

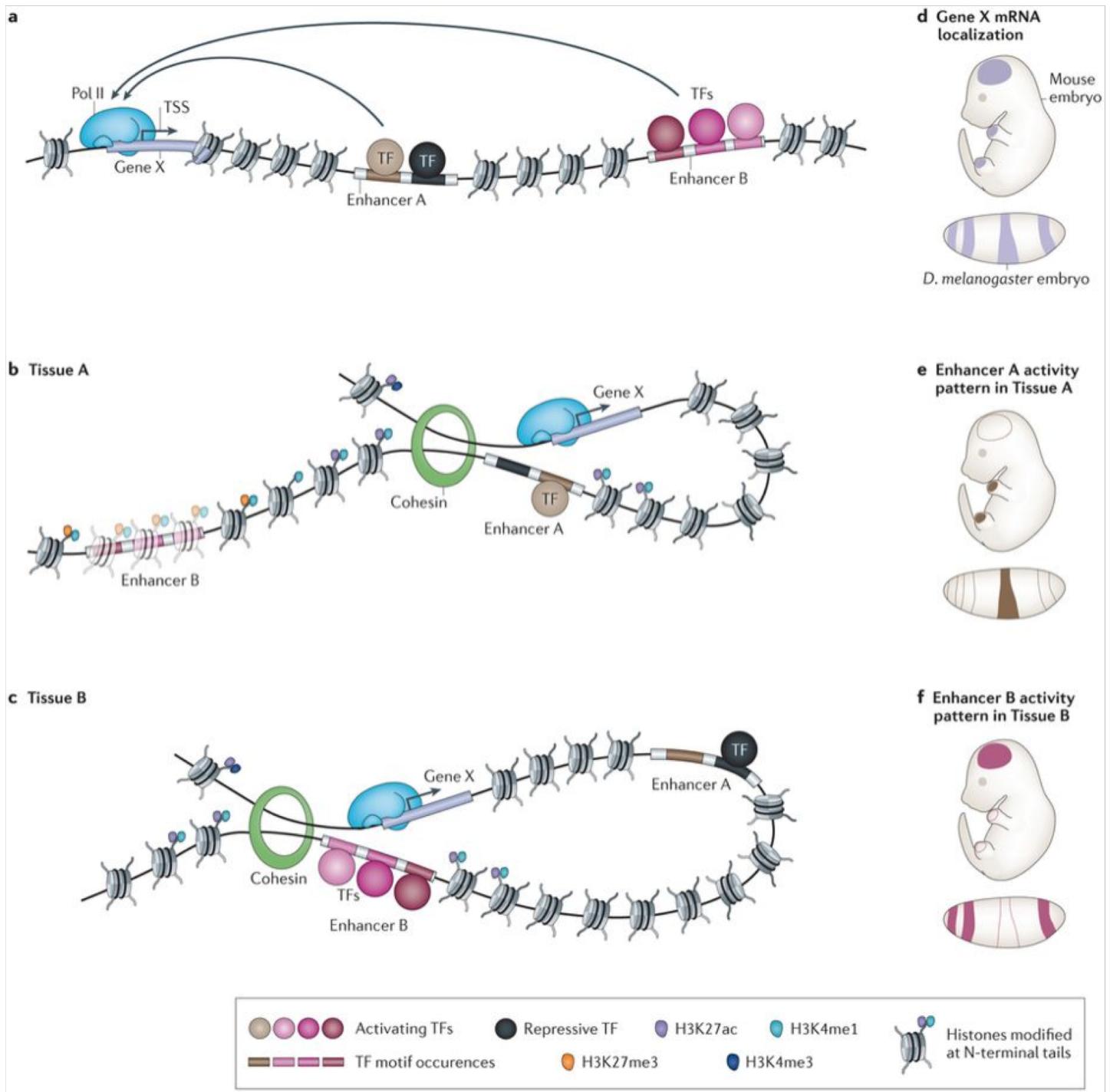
CHARACTERIZING EPMOTERS

Lucia Ramirez

Dra. Alejandra Medina Rivera

Outline

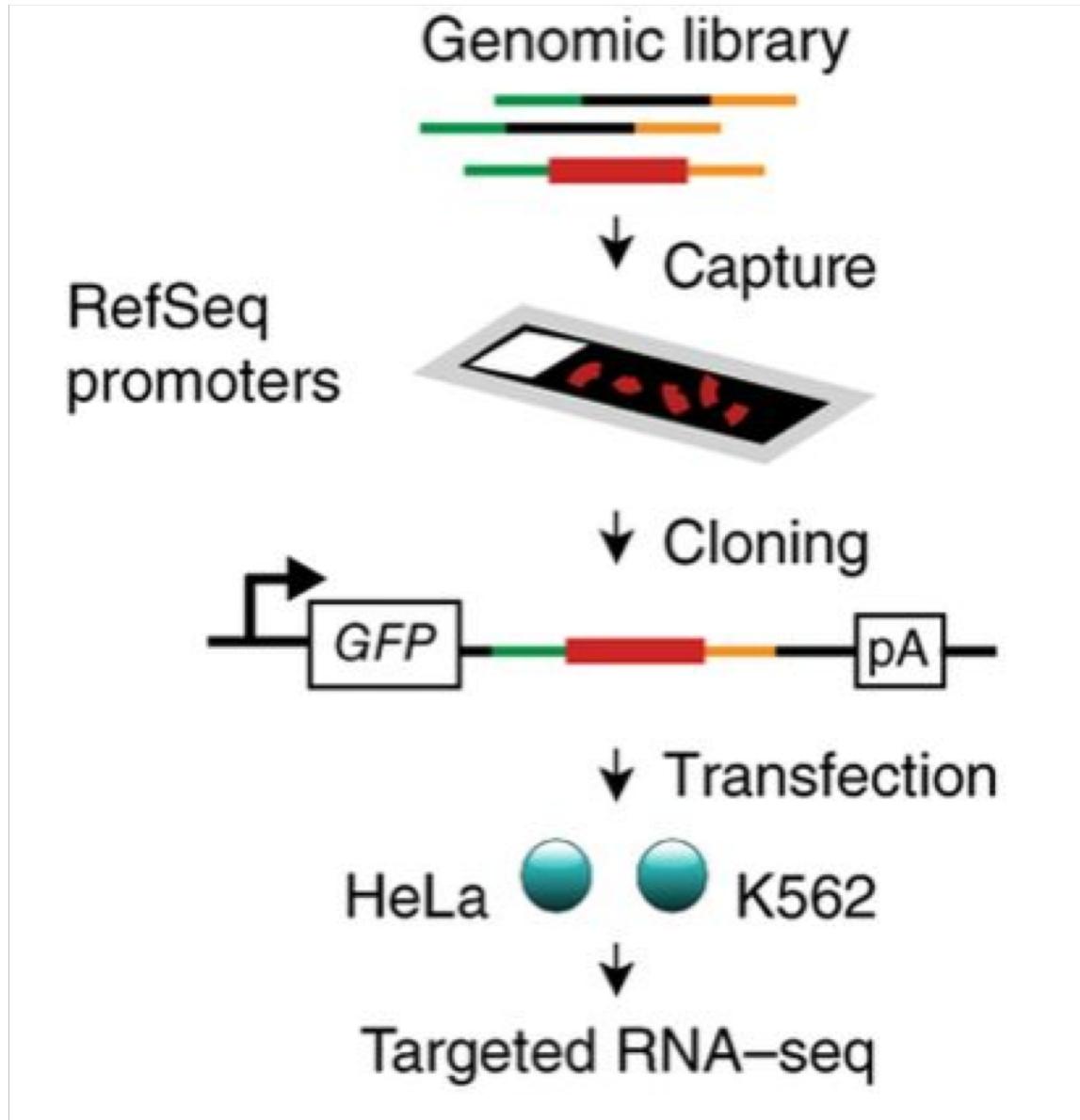
1. Introduction
2. Projects:
 - a. Characterization of the regulatory effect of genetic variants located in epromoters
 - b. Epromoters and their relationship with human diseases
 - c. Binding sites of the epromoters in response to IFN



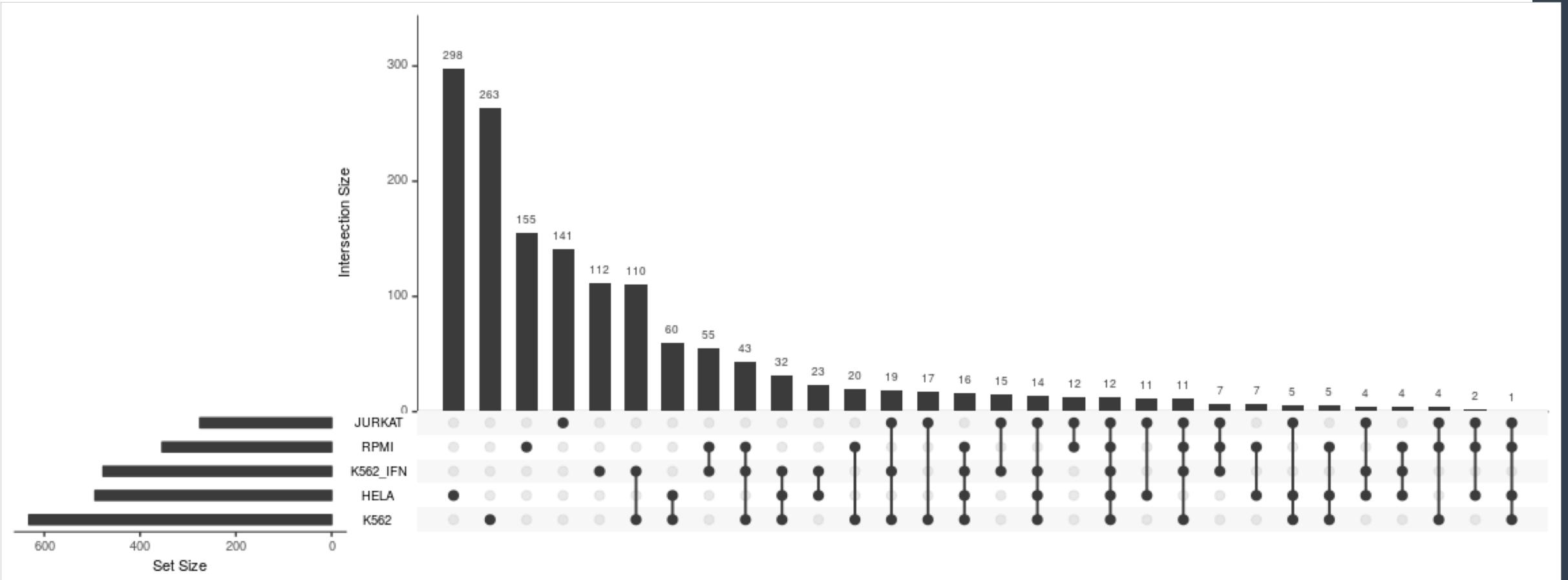
Features

Features (active elements)	Promoter	Enhancer
Intrinsic property	Induce transcription of a heterologous reporter gene	Activate a distal (heterologous) promoter
Transcription initiation	Unidirectional or divergent	Mainly divergent
Ratio between sense and antisense transcripts	Biased toward sense transcription	Equilibrated
Transcription elongation	Produce long polyadenylated transcripts	Some enhancers can produce low levels of polyadenylated transcripts
Histone modifications	H3K27ac (H3K4me1 < H3K4me3)	H3K27ac (H3K4me1 > H3K4me3)
RNAPII and GTF	Present	Present
GpG islands	Majority	Very rare

