

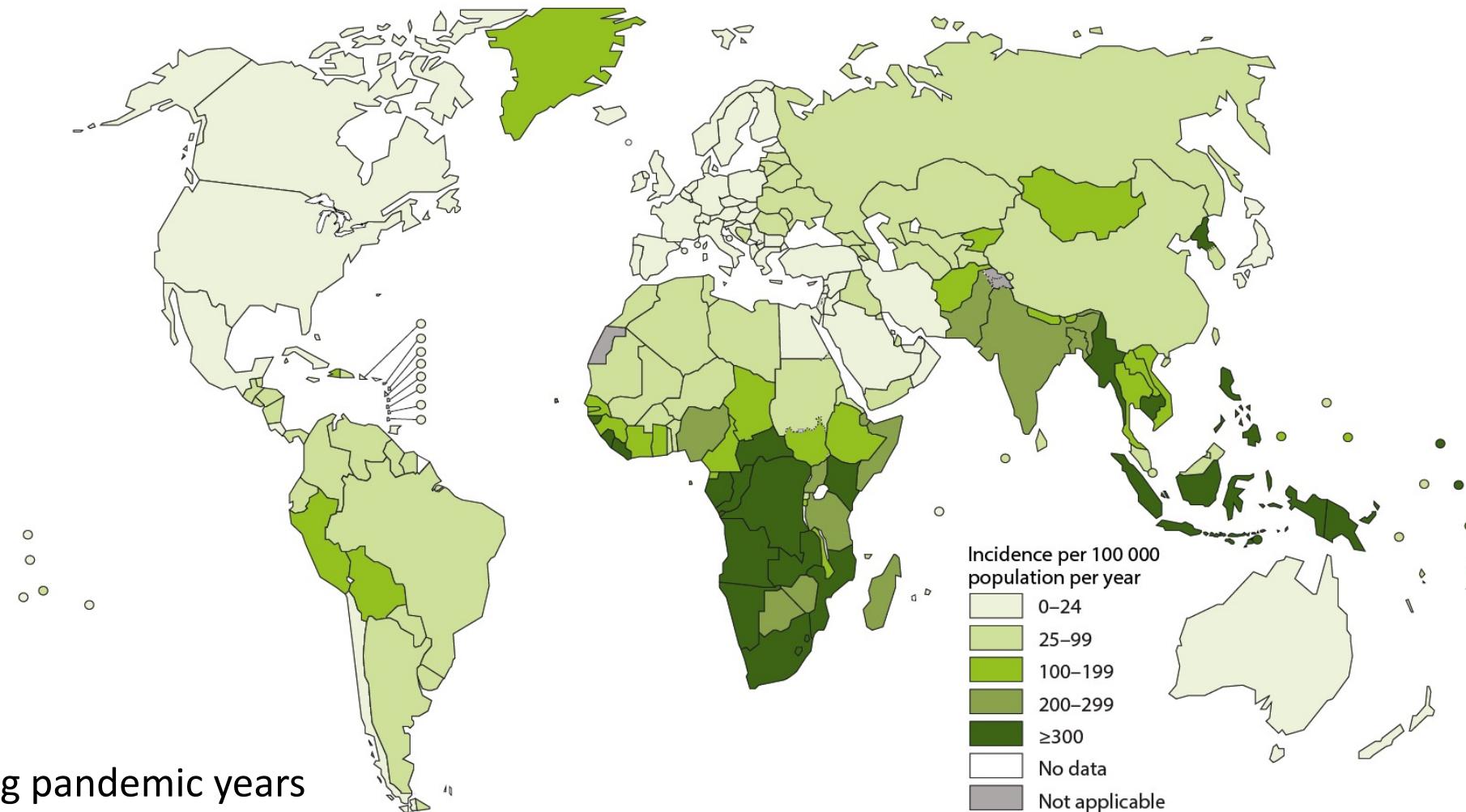
Differential methylation analysis: probes to regions to pathways

Kim Dill-McFarland

i4TB • 2021.06.08

@kdillmcfarland

M. tuberculosis is the leading cause of single agent infectious disease death... and has been for 100s of years.*

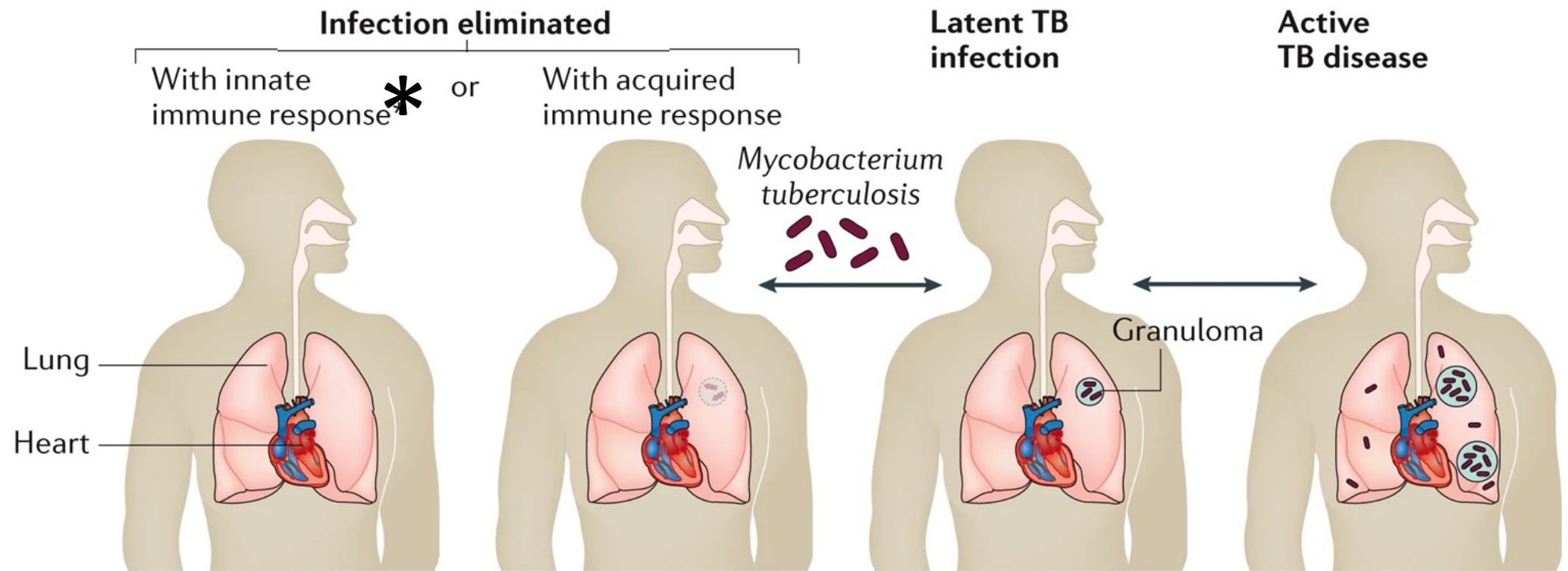


*Except during pandemic years such as 2019-2020 SARS-CoV-2

Image: World Health Org

A subset of the population appears to be
resistant* to *M. tuberculosis* infection.

