Epigenetics in life course and longitudinal cohort studies

Laura Barrett
Overview

Study Design
  Cross-sectional
  Longitudinal

Literature Search
  1 time point
  2 time points
  Nested case-control

Longitudinal Cohorts

Data Analysis

Summary
Study Designs

Cross-sectional

Most common

Advantages

✔ Low Cost
✔ One Time Point

Disadvantages

✘ Capture Dynamic Nature
✘ Direction of Causality
Global LINE-1 DNA methylation is associated with blood glycaemic and lipid profiles (Pearce et al. 2012)

**Cohort** – Newcastle Thousand Families Study
1142 individuals born to mothers resident in Newcastle in May-June 1947

228 study members – Blood sample taken at age 50
Clinical assessments – lipid analyses, total cholesterol, HDL, LDL, triglyceride concentrations, plasma glucose concentrations, serum insulin, height and weight

**Hypothesis** – Global LINE-1 methylation measured at age 49-51 is associated with traits indicative of early-stage metabolic disease
Global LINE-1 DNA methylation is associated with blood glycaemic and lipid profiles (Pearce et al. 2012)

Table 2: Results of linear regression analyses of relationships between log-transformed methylation and the listed dependent variables, all adjusted for sex.

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Coefficient (95% CI)</th>
<th>P-value</th>
<th>$R^2$ (%)</th>
<th>Direct $R^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (kg/m²)</td>
<td>2.30 (−9.99 to 14.59)</td>
<td>0.71</td>
<td>0.90</td>
<td>0.06</td>
</tr>
<tr>
<td>Waist/hip ratio</td>
<td>0.14 (−0.01 to 0.30)</td>
<td>0.07</td>
<td>65.42</td>
<td>0.51</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>5.14 (−19.01 to 29.30)</td>
<td>0.68</td>
<td>5.65</td>
<td>0.08</td>
</tr>
<tr>
<td>Fasting glucose (mmol/l)</td>
<td>2.80 (0.39 to 5.22)</td>
<td>0.02</td>
<td>9.05</td>
<td>2.12</td>
</tr>
<tr>
<td>Total cholesterol (mmol/l)</td>
<td>4.76 (1.43 to 8.10)</td>
<td>0.005</td>
<td>4.98</td>
<td>3.34</td>
</tr>
<tr>
<td>Total triglycerides (mmol/l)</td>
<td>3.83 (1.30 to 6.37)</td>
<td>0.003</td>
<td>9.65</td>
<td>3.57</td>
</tr>
<tr>
<td>HDL cholesterol (mmol/l)</td>
<td>−1.43 (−2.38 to −0.48)</td>
<td>0.003</td>
<td>8.96</td>
<td>3.54</td>
</tr>
<tr>
<td>LDL cholesterol (mmol/l)</td>
<td>5.38 (2.12 to 8.64)</td>
<td>0.001</td>
<td>7.02</td>
<td>4.37</td>
</tr>
<tr>
<td>HDL/LDL ratio</td>
<td>−1.06 (−1.76 to −0.36)</td>
<td>0.003</td>
<td>7.00</td>
<td>3.67</td>
</tr>
<tr>
<td>Insulin secretion</td>
<td>90.24 (−8.04 to 186.33)</td>
<td>0.07</td>
<td>1.57</td>
<td>1.54</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>2.34 (−2.68 to 7.36)</td>
<td>0.36</td>
<td>5.83</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Coefficients and corresponding 95% CIs indicate the change in outcome measure per unit increase in log-transformed LINE-1 methylation, after adjustment for sex. $R^2$ reflects the variance (%) in outcome measures accounted for by both sex and log-transformed LINE-1 methylation (i.e. the combined effect of both covariates). Direct $R^2$ reflects the variance (%) in outcome measures accounted for by log-transformed LINE-1 methylation alone (i.e. the direct effect of methylation).

Increasing LINE-1 methylation associated with:

- Increased – fasting glucose, total cholesterol, total triglycerides, LDL
- Decreased – HDL cholesterol

All biomarkers of CVD and/or Type 2 diabetes

Not possible to explore a potential association between LINE-1 methylation and subsequent disease phenotype or determine direction of effect (if one exists)
Study is cross-sectional – one time point

Potential to become longitudinal as participants followed up at age 60
Chance to assess whether these associations are biomarkers of CVD
Study Designs

