

Epigenetic Variability in Healthy Aging & Exceptional Longevity

Simone Ecker

UCL Cancer Institute

s.ecker@ucl.ac.uk

Centro Centroamericano de Población (CCP)

University of Costa Rica

25th September 2019

Every single individual is unique



- First cloned cat [*Shin et al. Nature. 2002*]
- Donor (left) and clone (right) are genetically identical
- Coat color is not the same

⇒ **Epigenetic variability**

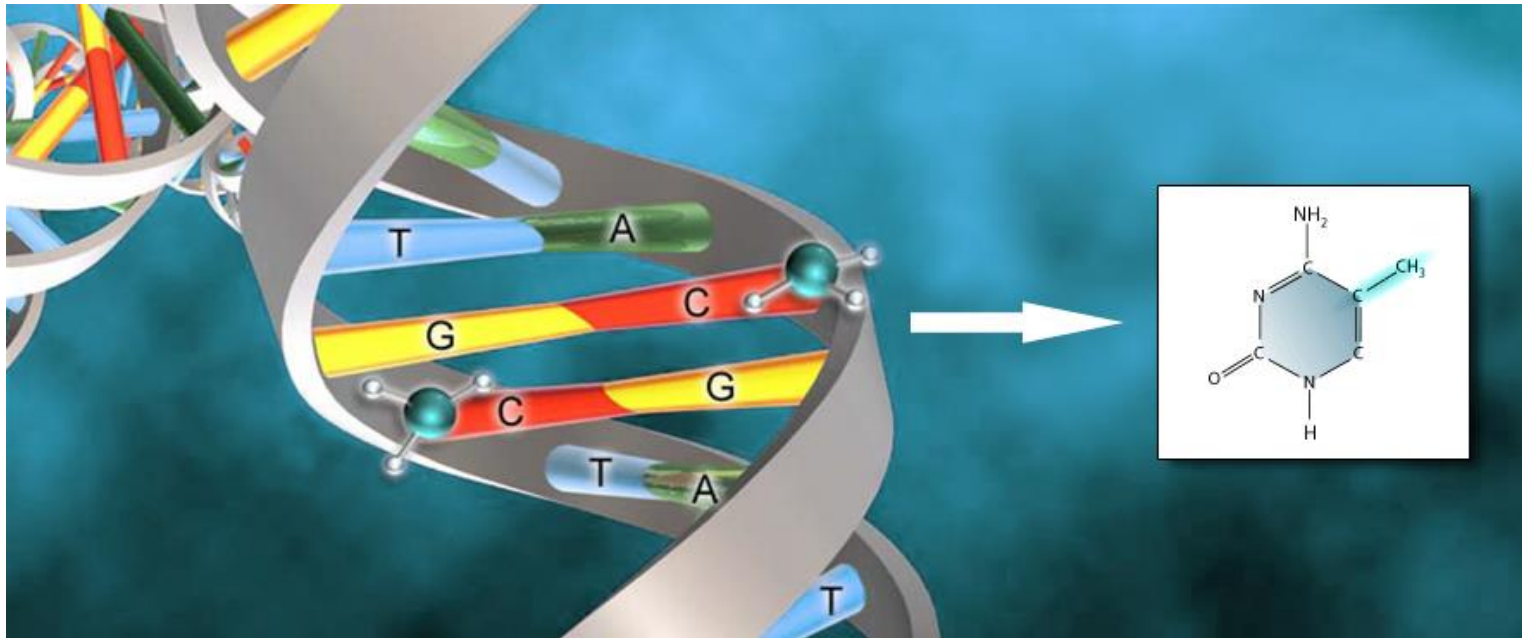
What is epigenetics?

- Epigenetics is the study of changes to the genome which are ***not*** caused by alterations in the DNA sequence
- Chemical compounds modify the genome “on top” of it
- Epigenetic modifications are highly dynamic and form the intersection between the genome and the environment
- They play an important role in controlling gene expression and genomic (in-)stability → phenotype

What do epigenetic modifications do?

- Many different cells in our body: brain cells, blood cells, bone cells...
- They all have essentially the same genome
- Epigenetic modifications influence which genes are active and which proteins are produced in a cell, giving rise to the phenotype
- **Multicellular organism has one genome but many epigenomes**

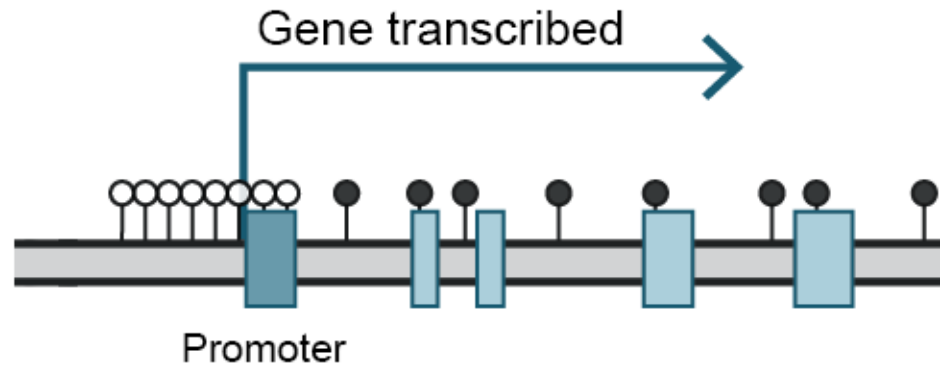
DNA methylation



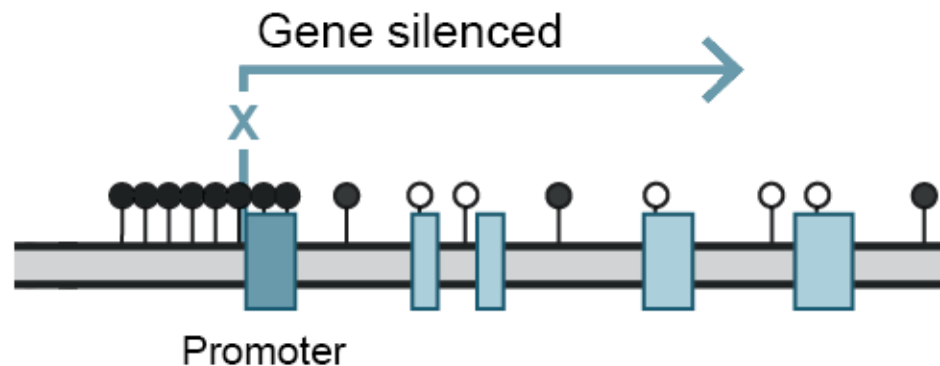
Adapted from jonliefmd.com

- DNA methylation is the most studied epigenetic modification
- Adds methyl group to DNA base cytosine → methylcytosine
- Regulates transcription, X chromosome inactivation, etc.

DNA methylation and transcription



● methylated
○ unmethylated



Example: Agouti gene variable methylation

- Agouti promoter methylated (right)
 - Gene not expressed
 - Brown healthy mouse
- Agouti promoter unmethylated (left)
 - Gene expressed
 - Yellow obese mouse with diabetes, cancer
- Depends on diet of mother during pregnancy: methyl-rich or standard diet
- Gene can be methylated to varying degrees
- Even differs from cell to cell



These two mice are genetically identical and have the same age

Source: learn.genetics.edu



Source: Morgan et al. Nat Genet. 1999

“Life is a study in contrasts between randomness and determinism”

Cell

Leading Edge
Review

Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences

Arjun Raj¹ and Alexander van Oudenaarden^{1,*}¹Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Correspondence: avano@mit.edu

DOI 10.1016/j.cell.2008.09.050

Gene expression is a fundamentally stochastic process, with randomness in transcription and translation leading to cell-to-cell variations in mRNA and protein levels. This variation appears in organisms ranging from microbes to metazoans, and its characteristics depend both on the biophysical parameters governing gene expression and on gene network structure. Stochastic gene expression has important consequences for cellular function, being beneficial in some contexts and harmful in others. These situations include the stress response, metabolism, development, the cell cycle, circadian rhythms, and aging.

