

# Exploiting public data

Bas Heijmans  
Molecular Epidemiology  
Leiden University Medical Center  
The Netherlands  
[bas.heijmans@lumc.nl](mailto:bas.heijmans@lumc.nl)

FOS course 'Molecular Epidemiology in Ageing Research – 30 June 2017.

# Public data

Genomics data is increasingly made publicly available upon publication

- Often requirement of journal
- 'Open science'
- Efficiency of research
- Transparency
- Reproducibility
- More citations

# New opportunities for data savvy scientists

Answering a research questions may not require the generation of new data!

(although a combination of new and old is preferred)

# Databases

Research data (up to many samples)

- Gene Expression Omnibus (GEO)  
<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48472>
- European Genome-phenome Archive (EGA)  
<https://www.ebi.ac.uk/ega/datasets/EGAD00001000733>

# Databases

## Reference data (few samples)

- Human reference genome  
<https://www.ncbi.nlm.nih.gov/projects/genome/guide/human/>
- Human reference epigenomes  
<http://epigenomesportal.ca/ihec/>
- Cancer genomes (reference & research data)  
<https://cancergenome.nih.gov>

# Databases

## Biological pathways

- Gene ontology  
<http://www.geneontology.org>
- Reactome  
<http://www.reactome.org>

# Databases

## Browsers

- Genome browsers like UCSC and EMBL  
<https://genome-euro.ucsc.edu/cgi-bin/hgGateway?redirect=manual&source=genome.ucsc.edu>  
<http://www.ensembl.org/index.html>
- QTL browsers (effect of SNPs on expression/methylation)  
<http://atlas.bbmrip3-lumc.surf-hosted.nl>

# Public data in practice

Slieker et al. *Genome Biology* (2016) 17:191  
DOI 10.1186/s13059-016-1053-6

Genome Biology

RESEARCH

Open Access



## Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms



Roderick C. Slieker<sup>1</sup>, Maarten van Iterson<sup>1</sup>, René Luijk<sup>1</sup>, Marian Beekman<sup>1</sup>, Daria V. Zhernakova<sup>2</sup>, Matthijs H. Moed<sup>1</sup>, Hailiang Mei<sup>3</sup>, Michiel van Galen<sup>4</sup>, Patrick Deelen<sup>2</sup>, Marc Jan Bonder<sup>2</sup>, Alexandra Zhernakova<sup>2</sup>, André G. Uitterlinden<sup>5</sup>, Ettje F. Tigchelaar<sup>2</sup>, Coen D. A. Stehouwer<sup>6</sup>, Casper G. Schalkwijk<sup>6</sup>, Carla J. H. van der Kallen<sup>6</sup>, Albert Hofman<sup>7</sup>, Diana van Heemst<sup>8</sup>, Eco J. de Geus<sup>9</sup>, Jenny van Dongen<sup>9</sup>, Joris Deelen<sup>1</sup>, Leonard H. van den Berg<sup>10</sup>, Joyce van Meurs<sup>5</sup>, Rick Jansen<sup>11</sup>, Peter A. C. 't Hoen<sup>4</sup>, Lude Franke<sup>2</sup>, Cisca Wijmenga<sup>2</sup>, Jan H. Veldink<sup>10</sup>, Morris A. Swertz<sup>12</sup>, Marleen M. J. van Greevenbroek<sup>6</sup>, Cornelia M. van Duijn<sup>13</sup>, Dorret I. Boomsma<sup>9</sup>, BIOS consortium, P. Eline Slagboom<sup>1</sup> and Bastiaan T. Heijmans<sup>1\*</sup>