Identification of epromoters

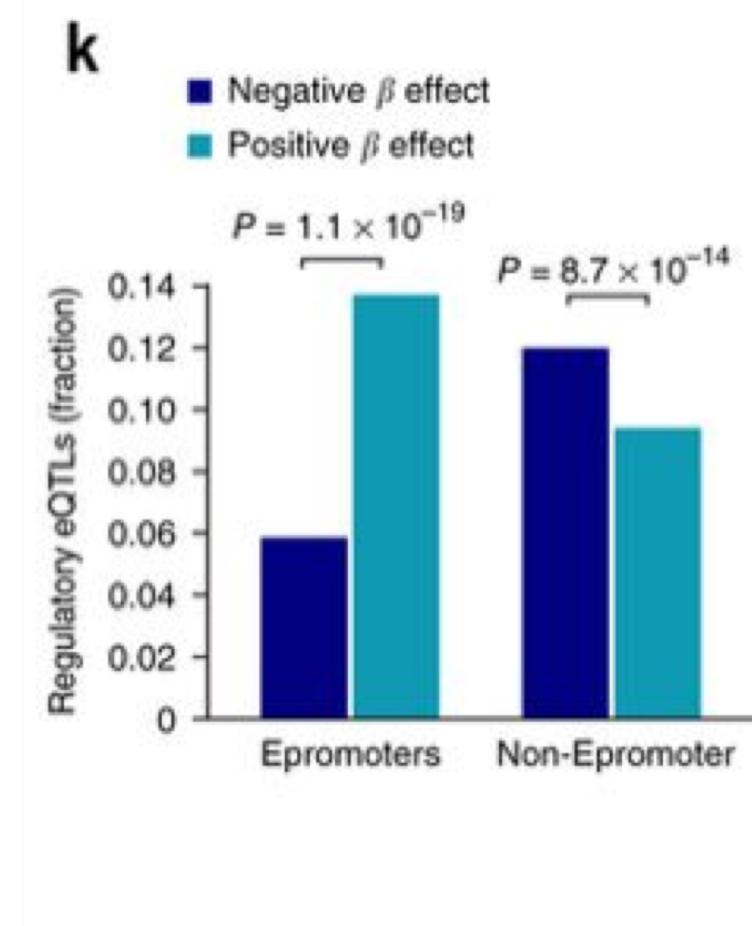
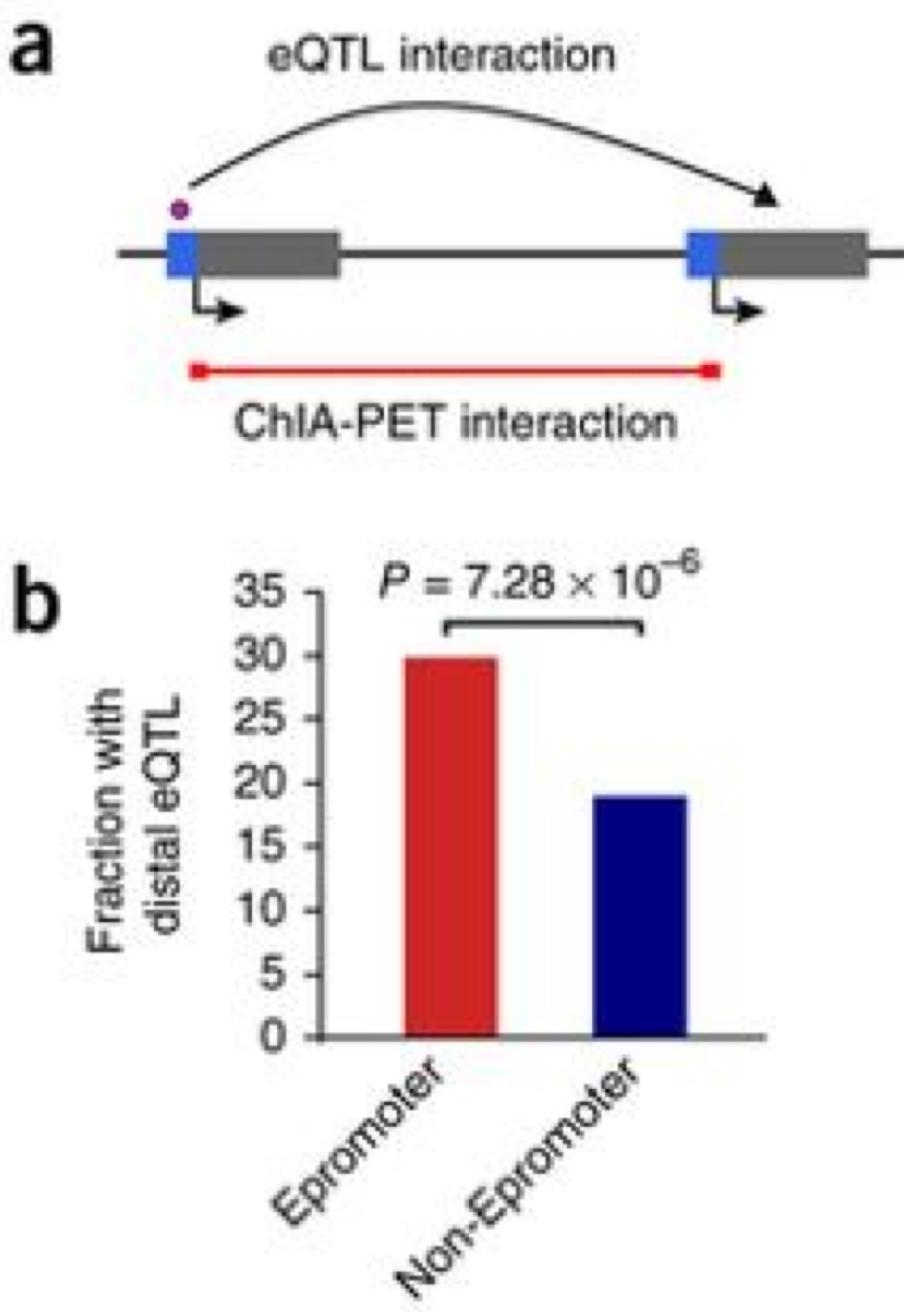


Data



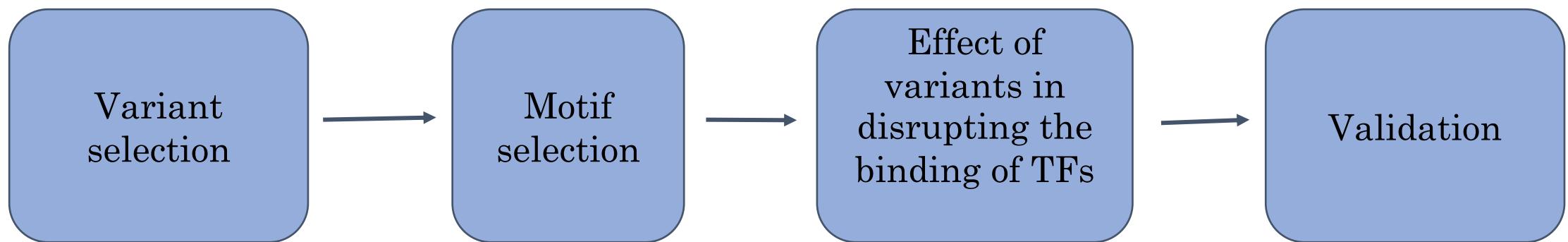
Projects

Characterization of the regulatory effect of genetic variants located in epromoters

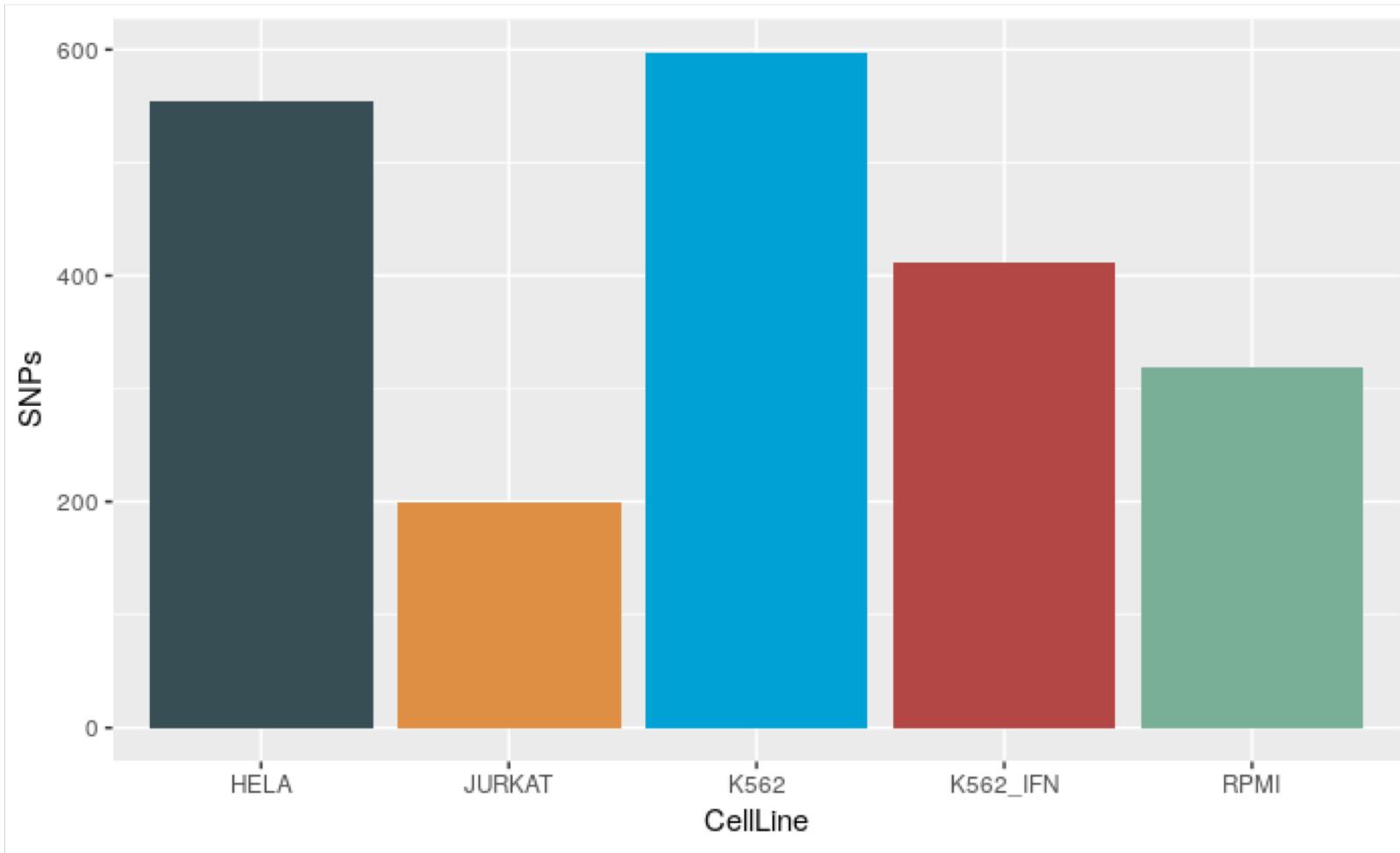


Objective: Functional characterization of the variants that are located in the epromoters, identifying the molecular mechanisms and their contributions to human diseases and traits.

