Differential loop analysis of phenotypically and genetically juxtaposed mouse strains using visualization tools

ELEANOR S CARR(1), Zeyuan Chen (2), Jonathan Flint (1), Jeffrey Chiang (4)
(1) BIG Summer Program, UCLA; (2) Dept of Computer Science, UCLA, (3) Dept of Psychiatry and Biobehavioral Sciences, UCLA; (4) Dept of Computational Medicine, UCLA

Introduction

Biological Framework:
- Previous study suggests promoters are key to gene regulation, can be regulated from far away(1)
- Mice strains C57BL/6 (BLK6) and DBA/2J (DBA) have differential gene expression in fear and anxiety genes
- Experiments to isolate these differences, however, are expensive and time-consuming

Computational Framework:
- Hi-C data and single-cell data modalities all describe differential expression in slightly different ways(2)
- Data collected using Hi-C bulk and single-cell sequencing

Methods

Single-Cell Sequencing(4):
- ATAC: Captures chromatin accessibility
- RNA: Captures transcription level
- BLK6 and DBA samples are pooled (n=1, 3 biological replicates, 2 technical replicates) and sequenced together
- De-multiplex/genotyped computationally
- Annotation with Leiden Clustering in low dimensional space
- Aggregated number of mapped reads to pseudobulk (stratified by strain, cell type and normalized by RPKM)

Hi-C(5):
- Captures long term interactions (loops)
- Collected Bulk Hi-C data for two strains separately
- Controlled for sequencing depth (KR normalized at 5kb resolution)
- Corrected for expected contact frequency
- Visualized by WUSTL Epigenome Browser

Results

Findings:
- ATAC, RNA, and Hi-C data demonstrate visually similar peaks
- Visible cell-type differences

Limitations:
- No one modality explains positive hits
- Tool demonstrates limited features for overlapping, multi-colored tracks

Discussion & Conclusions

- Further research is required to draw biologically relevant conclusions
- Accessible regions of a chromosome are distinguishable by all three types of data: RNA transcription, ATACseq, and Hi-C loops
- Promoter regions, are key sections in differential expression
- No distinct long-range connections; higher-resolution loop data may show more

References

(1) Smemo, S., et al. (2014).
(3) Tam, V., et al. (2019).
(4) Liu, N. et al. (2020).