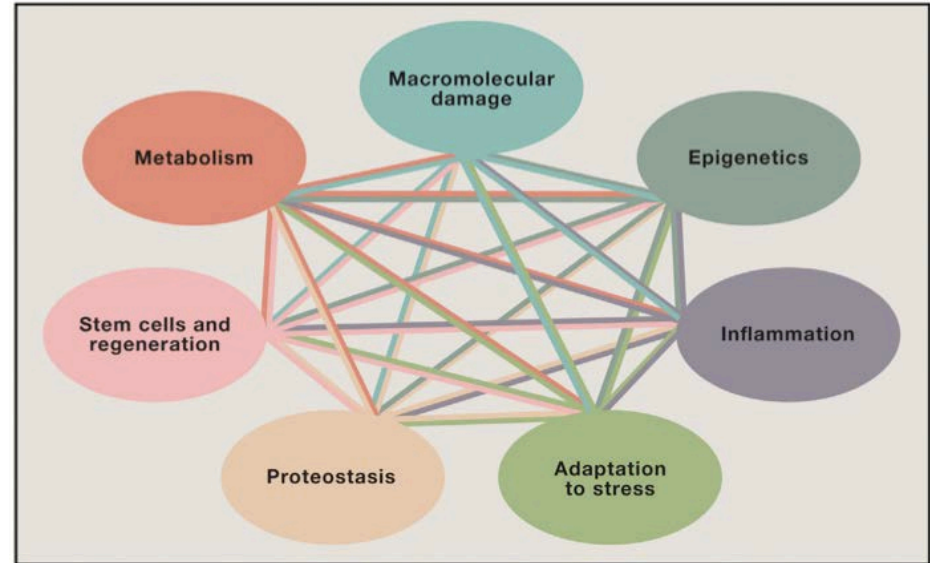


# Genomic dysregulation and ageing

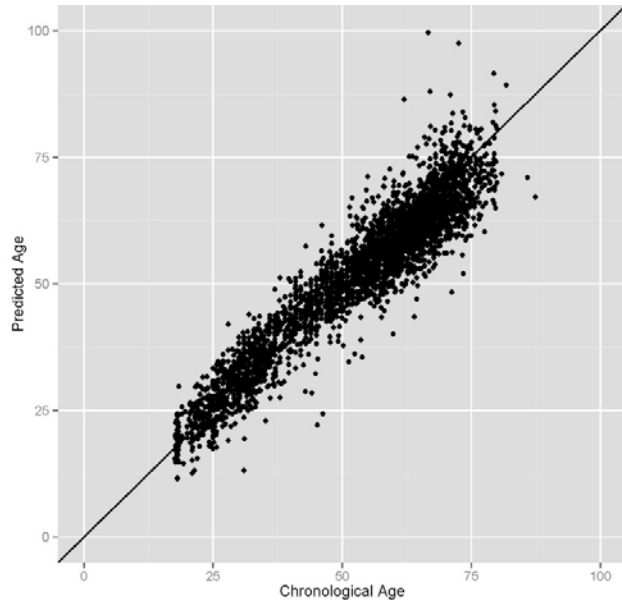
Lopes-Otin et al. *Cell* 2013:  
'9 hallmarks of ageing'



Kennedy et al. *Cell* 2014:  
'7 pillars of ageing'

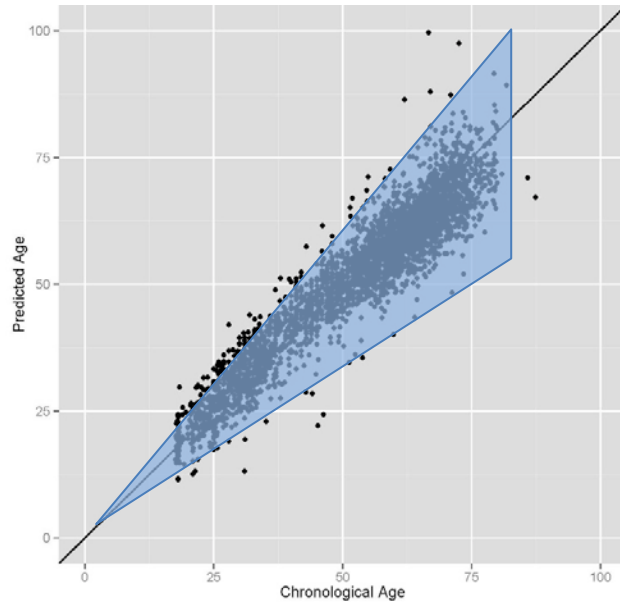


# DNA methylation tracks chronological age

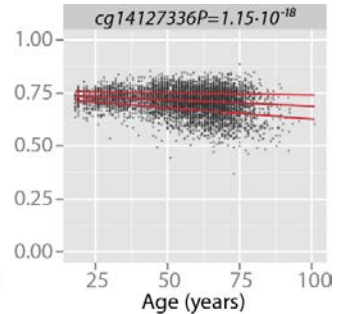
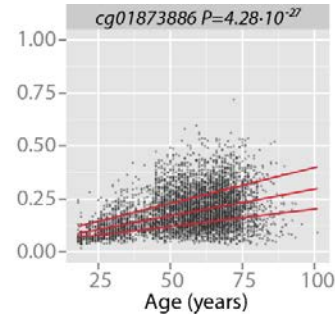
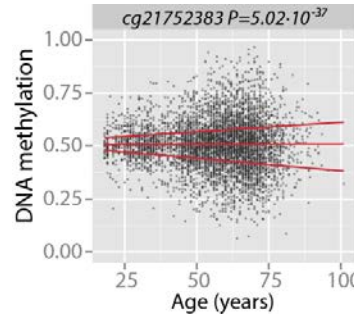


DNA meth clock of Steve Horvath of 353 CpGs (*Genome Biol* 2013) applied on our own data (N>3000).

