



PEREpigenomics Profile Explorer of Roadmap Epigenomics data

Guillaume Devailly

► To cite this version:

Guillaume Devailly. PEREpigenomics Profile Explorer of Roadmap Epigenomics data. UseR! 2019, Jul 2019, Toulouse, France. hal-02289519

HAL Id: hal-02289519

<https://hal.science/hal-02289519>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

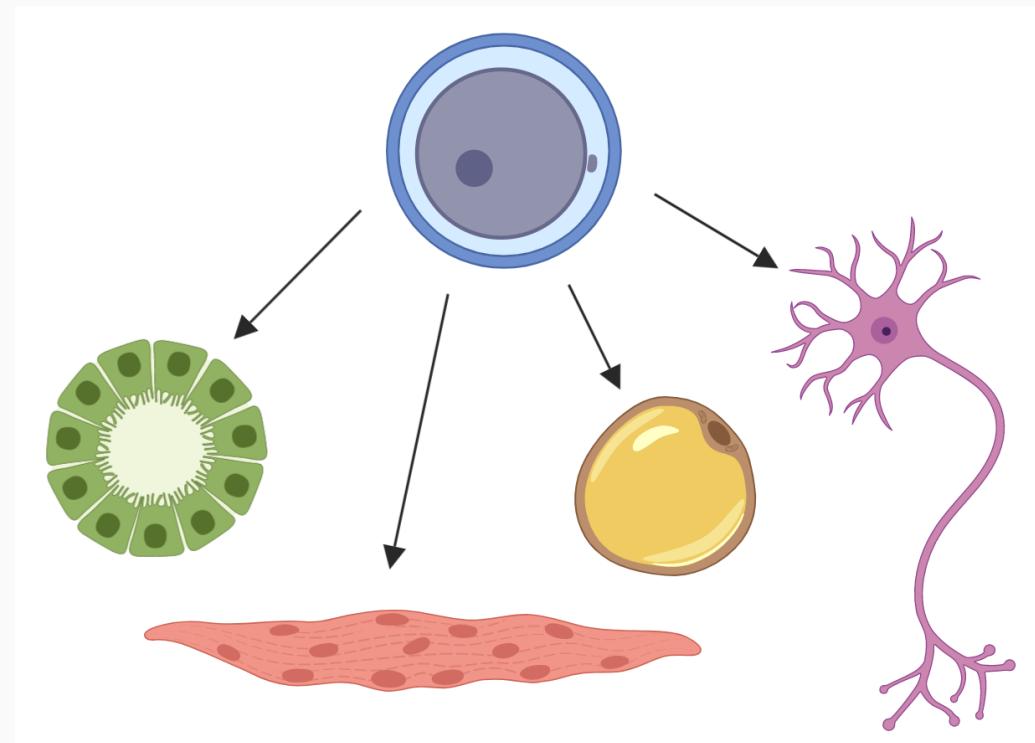
PEREpigenomics

Profile Explorer of Roadmap Epigenomics data

 @G_Devailly
Guillaume Devailly
useR!2019
2019/07/09



One genome, many cell types



Differential gene expression =
cellular environment + epigenetics

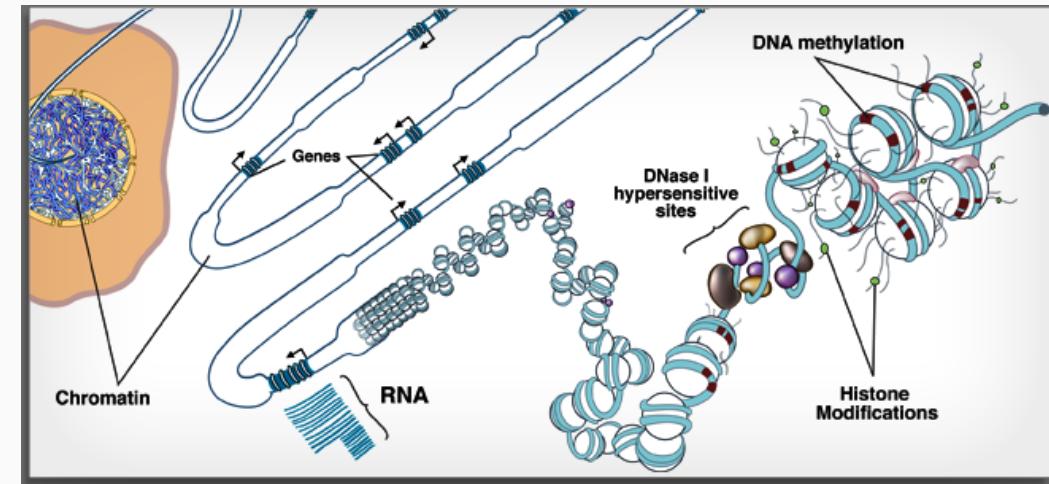
Epigenetics data

Many existing datasets and databases.



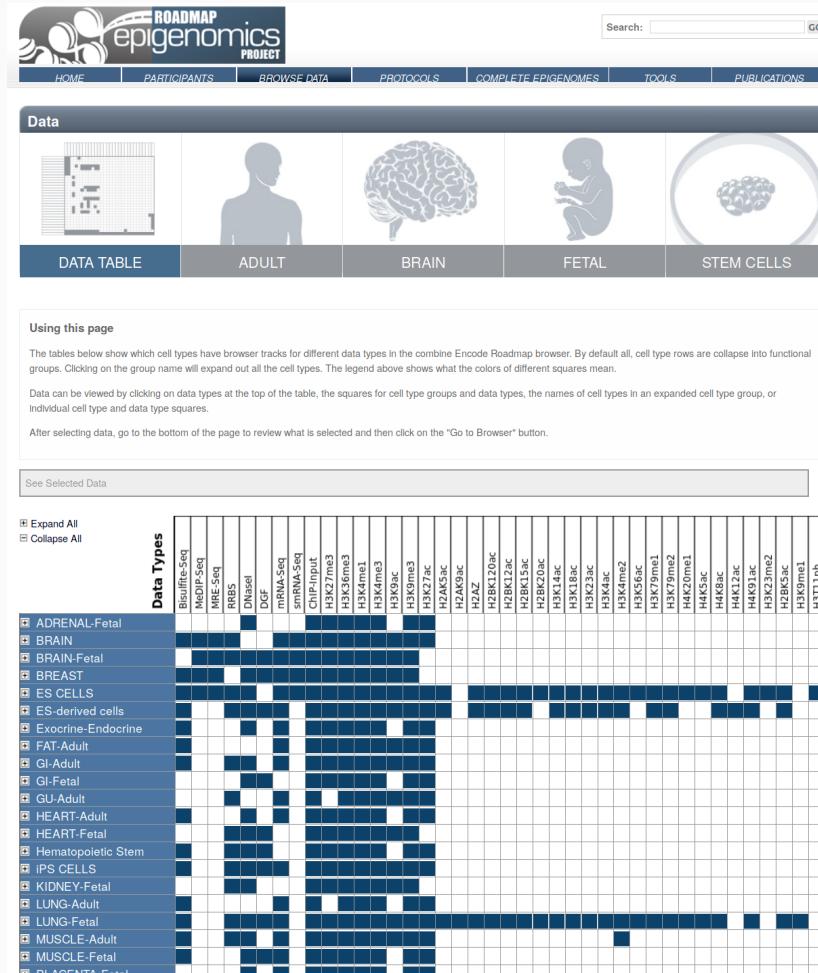
- RNA-seq
- DNase1
- WGBS,
- 10 different histone methylations
- 17 different histone acetylations

in 33 human cell lines & tissues,
uniformly processed.