TST: tuberculin skin test
IGRA: interferon gamma release assay



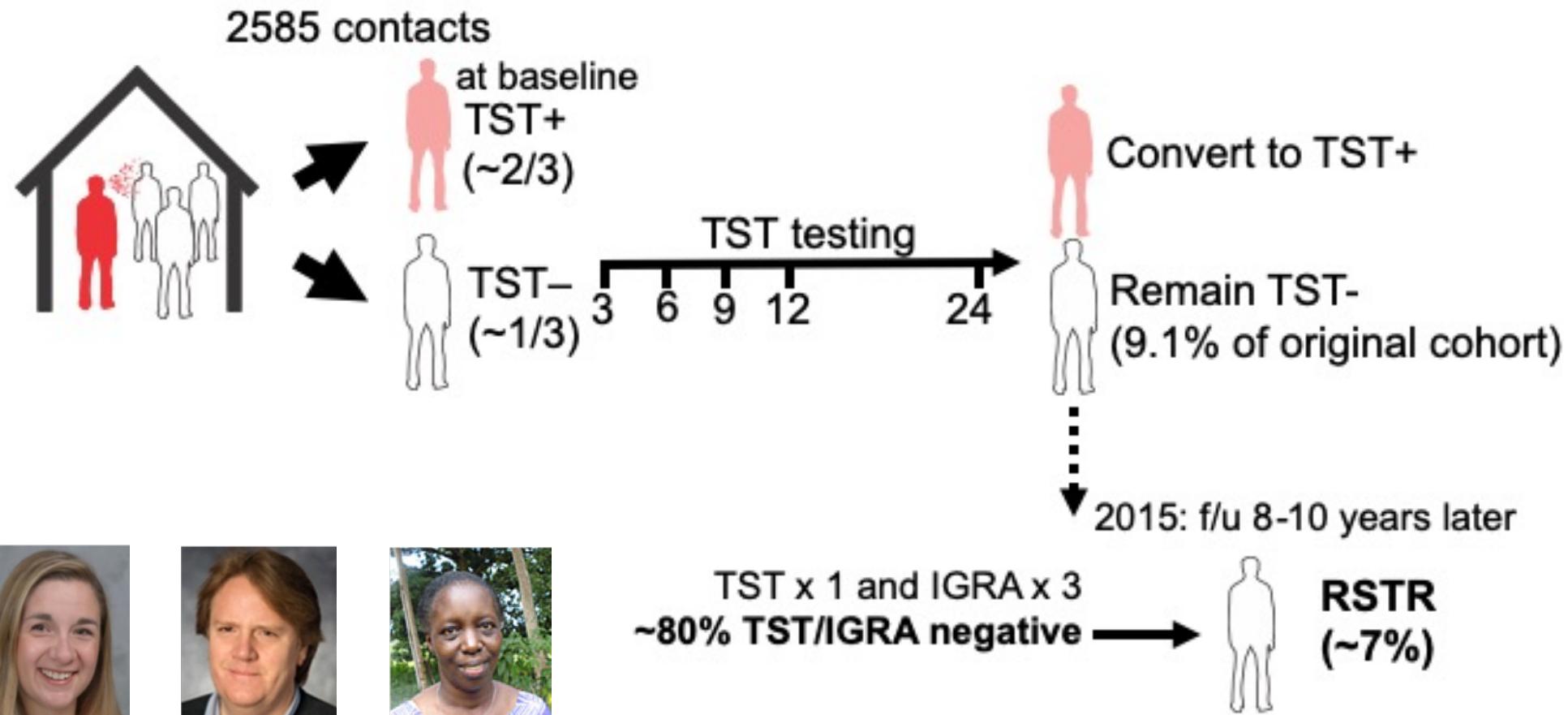
TST	Negative	Positive	Positive	Usually positive
IGRA	Negative	Positive	Positive	Usually positive
Culture	Negative	Negative	Negative	Positive

* Resister (RSTR): successfully clear *Mtb* infection through innate immune responses or acquired immune responses without T cell priming or memory

