Journal Club

Edmund Miller

2020-10-21 Wed

Roam Update

Total functional score of enhancer elements identifies lineage-specific enhancers that drive differentiation of pancreatic cells

Fig 1.

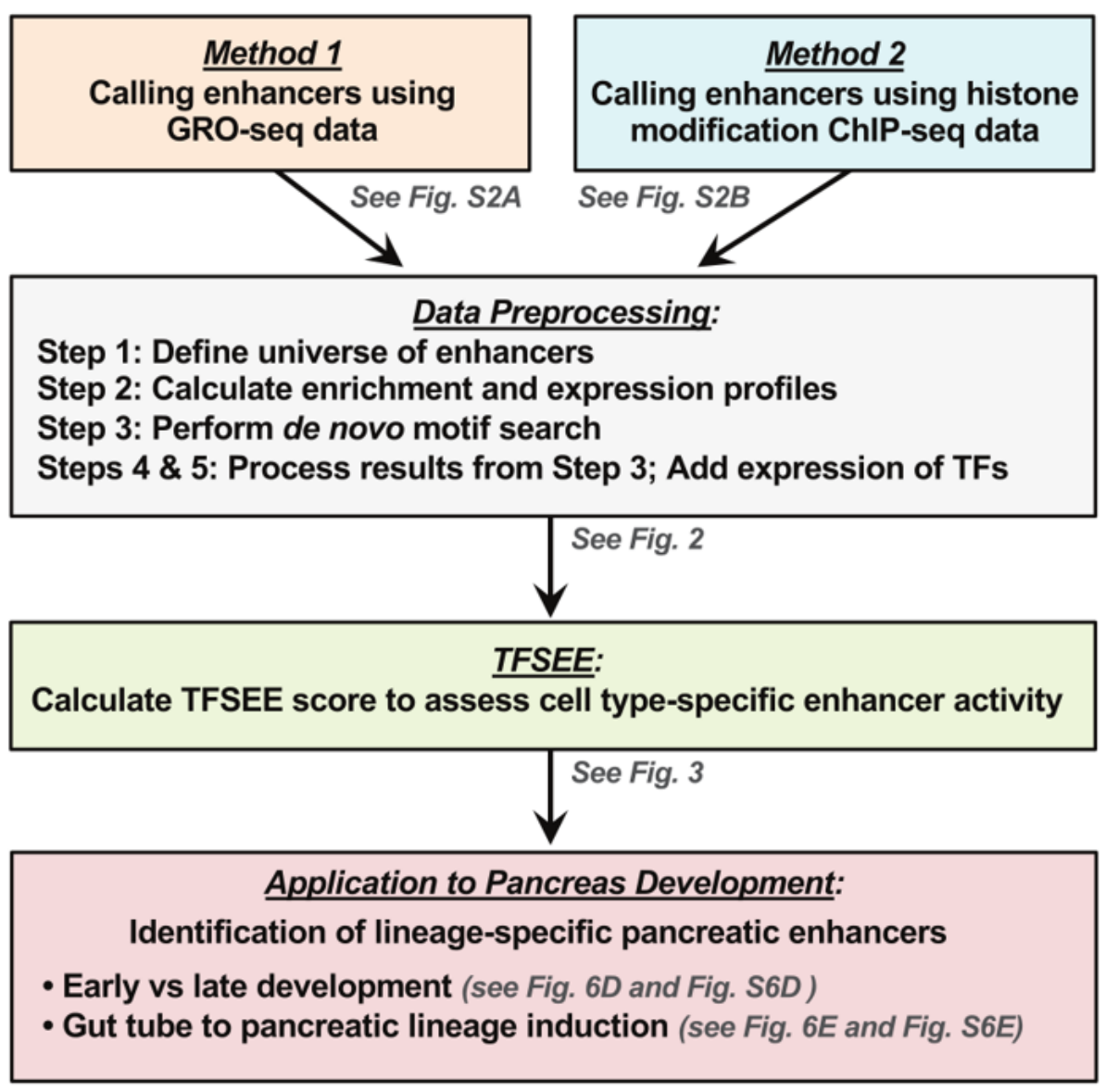

Method 1
Calling enhancers using GRO-seq data
Method 2
Calling enhancers using histone modification ChIP-seq data
See Fig. S2A
See Fig. S2B
Data Preprocessing:
Step 1: Define universe of enhancers
Step 2: Calculate enrichment and expression profiles
Step 3: Perform de novo motif search
Steps 4 & 5: Process results from Step 3; Add expression of TFs
See Fig. 2
TFSEE:
Calculate TFSEE score to assess cell type-specific enhancer activity
See Fig. 3
Application to Pancreas Development:
Identification of lineage-specific pancreatic enhancers
• Early vs late development (see Fig. 6D and Fig. S6D)
• Gut tube to pancreatic lineage induction (see Fig. 6E and Fig. S6E)

