

The track in which the computer comes first and the pipette second

Data @work

Dry-lab

Data Science

Data in Life Sciences

Quantitative Biomedical Sciences

Computational Biology

Data-driven Research

Data

Computational Life Sciences

Data-driven Life Sciences

Big Data

Research Methodology

Data track: from idea to conclusion

Data in Life Sciences

Bioinformatics

Exploiting public data

Bas Heijmans
Molecular Epidemiology
Leiden University Medical Center
The Netherlands
bas.heijmans@lumc.nl

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>
- ARCHS –RNA-seq gene counts from GEO ready for R
<https://amp.pharm.mssm.edu/archs4/index.html>

Databases

GWAS summary statistics

- GWAS catalog of genome-wide significant SNPs
<https://www.ebi.ac.uk/gwas/>
- MR-base: summary statistics with focus on Mendelian randomization
<http://www.mrbase.org/>
- UK Biobank GeneAtlas
<http://geneatlas.roslin.ed.ac.uk/>

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>
- Omics atlases (effect of SNPs on expression/methylation and much more)
<http://bbmri.researchlumc.nl/atlas/>

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



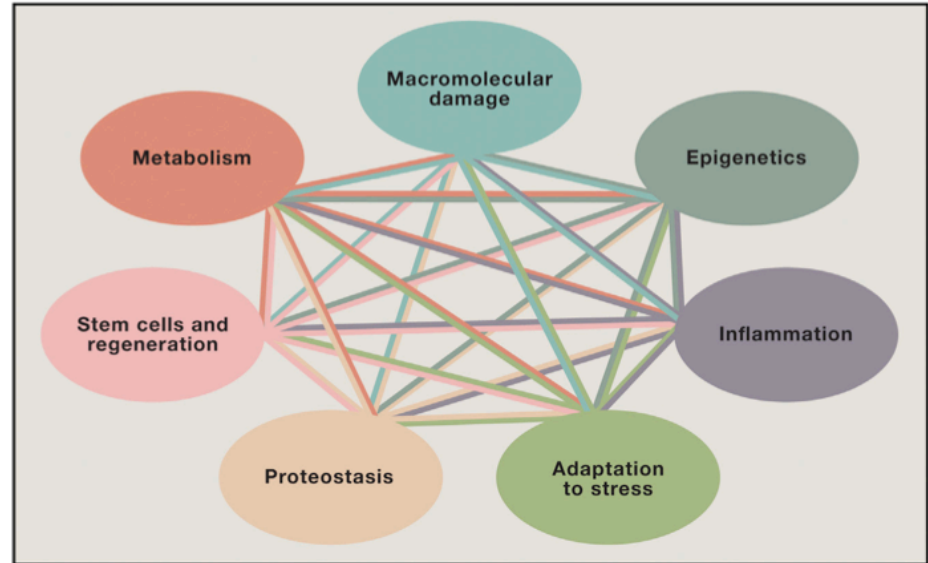
Roderick C. Slieker¹, Maarten van Iterson¹, René Luijk¹, Marian Beekman¹, Daria V. Zhernakova², Matthijs H. Moed¹, Hailiang Mei³, Michiel van Galen⁴, Patrick Deelen², Marc Jan Bonder², Alexandra Zhernakova², André G. Uitterlinden⁵, Ettje F. Tigchelaar², Coen D. A. Stehouwer⁶, Casper G. Schalkwijk⁶, Carla J. H. van der Kallen⁶, Albert Hofman⁷, Diana van Heemst⁸, Eco J. de Geus⁹, Jenny van Dongen⁹, Joris Deelen¹, Leonard H. van den Berg¹⁰, Joyce van Meurs⁵, Rick Jansen¹¹, Peter A. C. 't Hoen⁴, Lude Franke², Cisca Wijmenga², Jan H. Veldink¹⁰, Morris A. Swertz¹², Marleen M. J. van Greevenbroek⁶, Cornelia M. van Duijn¹³, Dorret I. Boomsma⁹, BIOS consortium, P. Eline Slagboom¹ and Bastiaan T. Heijmans^{1*}