Study design

TST: tuberculin skin test
IGRA: interferon gamma release assay

Uganda household contact study



Tom Hawn
UW



Cathy Stein
Case Western

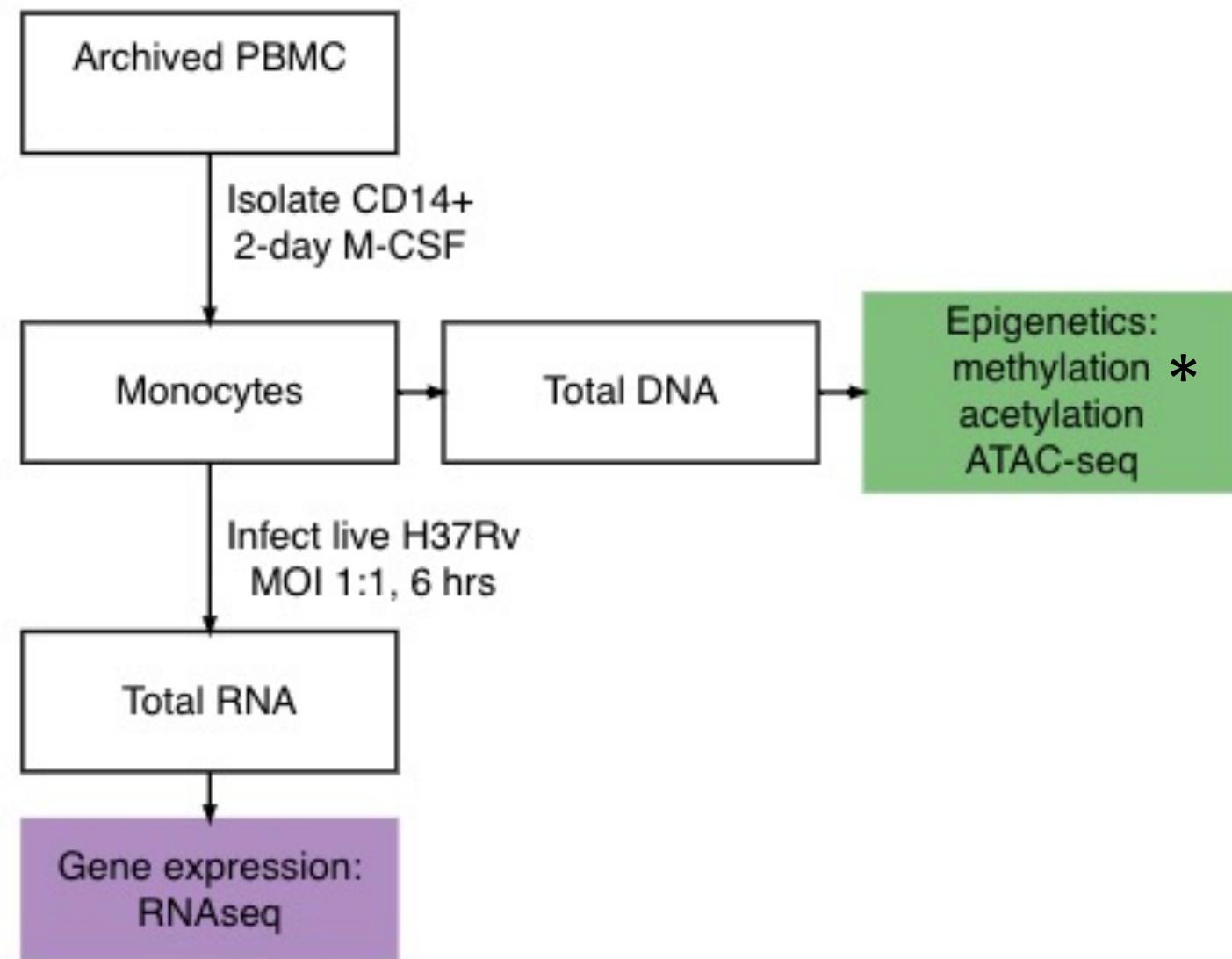


Henry Boom
Mayanja-Kizza
Makerere U



PBMC: peripheral blood
mononuclear cells

Design

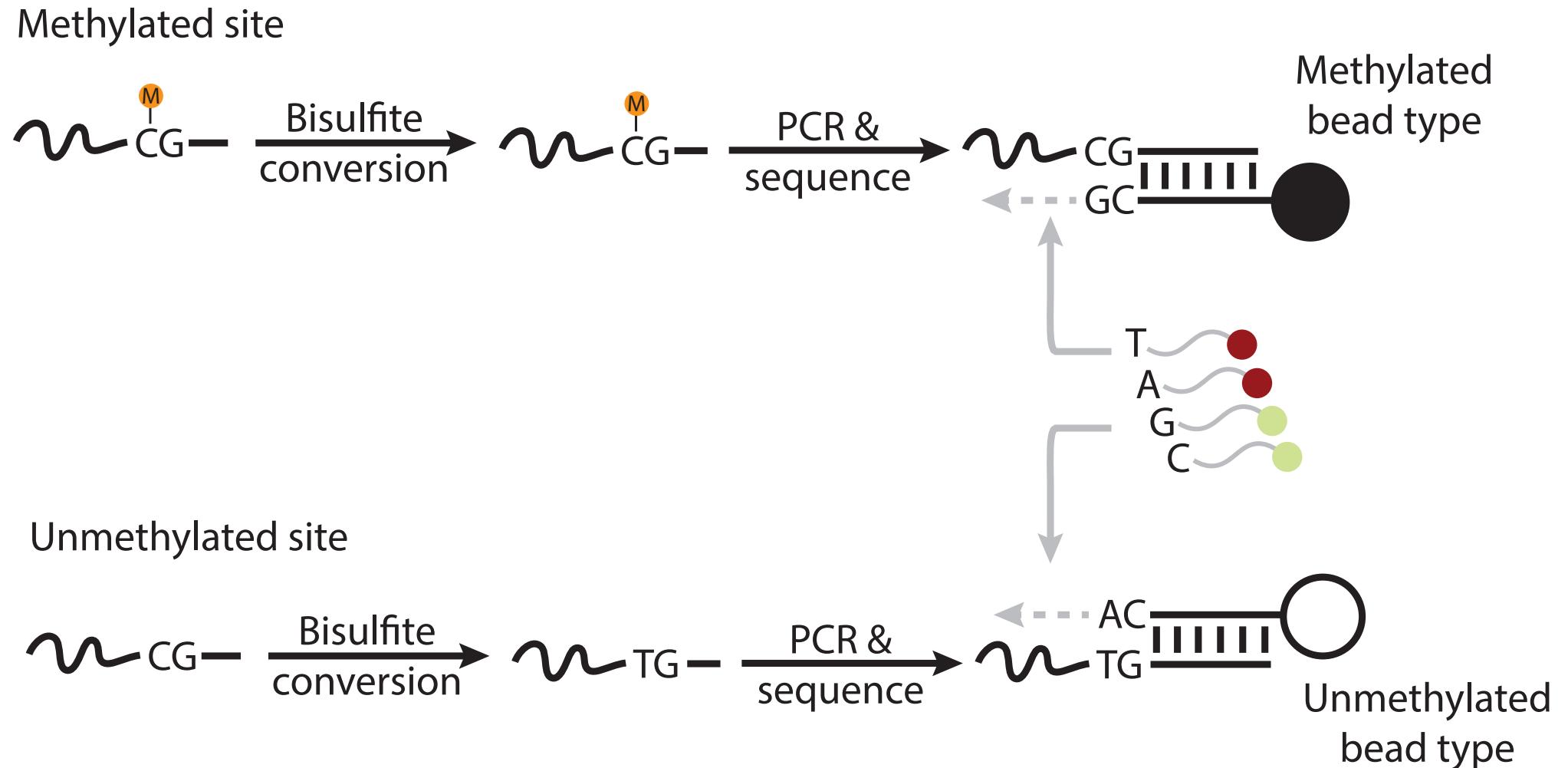


RSTR: resister
LTBI: latent TB infection

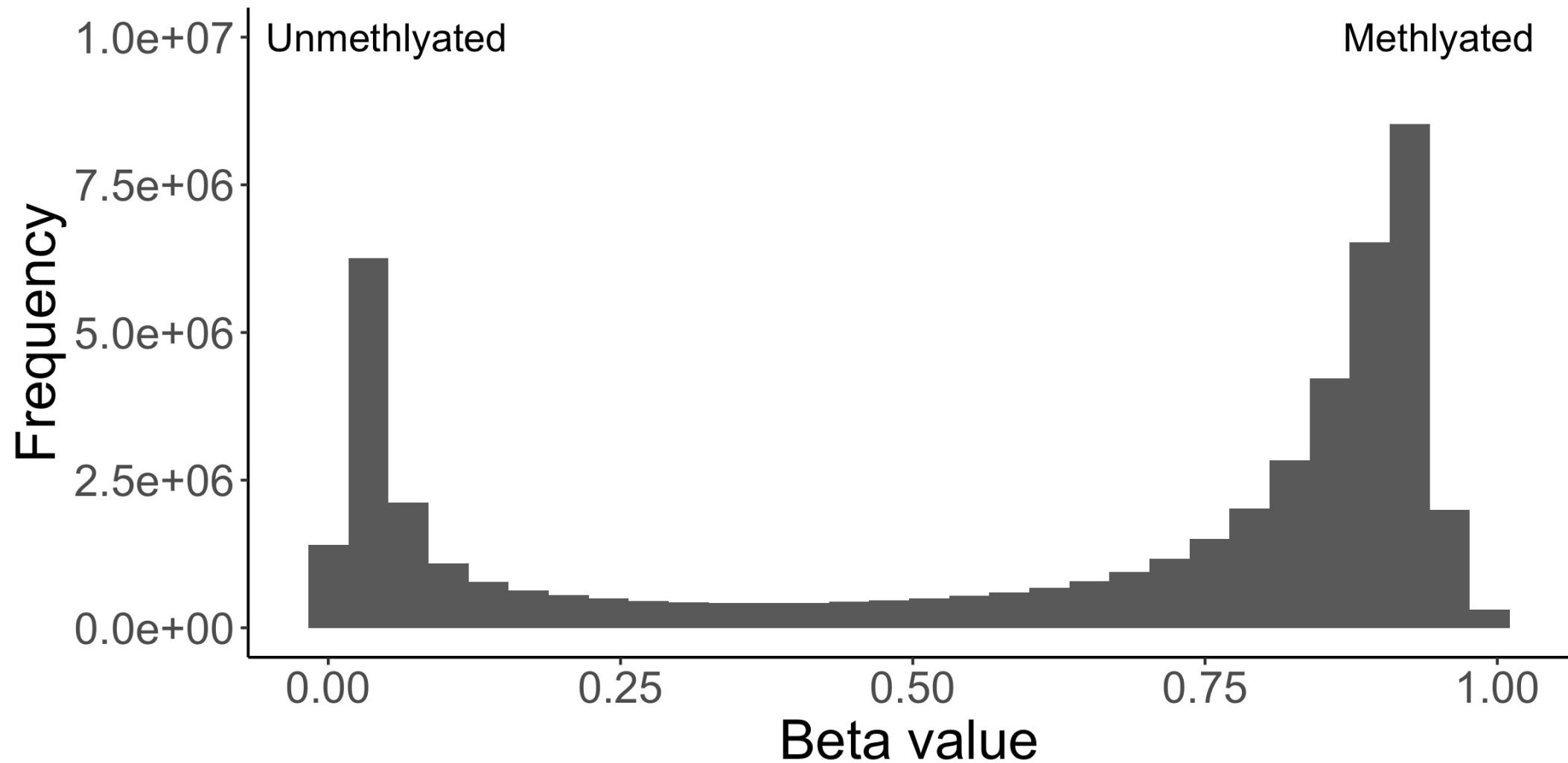
Demographics

	Methylation	RSTR vs LTBI	p-value
Donors with kinship	RSTR	29	
	LTBI	29	
Female, %	50.0		1
Age at enrollment, yrs	17.2 ± 12.6		0.439
Age at sampling, yrs	25.8 ± 12.3		0.318