Introduction

Life is a study in contrasts between randomness and determinism: from the chaos of biomolecular interactions to the precise coordination of development, living organisms are able to resolve these two seemingly contradictory aspects of their internal workings. Scientists often reconcile the stochastic and the deterministic by appealing to the statistics

et al. (1990). They examined the effect of different doses of glucocorticoid on the expression of a glucocorticoid-responsive transgene encoding beta-galactosidase and found that the cell-to-cell variability in the expression of the transgene was surprisingly large. Moreover, increasing the dose led to an increased frequency of cells displaying a high level of expression rather than a uniform increase in expression in every cell;

Biological variability

- Genetically identical cells or organisms display an incredible variety of phenotypes, even in homogenous environments
[Gaertner, Lab Anim, 1990]
- Arises from randomness and noise present in all biological systems and processes
- This variability
 - plays a key role in development and cellular differentiation in multicellular organisms allowing for selection and propagation of cell type specific expression
 - enables rapid adaptation to changing environmental conditions leading to benefits in survival, e.g. stress-response
 - leads to population robustness, e.g. allowing for the tight control of programmed cell death by graded responses of the population of cells

Different levels of variability

- Cell-to-cell variability in a population of cells
- Inter-individual variability of multicellular organisms
- Variability within and across populations and species
- Spatiotemporal variability

Different levels of variability are related to each other

- Correspondence between measuring variability at one time point in a population of 1,000 cells and measuring the variability of one cell across 1,000 time points
- Correlation between cell-to-cell variability and variability across cell populations, and even across species to a lesser extent

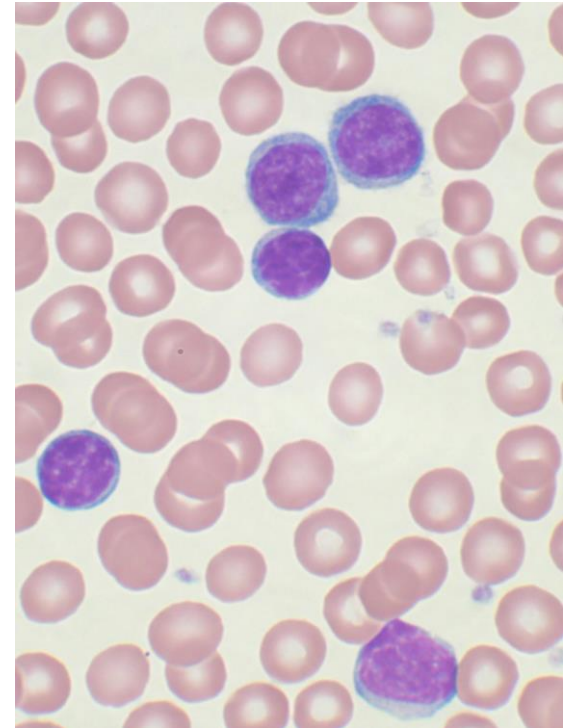
Importance of variability

- **Aging**
 - Increased variability with age
- **Disease**
 - Cancer!
 - Autoimmune diseases:
Type 1 Diabetes, Rheumatoid Arthritis...
 - ...
- **Therapy**
 - Fractional killing
 - Therapeutic resistance
 - Personalized medicine!



Chronic Lymphocytic Leukemia (CLL)

- Most frequent leukemia in adults
- Two subtypes based on the mutational status of IGHV region:
 - M-CLL
 - high level of IGHV mutations
 - favorable clinical outcome
 - U-CLL
 - no or low level of IGHV mutations
 - worse clinical outcome

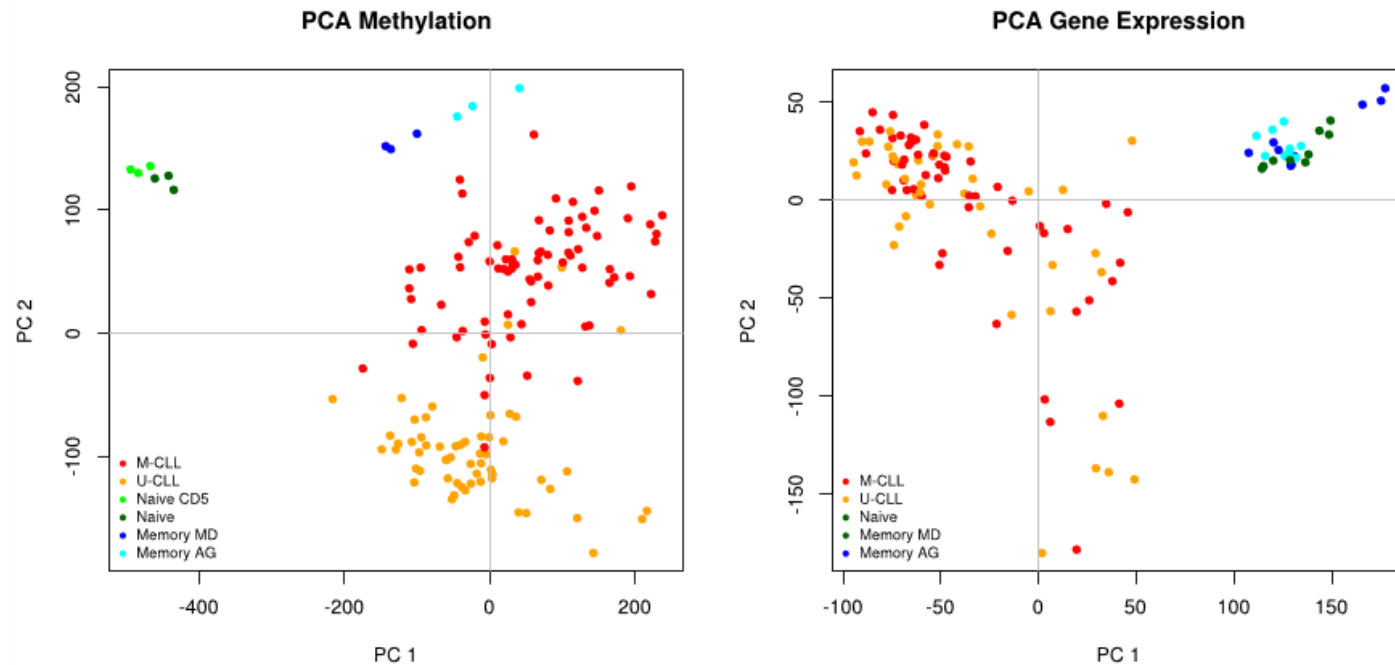


CLL cells

Source: Wikimedia.org

CLL differential expression variability study

- DNA methylation patterns directly related to known subtypes
- But no separation of M-CLL and U-CLL in gene expression data

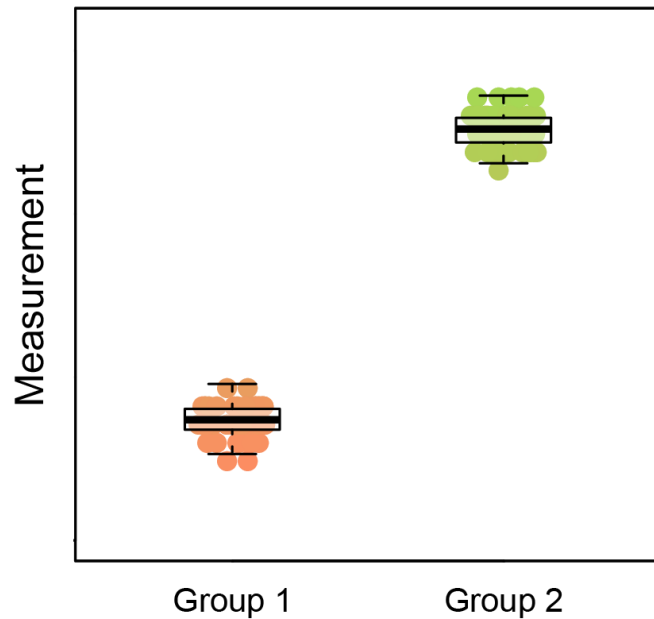


- Difference in variability between M-CLL and U-CLL?

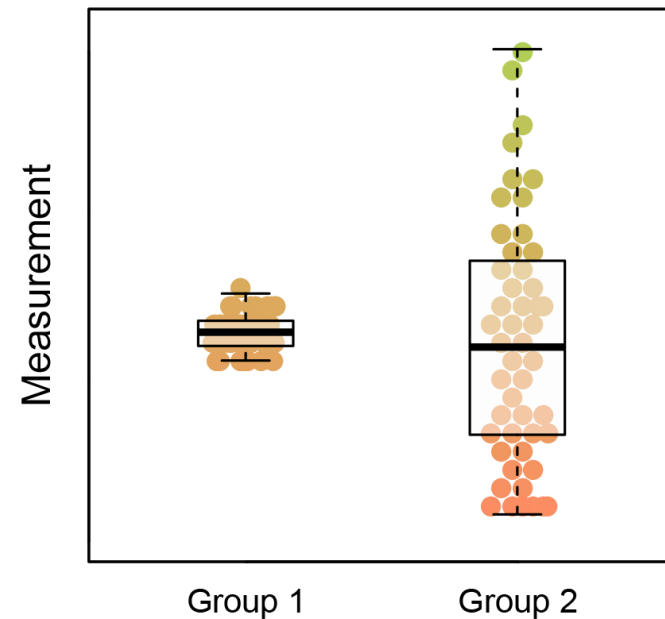
Differential variability

- Differential mean → different but consistent mean
- Differential variability → small vs large deviations from mean