Fig. S2A Pt1 Unbiased, genome-wide prediction of active enhancers

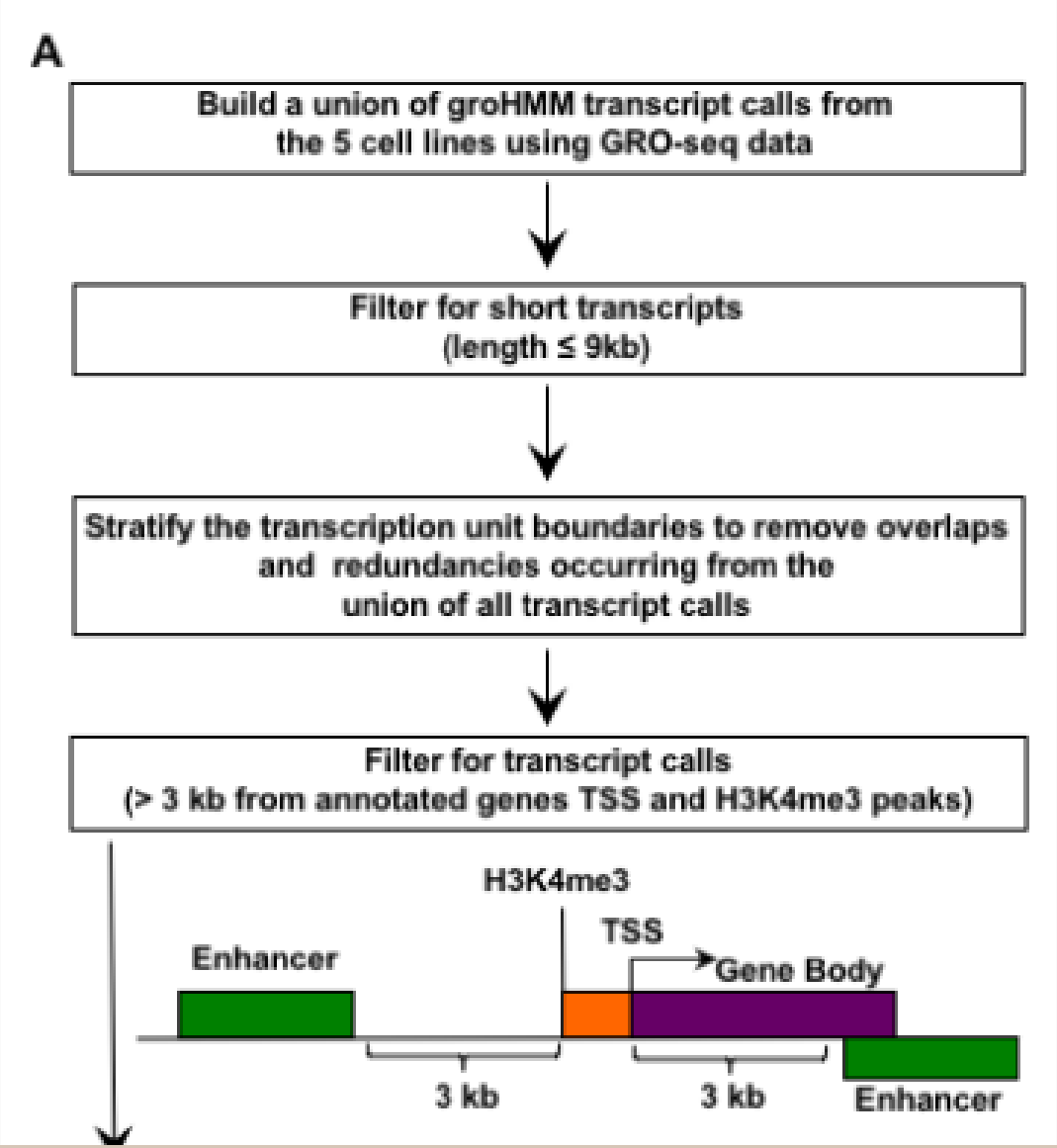


Fig. S2A Pt2 Unbiased, genome-wide prediction of active enhancers

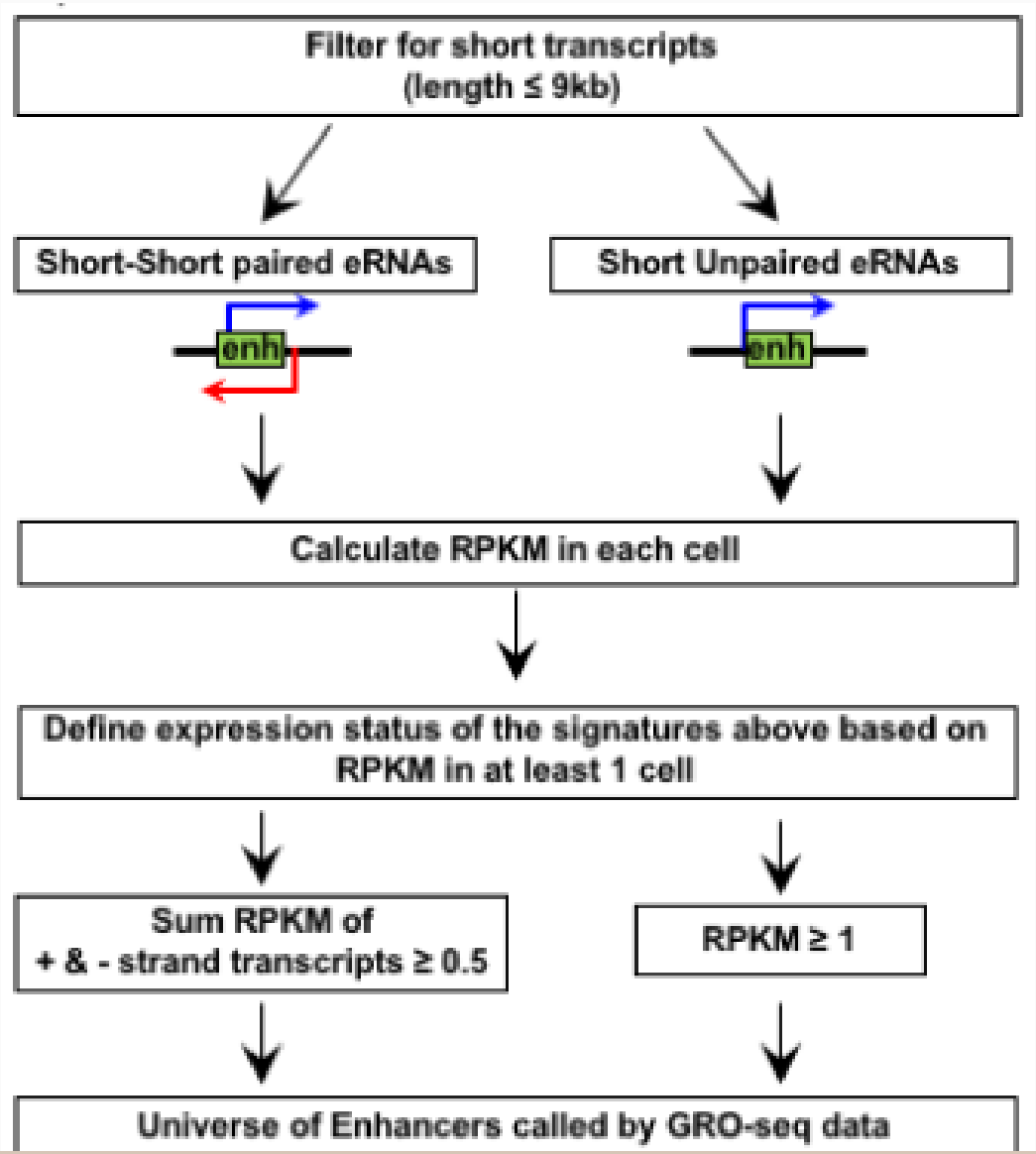