Infinium MethylationEPIC array: Bisulfite sequencing



Methylation data



Data cleaning

Methylation data cleaning (R::ChAMP)

1. Quality filtering

- Probes ≤ 3 beads in ≥ 95% of samples
- Probes failed ≥ 10% of samples
- Non-CpG, near SNP^a, multi-alignment^b, XY

2. Normalization

- Functional normalization^c
 - Quantile normalization based on control probes to estimate technical variation
- ComBat batch correction
- Log2 transformation

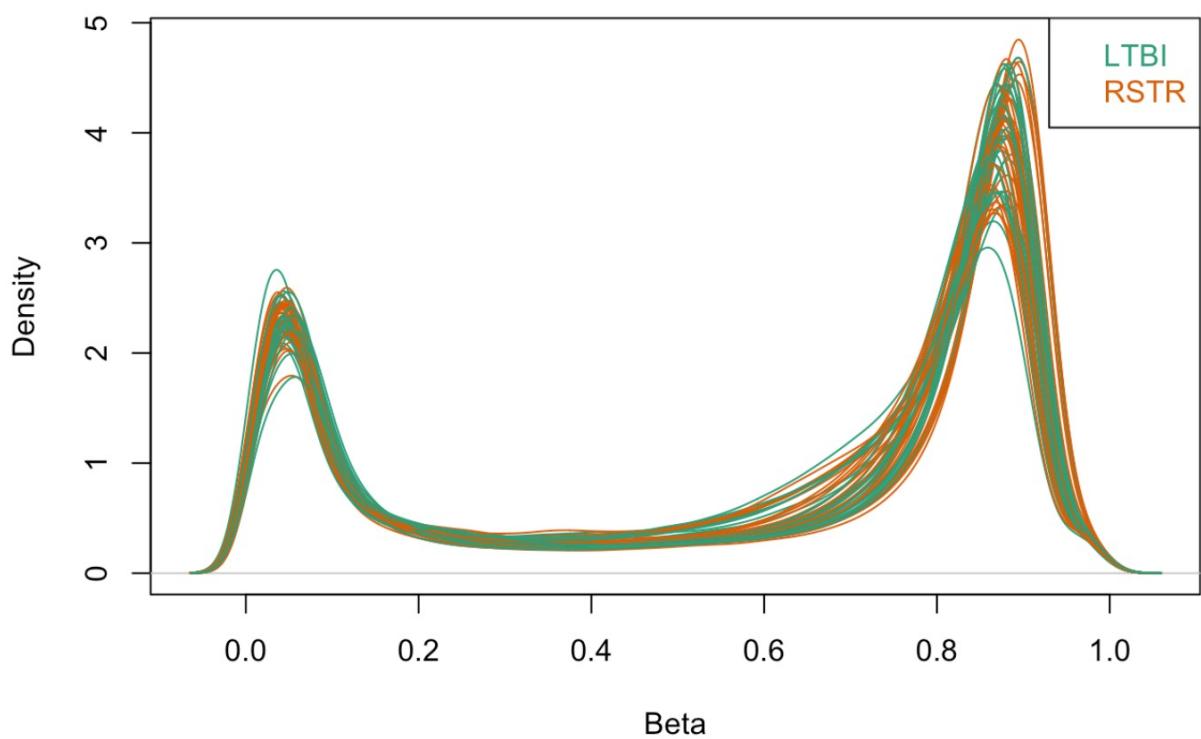
^aZhou *et al* 2017 doi: [10.1093/nar/gkw967](https://doi.org/10.1093/nar/gkw967)

^bNordlund *et al* 2013 doi: [10.1186/gb-2013-14-9-r105](https://doi.org/10.1186/gb-2013-14-9-r105)

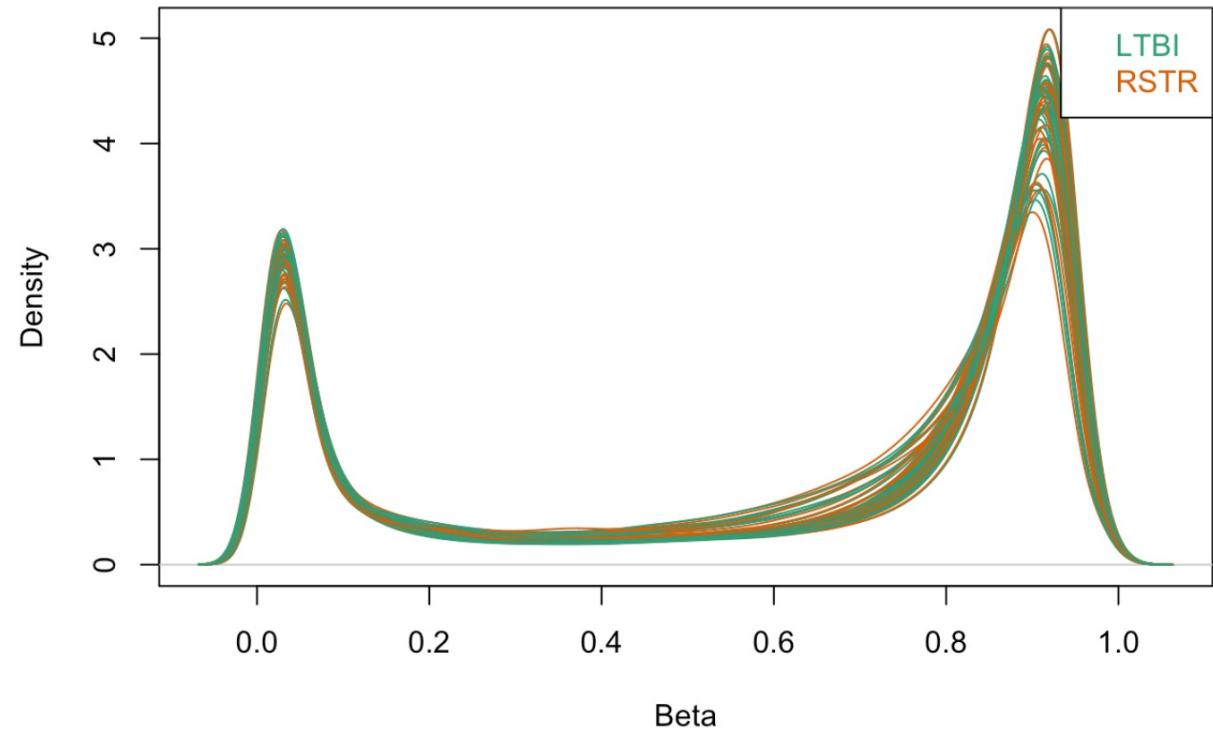
^cFortin *et al* 2014 doi: [10.1186/s13059-014-0503-2](https://doi.org/10.1186/s13059-014-0503-2)

