Evaluating the relationships between transcription factor binding and 3D genome structure

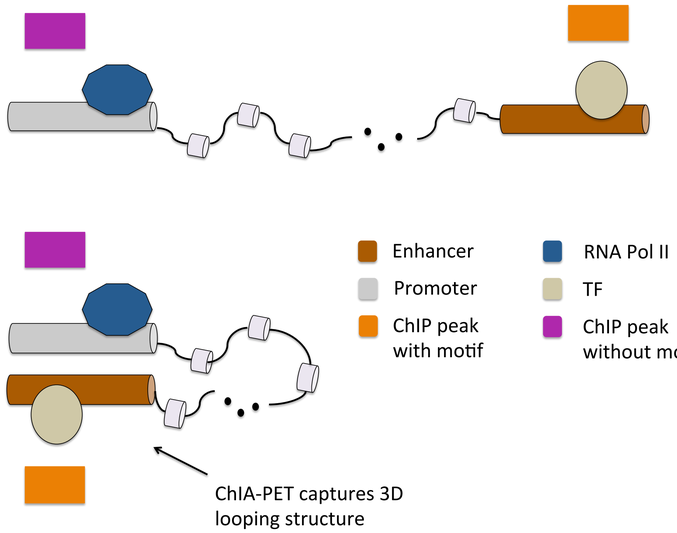
Joseph Sayad[1,2], Vikram Agarwal[3], Jay Shendure[1,3]

[1] Howard Hughes Medical Institute; Chevy Chase, MD, USA
[2] Hunter College of the City University of New York; New York, NY, USA
[3] Department of Genome Sciences, University of Washington; Seattle, WA, USA

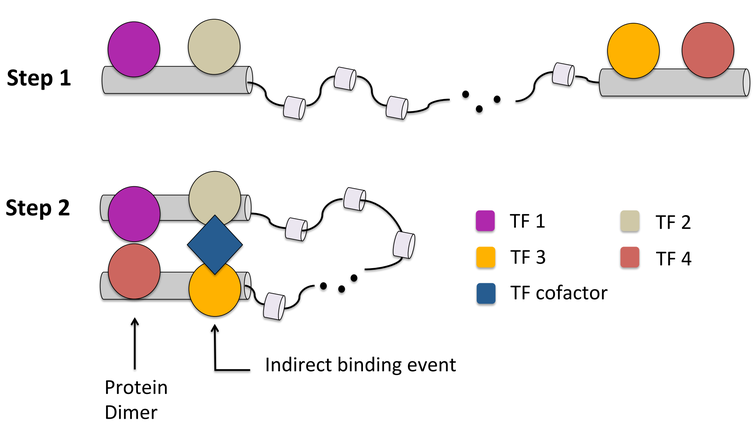
Introduction

ChIA-PET experiments identify regions of chromosomal contact from the perspective of a protein of interest, giving insight into 3D genome structure. ChIP-seq experiments reveal the genomic locations of transcription factor (TF) binding sites. Although TFs recognize DNA motifs (i.e., sequences of DNA that summarize the binding preferences of distinct TFs), many ChIP-seq peaks do not harbor known binding motifs. The goal of this project is two-fold: I) to evaluate whether ChIP-seq peaks without motifs are the consequence of chromatin looping interactions, and therefore hold information about 3D genome structure; and II) to evaluate which specific TF–TF interactions might explain observed chromatin looping.

(I) Does 3D chromatin structure explain the presence of ChIP peaks without a motif?



(II) Do specific TF–TF interactions govern the 3D spatial arrangement of the genome?



Methods

(I) Loop inference pipeline

ChIP peaks | ChIA-PET loops | Motif positions

- Filter datasets for K562 cell line data
- Partition ChIP-seq peaks in those with and without a known Factorbook motif
- Infer chromatin loops based on peaks and create consensus track
- Generate background model by permuting peak–motif labels