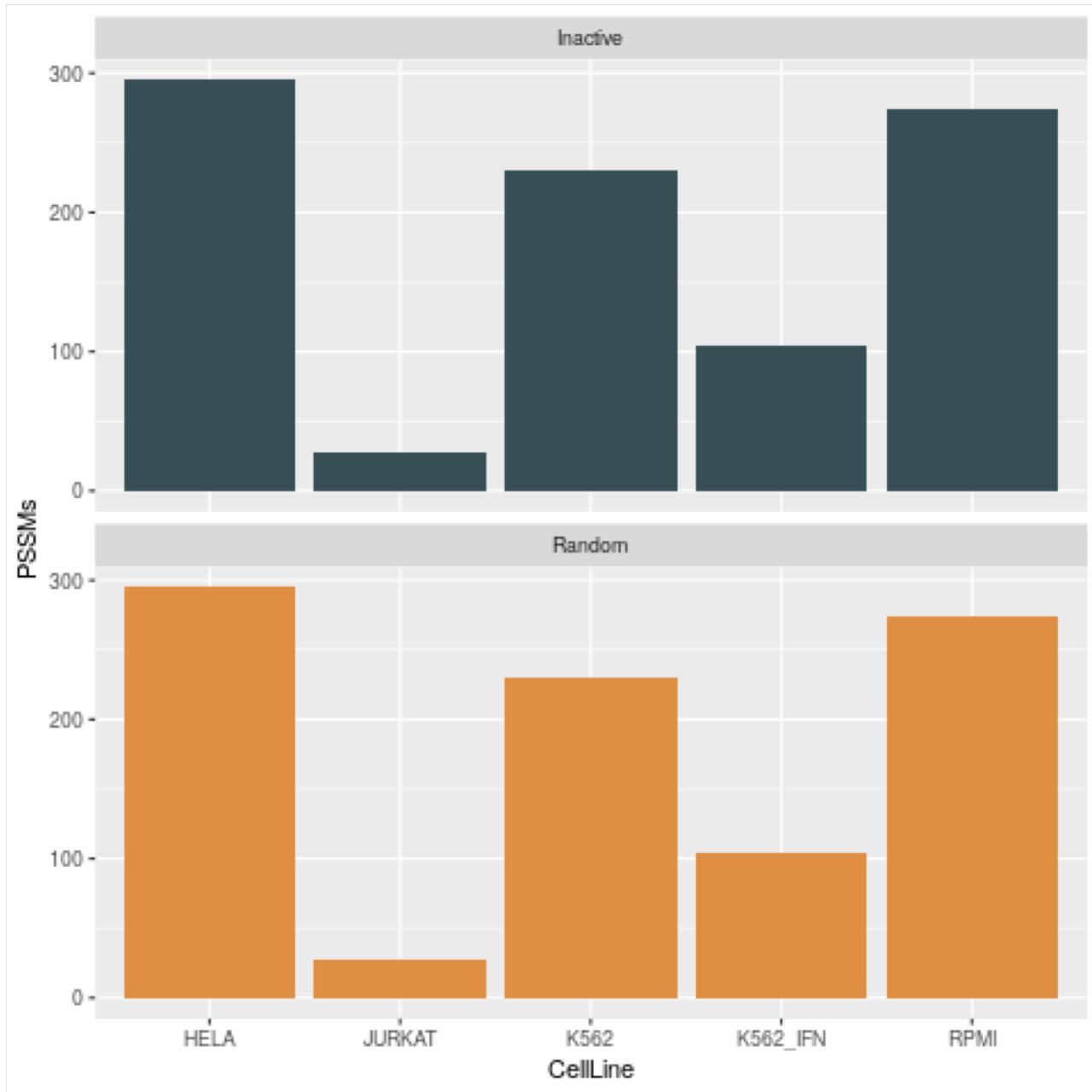
Pipeline



1. Variant selection



2. Motif selection



3. Effect of variants disrupting the binding site

Criteria:

- Opposite predicted effect between the reference and alternative allele
- P-value ratio greater than 10
- Expression of the specific TF in the cell line

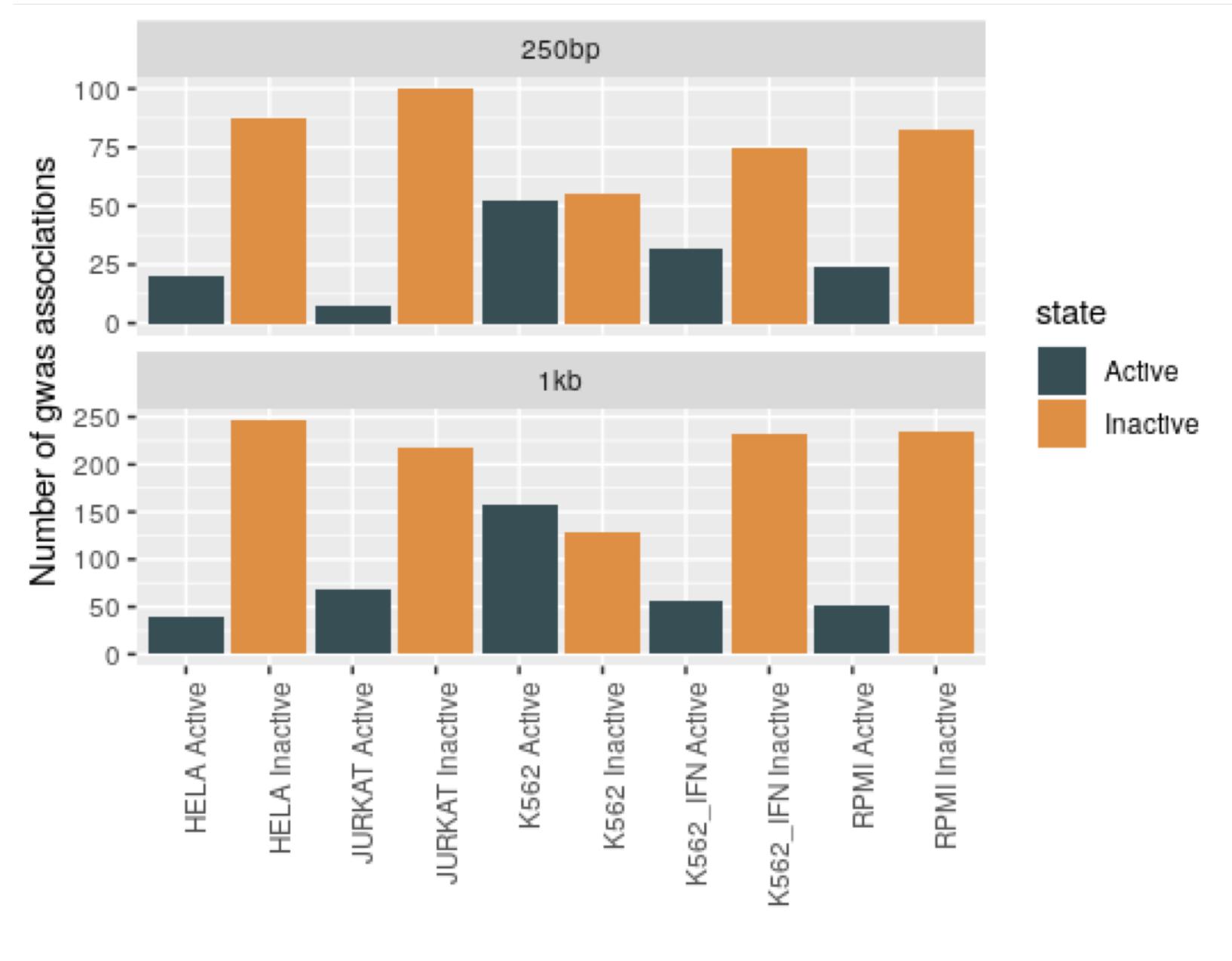
Highlights

- There are variants located in the promoters with enhancer activity that are predicted to have an effect in the binding sites of the TFs
 - Experimental validation
- Determine the contribution of the genetic variants to human phenotypes and diseases