Functional normalization

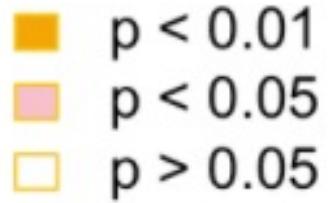
Before



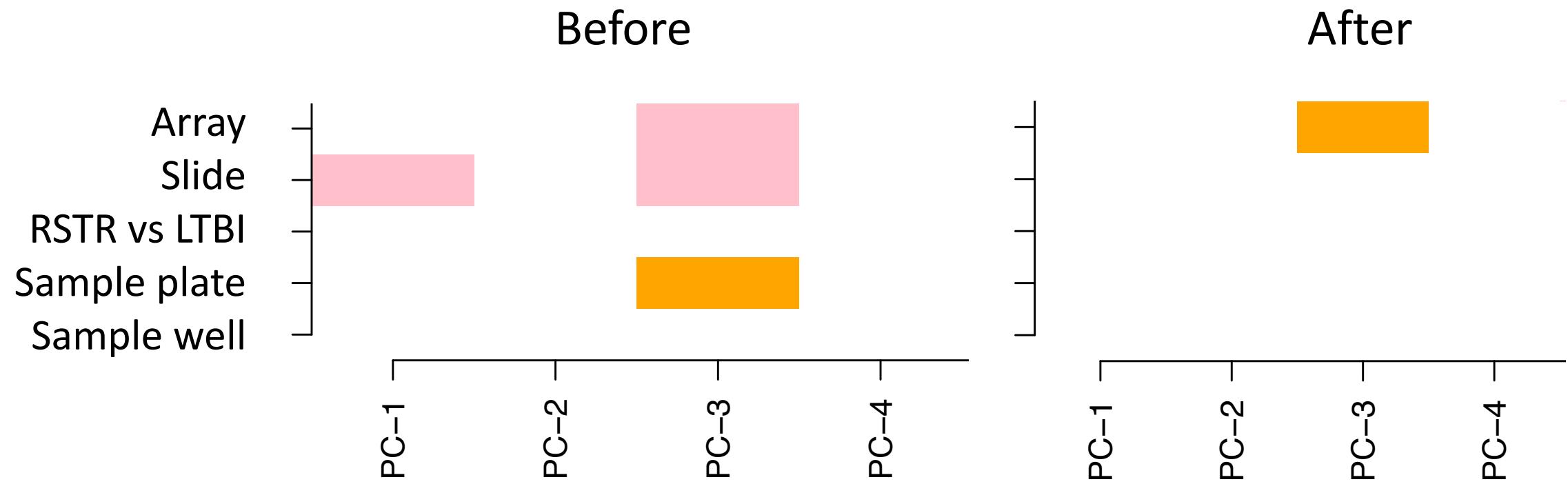
After



ComBat batch correction



- Singular value decomposition (SVD)
- PCs up to 80% total variation explained



Methylation in RSTR vs LTBI

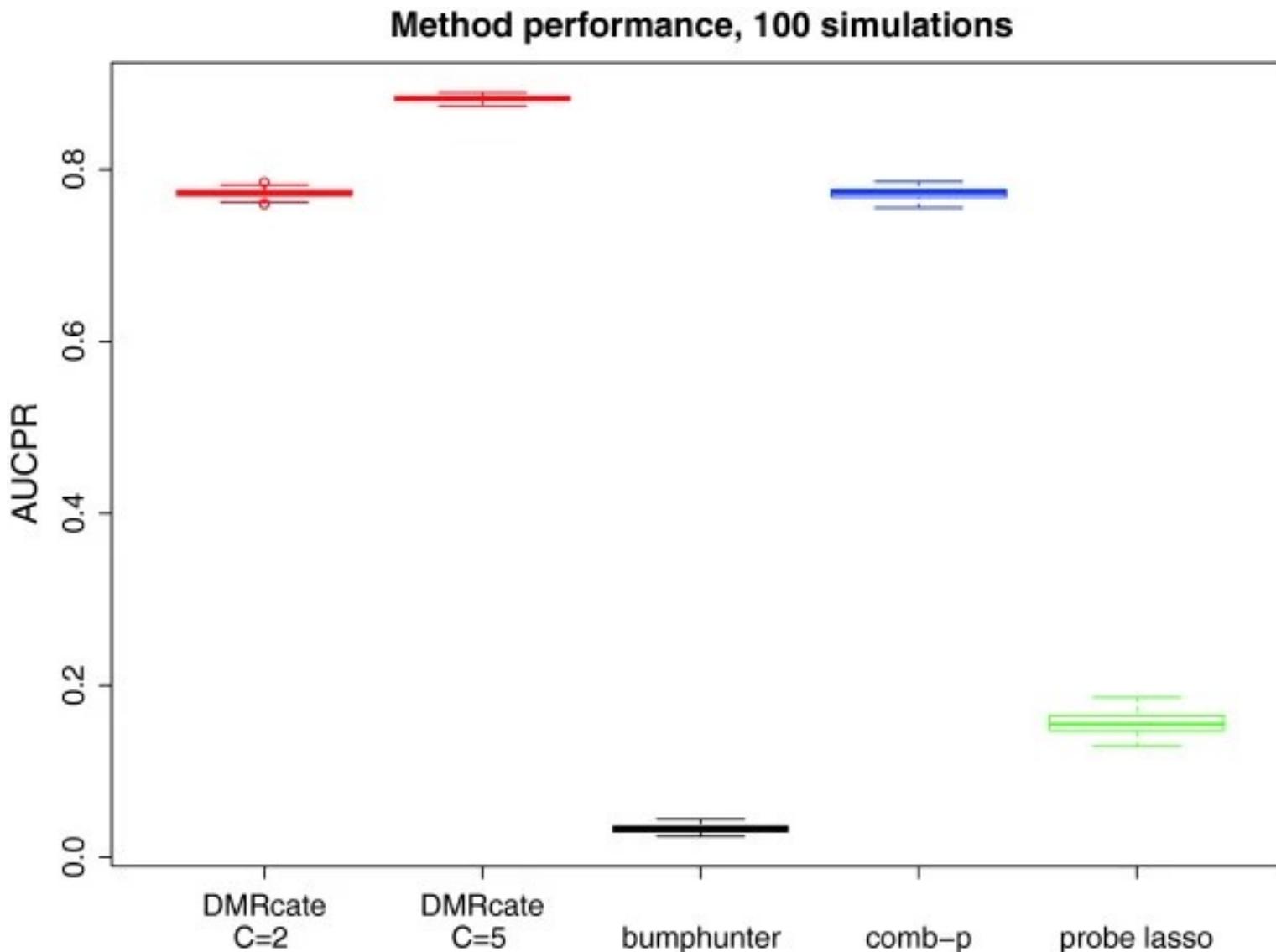
De novo differentially methylated regions (R :: DMRcate)

1. Linear model probe-level data and define significant probes
2. Apply Gaussian smoothing to probe test statistics using a given bandwidth (λ) and scaling factor (C) where $\lambda/C = \sigma$
3. Model smoothed test statistics using Satterthwaite approximation of sample variance
4. Estimate p-values and apply false discovery rate (FDR) correction
5. Re-define significant probes and collapse nearby probes using λ

Peters *et al* 2015 doi: [10.1186/1756-8935-8-6](https://doi.org/10.1186/1756-8935-8-6)

DMRcate

Area under the
precision-recall curve



Probe model fitting (R::lmeKin)