Differential mean

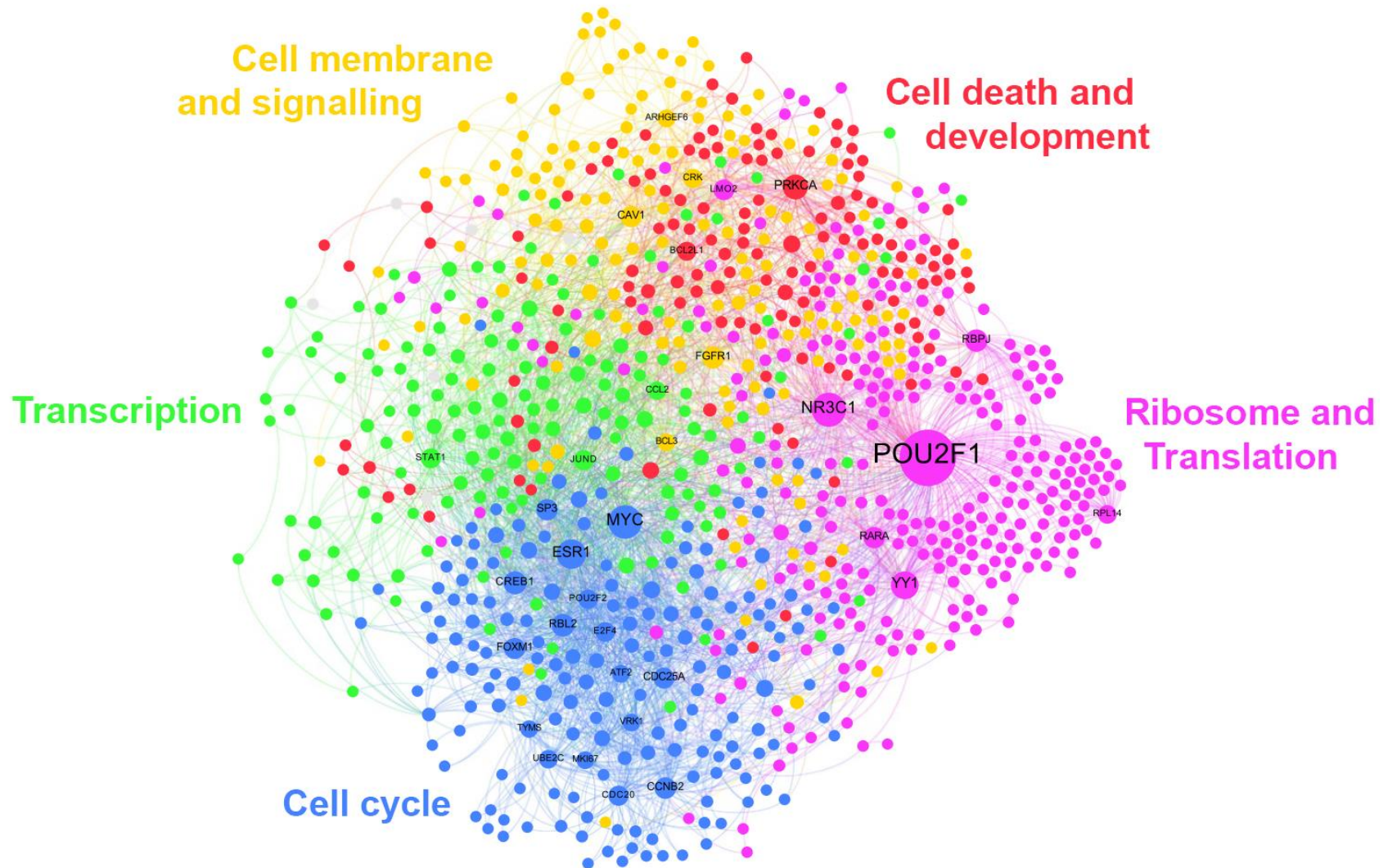


Differential variability



- No difference between M-CLL and U-CLL in gene expression
- Increased variability in U-CLL!

Variability significantly increased in U-CLL



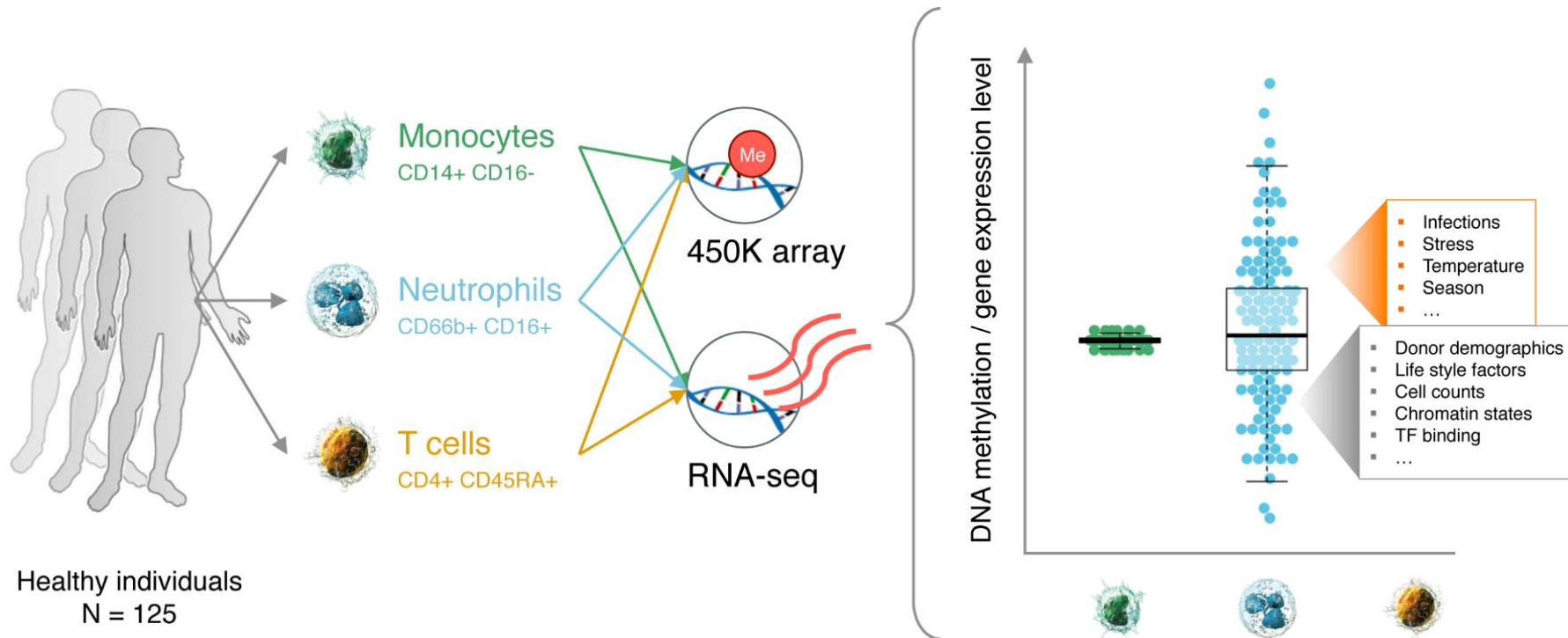
BLUEPRINT

- Large-scale European research effort of the International Human Epigenome Consortium (IHEC)
- Generation of ≥ 100 reference epigenomes of distinct normal human hematopoietic cells and their malignant counterparts
- Investigation of biological processes and mechanisms systematically linking epigenetic variation with phenotypic plasticity

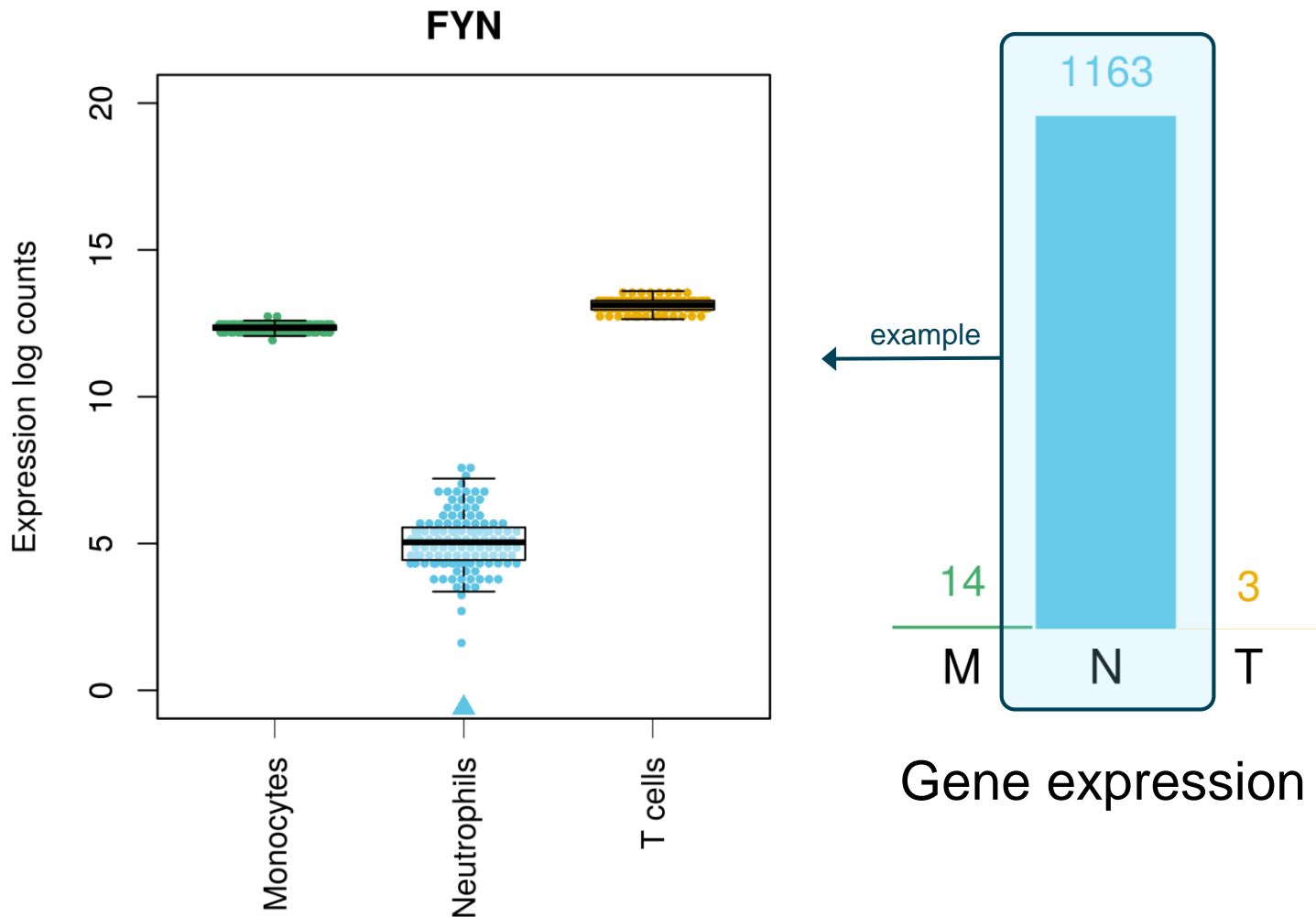
BLUEPRINT Human Variation Epigenome Project