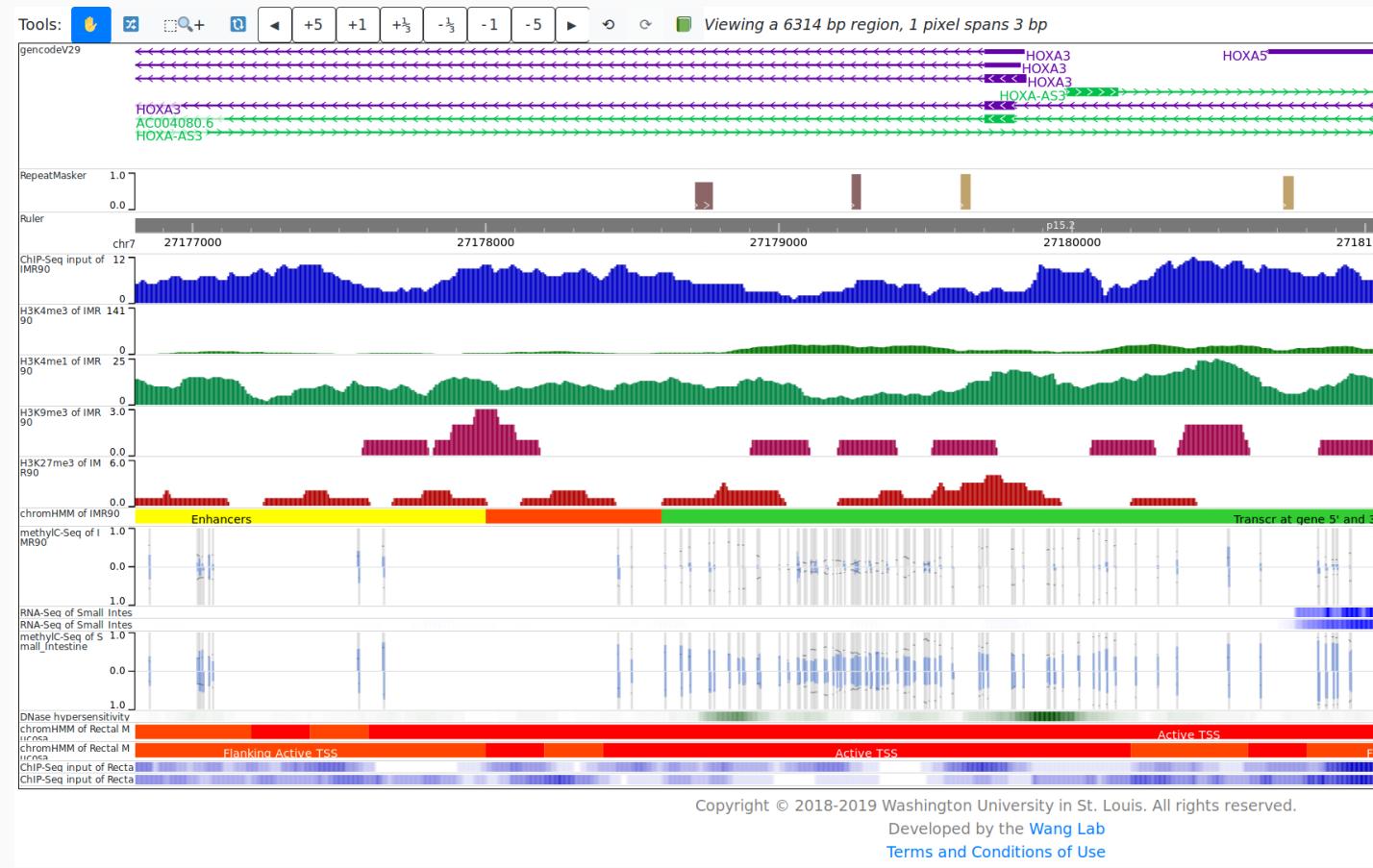
What to do with big dataset?

Build a data portal:



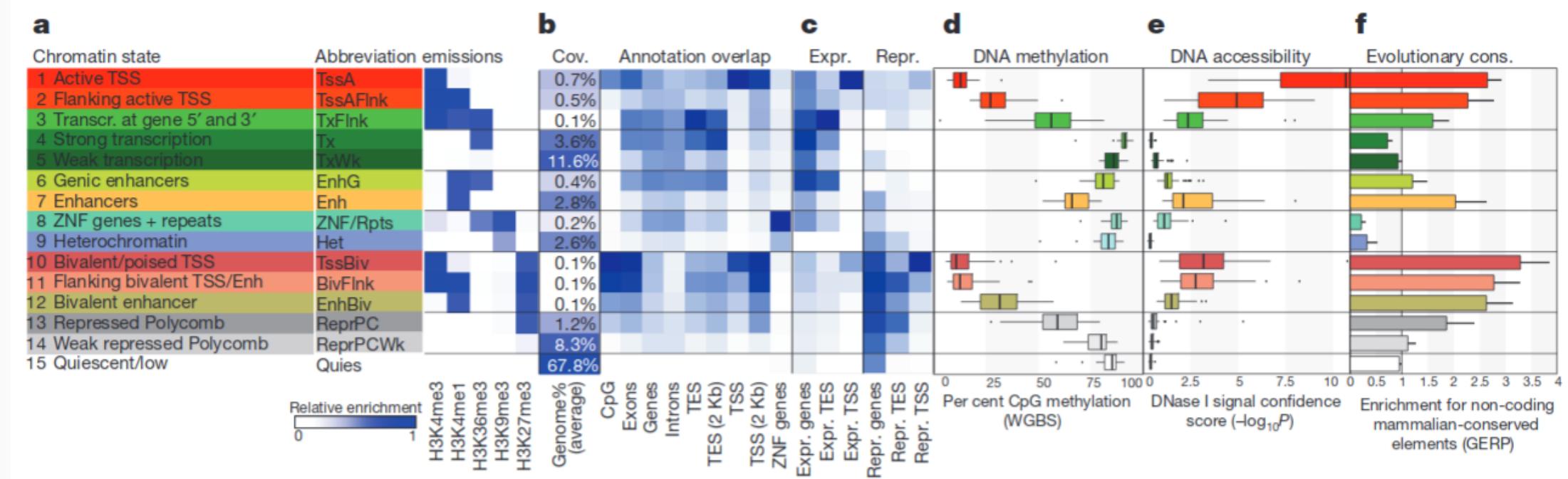
What to do with big dataset?