Main variable

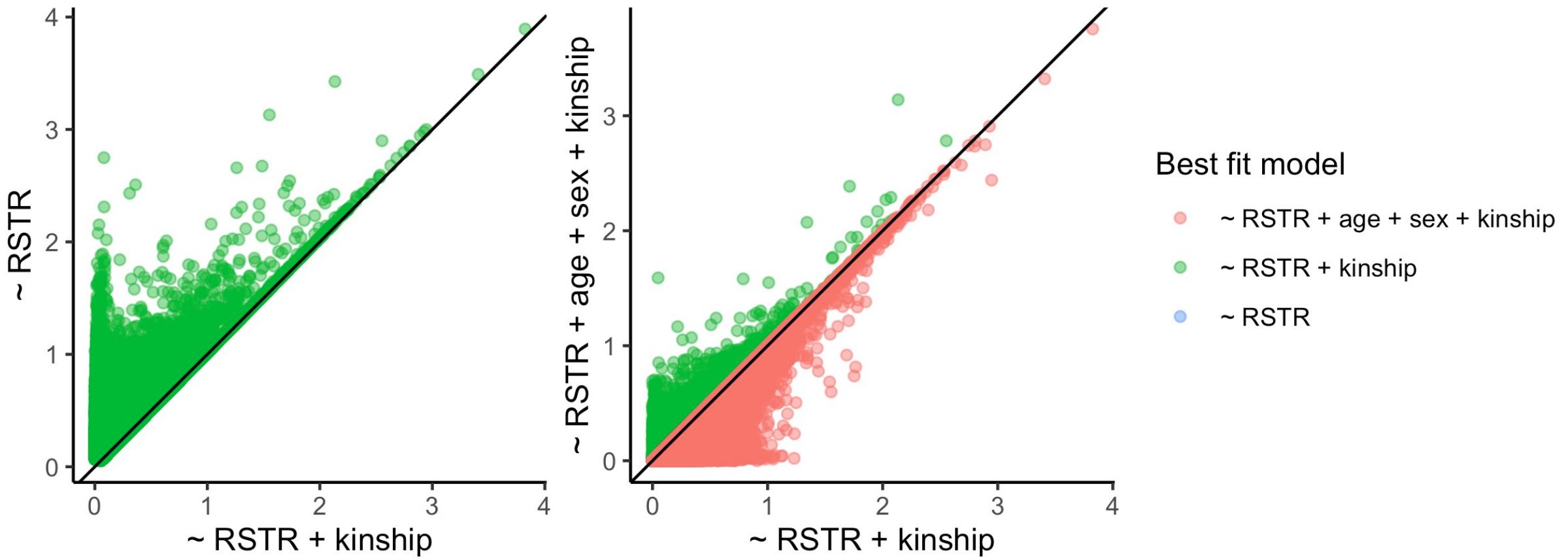
- RSTR vs LTBI

Potential co-variates

- Genetic relatedness (family blocks, kinship)
- Age
- Biological sex
- Body mass index (BMI)
- Composite TB risk score
- BCG vaccination

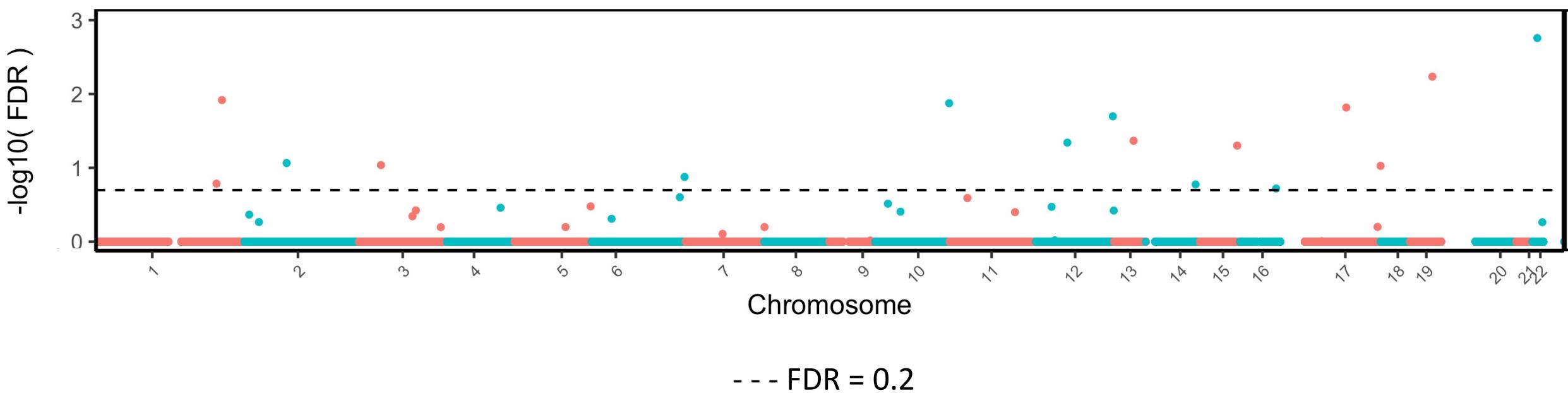
Probe model fitting

Mean residual standard deviation (sigma)



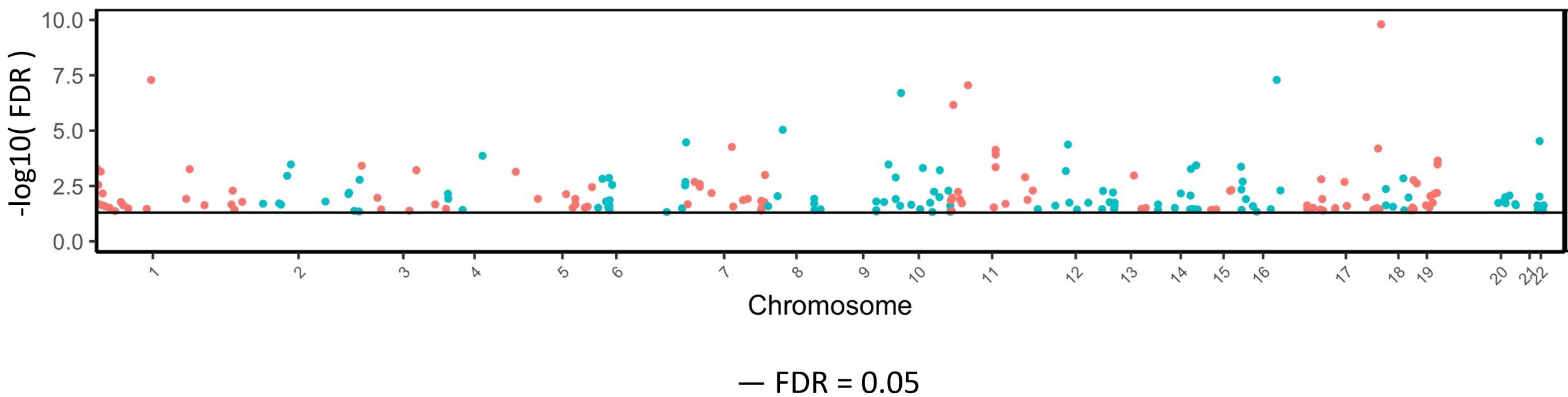
Differentially methylated probes (DMP)