Fig. S2B Unbiased, genome-wide prediction of active enhancers

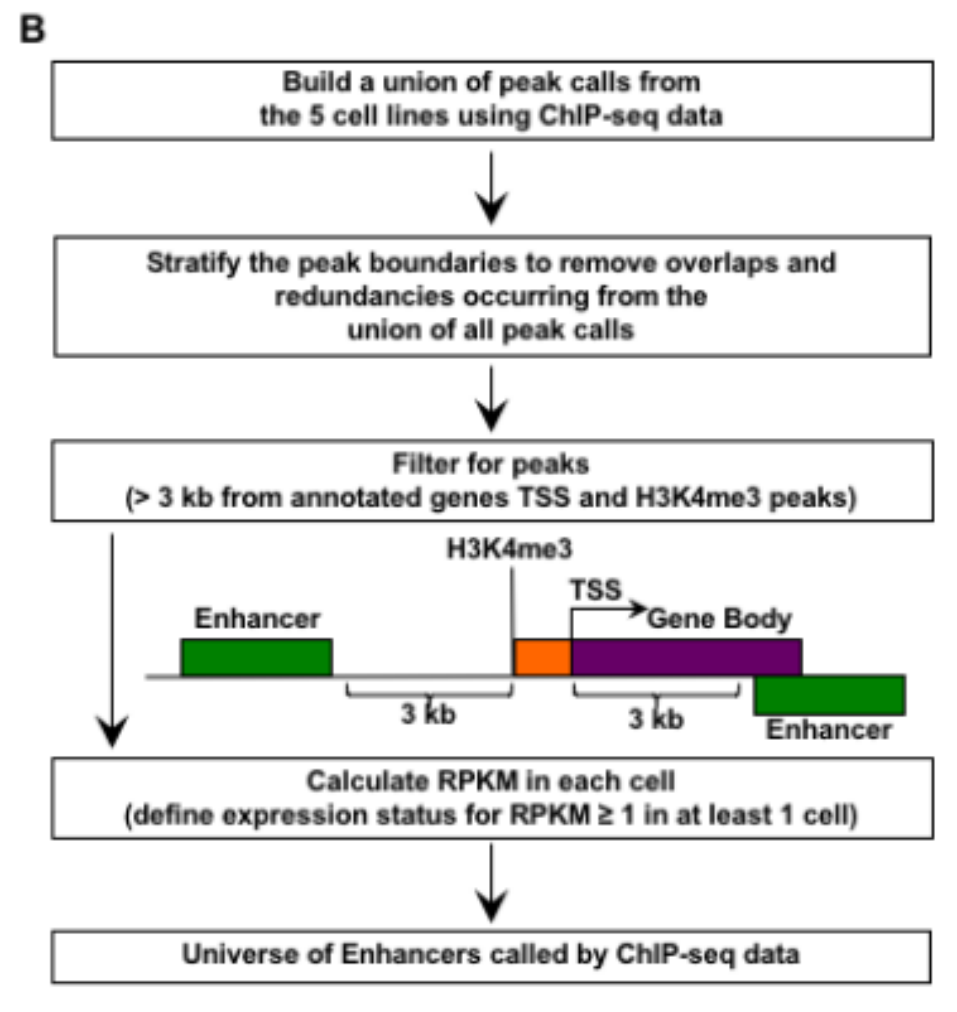


Fig 2. Data Processing for Total Functional Score of Enhancer Elements (TFSEE) Method