Longitudinal

Most suitable to analyse epigenetics of common complex disease

Advantages

✔ Establish temporal relationships
✔ Strengthen causality

Disadvantages

✘ Increased cost
Longitudinal and life course studies measuring DNA methylation

350 publications since 2010

34 measured DNA methylation in a longitudinal cohort

12 nested case-control studies

4 measured DNA methylation at 2 time points

30 measured DNA methylation at 1 time point

~300 studies that did not meet selection criteria
Study Designs

1 time point + longitudinal data relating to exposure/outcome

E.g. **Persistent epigenetic differences associated with prenatal exposure to famine in humans** (Heijmans et al. 2008)

**Cohort** – Dutch Hunger Winter 1944/45

**Hypothesis** – Periconceptional exposure to famine is associated with IGF2 methylation in adulthood (6 decades later)

60 conceived during the famine + same-sex unexposed siblings

Methylation measured in 5 CpG sites within IGF2 gene

<table>
<thead>
<tr>
<th>IGF2 DMR methylation</th>
<th>Mean methylation fraction (SD)</th>
<th>Relative change exposed</th>
<th>Difference in SDs</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exposed (n = 60)</td>
<td>Controls (n = 60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.488 (0.047)</td>
<td>0.515 (0.055)</td>
<td>-5.2%</td>
<td>5.9 x 10^-5</td>
</tr>
<tr>
<td>CpG 1</td>
<td>0.436 (0.037)</td>
<td>0.470 (0.041)</td>
<td>-6.9%</td>
<td>1.5 x 10^-4</td>
</tr>
<tr>
<td>CpG 2 and 3</td>
<td>0.451 (0.033)</td>
<td>0.473 (0.055)</td>
<td>-4.7%</td>
<td>8.1 x 10^-3</td>
</tr>
<tr>
<td>CpG 4</td>
<td>0.577 (0.114)</td>
<td>0.591 (0.112)</td>
<td>-2.3%</td>
<td>.41</td>
</tr>
<tr>
<td>CpG 5</td>
<td>0.491 (0.061)</td>
<td>0.529 (0.068)</td>
<td>-7.2%</td>
<td>1.4 x 10^-3</td>
</tr>
</tbody>
</table>

*P values were obtained using a linear mixed model and adjusted for age.*

Periconceptional exposure associated with a 5.2% lower IGF2 methylation
Study Designs

1 time point + longitudinal data relating to exposure/outcome
E.g. **Persistent epigenetic differences associated with prenatal exposure to famine in humans (Heijmans et al. 2008)**

62 exposed later in gestation + same-sex unexposed siblings

<table>
<thead>
<tr>
<th>IGF2 DMR methylation</th>
<th>Mean methylation fraction (SD)</th>
<th>Relative change exposed</th>
<th>Difference in SDs</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exposed (n = 62)</td>
<td>Controls (n = 62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.514</td>
<td>0.045</td>
<td>0.519</td>
<td>0.036</td>
</tr>
<tr>
<td>CpG 1</td>
<td>0.460</td>
<td>0.044</td>
<td>0.464</td>
<td>0.048</td>
</tr>
<tr>
<td>CpG 2 and 3</td>
<td>0.462</td>
<td>0.039</td>
<td>0.471</td>
<td>0.039</td>
</tr>
<tr>
<td>CpG 4</td>
<td>0.602</td>
<td>0.085</td>
<td>0.612</td>
<td>0.073</td>
</tr>
<tr>
<td>CpG 5</td>
<td>0.529</td>
<td>0.060</td>
<td>0.531</td>
<td>0.060</td>
</tr>
</tbody>
</table>

*P* values were obtained using a linear mixed model and adjusted for age.

No significant differences between exposed and unexposed siblings

**Conclusion** – Epigenome most vulnerable during early development

Exposures in early development can affect methylation status in later life
Longitudinal and life course studies measuring DNA methylation

350 publications since 2010

- ~300 studies that did not meet selection criteria
- 34 measured DNA methylation in a longitudinal cohort
- 4 measured DNA methylation at 2 time points
- 30 measured DNA methylation at 1 time point
- 12 nested case-control studies
Study Designs - Longitudinal

2 time points

E.g. A longitudinal study of epigenetic variation in twins (Wong et al. 2010)

Cohort – Environmental Risk (E-Risk) Longitudinal Twin Study – how genetic and environmental factors shape children’s development

Aim – Assess the contribution of heritable and environmental factors to variation in DNA methylation across 2 age ranges (5 and 10) during childhood development

3 genes – DRD4 (ADHD)
   SERT (depression)
   MAOA (antisocial behaviour)

associated with psychiatric disorders – risk is mediated by gene-environment interactions
2 time points
E.g. A longitudinal study of epigenetic variation in twins (Wong et al. 2010)