~ RSTR + age + sex + kinship

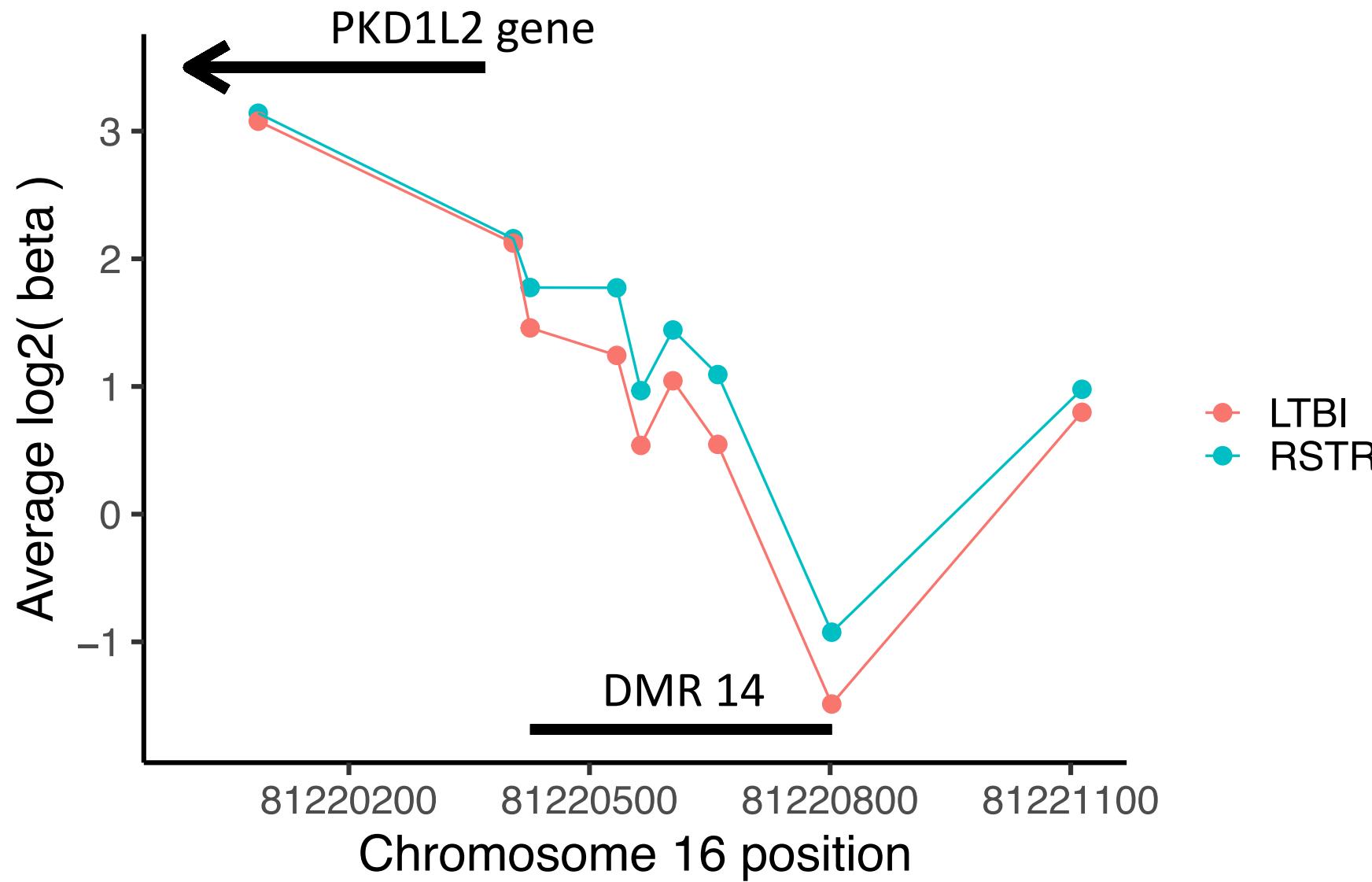


Differentially methylated regions (DMR)