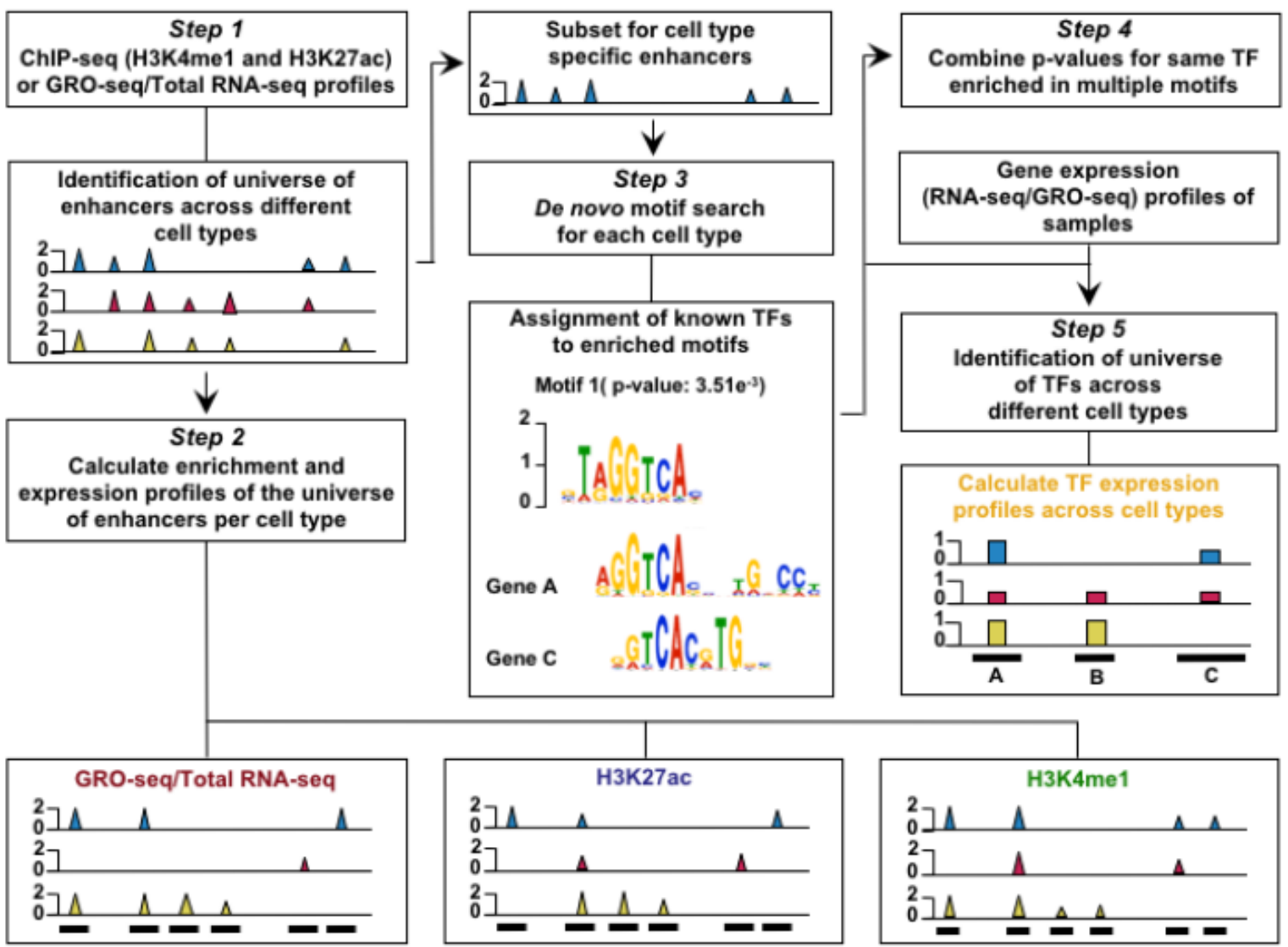


Fig 3. Overview of Total Functional Score of Enhancer Elements (TFSEE) Method

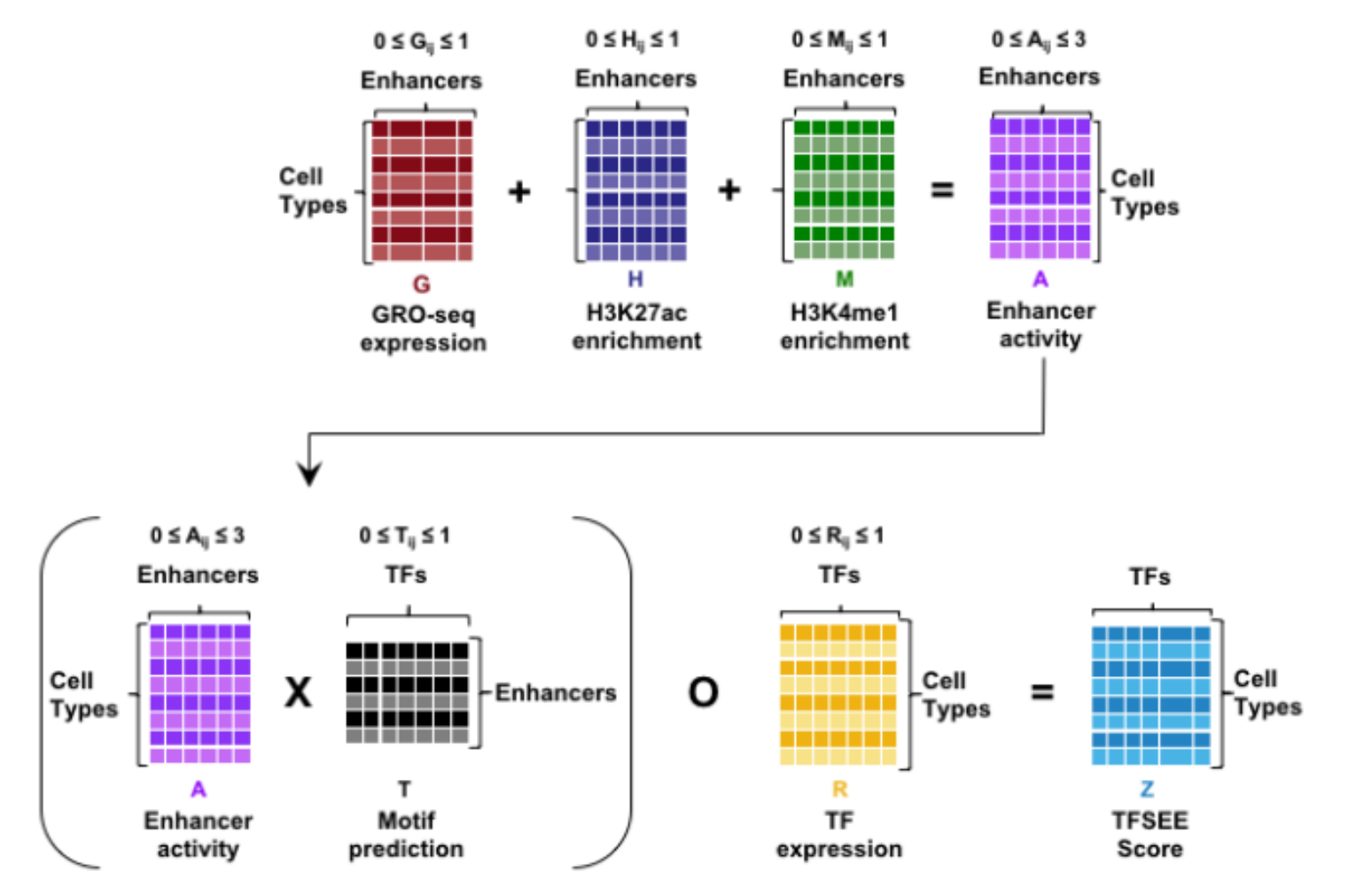


Fig 4A.