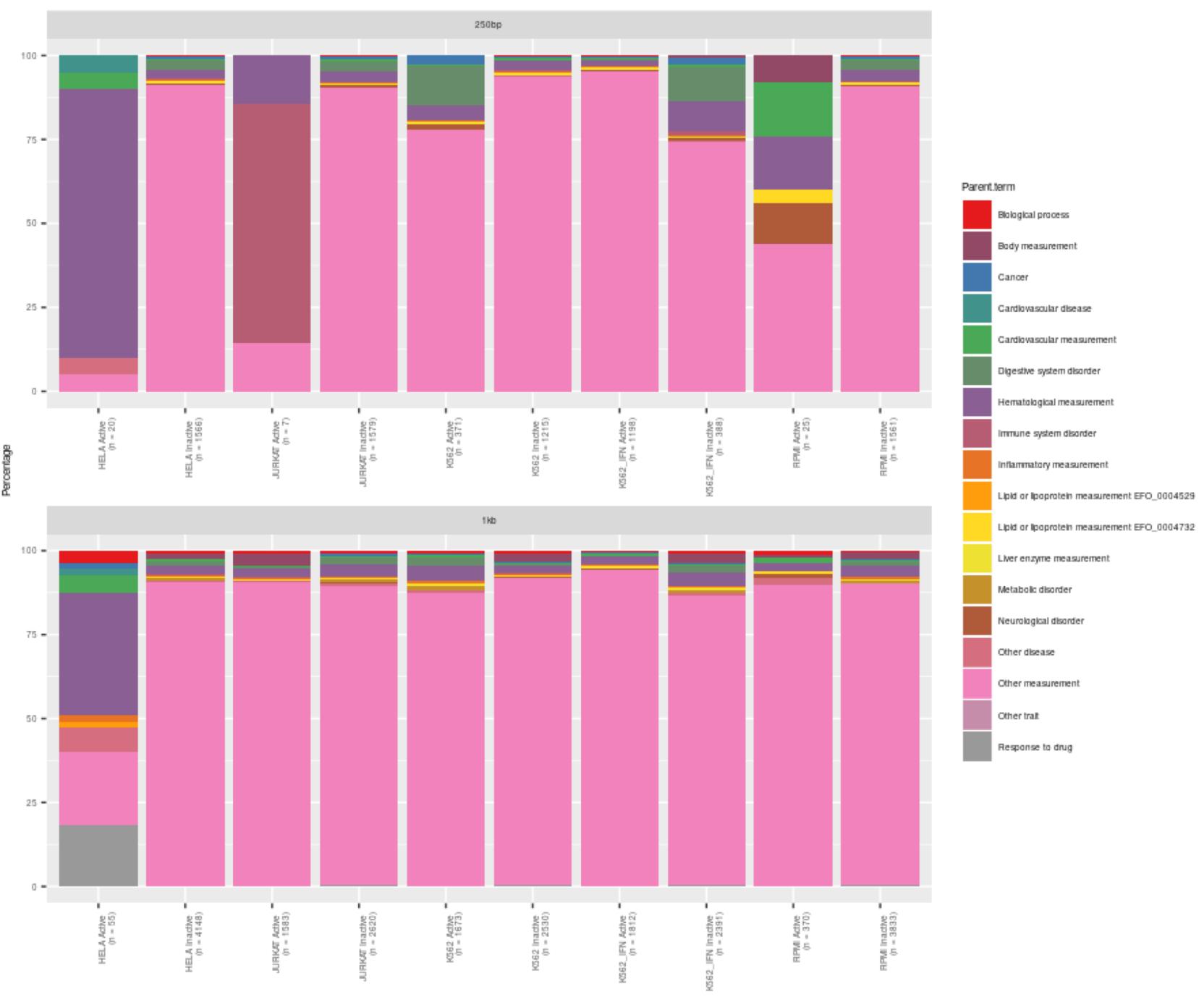
Epromoters and their relationship with human diseases

Objective: to know if there is an enrichment of variants with a particular characteristic or a greater impact in a specific tissue