Link data to genome browser:



What to do with big dataset?

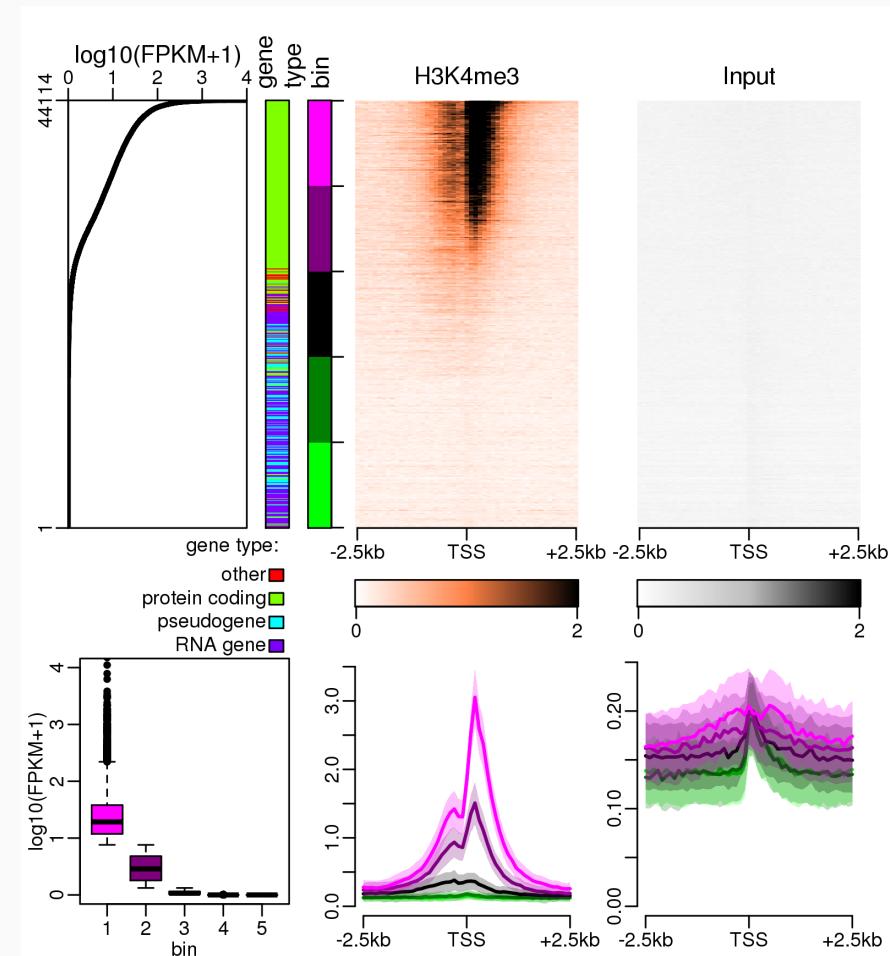
Machine learning:



Kundaje, Meuleman, Ernst, Bilenky et al., 2015, [10.1038/nature14248](https://doi.org/10.1038/nature14248)

Objectives

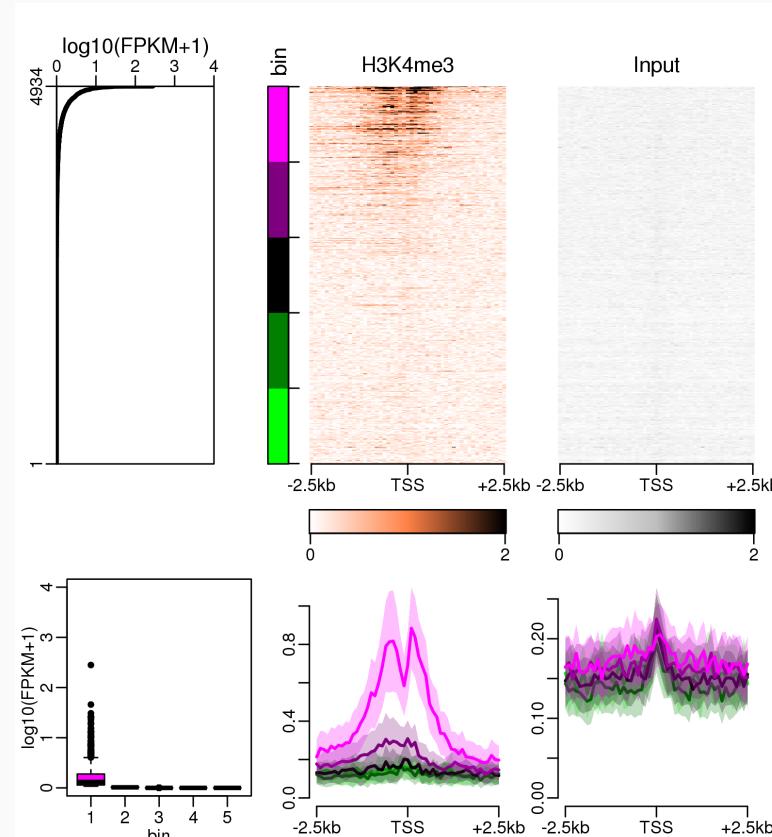
- Visualisation of marks, sorted by **gene** transcription level:
 - at TSS (gene start)
 - at TTS (gene end)
- Visualisation of marks at **middle exons** starts, sorted by:
 - transcription level
 - inclusion ratio
- For **all** genes/exons in **each** cell type.
- For **each** gene/exon in **all** cell types.