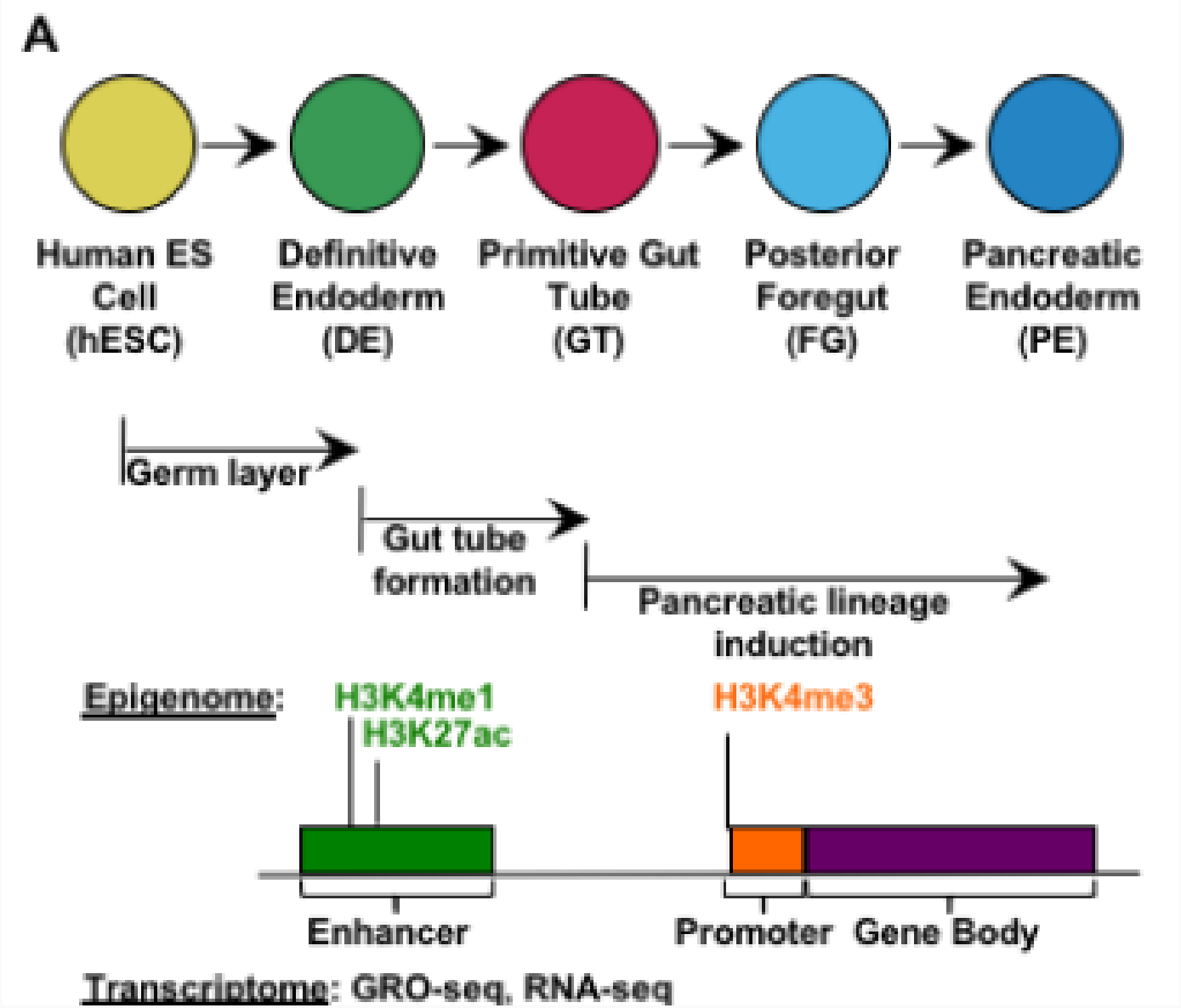

A
Human ES Cell (hESC)
Definitive Endoderm (DE)
Primitive Gut Tube (GT)
Posterior Foregut (FG)
Pancreatic Endoderm (PE)
Germ layer
Gut tube formation
Pancreatic lineage induction
Epigenome:
H3K4me1
H3K27ac
H3K4me3
Enhancer
Promoter
Gene Body
Transcriptome: GRO-seq, RNA-seq

Fig 4B.