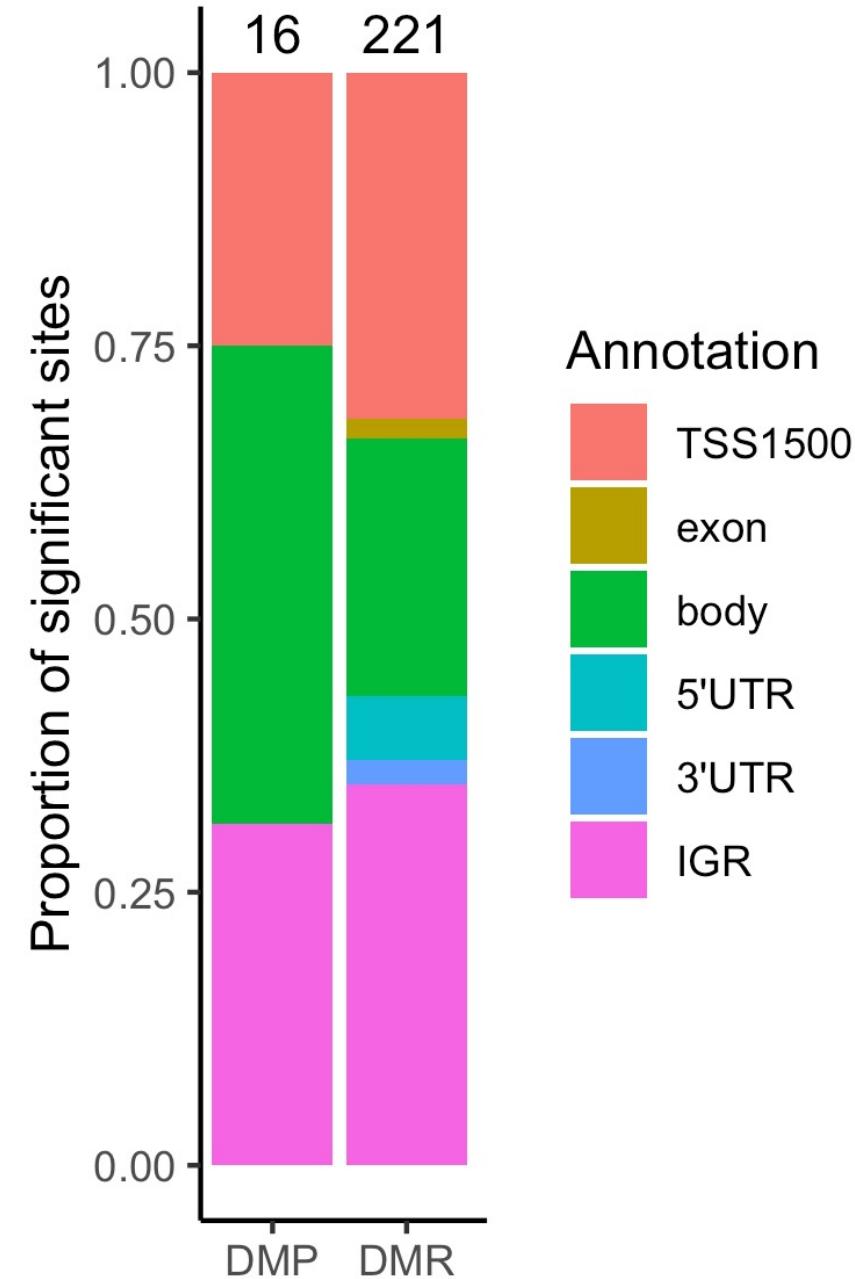
~ RSTR + age + sex + kinship



Example



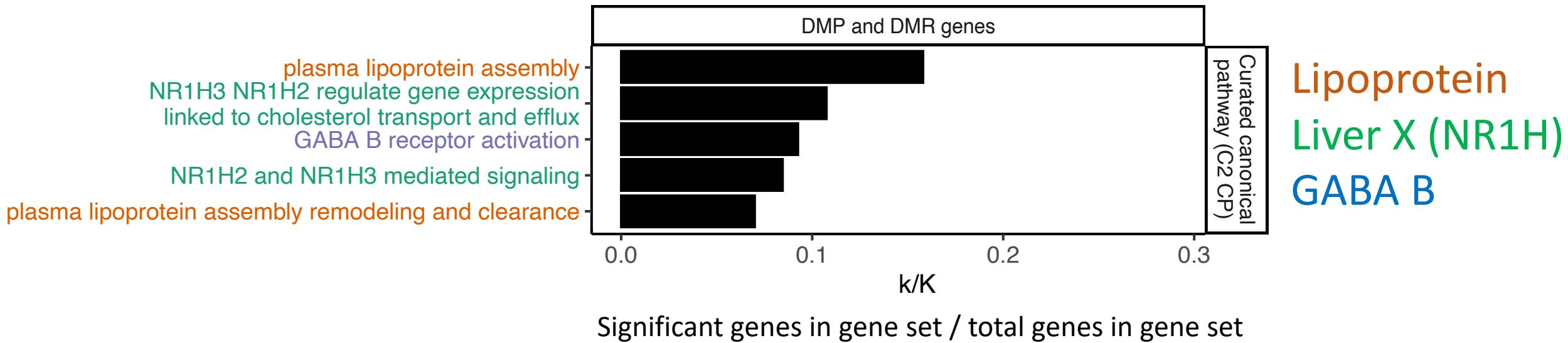
Annotation types



Genes with differential methylation (R::clusterProfiler)

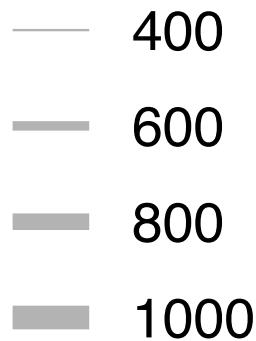
Enrichment of Broad gene sets (MSigDB)

FDR < 0.1

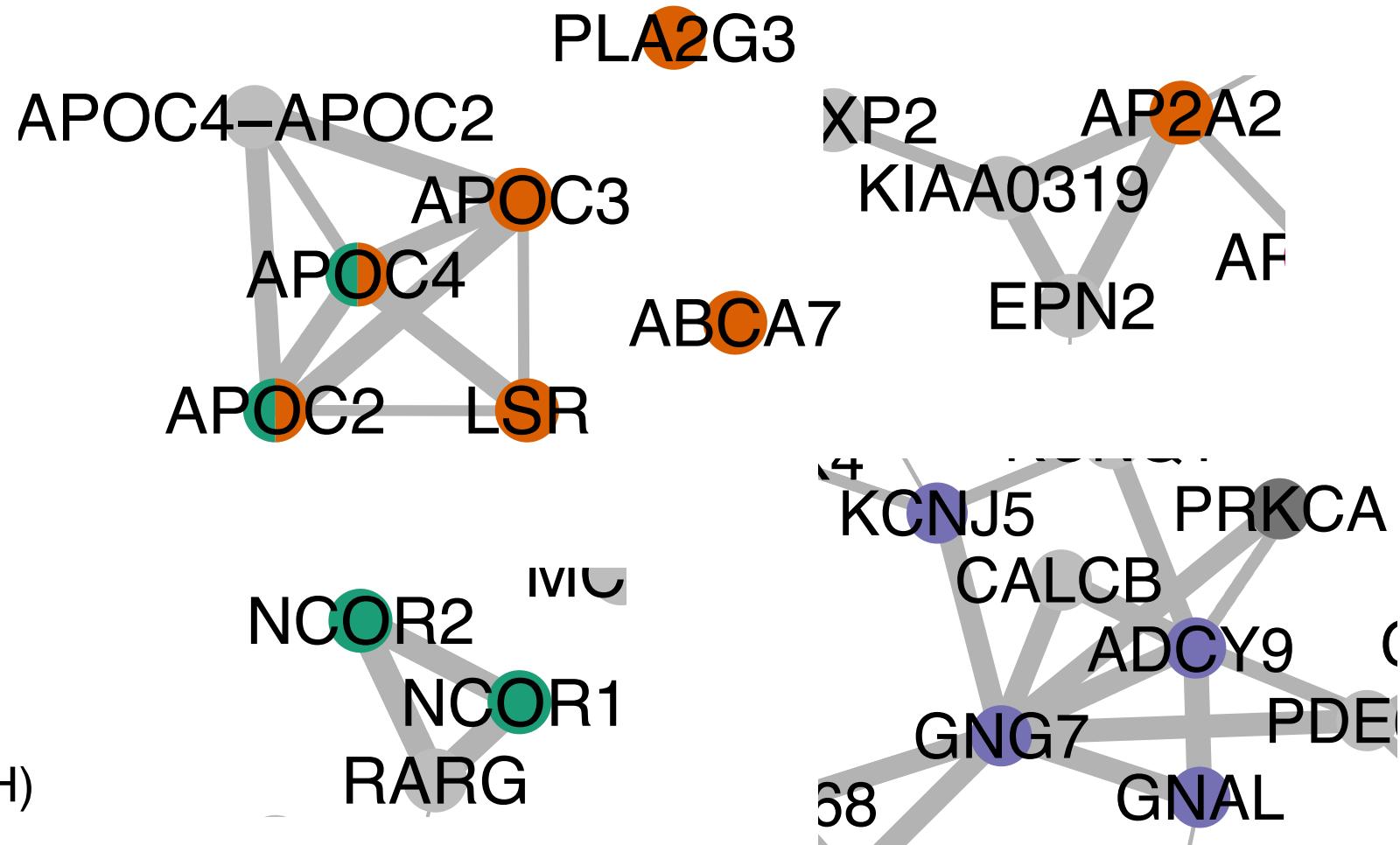
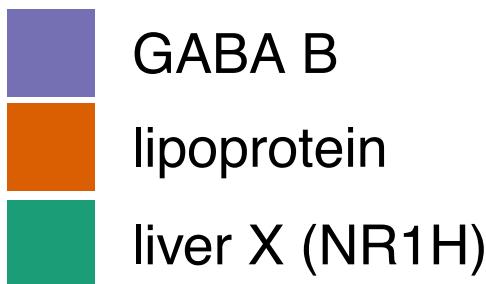


Genes with differential methylation

STRING score



Enrichment



Implications for TB

- *Mtb* utilizes cholesterol and fatty acids as major carbon sources
- *Mtb* promotes lipid accumulation in infected cells
- Lipid-laden (foamy) macrophages associated with TB granulomas
 - Cholestryl esters and triglycerides

Hypothesis

- RSTR are epigenetically programmed to better control excess cholesterol and lipids
- This prevents foam cell formation, thus “starving” *Mtb* and facilitating successful clearance of infection

Next steps

- Serum proteomics and lipidomics of high-density lipoproteins (HDL)
- Serum cholesterol efflux capacity



Tomas Vaisar
UW

- Expression quantitative trait methylation (eQTM)
 - Associate methylation with gene expression in response to *Mtb* infection *ex vivo*



BILL &
MELINDA
GATES
foundation



Thomas Hawn, MD

- **Felicia Nguyen**
- **Glenna Peterson**
- **Jason Simmons, MD**
- Max Segnitz, PhD
- Monica Campo, MD

RSTR project

- Cathy Stein, PhD, Case Western
- Henry Boom, MD, Case Western
- Harriet Mayanja-Kizza, MD, Makerere University
- Chetan Seshadri, MD, UW