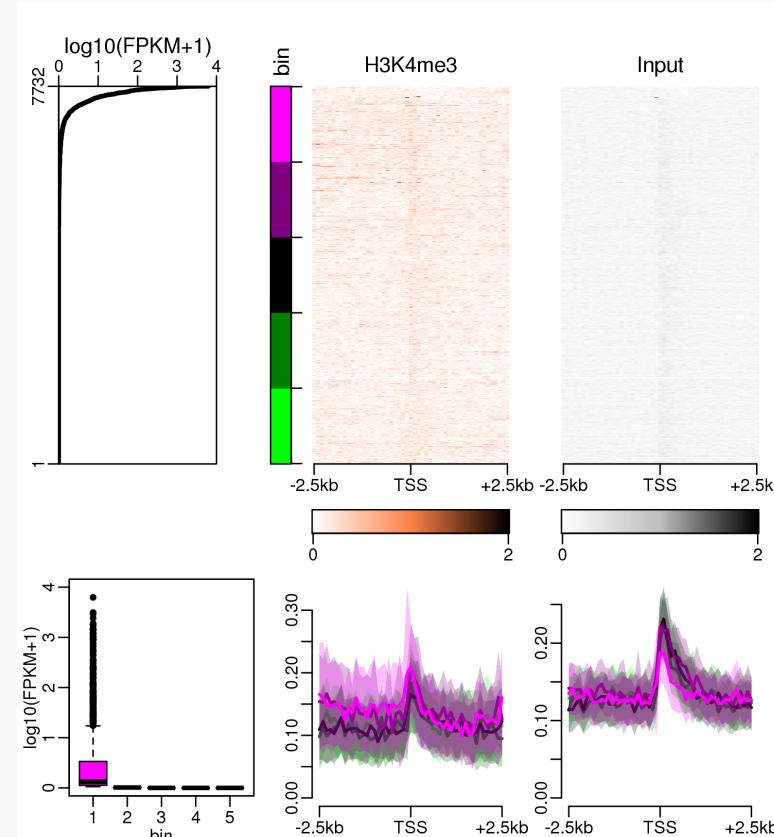
H3K4me3, TSS, small intestine

Different gene types, different associations

lincRNA

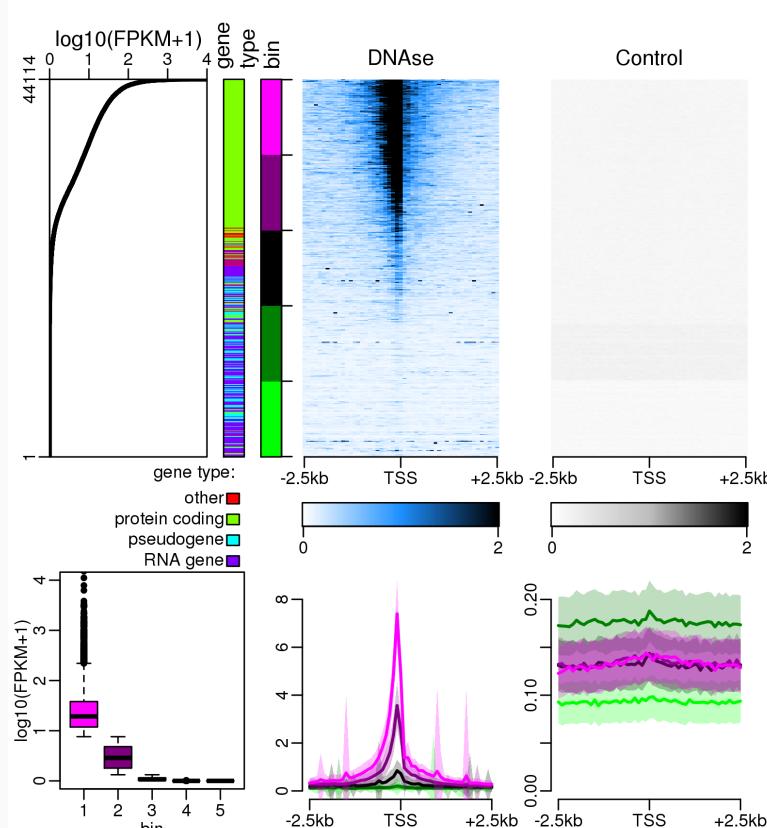


processed pseudogenes

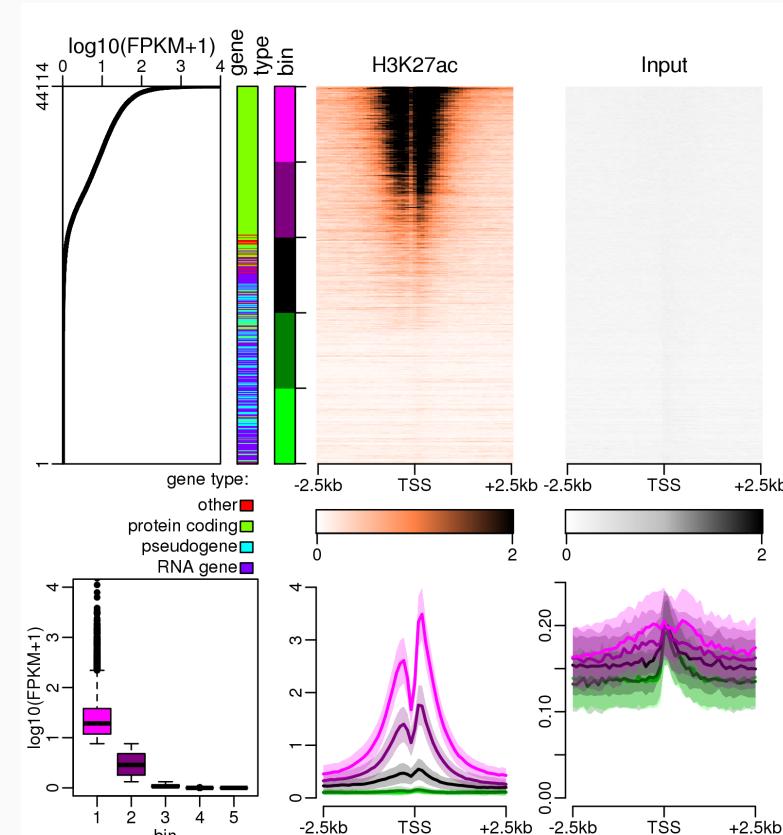


Position of signal matters

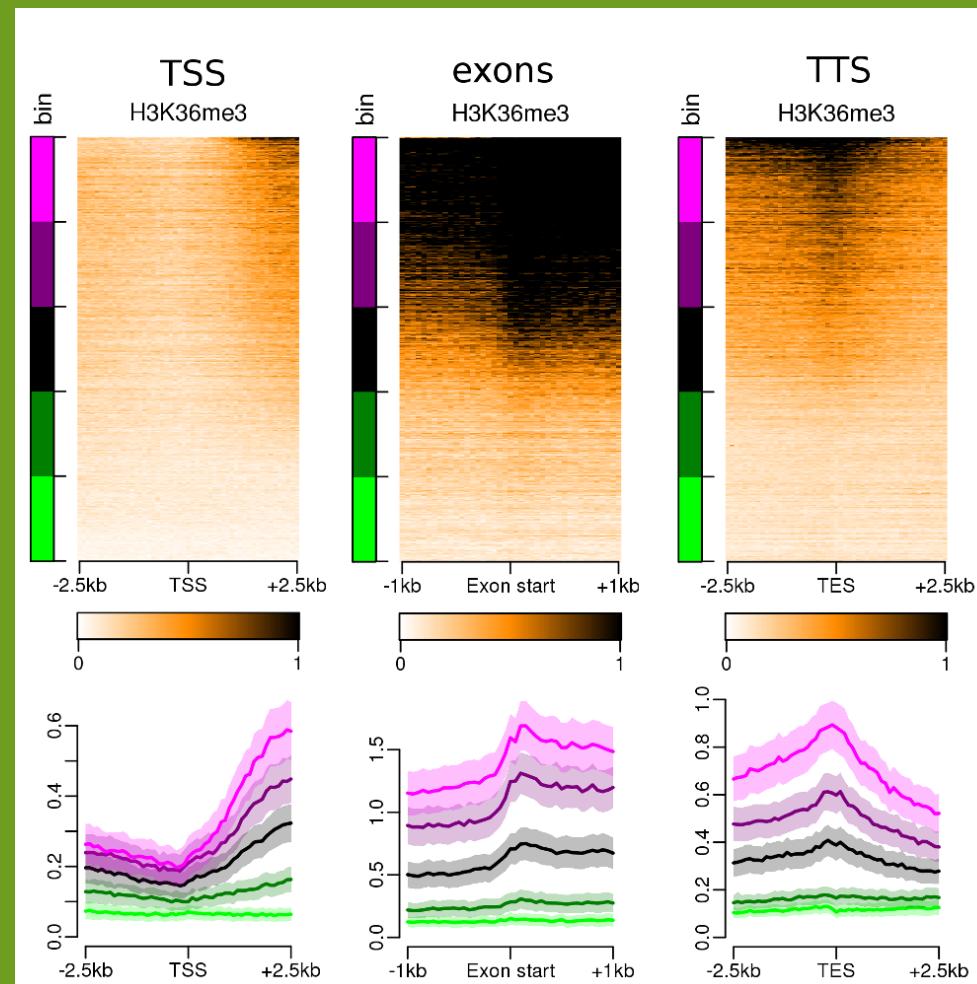
DNase 1 (accessible chromatin)