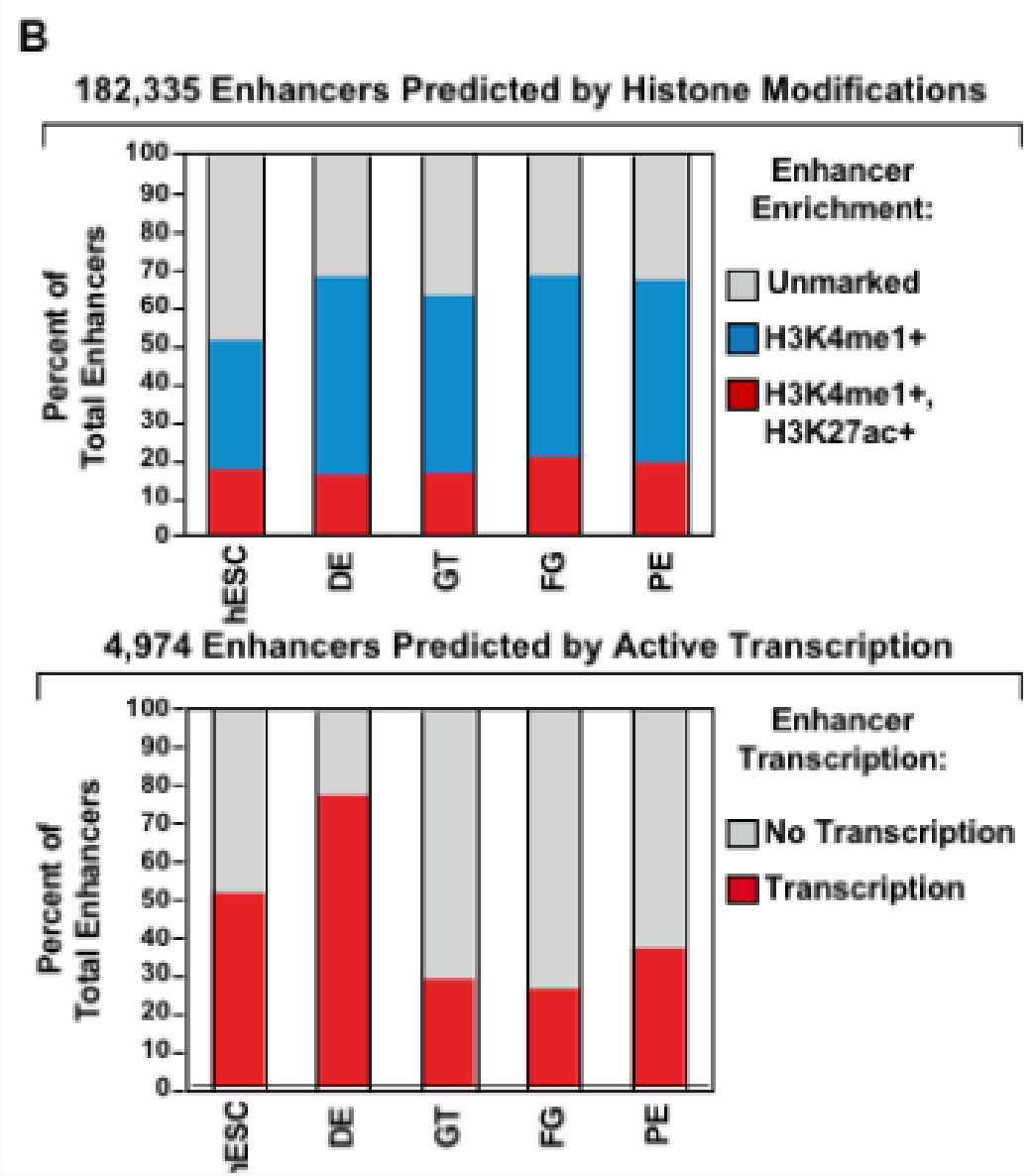

B
182,335 Enhancers Predicted by Histone Modifications
Percent of Total Enhancers
100
90
80
70
60
50
40
30
20
10
0
hESC
DE
GT
FG
PE
Enhancer Enrichment:
Unmarked
H3K4me1+
H3K4me1+, H3K27ac+
4,974 Enhancers Predicted by Active Transcription
Percent of Total Enhancers
100
90
80
70
60
50
40
30
20
10
0
hESC
DE
GT
FG
PE
Enhancer Transcription:
No Transcription
Transcription

Fig 4C.

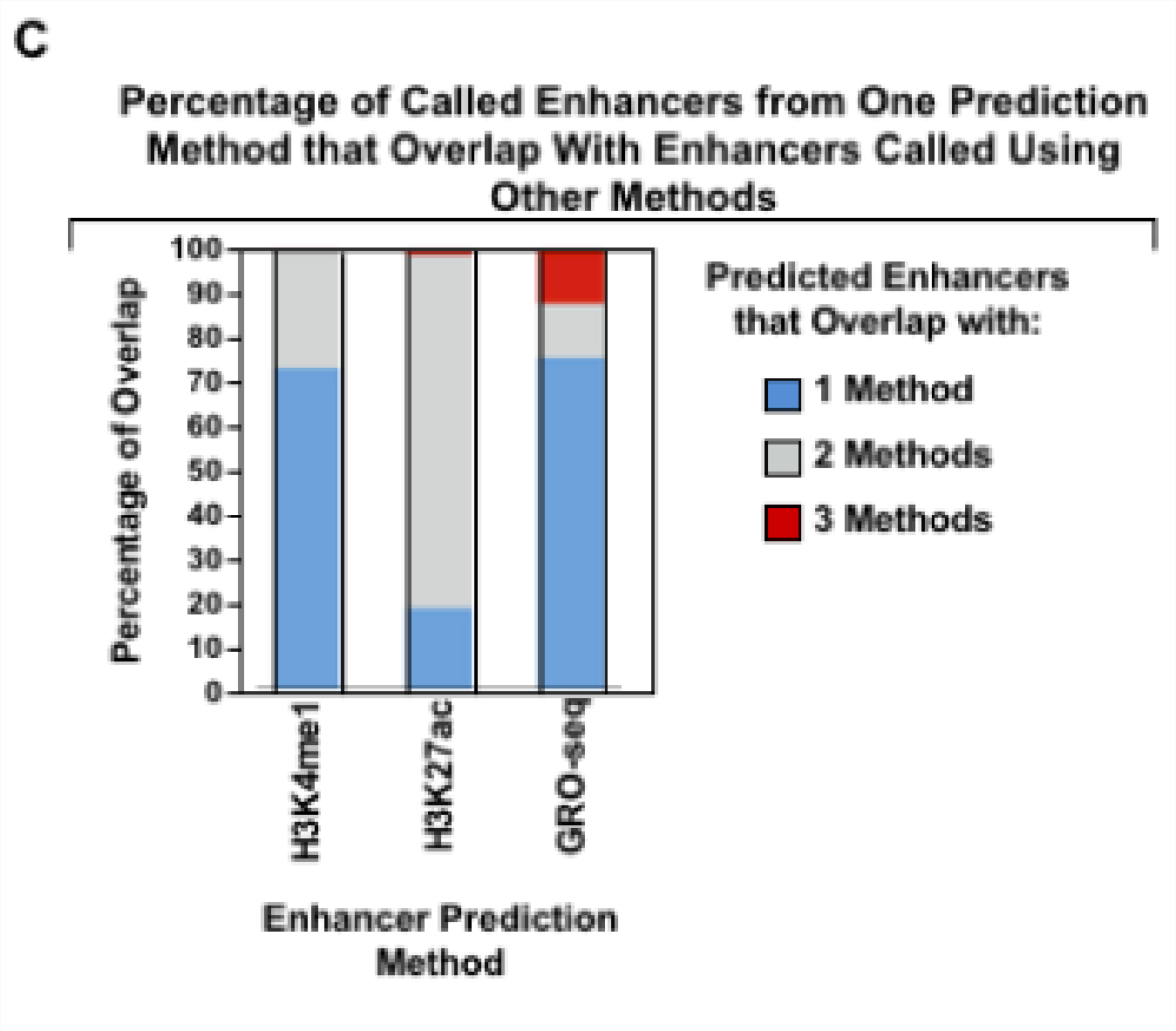

C
Percentage of Called Enhancers from One Prediction Method that Overlap With Enhancers Called Using Other Methods
Percentage of Overlap
100
90
80
70
60
50
40
30
20
10
0
H3K4me1
H3K27ac
GRO-seq
Enhancer Prediction Method
Predicted Enhancers that Overlap with:
1 Method
2 Methods
3 Methods

Fig 4D.