Methylation changes due to shared environmental factors

Methylation changes due to unique environmental factors

Conclusion – Environmental influences are important factors accounting for inter-individual DNA methylation differences
The observation of dynamic changes over time highlights the importance of longitudinal research
2 time points
E.g. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology (Talens et al. 2010)

Cohort – Netherlands Twin Register
Aim – To study stability of DNA methylation over time

34 individuals with blood samples taken 11-20 years apart
Previously identified differences in LEP methylation in unrelated individuals – change over time?
Methylation measured at 8 loci – cardiovascular and metabolic related genes

DNA methylation was similar at the two time points

DNA methylation did not differ in the CpG sites studied in these 8 loci in blood
Study Designs - Longitudinal

2 time points
E.g. Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men (Madrigano et al. 2011)

Cohort – Normative Aging Study
Hypothesis – prolonged exposure to particulate pollution is associated with hypomethylation of repetitive DNA elements

1406 blood samples from 712 elderly (1-3 samples from each participant)
Exposures to particulate matter (PM$_{2.5}$), black carbon (BC) and sulphates (SO$_4$) at a range of times between 28 and 180 days
Methylation analysis of repetitive elements - LINE-1 and Alu

Findings – BC – 45-90 days exposure = Alu methylation decreased
28-60 days exposure = LINE-1 methylation decreased
SO$_4$ – 90 days exposure = LINE-1 methylation decreased

Conclusion – Dynamic nature of epigenetic mechanisms can even be seen in short periods of time
Study Designs - Longitudinal

2 time points
E.g. Aging and epigenetics: Longitudinal changes in gene-specific DNA methylation (Madrigano et al. 2012)

Cohort – Normative Aging Study
Aim – to describe the intra-individual change in gene-specific methylation in a cohort of over 700 men

Candidate gene approach – 9 genes (GCR, iNOS, ICAM, IL6, TLR2, IFNγ, F3, CRAT, OGG) related to processes (inflammation, endothelial function, oxidative stress) known to be associated with aging and age-related diseases
1-5 blood samples collected from each participant

Findings – Change in age associated with decreased methylation of GCR, iNOS, TLR2
Change in age associated with increased methylation of IFNγ, F3, CRAT, OGG

Conclusion – continued DNA methylation changes throughout the life course stress the importance of longitudinal cohort studies
Longitudinal and life course studies measuring DNA methylation

- 350 publications since 2010
- 34 measured DNA methylation in a longitudinal cohort
- 4 measured DNA methylation at 2 time points
- 30 measured DNA methylation at 1 time point
- 12 nested case-control studies
- ~300 studies that did not meet selection criteria
E.g. Intragenic ATM methylation in peripheral blood DNA as a biomarker of breast cancer risk (Brennan et al. 2012)

640 incident breast cancer cases and 741 controls were taken from 3 studies with blood samples taken before diagnosis

KConFab – Familial samples (12,747) from 1,395 families in New Zealand and Australia

BGS – general population cohort consisting of 110,000 women from the UK

EPIC – general population cohort consisting of 520,000 Europeans

Methylation of 2 intragenic ATM loci (ATMmvp2a and ATMmvp2b) and LINE-1 repetitive elements measured
E.g. Intragenic ATM methylation in peripheral blood DNA as a biomarker of breast cancer risk (Brennan et al. 2012)

Cases exhibited a higher median methylation than controls in 2 of the cohorts

ATM methylation – potential biomarker?

A biomarker for risk should ideally be stable over time

Measured stability of ATMMvp2a and LINE-1 methylation in a control pop, where blood samples were taken 6 years apart (BGS cohort, n=92 pairs)

Observed no significant change at either locus– ATM methylation stable for at least 6 years

Conclusion – High levels of ATM methylation might be a biomarker for breast cancer risk

Methylation status before disease onset may predict later development of disease
Longitudinal Cohorts

Ideally:
Extensive prospectively collected data
Biological samples at multiple time points

Population Based
Often Thousands of Participants

Specific subgroup

Birth Cohorts Transgenerational and across the life course
Life Course Epidemiology

Well-established field

It allows us to model change over time

Assesses how health in later life may be modified by early and/or later life exposures
Data Analysis

• Longitudinal data analysis is complex

• We can borrow approaches from gene expression studies and from life course epidemiology

• Multiple strategies will need to be adopted

• These methods are in their early development as very few studies have longitudinal epigenetic data as yet
Why are longitudinal studies valuable?

- Dynamic nature of epigenetic patterns can be captured
- Epigenetic patterns that change in response to environmental, lifestyle and behavioural factors can be identified
- Temporal relationships can assist in defining causality
- Potentially useful in identifying predictive epigenetic biomarkers