- Chen *et al.* Genetic drivers of epigenetic and transcriptional variation in human immune cells. *Cell*. 2016.
- Ecker *et al.* Genome-wide analysis of differential transcriptional and epigenetic variability across human immune cell types. *Genome Biol.* 2016.

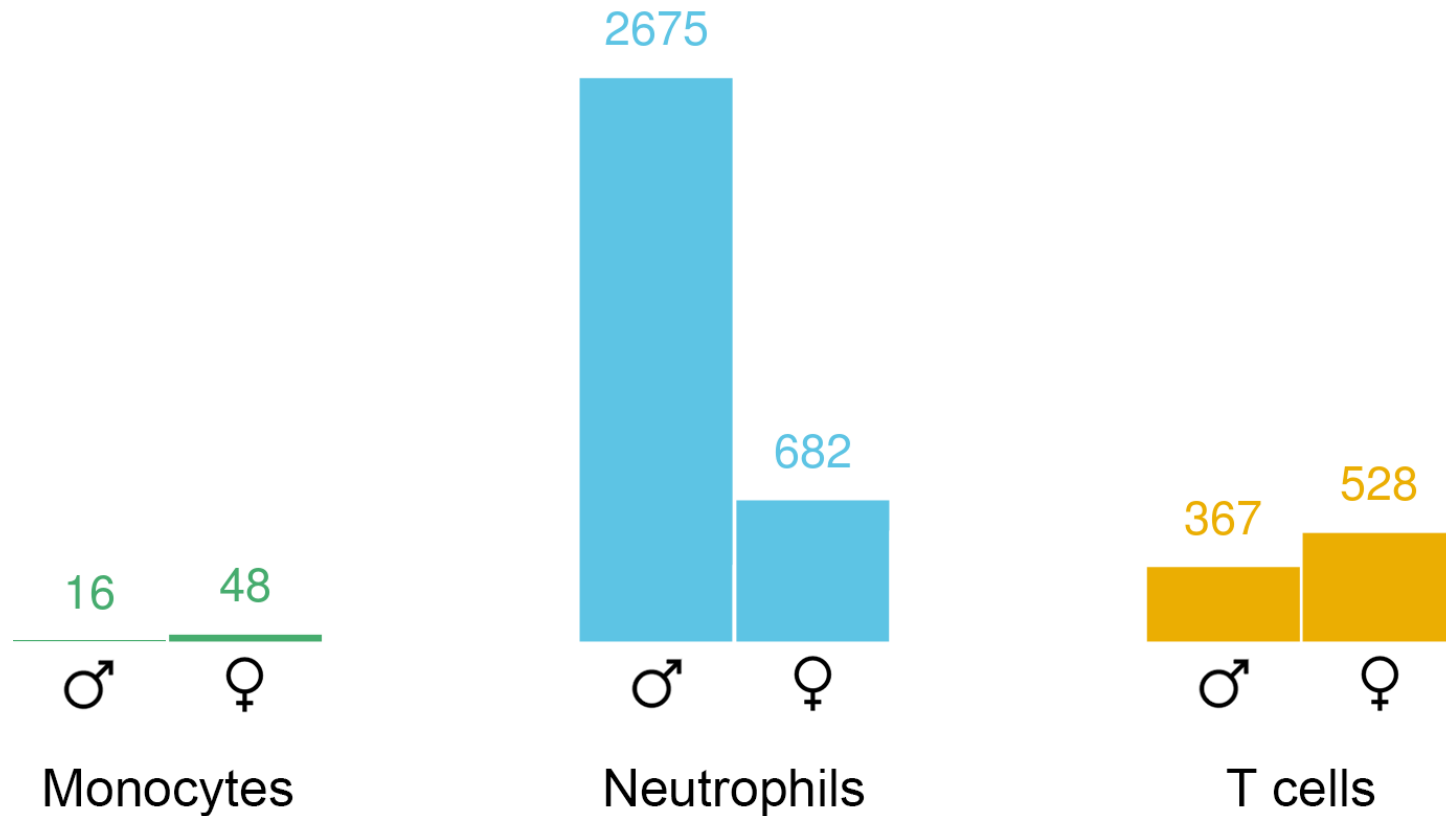
Differential variability in normal blood cells



Increased variability in neutrophils



Sex-specific differential expression

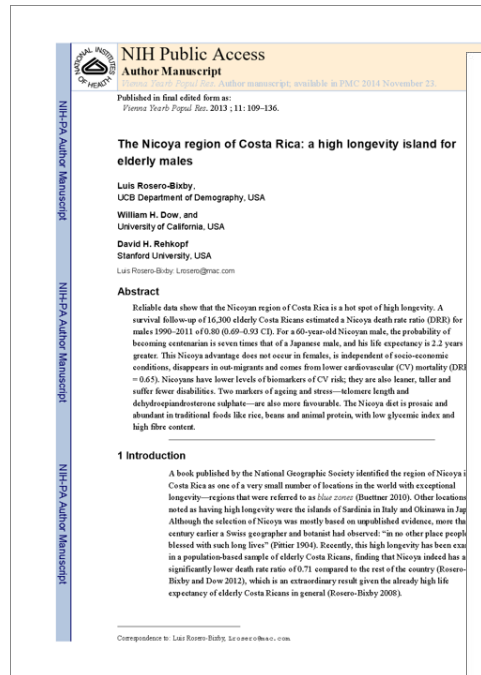


Blue Zones: Highest longevity in the world



Blue Zone of Nicoya

- Nicoya peninsula is the largest Blue Zone in the world
- Nicoyans 60 years old have seven times the probability of reaching 100 years compared to the rest of Costa Rica



1. Introduction
Telomeres are repetitive canonical sequences of DNA at the ends of chromosomes that, together with associated proteins, protect chromosome ends and also prevent the degradation of coding regions of DNA that would otherwise result from the inability of DNA replication enzymes to copy the end of a DNA strand. Although the functional importance of telomeres has been understood for decades, their implication in the human aging process has begun to emerge more recently. Shorter telomeres were first found to occur at older ages (Lee et al., 2002), and more recent work has shown associations with chronic disease (Bouillon-Lé et al., 2007) and mortality, independent of age (Bakaya et al., 2007; Cawthon et al., 2003; Hong et al., 2006). However, associations with mortality have not been consistent, especially for individuals at older ages (Boonekamp et al., 2013; Mather et al., 2011). Furthermore, the social and economic determinants of telomere length are not yet clear. There is some evidence that shorter telomeres are

associated with chronic stress (Epel et al., 2004; Kiecolt-Glaser, 2009; Kiecolt-Glaser et al., 2006; Cherkas et al., 2006; Steptoe et al., 2006) and that associations are stronger with earlier life position, such as education (Hendelman et al., 2002). The Nicoya Peninsula region in Costa Rica accreted as a region with exceptionally high longevity (Rosero-Bixby et al., 2010). Mortality rates of elderly people lower than in the rest of Costa Rica across up to a population-based sample of close to 60 years and over (Rosero-Bixby and Dow 2012) or three years of additional life expectancy (Rosero-Bixby et al., 2010) have already been reported. The present article aims to examine health in Nicoya by studying a representative population-based sample of elderly Nicoyans. Understanding whether life in Nicoya differs from other Costa Rican



DNA methylation pilot study

McEwen et al. *Epigenetics & Chromatin* (2017) 10:21
DOI 10.1186/s13072-017-0128-2

Epigenetics & Chromatin

RESEARCH

Open Access



Differential DNA methylation and lymphocyte proportions in a Costa Rican high longevity region

Lisa M. McEwen¹ , Alexander M. Morin¹, Rachel D. Edgar¹, Julia L. MacIsaac¹, Meaghan J. Jones¹, William H. Dow², Luis Rosero-Bixby³, Michael S. Kobor¹ and David H. Rehkopf^{4*}