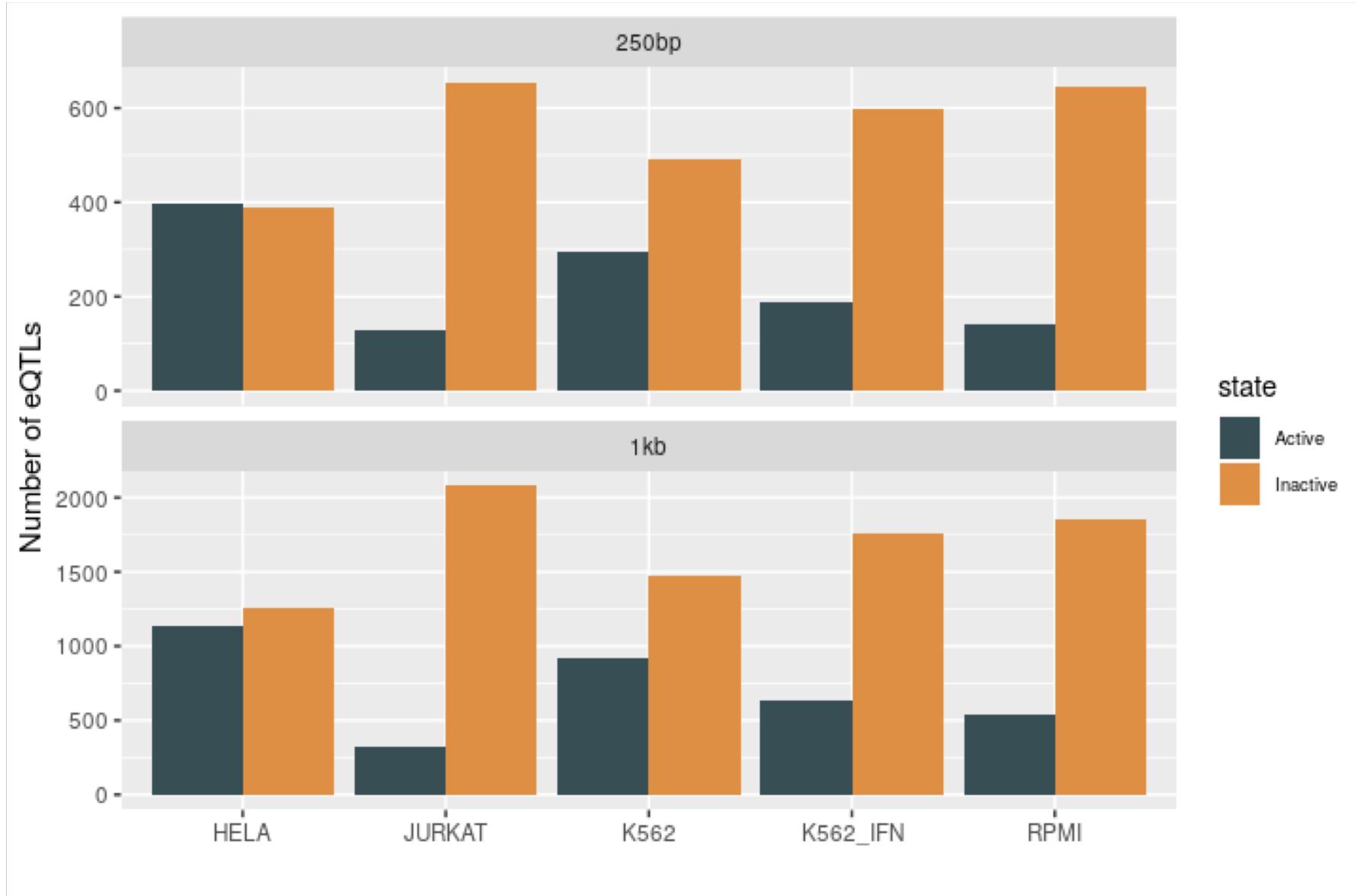
1. Gwas associations



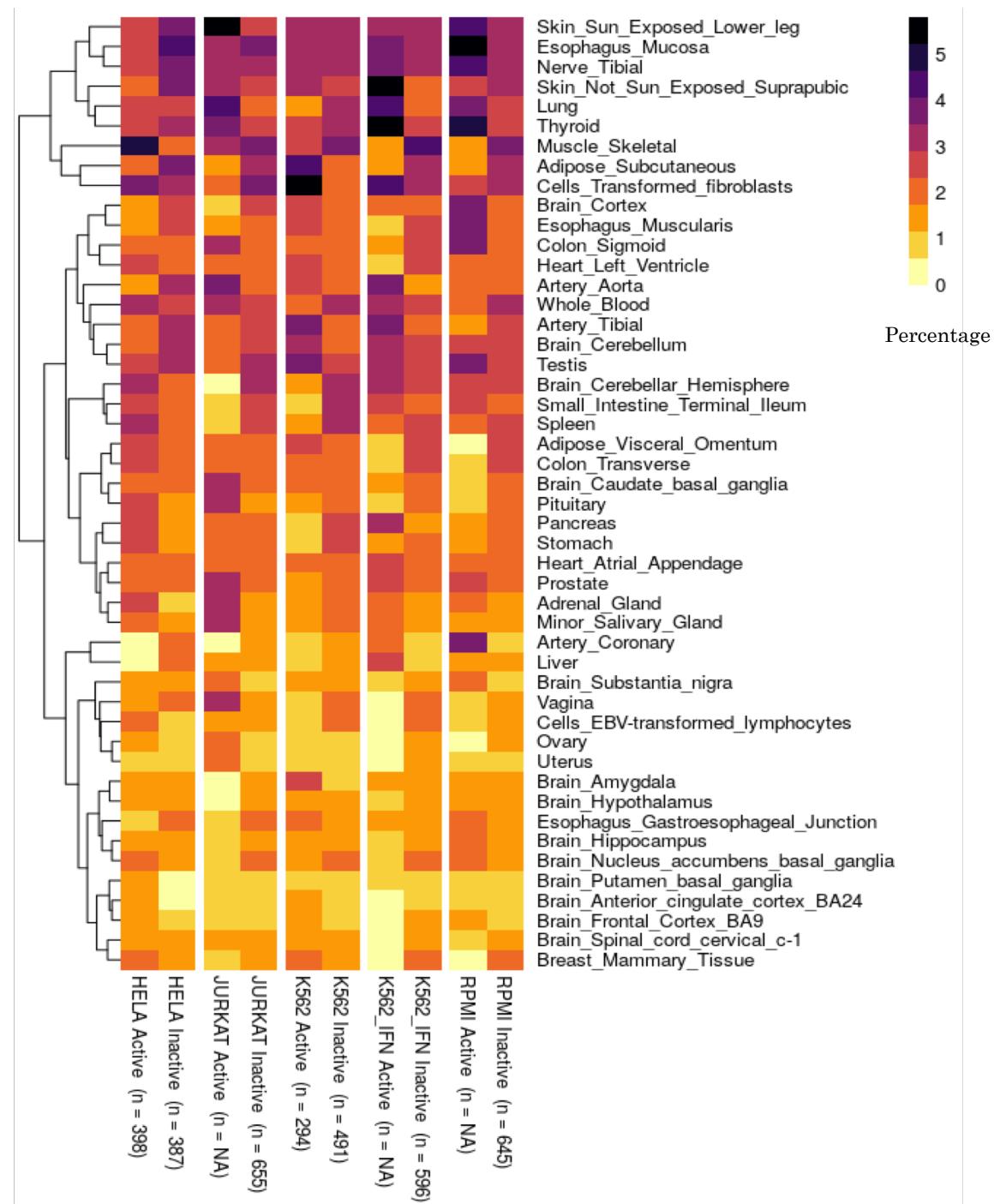
1. Gwas associations



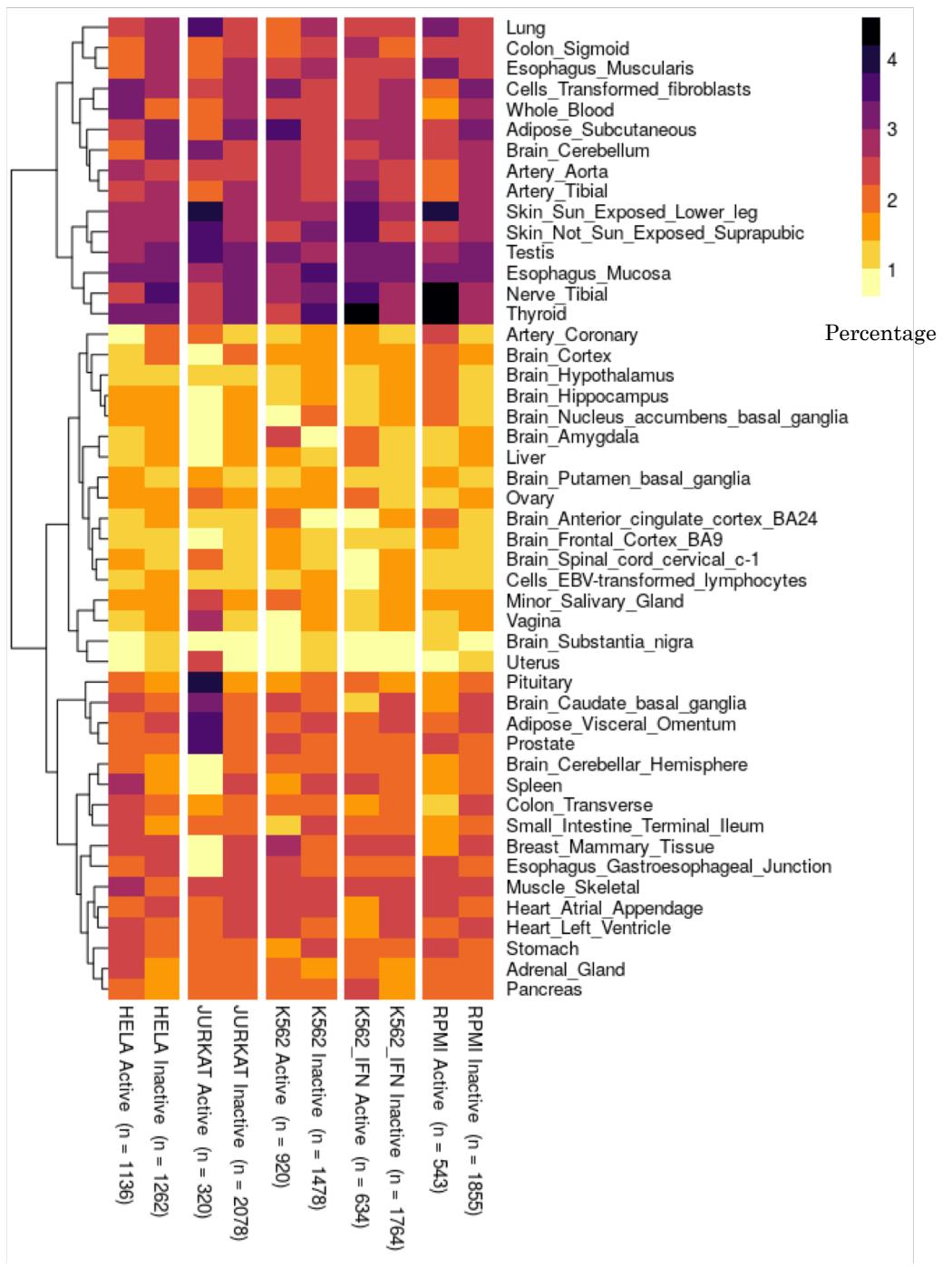
2. eQTLs



2. eQTLs (250 bp)



2. eQTLs (1 kb)



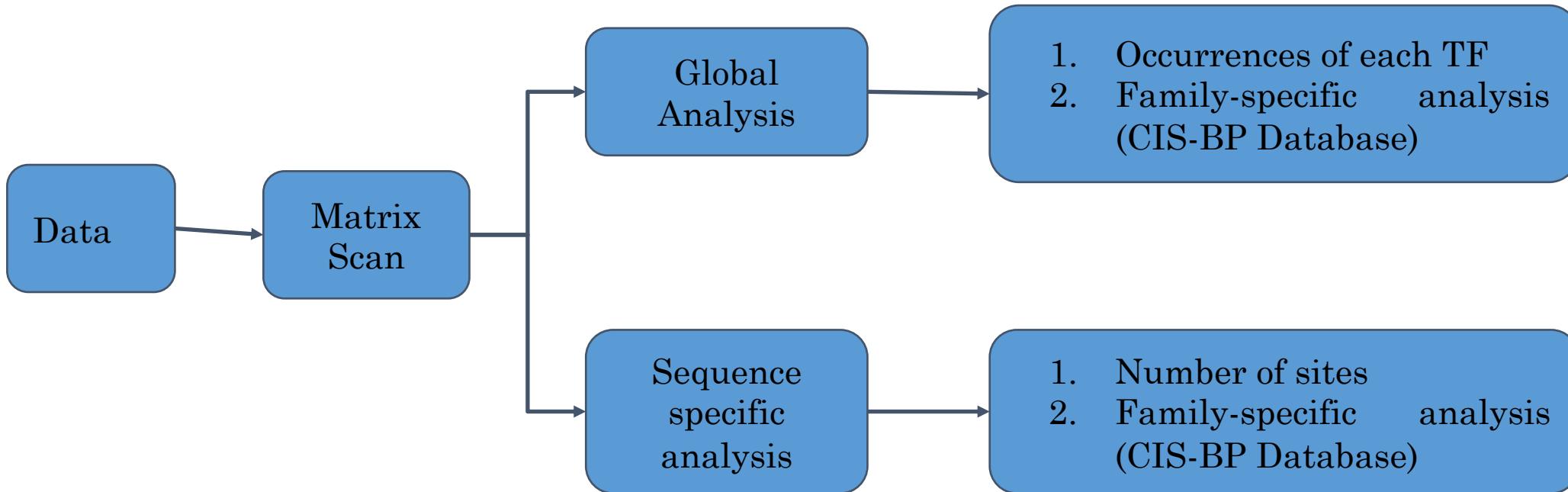
Highlights

- Some datasets present differences between active and inactive epromoters in the enrichment of tissue-specific eQTLs and parent terms obtained from the gwas catalog
 - Statistics?

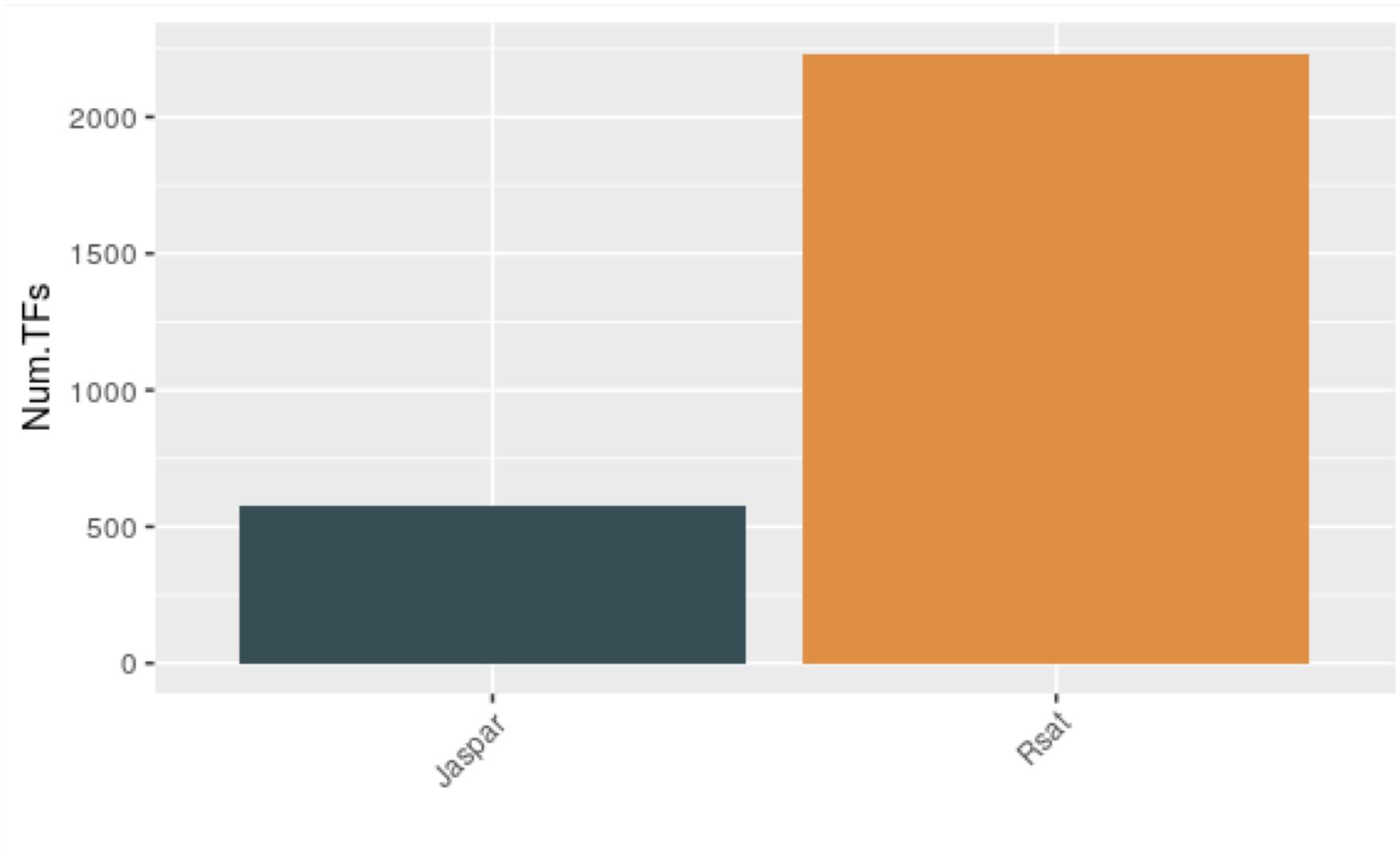
Binding sites of the
epromoters in
response to IFN

Objective: Determine if there is an increasing amount of TFs for the epromoters in cell lines that are induced in IFN