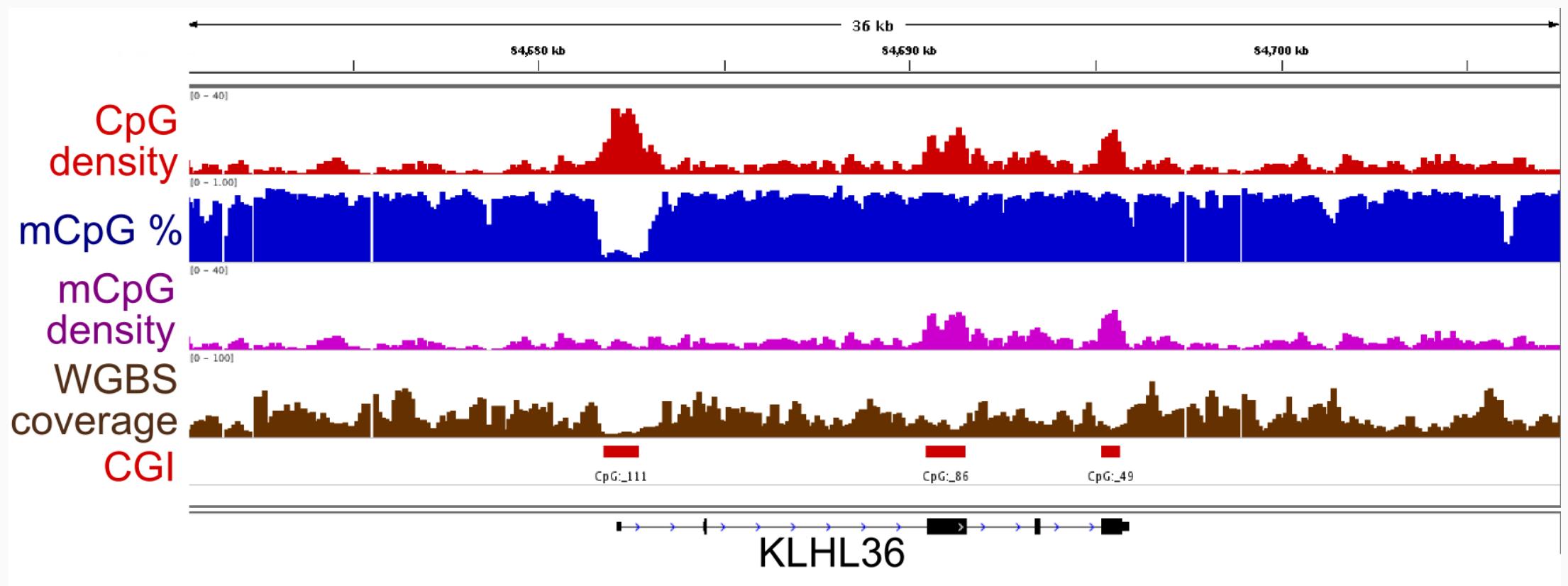
H3K27ac



Exonic mark: H3K36me3, fetal large intestine

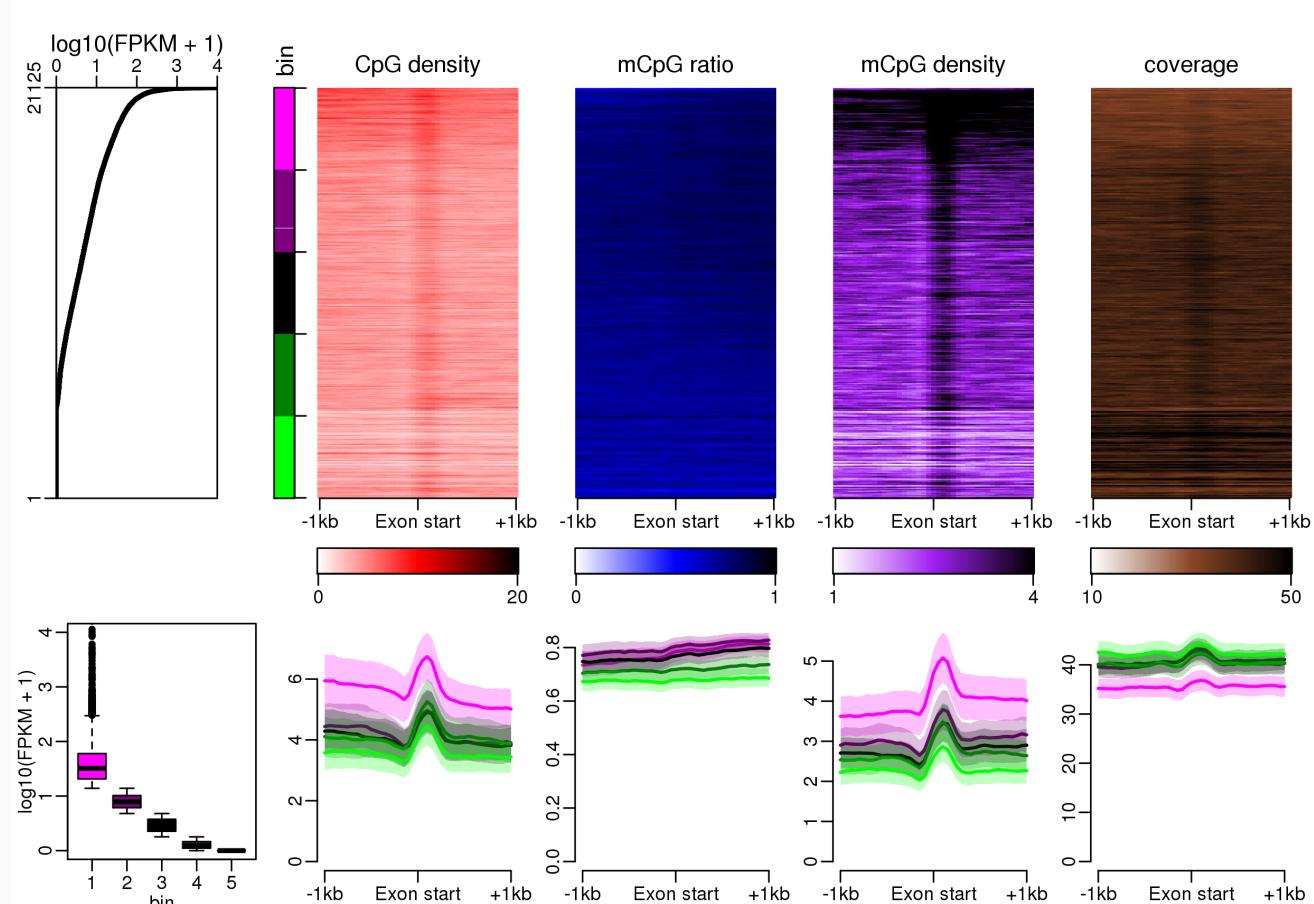


DNA methylation: ratio and density