Abstract

Background: The Nicoya Peninsula in Costa Rica has one of the highest old-age life expectancies in the world, but the underlying biological mechanisms of this longevity are not well understood. As DNA methylation is hypothesized to be a component of biological ageing, we first used an existing methylome dataset to determine the associations

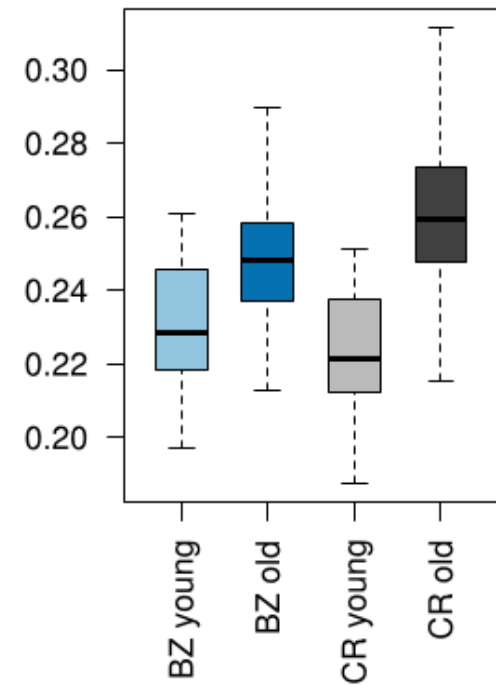
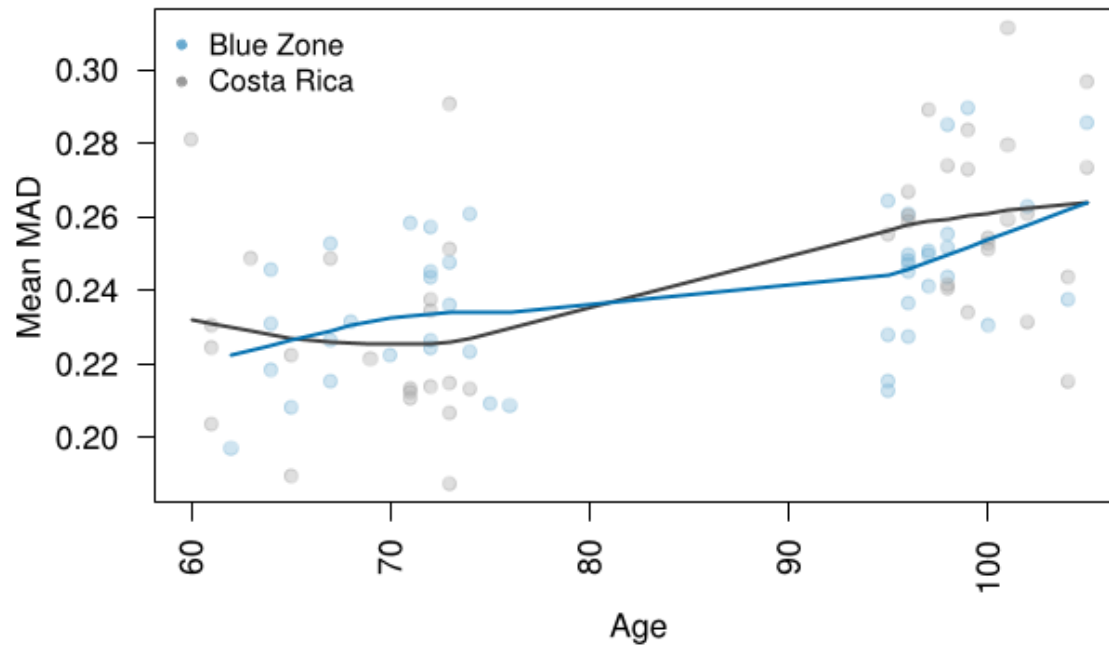
Pilot data set

- Cross-sectional subsample (n=95) of Costa Rican Longevity and Healthy Aging Study (CRELES, n>2500)
- Age-matched whole blood DNA methylation (450K array) of Nicoyans and non-Nicoyan Costa Ricans in two age groups

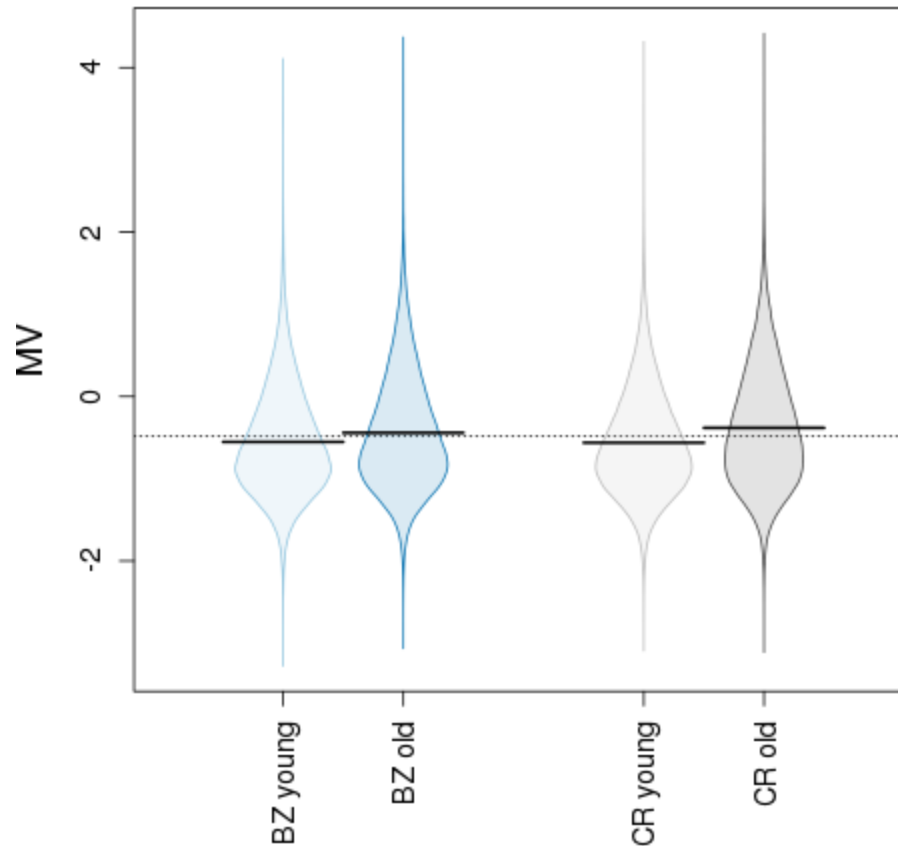
Table 1 Cohort characteristics (means and percents), Nicoyans and non-Nicoyans

Characteristics	Nicoya (n = 48)	Non-Nicoya (n = 47)
Age (mean in years)	83 (14)	85 (16)
Female (%)	57	55
Low education (%)	80	68
Low wealth (%)	35	21
Currently smoke (%)	4	6
Systolic blood pressure (mean mmHg)	139 (23)	140 (25)
Diastolic blood pressure (mean mmHg)	78 (12)	78 (13)
Body mass index (mean)	24 (7.1)	25 (5.8)