Visualize candidate loci with R Gviz module

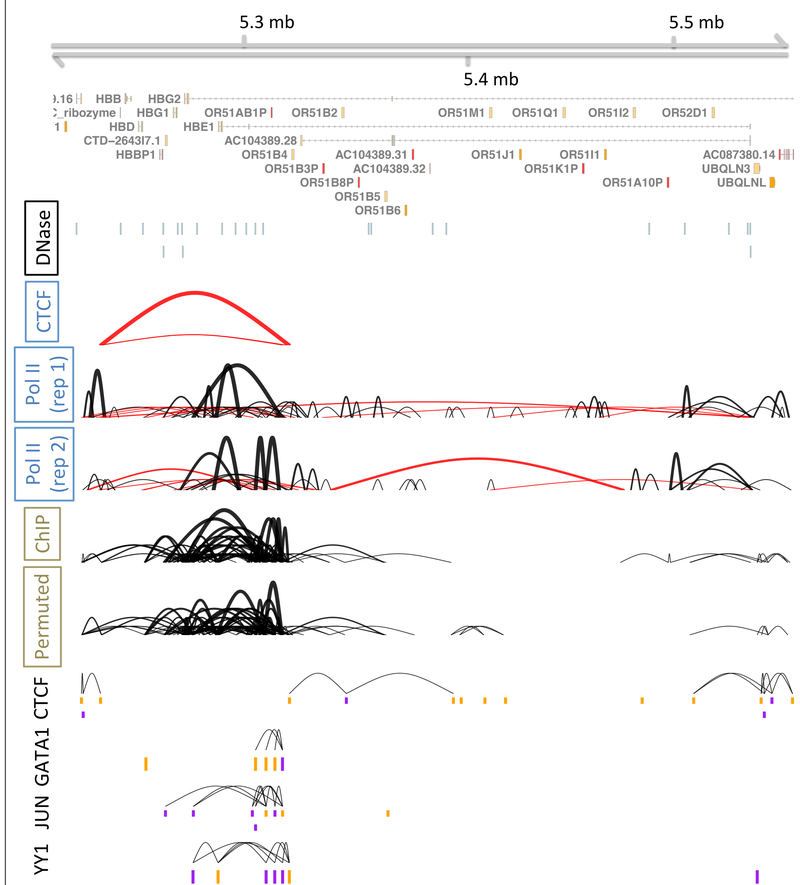
(II) TF–TF interaction pipeline

ChIP peaks | ChIA-PET loops

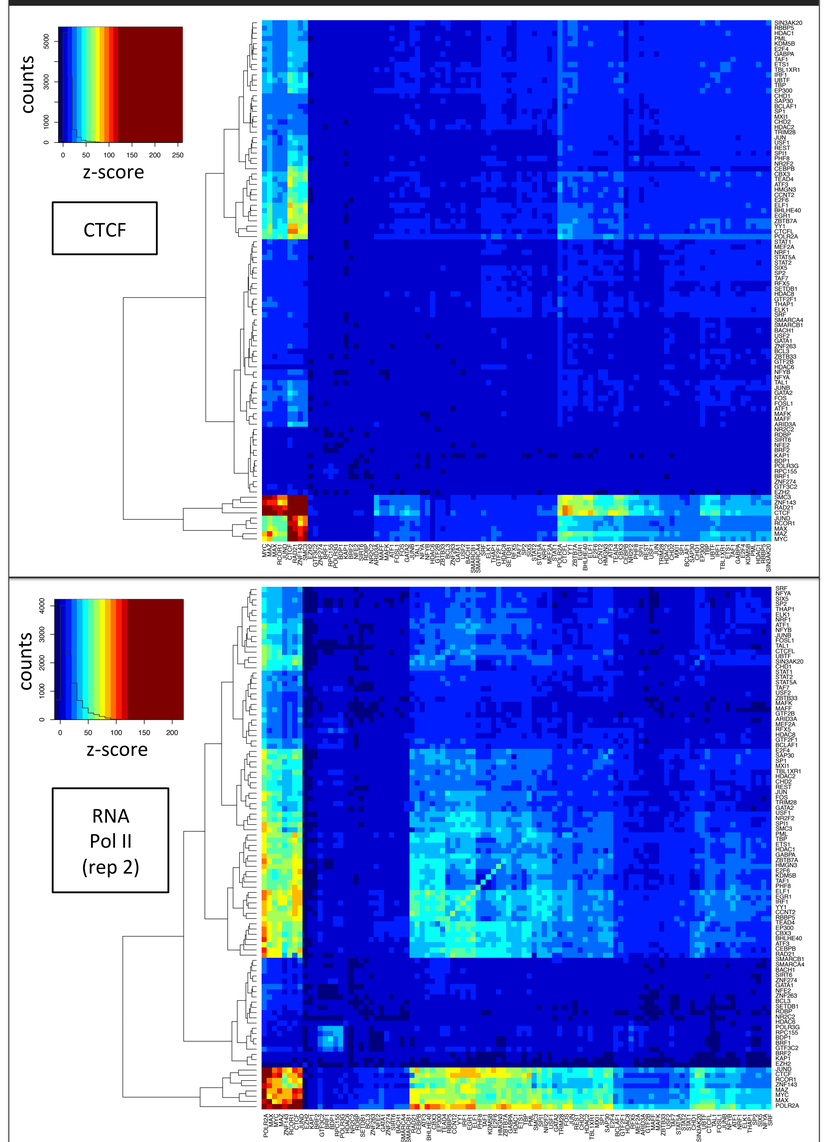
- Intersect ChIP peaks with loops
- Record TF–TF interaction frequencies
- Generate 100 background datasets by permuting TF–loop labels
- Compute z-scores for TF–TF interaction frequencies relative to background distributions

Display heatmap of clustered TF–TF interaction z-scores

Loop inferences in the β-globin locus



Enriched TF–TF interactions in ChIA-PET loops



Conclusions

(I) Loop inference plots were generated in parallel for three selected genomic loci (β-globin, MYC, and GATA1) and seven randomly selected loci. The permuted consensus tracks largely recapitulated the ChIP consensus tracks, indicating that the ChIA-PET loops cannot explain the arrangement of ChIP peaks with and without motifs moreso than the inherent spatial biases of the peaks.

(II) CTCF ChIA-PET data was used as a positive control. The analysis uncovered known TF–TF interactions, with strong enrichments between CTCF and cohesion proteins. Extension of this analysis to RNA Pol II ChIA-PET data discovered interactions between Pol II and a number of factors, especially MYC and MYC-associated factors.

Acknowledgments

This project is sponsored by HHMI as a part of the Exceptional Research Opportunities Program (EXROP). I want to thank V. Agarwal and J. Shendure for their mentorship, and members of the Shendure lab for providing additional insights.