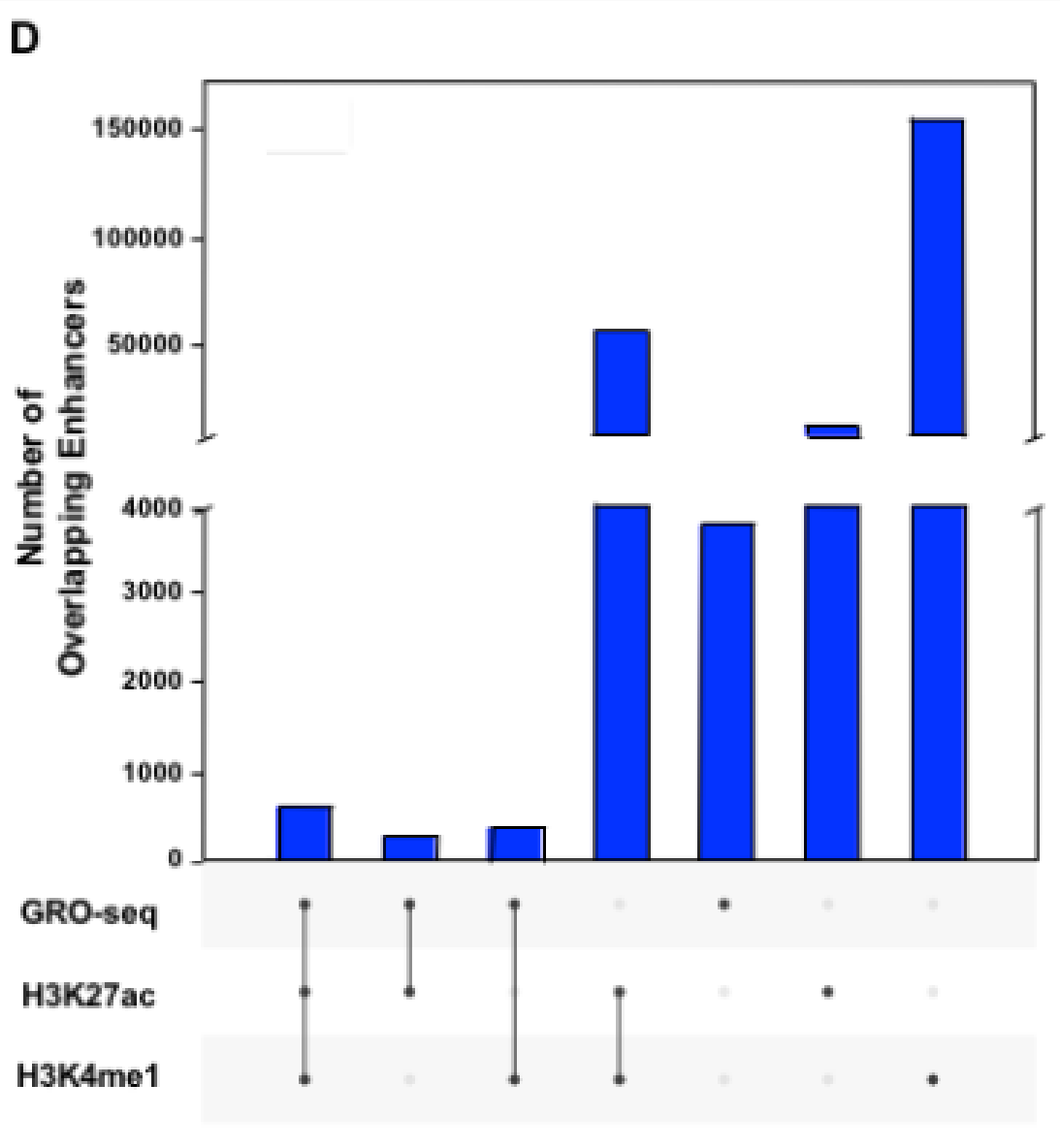

D
150000
100000
50000
4000
3000
2000
1000
0
Number of Overlapping Enhancers
GRO-seq
H3K27ac
H3K4me1

Figure 5. TFSEE identifies cell type–specific enhancers and their cognate TFs that drive gene expression during pancreatic differentiation