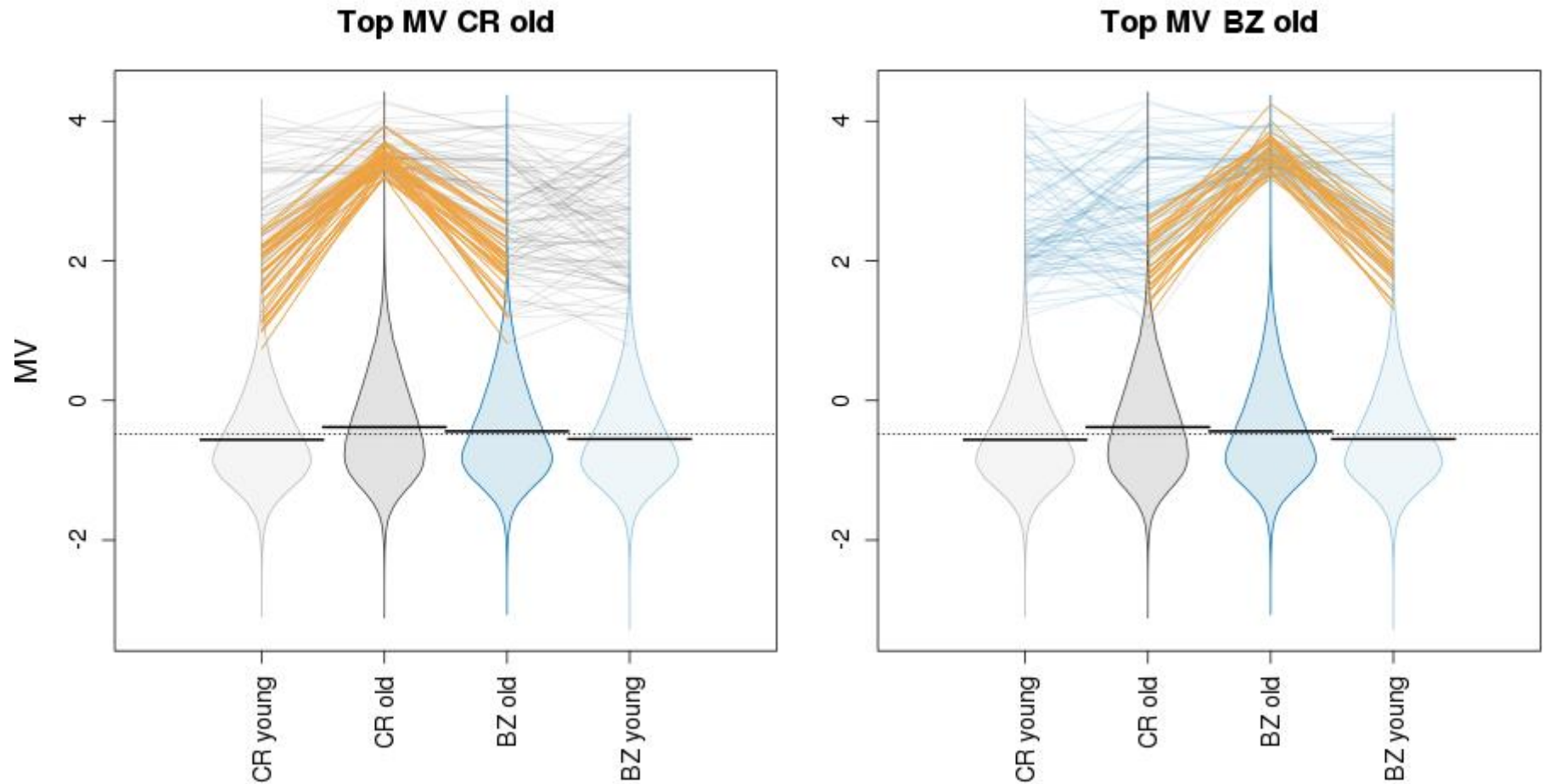
Global patterns of variability (sample-wise)



Global patterns of variability (CpG-wise)



Top 100 variable CpGs



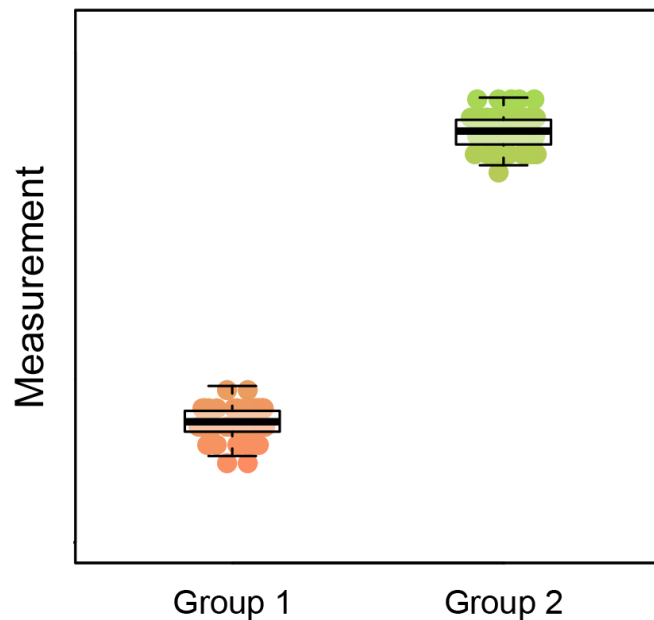
Top 100 CpGs with high variability in BZ

- **32 CpGs in 22 genes (9 in promoters, 17 in bodies, 6 in IGRs)**
 - **GFPT2 (4 CpGs):** Glucose flux in hexosamine pathway
 - *MIR885* (3 CpGs)
 - **RPTOR:** Control of **rapamycin complex** activity in response to nutrient and hormonal signals
 - *KCNG2:* Potassium channel activity
 - *ALOXE3:* Lipoxygenase
 - *FN3K:* Deglycation of proteins and fructoselysine
 - *CCS:* Copper delivery, copper chaperone
 - ...
- **GREAT functional enrichment (> 20 significant results), e.g.:**
 - Glutamine metabolic process
 - Sugar biosynthetic process
 - Regulation of fatty acid biosynthetic process

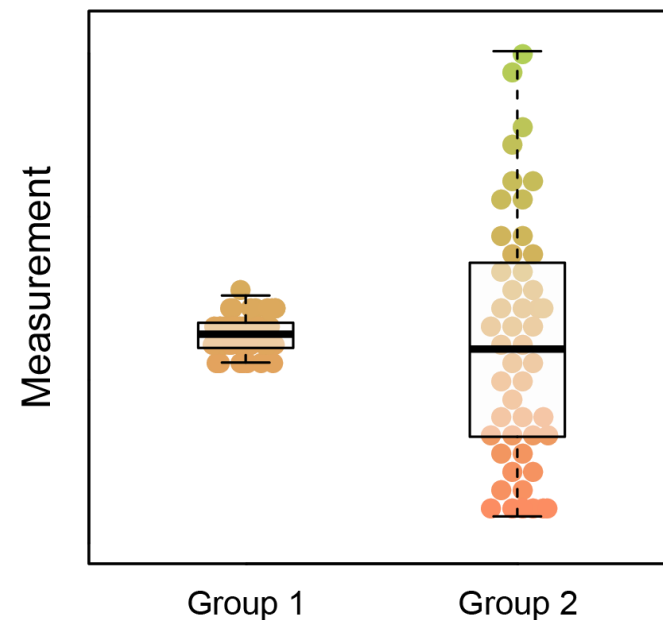
Differential variability

- Differential mean → different but consistent mean
- Differential variability → small vs large deviations from mean

Differential mean



Differential variability



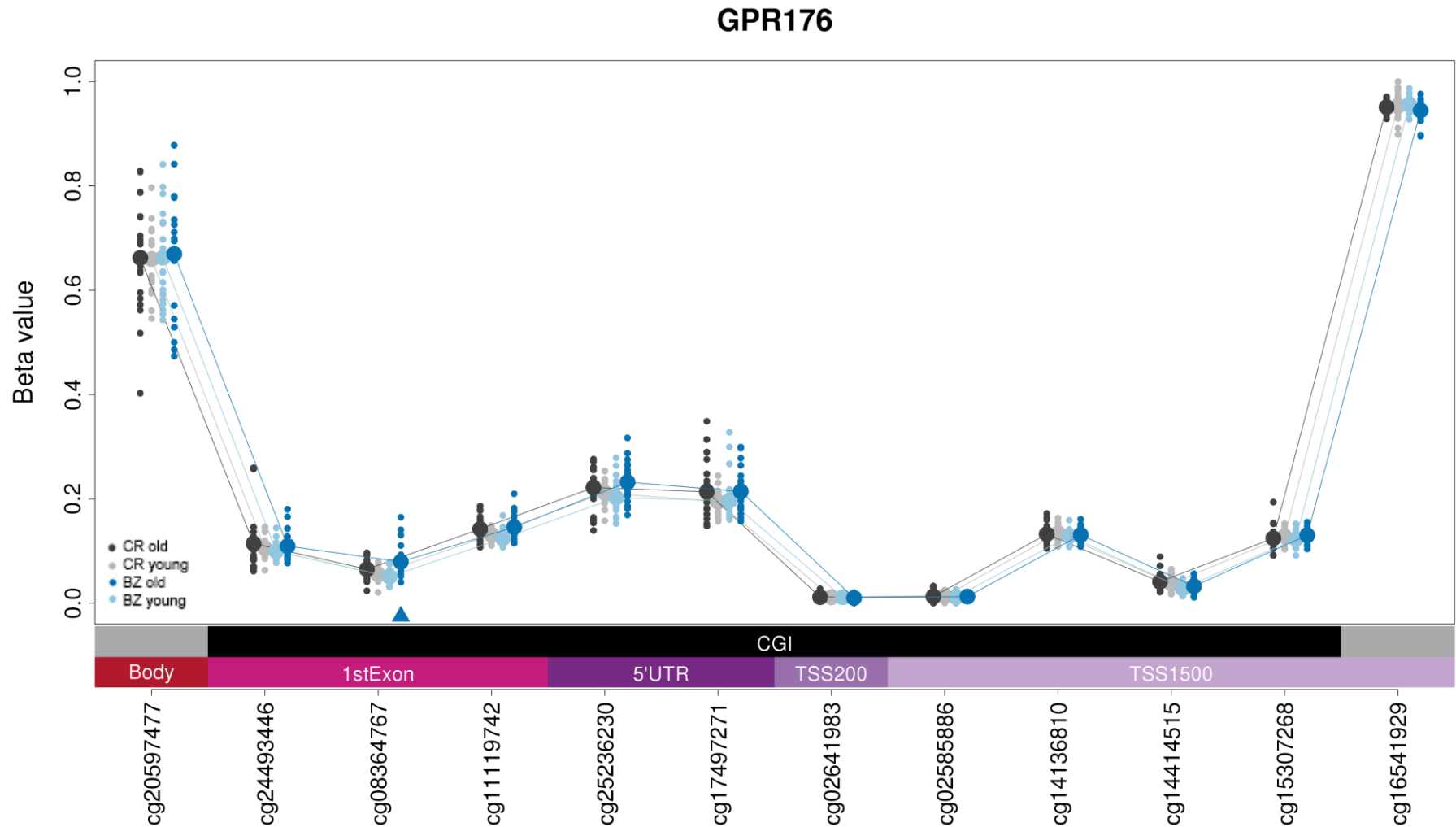
Statistically significant differential variability