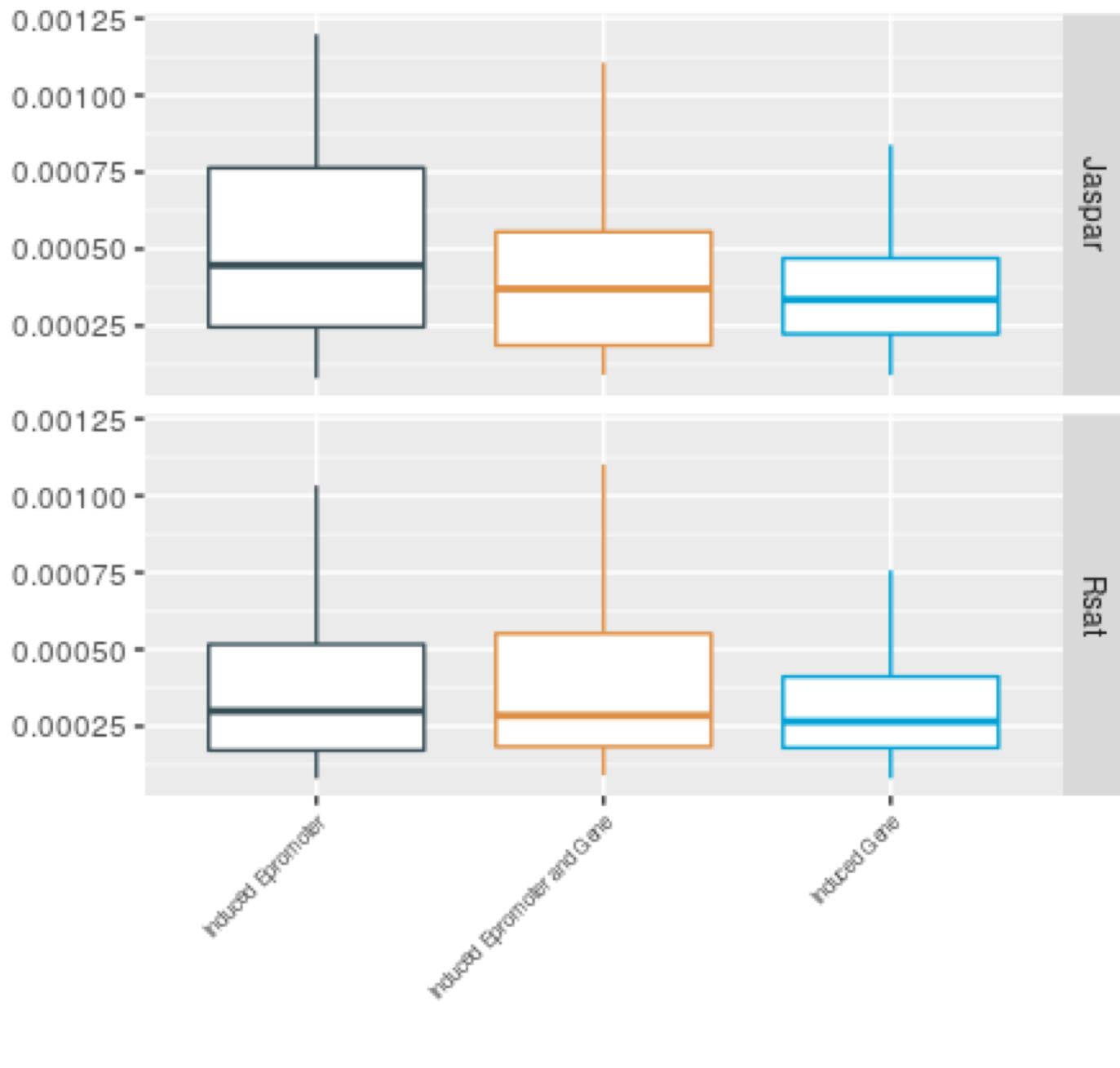
Pipeline



Matrices used



Relative occurrences for each TF



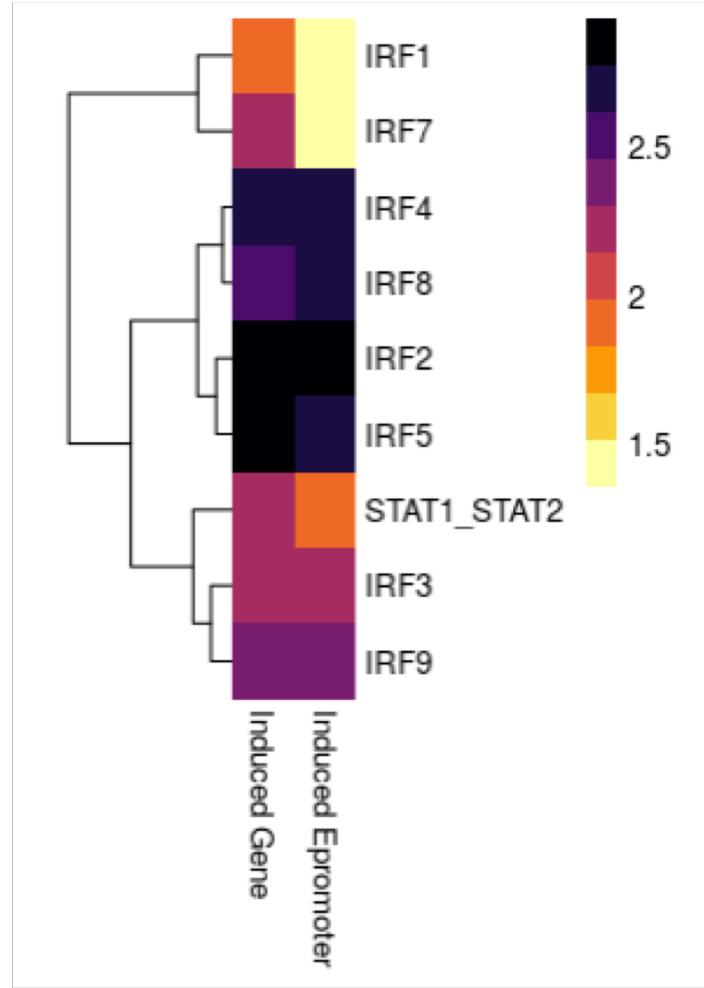
Relative occurrences =

$$\frac{O}{(L_p - L_m + 1) * 2 * N}$$

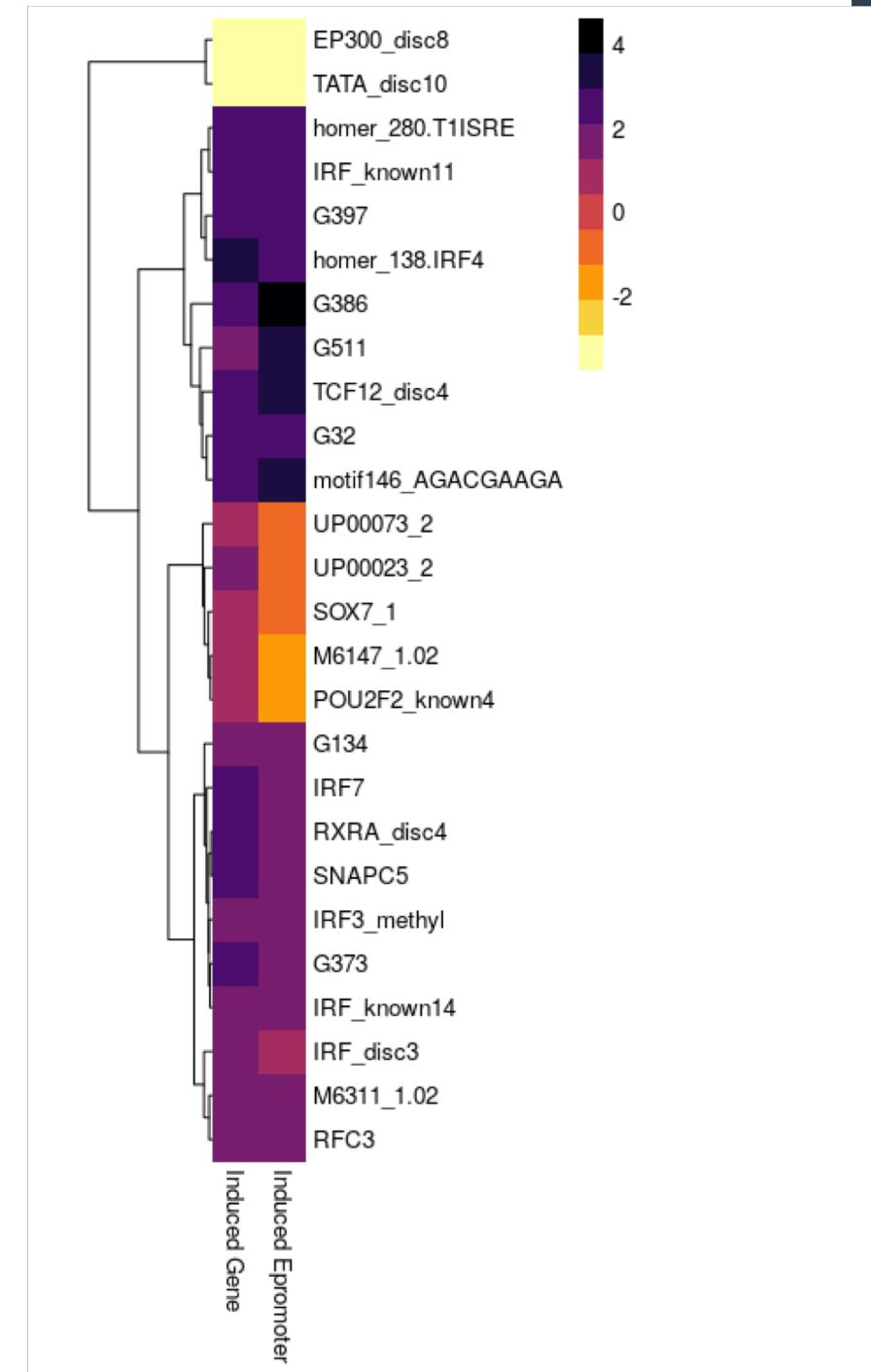
where:

- O: Occurrences of each TF
- L_p: Length of the promoter (250 bp)
- L_m: Length of the motif
- N: Number of promoters