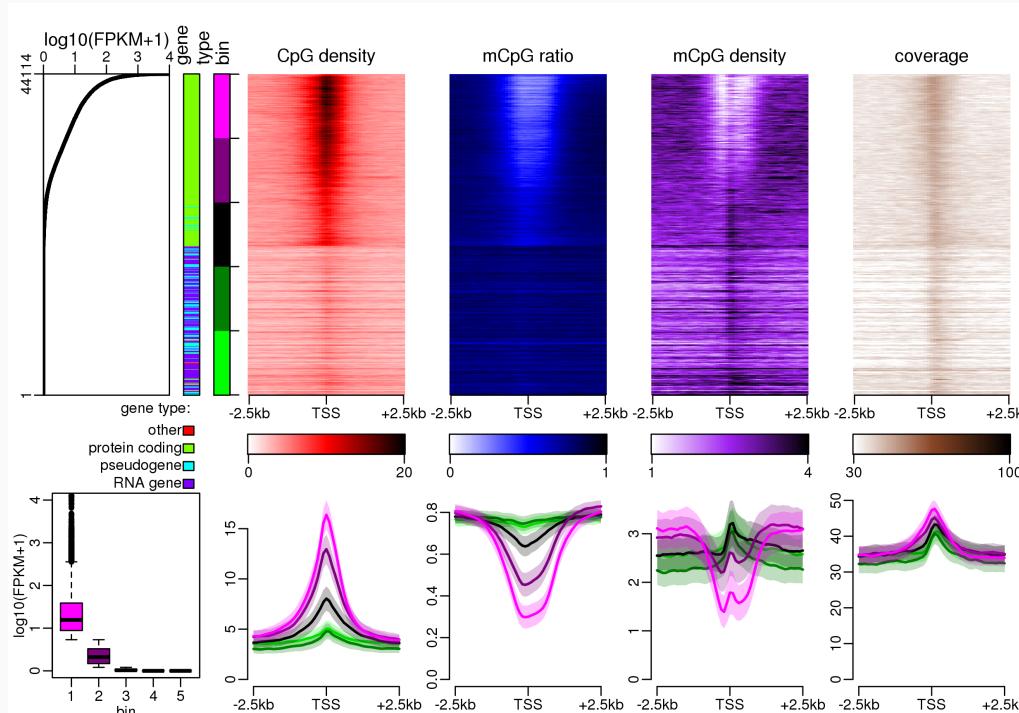
Exonic DNA methylation

WGBS, pancreas

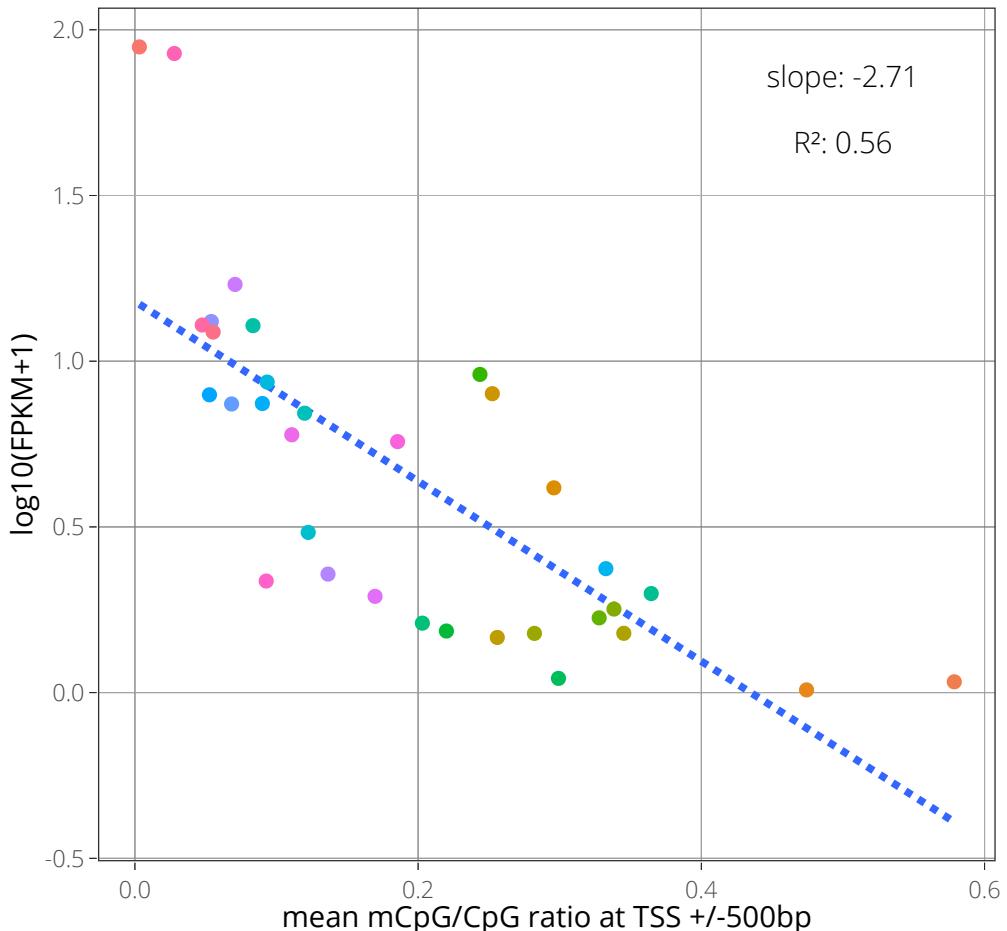


Cell by cell vs gene by gene

WGBS, adult liver