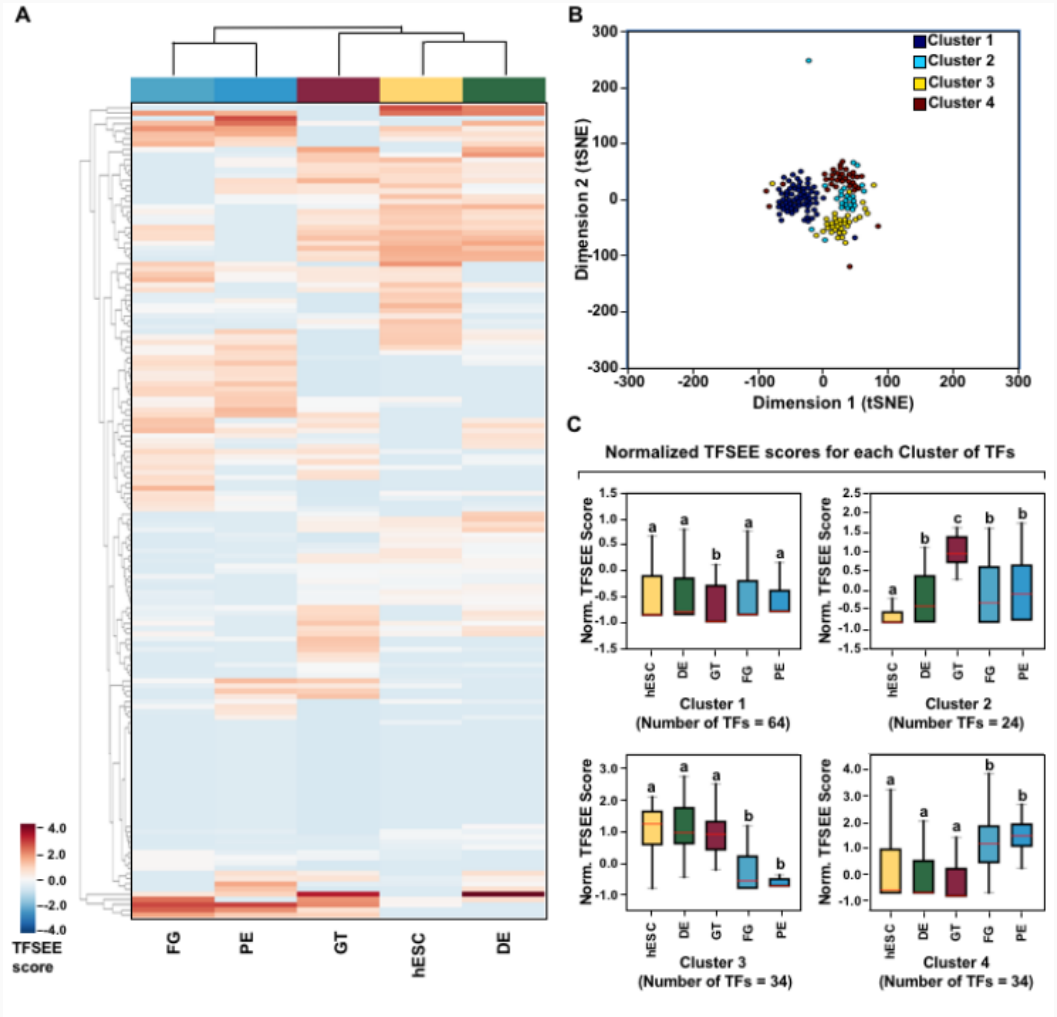
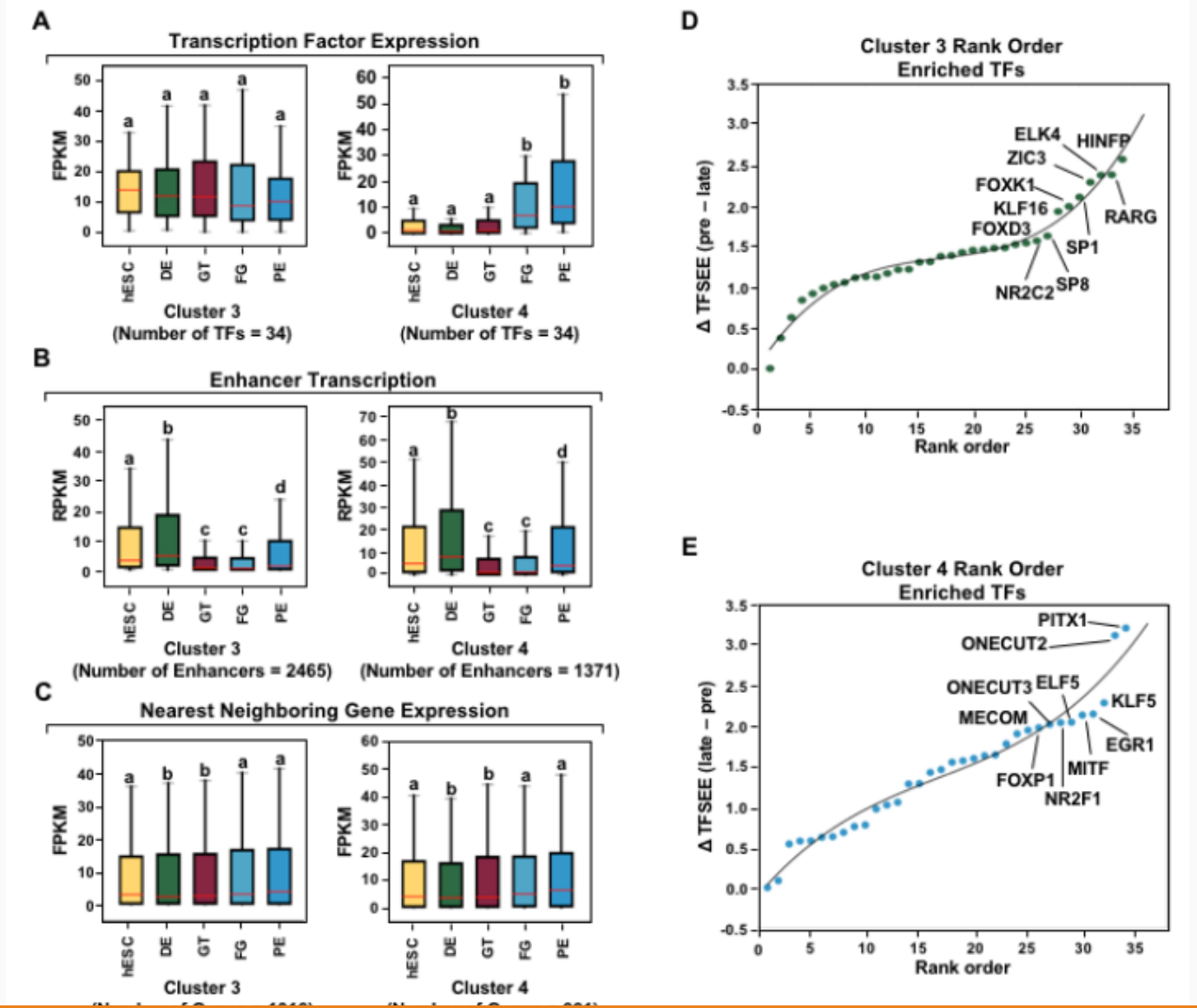


Figure 6. TFSEE-predicted TFs are enriched in pre and late pancreatic differentiation



Strengths and Limitations

- TFSEE enables analysis of driver TFs using a limited amount of data
- The TFSEE model was able to identify lineage-specific TFs with as little as 5 cell types and with only 2 data types, RNA-seq and ChIP-seq (for H3K4me3, H3K4me1, and H3K27ac)
- A limitation of the TFSEE method is that while the model can be used with a reduced number of data types for enhancer identification, it fails to identify additional subtype- or stage-specific drivers with reducde data input

Integrating additional genomic data into TFSEE

- Integrate genomic data indicating open regions of chromatin (ATAC-seq, DNase-seq, or MNase-seq)
- ChromHMM could be used to annotate alternate chromatin states with additional histone modifications
- Chromatin Looping data for enhancer-promoter interactions (as measured by 4C, ChIA-PET, or Hi-C)