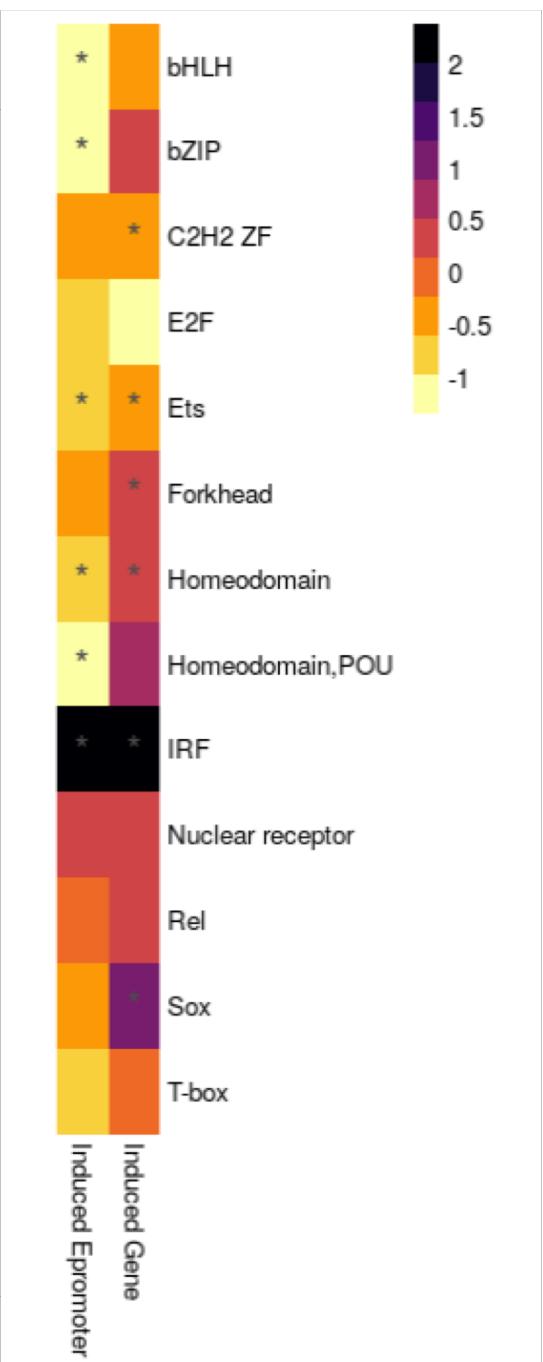
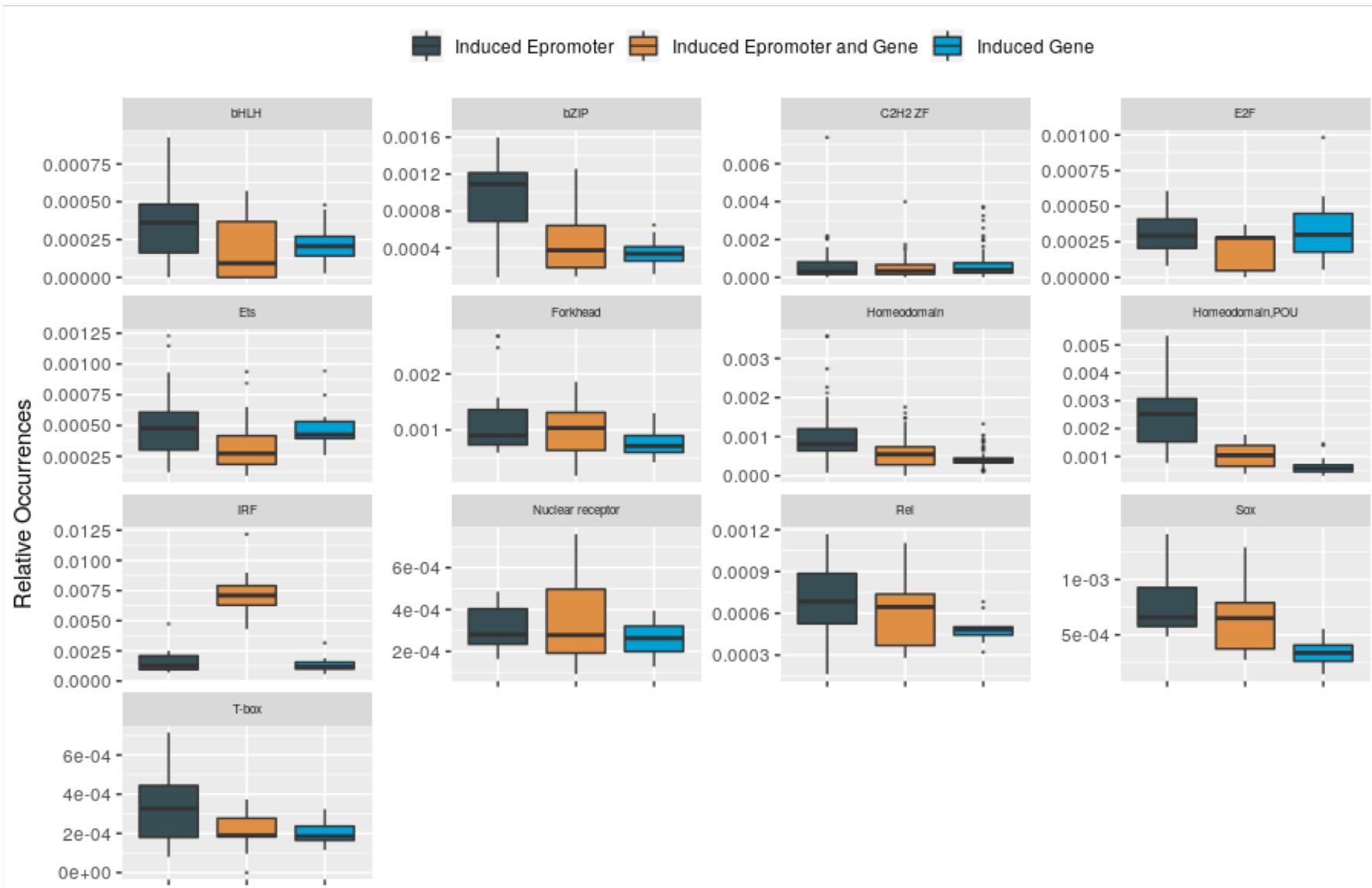
Global analysis:
Log2FC of the relative occurrences for each TF