Genomic dysregulation and ageing

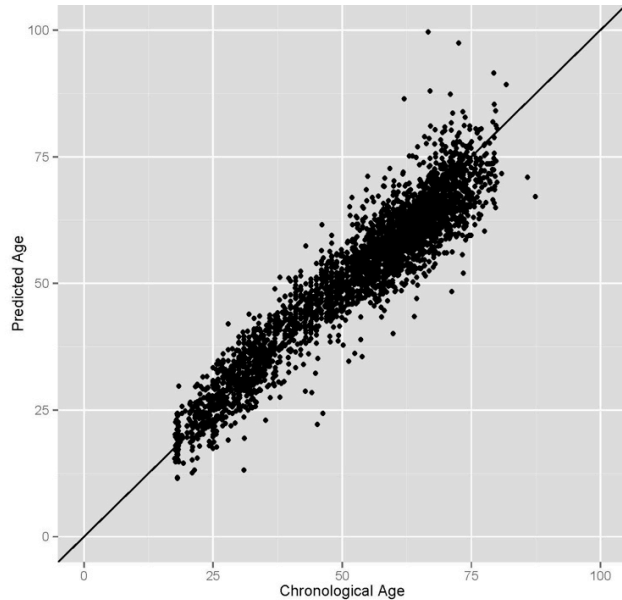
Lopes-Otin et al. *Cell* 2013:
'9 hallmarks of ageing'



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

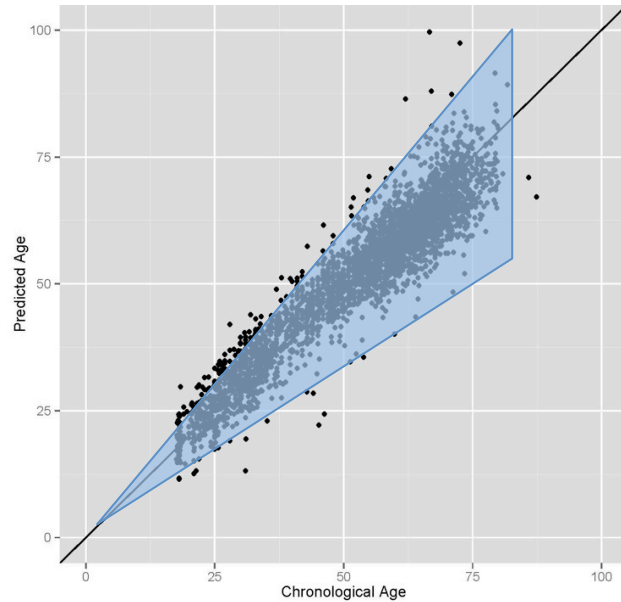


DNA methylation tracks chronological age

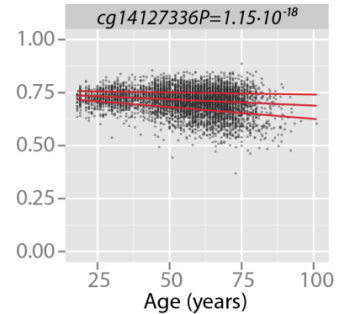
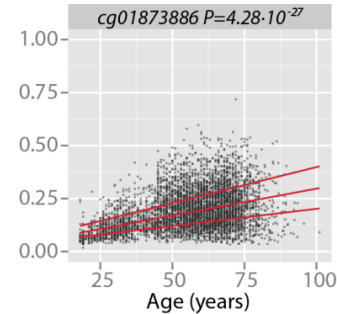
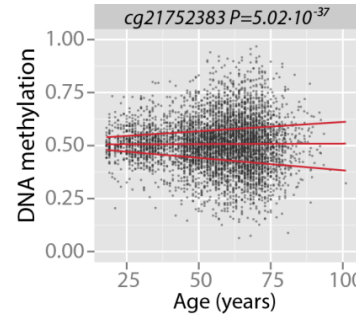


DNA meth clock of Steve Horvath of 353 CpGs (*Genome Biol* 2013) applied on our own data (N>3000).

