</td>
<td>MZ</td>
<td>MC</td>
<td>F</td>
<td>F</td>
<td>2813</td>
<td>2105</td>
<td>25.2%</td>
<td>33</td>
<td>H</td>
</tr>
<tr>
<td>3068</td>
<td>MZ</td>
<td>MC</td>
<td>M</td>
<td>M</td>
<td>2392</td>
<td>2534</td>
<td>5.6%</td>
<td>38</td>
<td>P, C</td>
</tr>
</tbody>
</table>
Genome-wide methylation data (MZ twins)
- evidence of environmental effects

27,000 data points
Highly variable within pairs

Clear evidence of variability within MZ twins
Strength of evidence for specific exposures
Maternal smoking and the early life epigenetic profile in humans

450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking During Pregnancy

Bonnie R. Joubert, Siri E. Håberg, Roy M. Nilsen, Xuting Wang, Stein E. Vollset, Susan K. Murphy, Zhiqing Huang, Cathrine Hoyo, Øivind Midttun, Lea A. Cupul-Uicab, Per M Ueland, Michael C. Wu, Wenche Nystad, Douglas A. Bell, Shyamal D. Peddada, Stephanie J. London

2012
Prolonged exposure is necessary
Différence en méthylation - non-fumeurs vs fumeurs à la naissance/18 mois

Effet de la cigarette maternelle persiste postnatallement

Boris Novakovic
Genetic variation plays a role

A. DZ pairs assay A CpG 7

B. MZ pairs assay A CpG 7

C. DZ pairs assay A mean

D. MZ pairs assay A mean

Postnatal stability and tissue- and time-specific effects of AHRR methylation change in response to maternal smoking in pregnancy

Boris Novakovic1, Joanne Ryan1, Natalie Pereira1, Berin Boughton1, Jeffrey M Craig1, and Richard Saffery1,2
A total of 13 cohorts
6,685 newborns, 897 (13%) exposed to sustained maternal smoking and 1,646 (25%) to any maternal smoking during pregnancy
6,073 CpGs with FDR significance, 568 met Bonferroni threshold
- Dose dependent effects
- Gene specific stability postnatally
- mediates the effects of maternal smoking on infant birthweight
From smoking to diet
Pilot: Folic acid supplementation and infant DNA methylation

- Maternal folate levels and infant blood DNA methylation profile

- n = 222 mother-infant pairs in main folic acid intervention study

- 12 Highest and lowest maternal folate levels used

- APC and CD4+ T-cells from cord blood

- DNA methylation profiled using Illumina Infinium HM450K platform
Differentially methylated probes

- 13,424 significant sites (3.5% of total probes)
- 5048 hypermethylated, 8376 hypomethylated sites.

Differentially methylated regions (coordinated methylation)

<table>
<thead>
<tr>
<th>DMR</th>
<th>Chromosome</th>
<th>Start</th>
<th>End</th>
<th>Size (bp)</th>
<th>Probes</th>
<th>Average</th>
<th>Maximum</th>
<th>HE vs. LF</th>
<th>Gene association</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>29,648,161</td>
<td>29,649,084</td>
<td>923</td>
<td>19</td>
<td>0.19</td>
<td>0.28</td>
<td>Hypomethylated</td>
<td>ZFP57 (-3692)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>37,123,638</td>
<td>37,123,949</td>
<td>311</td>
<td>8</td>
<td>-0.08</td>
<td>-0.12</td>
<td>Hypermethylated</td>
<td>LASP1 (+97,682)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>76,189,707</td>
<td>76,190,008</td>
<td>301</td>
<td>5</td>
<td>-0.13</td>
<td>-0.23</td>
<td>Hypermethylated</td>
<td>ACADM (-185)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>144,120,106</td>
<td>144,120,706</td>
<td>600</td>
<td>7</td>
<td>0.08</td>
<td>0.10</td>
<td>Hypomethylated</td>
<td>LY6E (+20,504)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>228,075,423</td>
<td>228,075,749</td>
<td>326</td>
<td>5</td>
<td>-0.10</td>
<td>-0.13</td>
<td>Hypermethylated</td>
<td>WNT9A (+60,909)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>47,604,052</td>
<td>47,604,654</td>
<td>602</td>
<td>5</td>
<td>0.09</td>
<td>0.11</td>
<td>Hypermethylated</td>
<td>C21orf56 (+20)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>202,901,045</td>
<td>202,901,470</td>
<td>425</td>
<td>5</td>
<td>-0.08</td>
<td>-0.11</td>
<td>Hypermethylated</td>
<td>FZD7 (+1948)</td>
<td></td>
</tr>
</tbody>
</table>

Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans

Manori Amaresekera,* David Martinez,† Sarah Ashley,‡ Hani Harb,§ Dorthe Kesper,∥ Deborah Strickland,* Richard Saffrey,†,‡ and Susan L. Prescott*†

*School of Paediatrics and Child Health, University of Western Australia, Perth, Western Australia, Australia; †Cancer and Disease Epidemiology Group, Murdoch Children’s Research Institute, Melbourne, Victoria, Australia; ‡Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia; and †Te勒ton Institute for Child Health Research, Perth, Western Australia, Australia
Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns

Bonnie R. Joubert1*, Herman T. den Dekker2,3,4,*, Janine F. Felix2,4,5, Jon Bohlin6, Symen Ligthart4, Emma Beckett7,8, Henning Tiemeier4,9, Joyce B. van Meurs10, Andre G. Uitterlinden4,10, Albert Hofman4, Siri E. Häberg6, Sarah E. Reese1, Marjolein J. Peters10, Bettina Kulle Andreasen11, Eric A.P. Steegers12, Roy M. Nilsen13, Stein E. Vollset6,14, Øivind Midttun15, Per M. Ueland16,17, Oscar H. Franco4, Abbas Dehghan4, Johan C. de Jongste3, Michael C. Wu18, Tianyuan Wang1, Shyamal D. Peddada1, Vincent W.V. Jaddoe2,4,5, Wenche Nystad6, Liesbeth Duijts2,3,4,*** & Stephanie J. London1***

In the only previous study using the 450K platform, Amarasekera et al12 reported differential methylation in relation to maternal folate in a 923-bp region on chromosome 6, 3-kb upstream of ZFP57. Our studies differ in sample sizes, design and analysis methods. However, when we evaluate the 20 CpGs that map to ZFP57, we find 5 with uncorrected $P$ values of 0.05 or smaller—more than would be expected by chance alone. Thus, our data provide support for association at this locus.

• MoBa. 1275 participants with plasma folate and complete covariate data. Blood at 18 weeks gestation.
• 713 children from GenR with methylation, plasma folate and complete covariate data. Blood at 12.9
Considerations/questions:

5. Environmental influence on early DNA methylation profile is pervasive.

6. Unclear how stable the effects of specific exposures are at specific timepoints in specific cell types?

7. Can transient exposures at specific timepoints induce transient epigenetic change with long lasting impact on phenotype?
Fact 4:

Epigenetic profile is sensitive to genetic and sex-specific influence
Unsupervised clustering of 27,000 DNA methylation values

11/16 MZ cluster – 69%
5/10 DZ cluster – 50%

M, male; F, female, ■, MZ, □, DZ, [ ] clustering twin pair
HUVECs

Unsupervised clustering of 27,000 DNA methylation values

6/13 MZ cluster – 46%
0/8 DZ cluster – 0%

M, male; F, female, MZ, DZ, clustering twin pair

Lavinia Gordon
Unsupervised clustering of 27,000 DNA methylation values

7/8 MZ cluster – 88%

3/6 DZ cluster – 50%

M, male; F, female, □, MZ, ■, DZ, □□ clustering twin pair
Genetics and DNA methylation

The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes

Ai Ling Teh,1,10 Hong Pan,1,2,10 Li Chen,1 Mei-Lyn Ong,1 Shaillay Dogra,1 Johnny Wong,1 Julita L. MacIsaac,3 Sarah M. Mah,3 Lisa M. McEwen,5 Seang-Mei Saw,4 Keith M. Godfrey,5 Yap-Seng Chong,1,6 Kenneth Kwek,7 Chee-Keong Kwoh,2 Shu-E. Soh,4,6 Mary F.F. Chong,1,6 Sheila Barton,5 Neerja Karnani,1 Clara Y. Cheong,1 Jan Paul Buschdorf,1 Walter Stunkel,1 Michael S. Kober,3 Michael J. Meaney,1,8 Peter D. Gluckman,1,9 and Joanna D. Holbrook1,11

- Heritability of DNA methylation profile estimated at 15-20%
Systematic identification of genetic influences on methylation across the human life course


Genetic and environmental influences interact with age and sex in shaping the human methylome

Summary – Genetic Influence

Genetic considerations are paramount in all Epigenetic studies.
Considerations/unanswered questions:

8. Genetic variation impacts on DNA methylation early, potentially prior to fertilisation
9. Genetic variation moderates the impact of environmental exposures on DNA methylation
10. context (age, sex, dose, timing) is very important in regulating this interaction
11. Genetic effects on other epigenetic processes remain unclear
Fact 5:

Epigenetic variation does not exist in isolation
Big data – multidimensional 1

EXPOSOME

Phenome

Metabolome (>10^6)

Proteome (>10^6)

Transcriptome (>10^5)

Epigenome (>3x10^7 methylation)

Genome

(6x10^8 nucleotides)

(2x10^4 genes)

(2x >5x10^6 genetic variants)
Big data – multidimensional 2 – cell/tissue effects

‘Average signature sample 1’

‘Average signature sample 2’
Big data – multidimensional 3 – timing effects

- Change over time has multiple inputs and implications for causal inference
- Fewer implications for biomarker discovery
‘omic association studies - considerations

• **Select appropriate tissue** – biological relevance to phenotype?

• **Consider cell heterogeneity** – most tissues are not uniform!

• **Consider population structure** – methylation status at many sites is (partly) genetically determined.

• **Validation and replication** – are the primary results reproducible?

• **Issues of causation** – epigenetic marks vary temporally

---

**Recommendations for the design and analysis of epigenome-wide association studies**

Karin B Michels$^{1,2}$, Alexandra M Binder$^2$, Sarah Dedeurwaerder$^3$, Charles B Epstein$^4$, John M Greally$^5$, Ivo Gut$^6$, E Andres Houseman$^7$, Benedetta Izzi$^8$, Karl T Kelsey$^9,10$, Alexander Meissner$^{4,11}$, Aleksandar Milosavljevic$^{12}$, Kimberly D Siegmund$^{13}$, Christoph Bock$^{14-16,18}$ & Rafael A Iriarráz$^{17,18}$
Fact 6:

All of our current approaches are subject to limitations (suboptimal)
Modifiable considerations

- Generally insufficient sample size
- Inadequate/limited sampling of exposure and outcome data and/or appropriate biospecimens
- Suboptimal (and expensive) ‘omics data generation approaches
- Immature analytical approaches for integration of multidimensional datasets
Solutions?
The power of pooling
Start with the right tools!!

- **MEASURES**
  - Accurate, detailed and frequent exposure outcome measures are critical

- **BIOSPECIMENS**
  - Collect as many non-life threatening biosamples as possible
  - Collect these as frequently as possible

- Harmonisation is desirable

- **TIMING**
  - Start as early as possible and never stop
Lifecourse birth/pre-birth/intergenerational cohorts

start here

end this way
Fact 7: There are two certainties in human DOHaD research.

- Costs associated with cohort-based studies will continue to rise
- The cost of ‘omics technologies will continue to fall
Final Consideration/Question…

• Given the inevitable limited research $$$, are our current priorities and plans appropriate?
Acknowledgements

Cancer and Disease Epigenetics Group
Nick Wong, David Martino, Eric Joo
Zac Chatterton, Leah Morenos, Joanne Ryan
Mandy Sibson, Anna Czajko, Jane Ng, Boris Novakovic,
Anna Chen, Alex Sexton-Oates

Early Life Epigenetics Group
Jeff Craig  Mark Cruickshank, Xin Li, Jane Loke

Barwon Infant Study
Anne-Louise Ponsonby, Peter Vuillermin

Bioinformatics and CEBU
Jovana Maksimovic
Alicia Oshlack
Lavinia Gordon
John Galati
John Carlin
Ruth Morley

Allergy Research
Katie Allen

University of Western Australia
Susan Prescott, Deborrah Strickland

Prince Henry's Research Institute
Eva Dimitriadis

Cambridge University
Andrew Sharkey, Ashley Moffett

University of Helsinki
Miina Ollikainen, Jussi Naukkarinen

University of British Columbia
Wendy Robinson
Ryan Yuen

University of Graz
Gernot Desoye, Silvija Cvitic, Ursula Hiden

KCL London
Lucilla Poston, Jordana Bell, Tim Spector

University of Southampton
Keith Godfrey

QIMR
Joseph Powell, Peter Visscher

Inserm Paris
Daniele Evain Brion, Thierry Fournier

University of Munich
Bertold Koletzko, Peter Rzehak

Funding
Generation Victoria (GenV) Universal – a population wide cohort study.

GenV Deep – a sub population birth cohort with extensive biospecimen collection.
Summary of evidence at present

Prenatal risk
Exposure
* measured
Genetic variation *

Adulthood
Epigenetic variation
* measured
Altered tissue development/function
Birth
Epigenetic variation *
Birth Phenotype *
Pre-disease phenotype *

Childhood
Epigenetic variation *
"Adult" onset disease * measured
Genetic variation *

associations