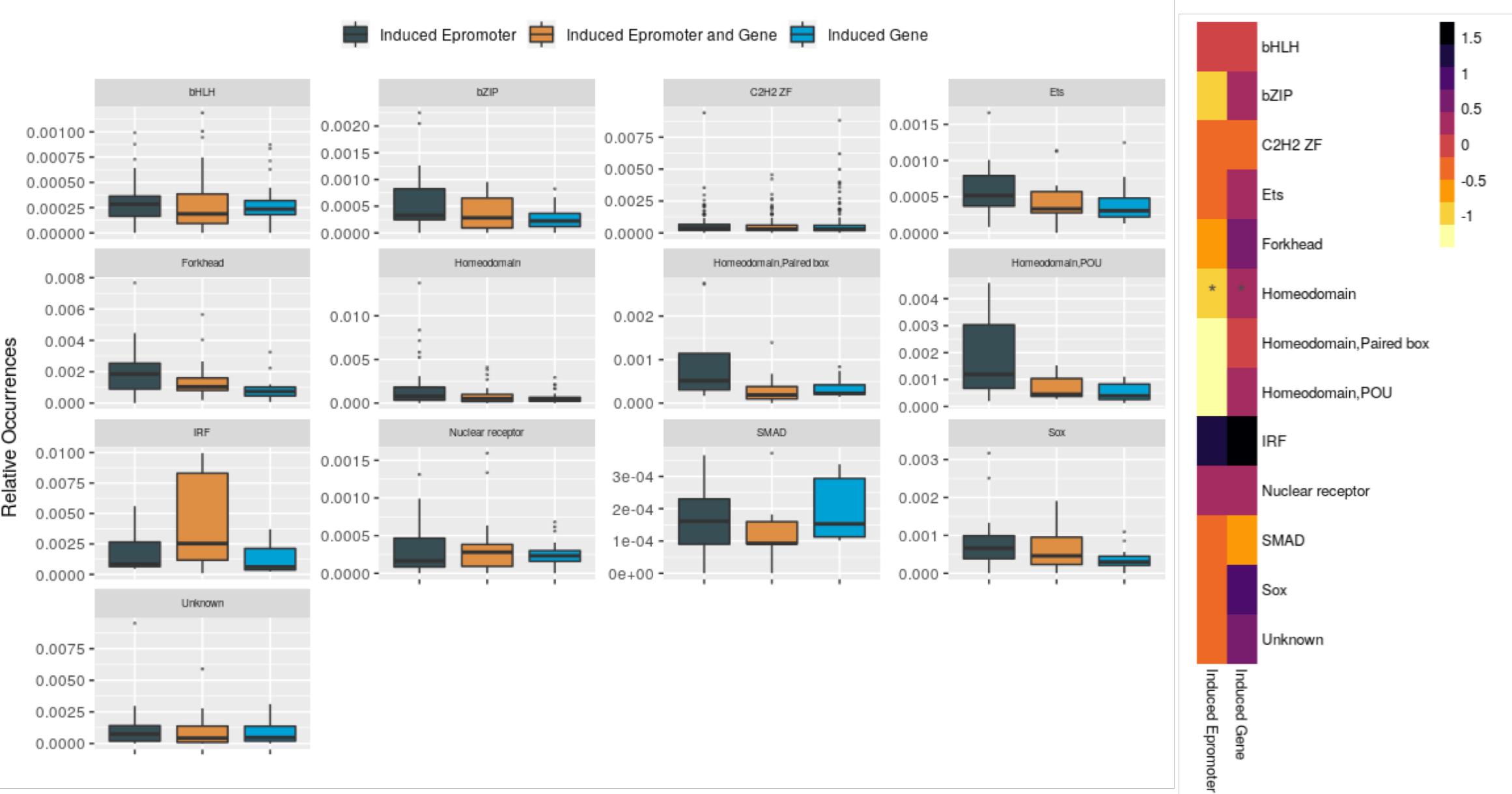
Jaspar
Rsat



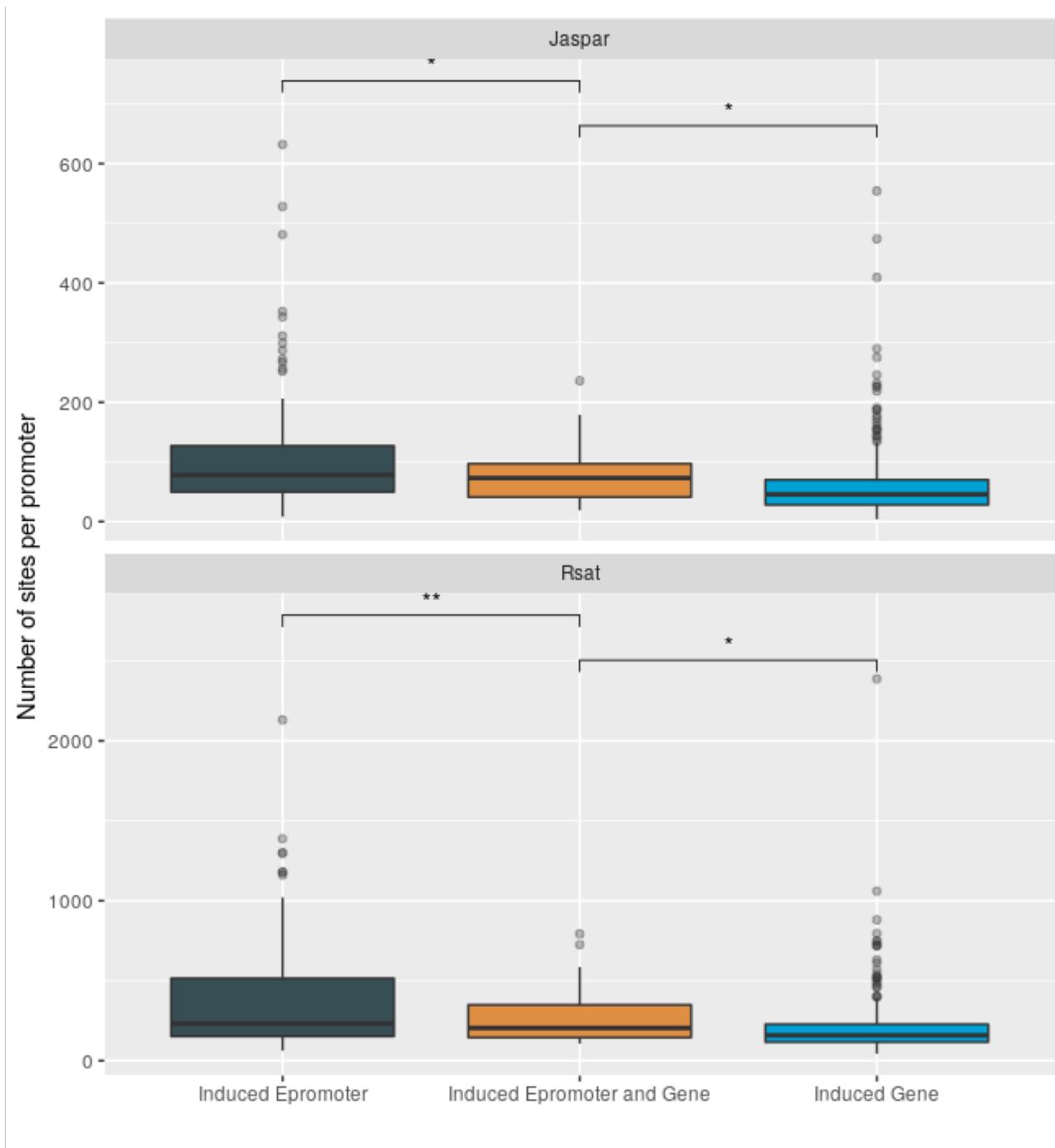
Global analysis: Family-specific analysis (Jaspar)



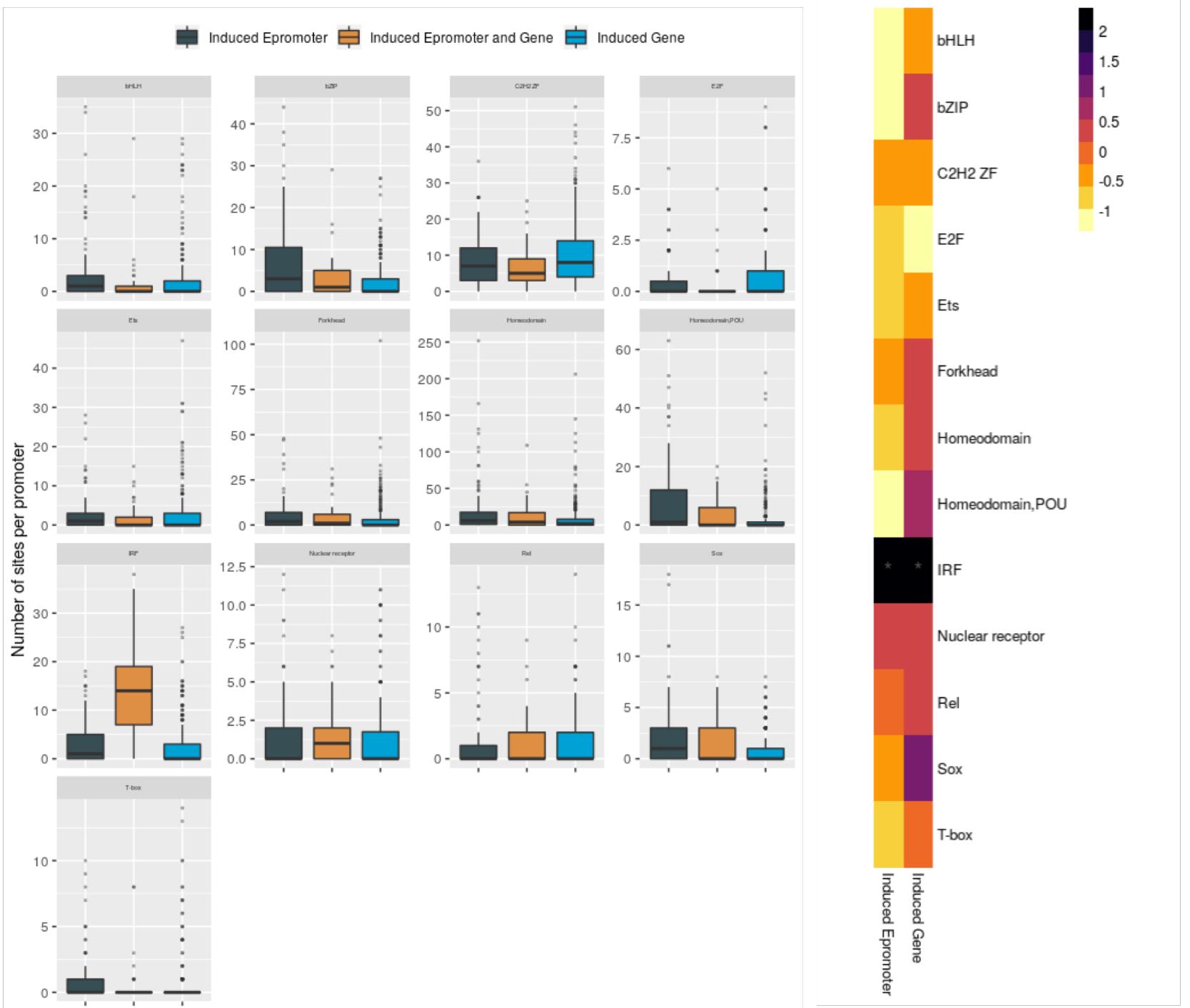
Global analysis: Family-specific analysis (Rsat)



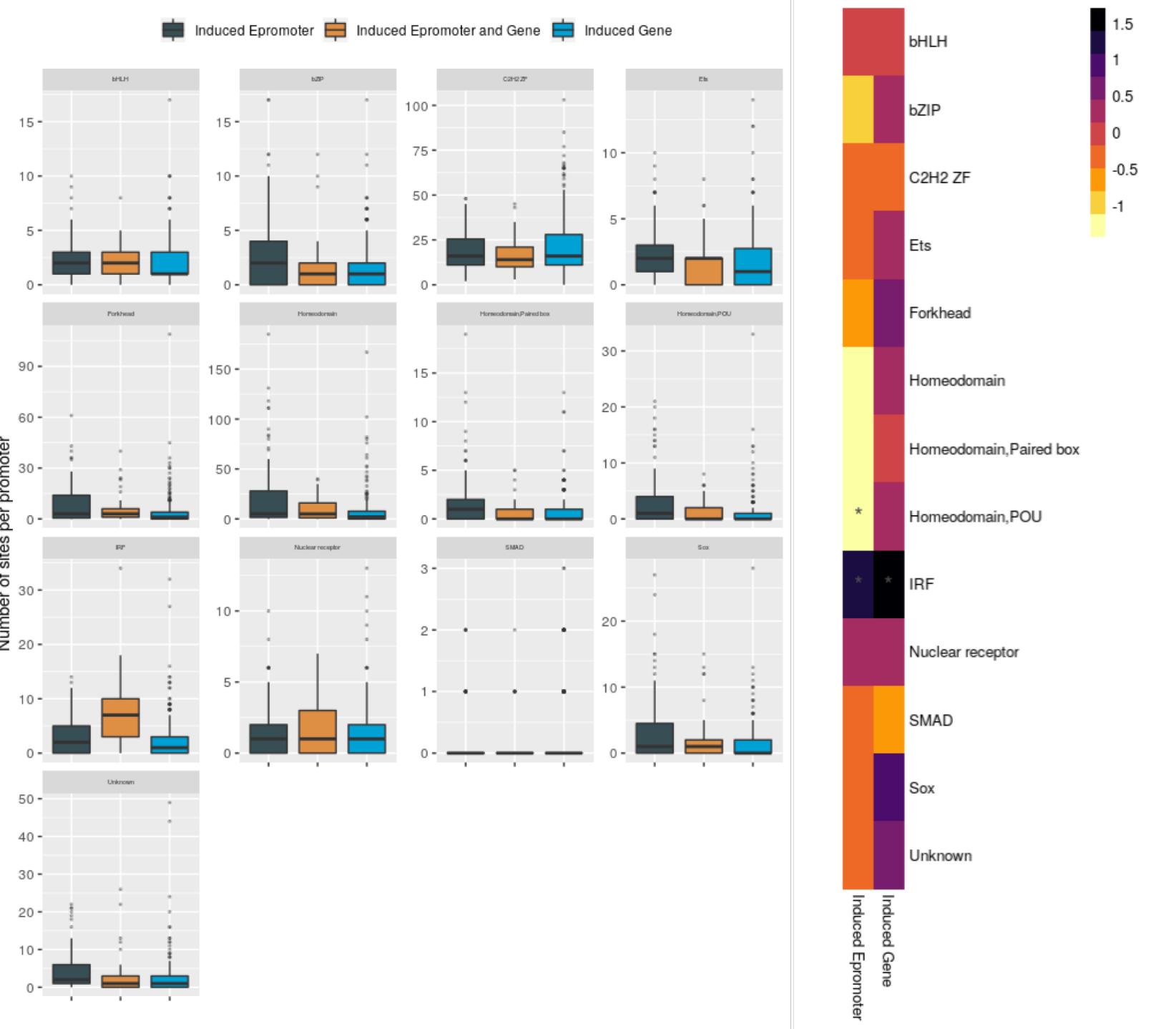
Sequence specific analysis: Number of sites



Sequence specific analysis: Family-specific (Jaspar)



Sequence specific analysis: Family-specific (Rsat)



Highlights

There is an increase of binding sites for the members of the IRF family
in the datasets that have induced genes and promoters, compared to
the ones that only have one of them induced

Thanks 😊