# Loss of epigenetic control with age

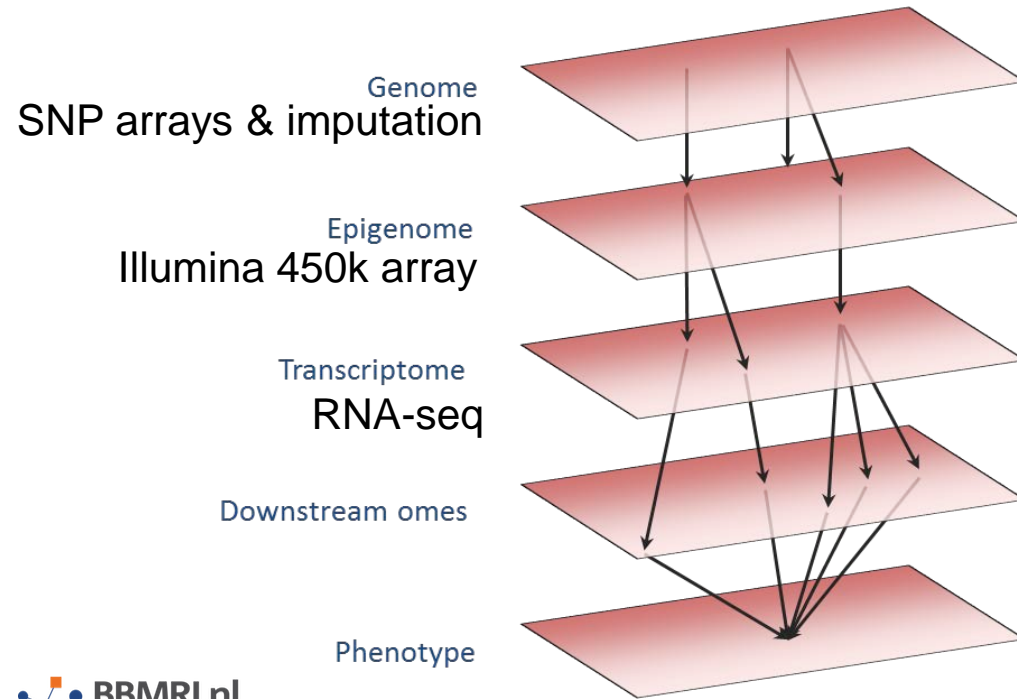


age-related Variably Methylated Positions  
aVMPs

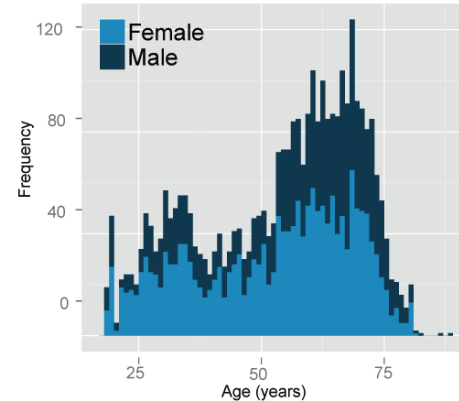


- Biology remains elusive.
- No link with gene expression.
- We already know chronological age.

# Large scale: Biobank-based Integrative Omics Study (BIOS)



- Whole blood samples.
- 3295 methylome profiles: 450k array.
- 2044 transcriptome profiles: RNA-seq.



# Genomics papers

1. Discovery
2. Validation
3. Interpretation

And often public data helps a lot!

# Get a flavour of the possibilities

Slieker et al. *Genome Biology* (2016) 17:191  
DOI 10.1186/s13059-016-1053-6

Genome Biology

RESEARCH

Open Access



## Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms

Roderick C. Slieker<sup>1</sup>, Maarten van Iterson<sup>1</sup>, René Luijk<sup>1</sup>, Marian Beekman<sup>1</sup>, Daria V. Zhernakova<sup>2</sup>, Matthijs H. Moed<sup>1</sup>, Hailiang Mei<sup>3</sup>, Michiel van Galen<sup>4</sup>, Patrick Deelen<sup>2</sup>, Marc Jan Bonder<sup>2</sup>, Alexandra Zhernakova<sup>2</sup>, André G. Uitterlinden<sup>5</sup>, Ettje F. Tigchelaar<sup>2</sup>, Coen D. A. Stehouwer<sup>6</sup>, Casper G. Schalkwijk<sup>6</sup>, Carla J. H. van der Kallen<sup>6</sup>, Albert Hofman<sup>7</sup>, Diana van Heemst<sup>8</sup>, Eco J. de Geus<sup>9</sup>, Jenny van Dongen<sup>9</sup>, Joris Deelen<sup>1</sup>, Leonard H. van den Berg<sup>10</sup>, Joyce van Meurs<sup>5</sup>, Rick Jansen<sup>11</sup>, Peter A. C. 't Hoen<sup>4</sup>, Lude Franke<sup>2</sup>, Cisca Wijmenga<sup>2</sup>, Jan H. Veldink<sup>10</sup>, Morris A. Swertz<sup>12</sup>, Marleen M. J. van Greevenbroek<sup>6</sup>, Cornelia M. van Duijn<sup>13</sup>, Dorret I. Boomsma<sup>9</sup>, BIOS consortium, P. Eline Slagboom<sup>1</sup> and Bastiaan T. Heijmans<sup>1\*</sup>

# Reference epigenomes inform on biological function

