RNA-seq and ChIP-seq data integration

(answer to questions)

Stéphanie Le Gras
(slegras@igbmc.fr)

Exercise

- 1.
  - Retrieve Gene symbols of up-regulated genes (use the file siMitfvssiLuc.up.txt you annotated with BioMart)
  - Download the annotated peaks (dataset generated with HOMER). Use the Gene Name column.

Up (RNA-seq) 3791

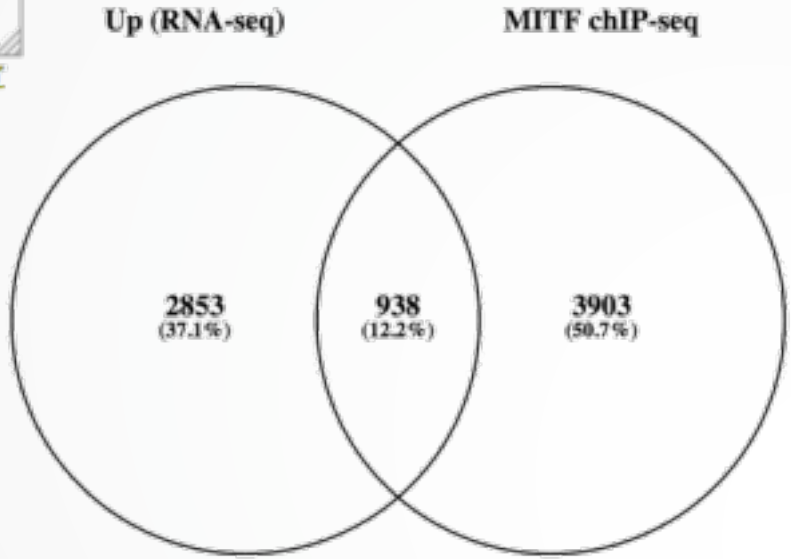
```
COX20P2
MCRS1
HERC4
RP11-39C10.1
CEP85L
```