- **Applied combination of DiffVar and MV-score**

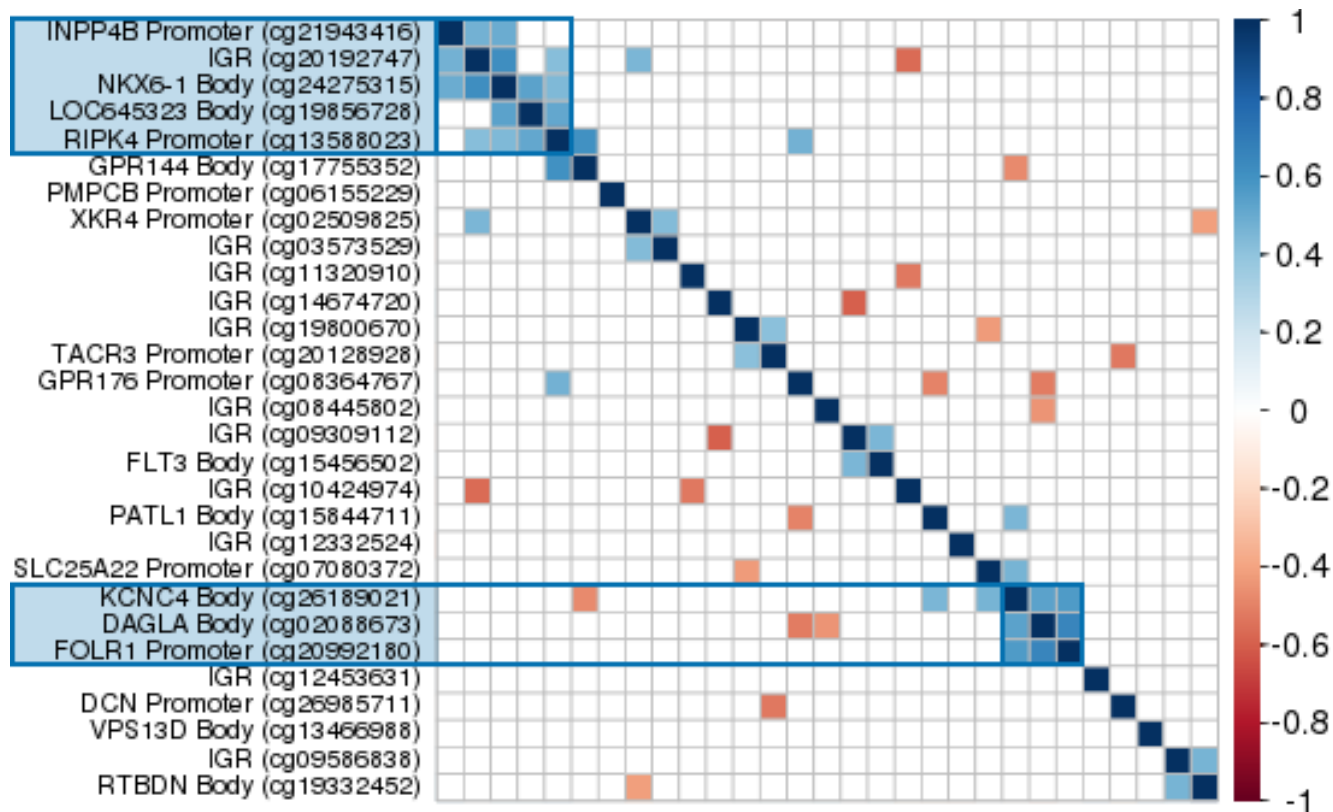
Adjusted p-value (BH) < 0.05 and MV-difference \geq 10%

- CR old versus young: 0 CpGs \uparrow and 0 CpGs \downarrow
- BZ old versus young: **29 CpGs** \uparrow and 5 CpGs \downarrow
18 genes (9 in promoter, 9 in body, 11 in IGR), e.g.:
 - *NKX6-1*: Insulin secretion, glucose detection, response to nicotine
 - *FOLR1*: Folic acid receptor, drug binding
 - *GPR176*: G-prot coupled receptor, response to hormones
 - *DAGLA*: Lipid metabolism, metal binding, G-prot coupled receptor
 - *FLT3*: ATP binding, glucocorticoid receptor binding
 - *RIPK4*: Serine/Threonine Kinase, ATP binding
 - *KCNC4*: Ion and potassium channel activity
 - *TACR3*: Regulation of heart rate, blood pressure

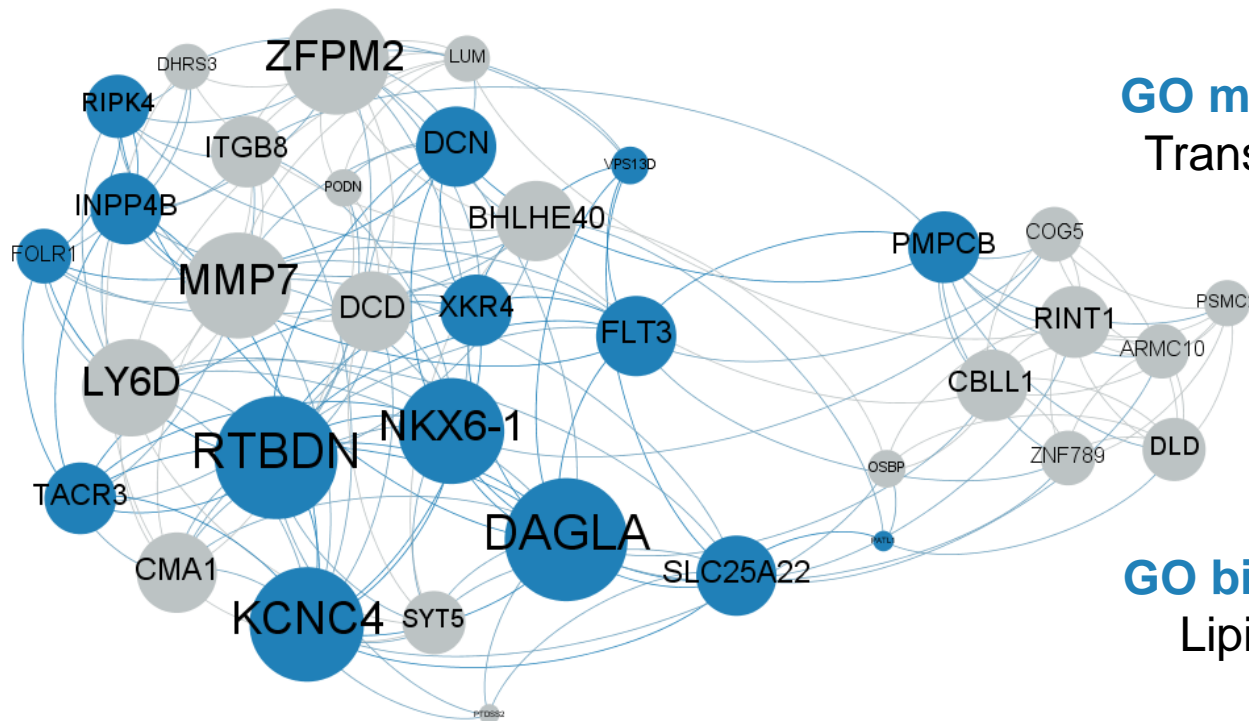
Increased variability in BZ old versus young



Correlation between hypervariable CpGs



Network of genes with sig increased variability



GO cellular compartments:

Extracellular matrix

GO molecular functions:

Transcriptional regulation

GO biological functions:

Lipid metabolic process

Aging

Cardiovascular and circulatory system

JAK-STAT cascade

Organ morphogenesis

Response to cytokine stimulus

- Significantly increased DNAm variability
- Neighboring genes

Sex-specific differential methylation

- **Applied Bumphunter and DMRCate**

Adjusted p-value (BH) < 0.05 and region ≥ 3 neighboring CpGs

- BZ male versus female: 2 CpGs \uparrow and **10 CpGs** \downarrow

8 genes (5 in promoter, 3 in body, 2 in IGR), e.g.:

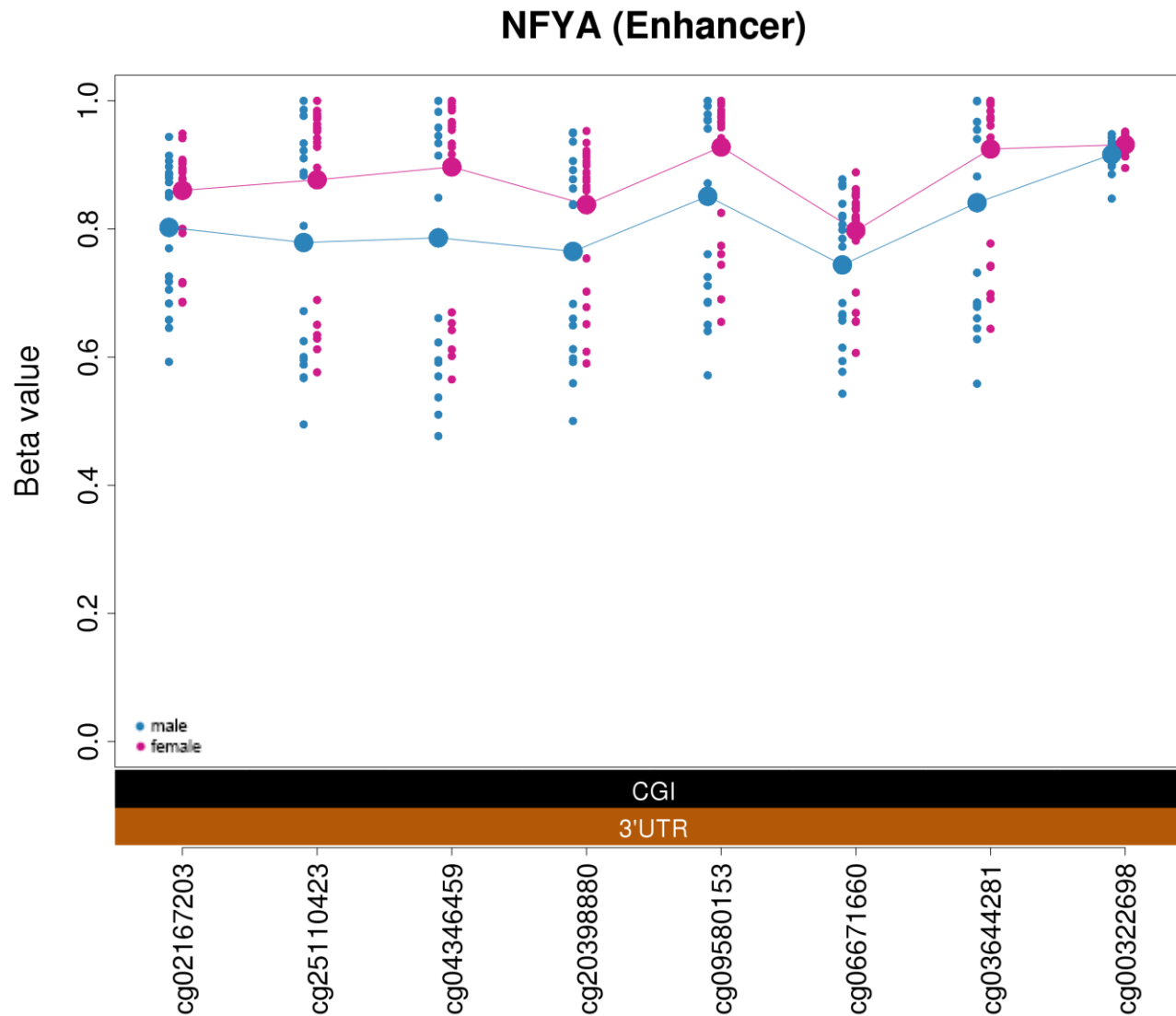
- **NFYA**: TF subunit, enhancer, regulation of cholesterol biosynthesis
- *OR2L12*: Olfactory receptor, G-prot coupled receptor activity
- *DOX43*: RNA binding, ATP binding
- *PSM4A8*: Proteasome subunit, hydrolase activity, spermatogenesis
- *PRRT1*: Proline rich transmembrane protein, response to stimulus
- *ASCL2*: TF activity, multicellular organism development

- CR male versus female: **3 CpGs** \uparrow and 6 CpGs \downarrow

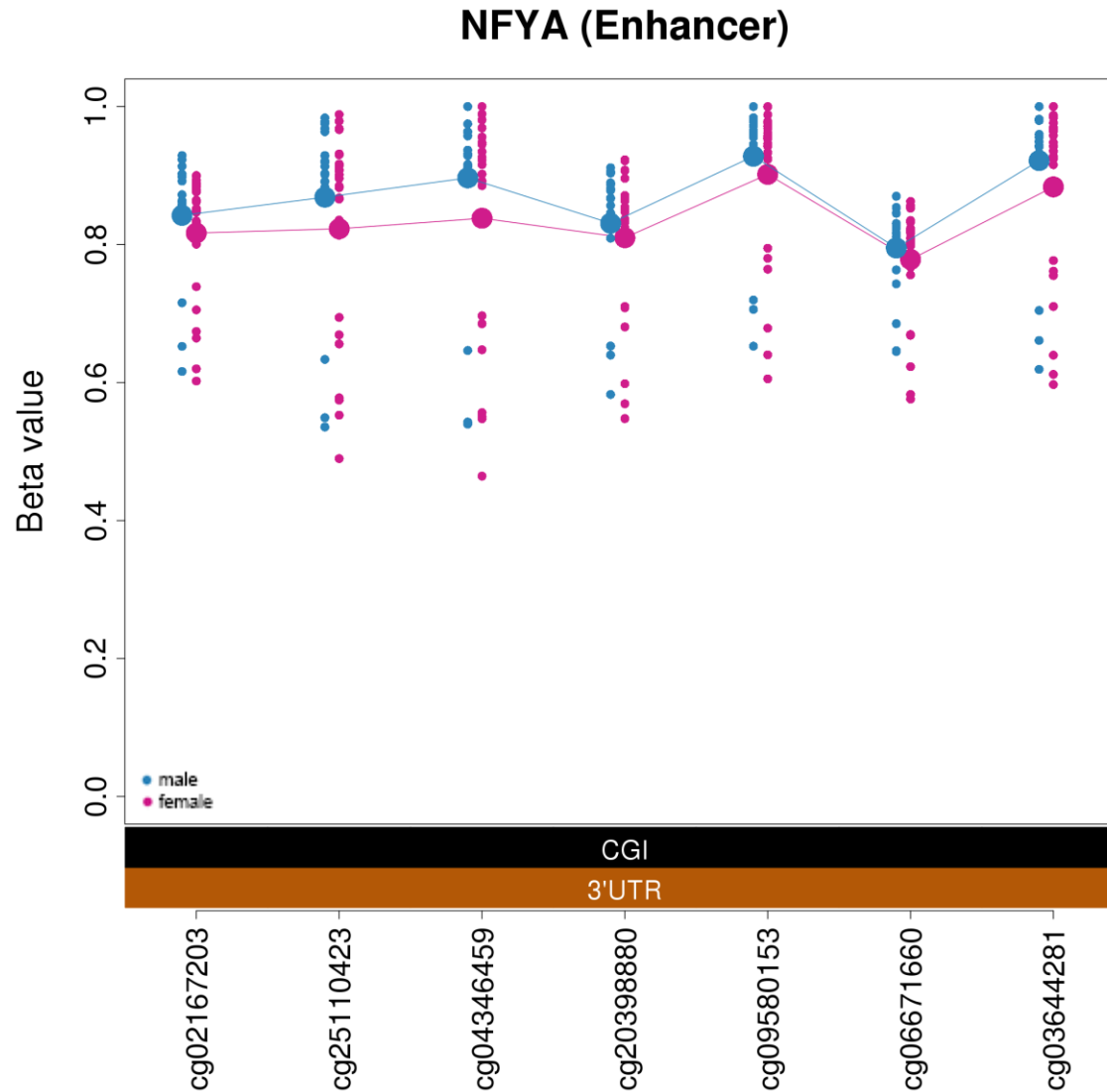
3 genes (2 in promoter, 1 in body), e.g.:

- **NFYA**: TF subunit, enhancer, regulation of cholesterol biosynthesis

NFYA hypomethylation in males of BZ



NFYA hypermethylation in males of CR



Conclusion and discussion

- Pathways and functions involved in all BZ-specific patterns:
 - Nutritional factors, metabolism, hormone signalling
(*GFPT2*, *RPTOR*, *FOLR1*, *NKX6-1*, *NFYA*)
 - Serine/threonine phosphatase
 - G-protein coupled receptor activity
- Discussion:
 - Possible technical and biological confounders
 - Cell subpopulations
 - Relation single cell variability and interindividual variability
 - Larger dataset necessary
- Future plan:
 - 1000 methylomes and matched genotypes currently processed
 - Integration of further healthy aging cohorts
 - Longitudinal follow-up of CRELES participants

Acknowledgments

UCL Medical Genomics Group

Stephan Beck

Andy Feber

Amy Webster

Ismail Moghul

ICGC (CLL) & IHEC (BLUEPRINT)

Iñaki Martín-Subero (University of Barcelona)

Vera Pancaldi (Cancer Research Center of Toulouse)

Dirk Paul (University of Cambridge)

CRELES

David Rehkopf (University of Stanford)

Lisa McEwen (University of British Columbia)

Luis Rosero-Bixby (University of Costa Rica)

Allan Orozco (University of Costa Rica)

Juan Porras (University of Costa Rica)

Michael Kobor (University of British Columbia)

Will Dow (Berkeley University of California)



UNIVERSIDAD DE
COSTA RICA

THANK YOU!