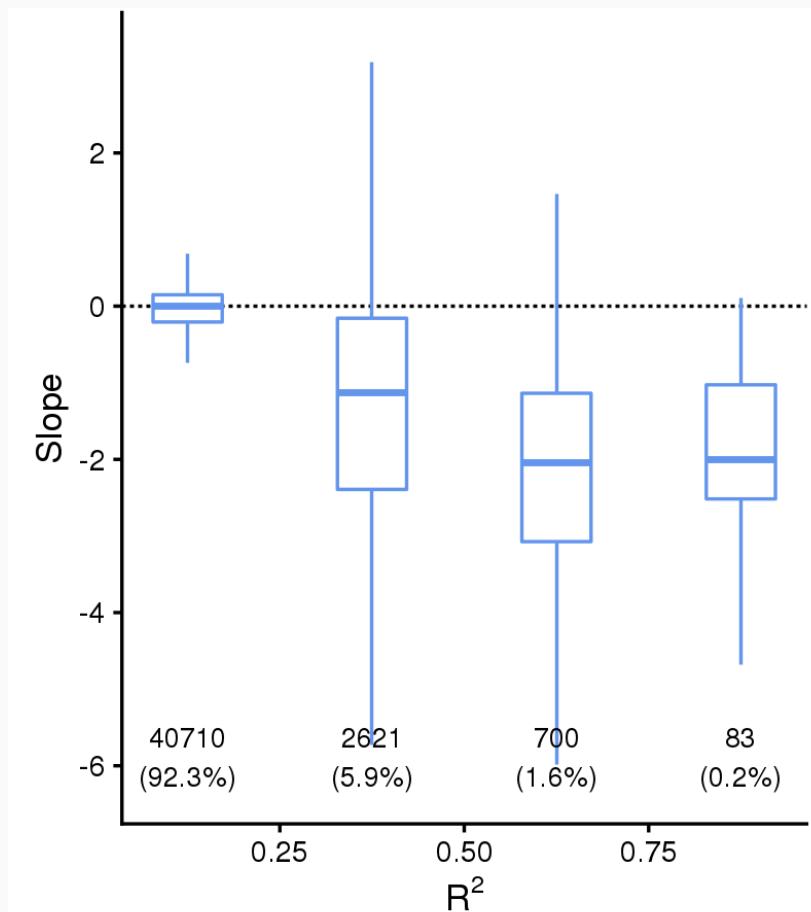


LYL1 (ENSG00000104903)

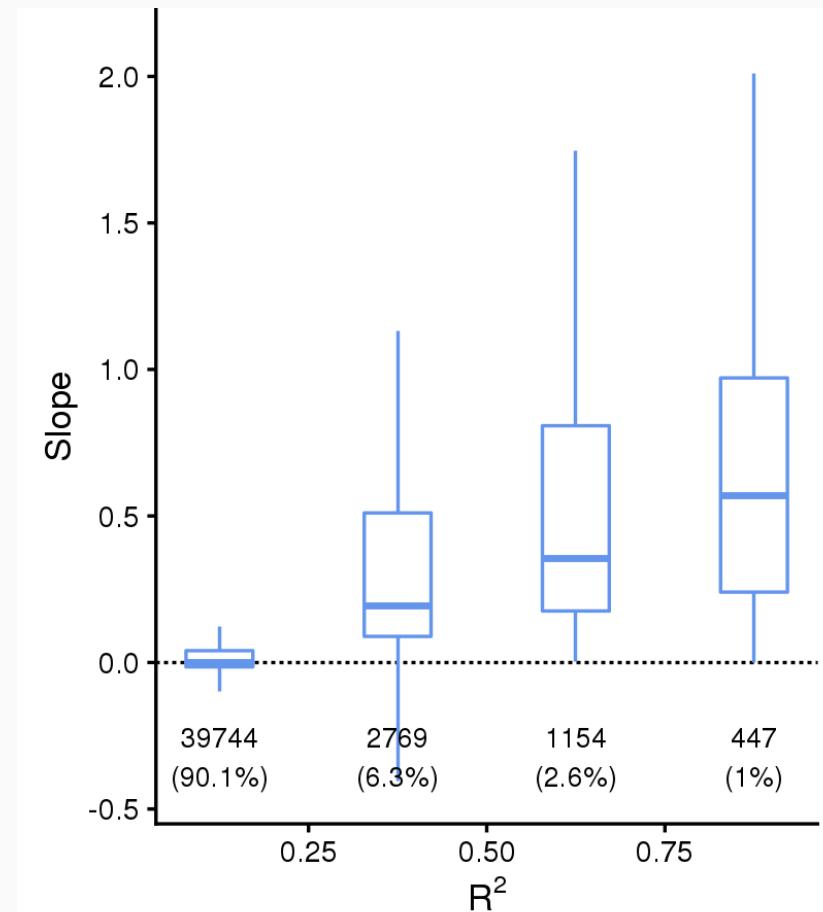


Repressing and activating marks

mCpG ratio



H3K4me3



Too many plots?