clear

MITF chIP-seq 4841

```
PBRM1
SIM2
FAM183A
SNRNP200
PPP4R1-AS1
```

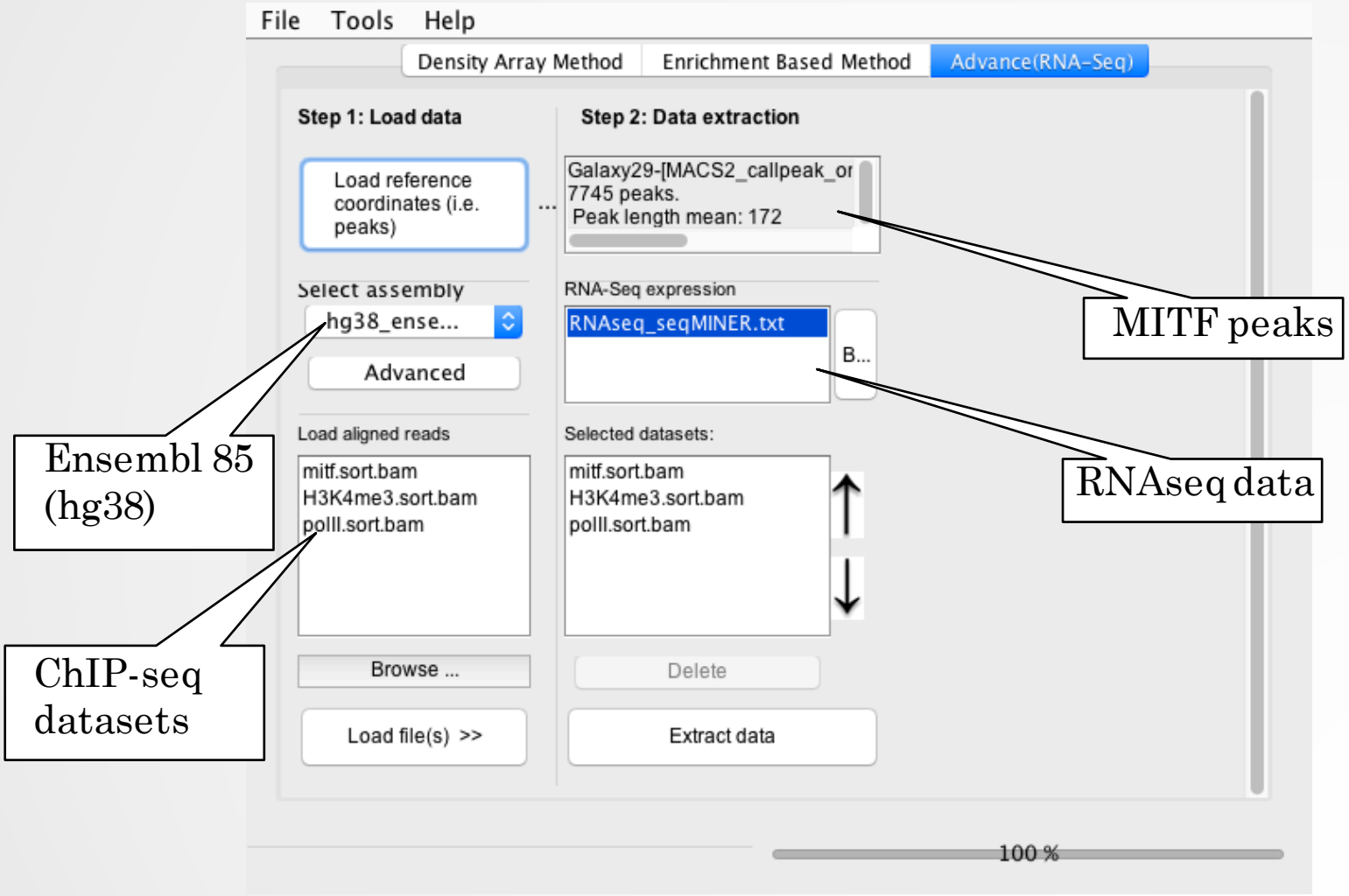
clear



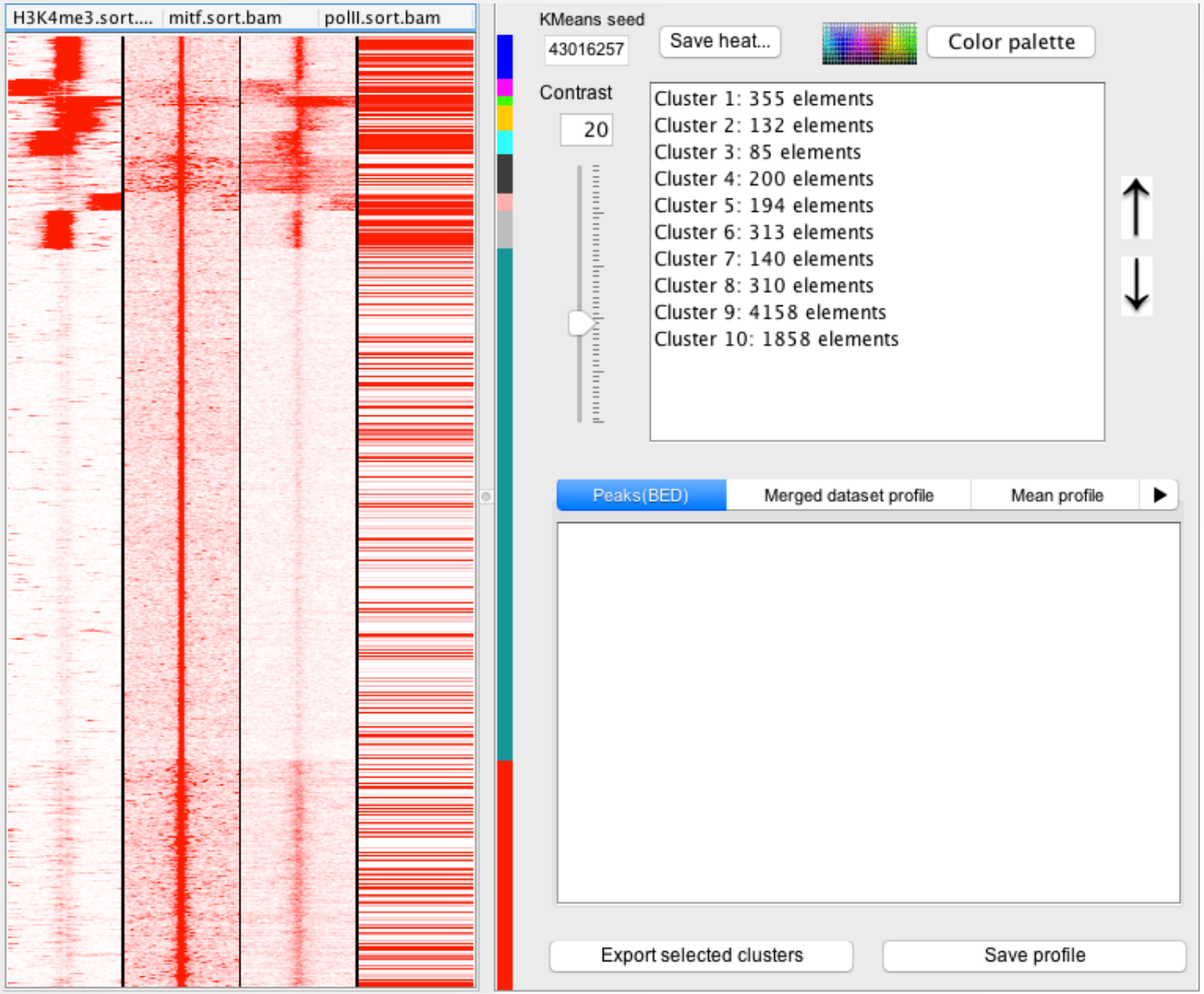
Exercise

- 2.
  - Download MITF peaks (Output of MACS2 narrow peaks) -> Use it as reference coordinates in seqMINER
  - Create the RNAseq file with excel starting from the file (siMitfvssiLuc.up.txt annotated with Ensembl):
    - 1st column : Ensembl Gene IDs
    - 2nd column : normalized siMITF divided by gene length in Kb
    - (save the file as tabulated text file)
  - Go to seqMINER

Exercise


File Tools Help
Density Array Method
Enrichment Based Method
Advance(RNA-Seq)
Step 1: Load data
Step 2: Data extraction
Load reference coordinates (i.e. peaks)
Galaxy29-[MACS2_callpeak_or
7745 peaks.
Peak length mean: 172
Select assembly
hg38_ense...
Advanced
RNA-Seq expression
RNAseq_seqMINER.txt
B...
Load aligned reads
mitf.sort.bam
H3K4me3.sort.bam
polII.sort.bam
Selected datasets:
mitf.sort.bam
H3K4me3.sort.bam
polII.sort.bam
Browse ...
Delete
Load file(s) >>
Extract data
100 %
MITF peaks
RNAseq data
Ensembl 85 (hg38)
ChIP-seq datasets

Exercise


H3K4me3.sort....
mitf.sort.bam
polII.sort.bam
KMeans seed
43016257
Save heat...
Color palette
Contrast
20
Cluster 1: 355 elements
Cluster 2: 132 elements
Cluster 3: 85 elements
Cluster 4: 200 elements
Cluster 5: 194 elements
Cluster 6: 313 elements
Cluster 7: 140 elements
Cluster 8: 310 elements
Cluster 9: 4158 elements
Cluster 10: 1858 elements
Peaks(BED)
Merged dataset profile
Mean profile
Export selected clusters
Save profile