Loss of epigenetic control with age

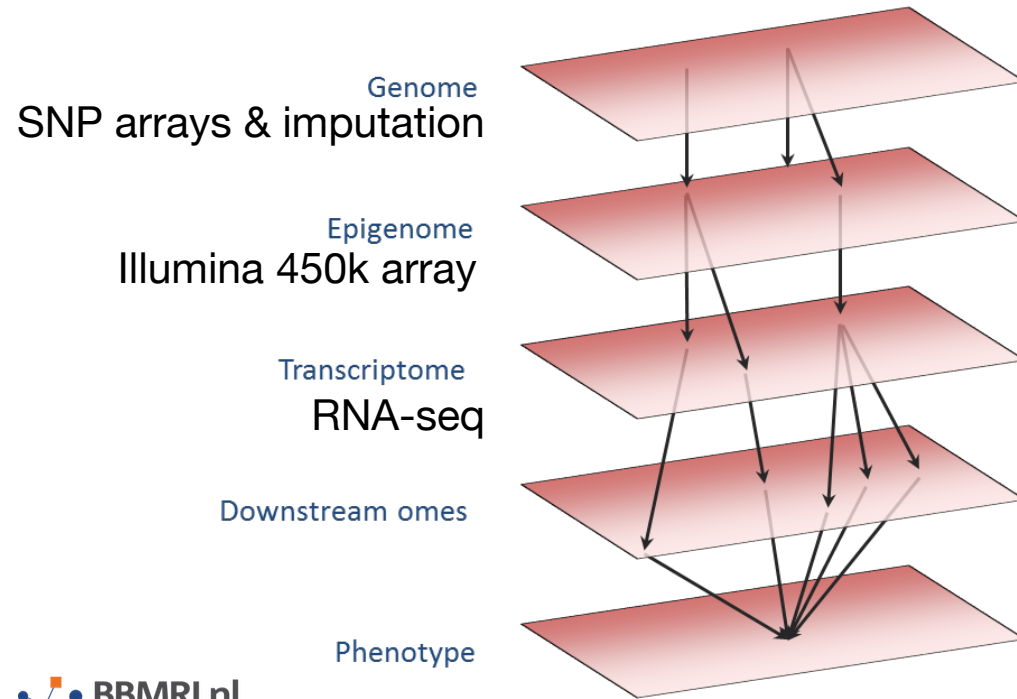


age-related Variably Methylated Positions aVMPs

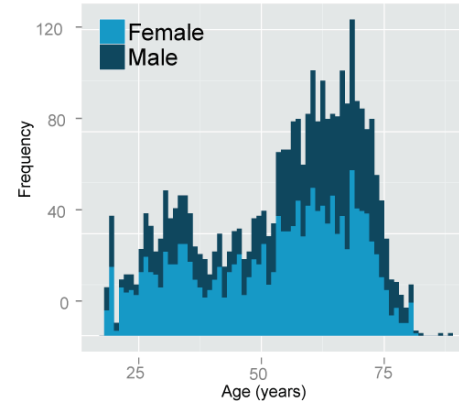


- Biology remains elusive.
- No link with gene expression.
- We already know chronological age.

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. Slieker¹, Maarten van Iterson¹, René Luijk¹, Marian Beekman¹, Daria V. Zhernakova², Matthijs H. Moed¹, Hailiang Mei³, Michiel van Galen⁴, Patrick Deelen², Marc Jan Bonder², Alexandra Zhernakova², André G. Uitterlinden⁵, Ettje F. Tigchelaar², Coen D. A. Stehouwer⁶, Casper G. Schalkwijk⁶, Carla J. H. van der Kallen⁶, Albert Hofman⁷, Diana van Heemst⁸, Eco J. de Geus⁹, Jenny van Dongen⁹, Joris Deelen¹, Leonard H. van den Berg¹⁰, Joyce van Meurs⁵, Rick Jansen¹¹, Peter A. C. 't Hoen⁴, Lude Franke², Cisca Wijmenga², Jan H. Veldink¹⁰, Morris A. Swertz¹², Marleen M. J. van Greevenbroek⁶, Cornelia M. van Duijn¹³, Dorret I. Boomsma⁹, BIOS consortium, P. Eline Slagboom¹ and Bastiaan T. Heijmans^{1*}

Reference epigenomes inform on biological function