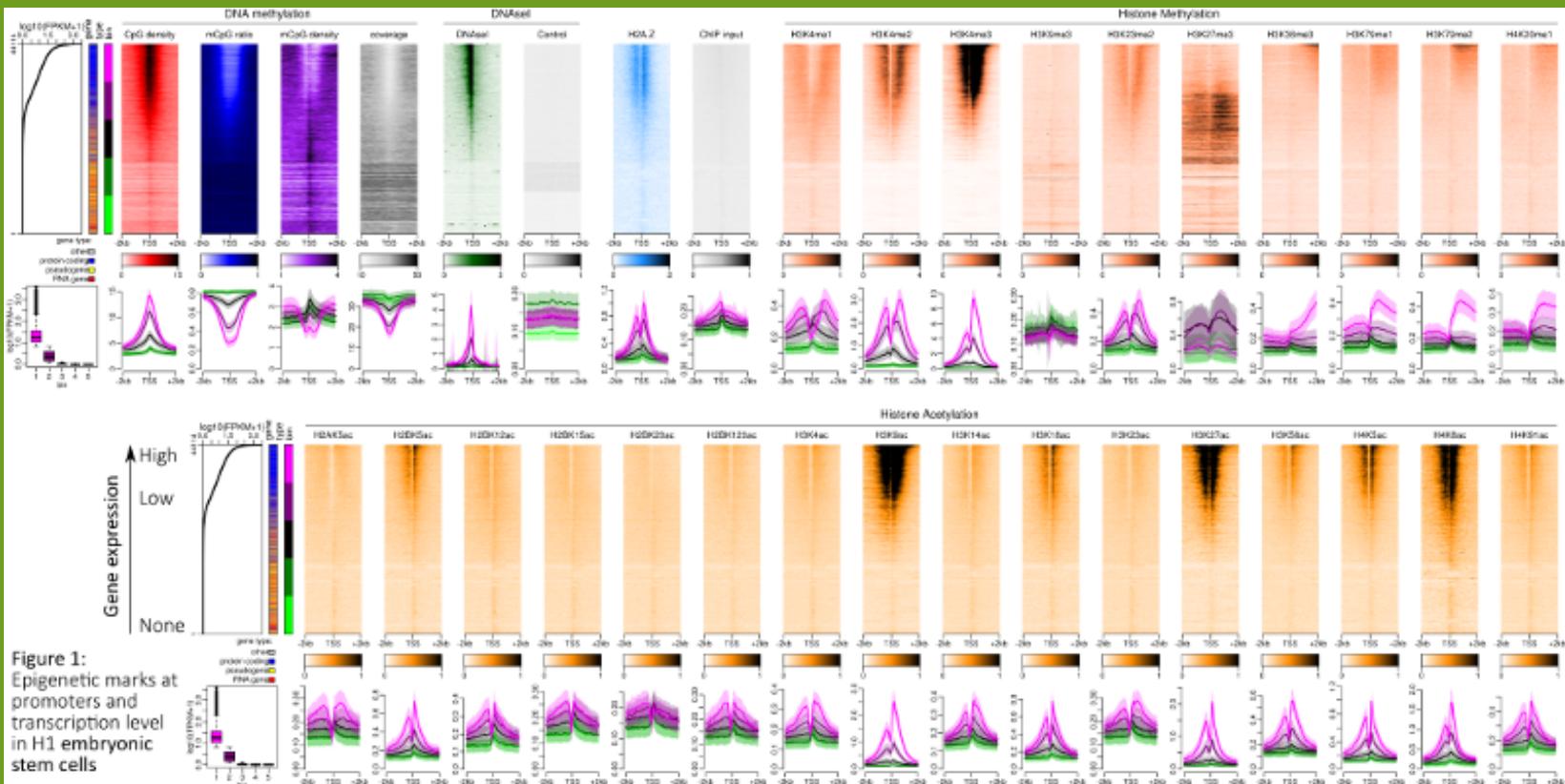
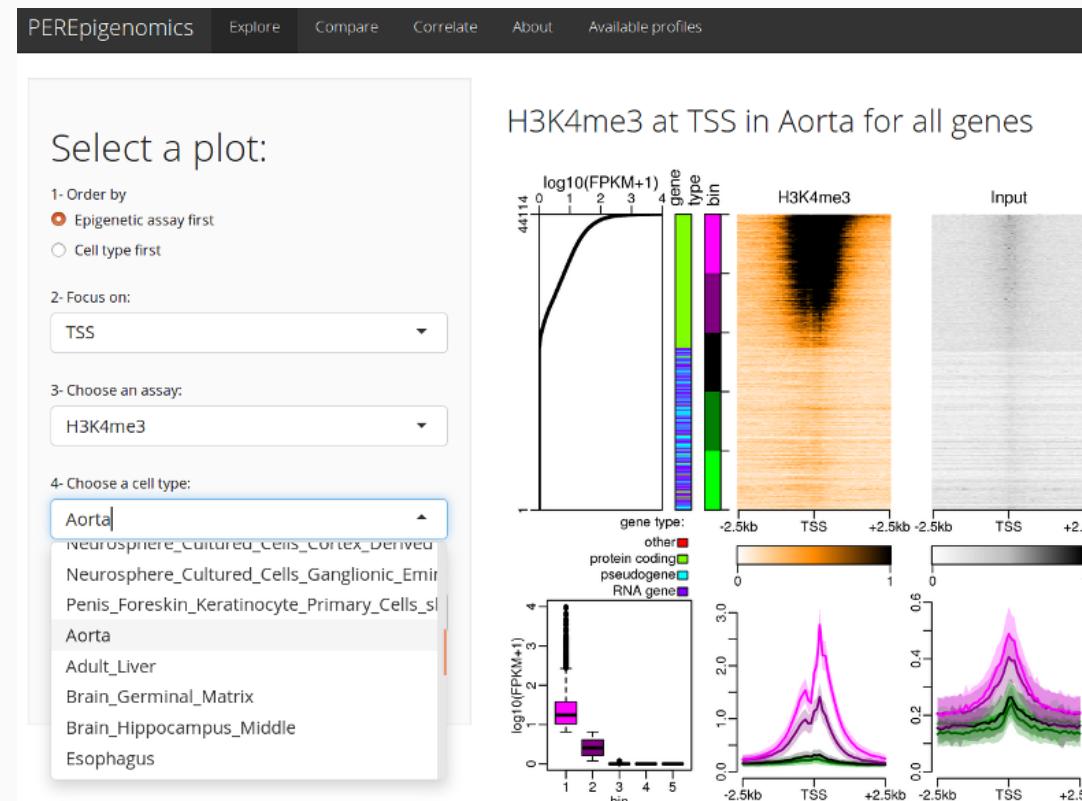


Figure 1:
Epigenetic marks at
promoters and
transcription level
in H1 embryonic
stem cells

PEREpigenomics

Profile Explorer of Roadmap Epigenomic data

www.perepigenomics.roslin.ed.ac.uk



Summary of results

mark	cell type by cell type	gene by gene	center on TSS
WGBS	negative	negative	
DNase	positive	positive	
H2A.Z	positive	balanced	
H3K4me1	positive	positive	
H3K4me2	positive	positive	
H3K4me3	positive	positive	
H3K9me3	unclear	unclear	no
H3K23me2	positive	NA	
H3K27me3	negative – variable	negative	
H3K36me3	positive	positive	no
H3K79me1	positive	positive	no
H3K79me2	positive	balanced	no

mark	cell type by cell type	gene by gene	center on TSS
H2AK5ac	positive	positive	
H2BK120ac	positive	positive	
H2BK12ac	positive	positive	
H2BK15ac	positive	balanced	
H2BK20ac	neutral	NA	
H2BK5ac	positive	positive	
H3K4ac	positive	positive	
H3K9ac	positive	positive	
H3K14ac	positive	positive	
H3K18ac	positive	positive	
H3K23ac	positive	positive	
H3K27ac	positive	positive	
H3K56ac	positive	NA	
H4K8ac	positive	positive	
H4K12ac	positive	NA	
H4K91ac	positive	positive	

Conclusions

PEREpigenomics offers interesting visualisations of epigenetic data gathered by Roadmap Epigenomics.

Perspectives

- update with new annotation data (ongoing)
- preprint & source code available soon
- develop similar approach for FAANG (Functional Annotation of the Animal Genomes) data

Thanks

Anagha Joshi



Anna Mantsoki



Deepti Vipin



THE UNIVERSITY
of EDINBURGH



CREATED WITH
bio
RENDER
WWW.BIORENDER.COM