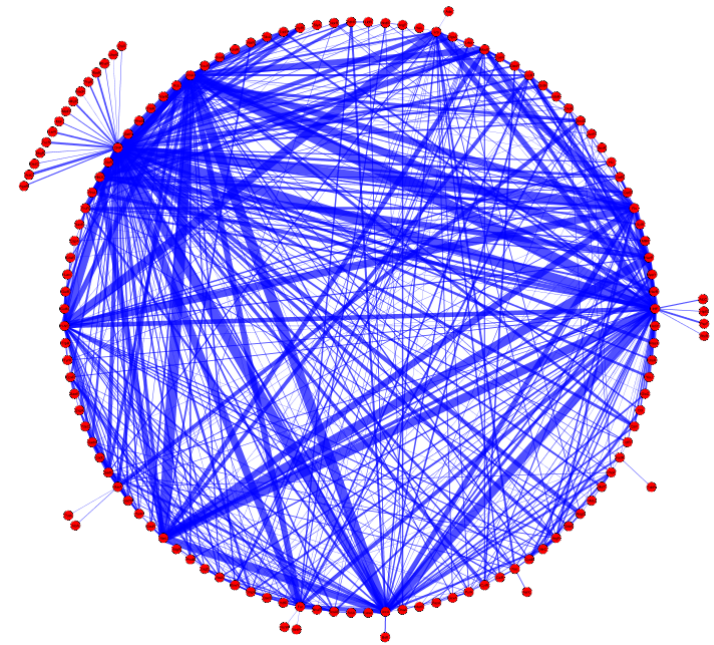# Materials for Integrative and Networks Papers

(With Nitin, Ben Brown, Nathan Boley and Peter Bickel)

Overview

- TF co-association
  - Global view
  - Clusters
  - Region specificity
  - Cell-line specificity
  - Properties of co-associating TF
- Networks
  - Co-association vs. TF regulation hierarchy
  - Regulation at non-promoter regions

Weighted co-association network

TF co-association

- Global view
  - Base overlap matrix (supp. figure, ready)
  - Unsegmented block sampling z-scores (supp. figure, ready)
  - DNase-segmented block sampling z-scores (figure, ready)
  - Gaussianity tests – skewness, robust kurtosis, "optimal" segmentation (supp. figure)
  - Weighted network of co-associating TFs (figure)
  - List of co-associating TFs (supp. table)

TF co-association

- Clusters
  - Global two-way clustering (figure)
  - Bi-clusters (supp. table)
- Region specificity
  - Differences of matrices and clusters in whole-genome, promoters and distal regions (figure)
- Cell-line specificity
  - Differences of matrices and clusters in different cell lines (figure)

TF co-association

- Properties of co-associating TFs (big matrix) (figure)
  - Tissue specificity
  - PPI degree
  - Number of kinase interaction
  - Number of miRNA interaction
  - Number of targets in regulatory network
  - Number of TF targets in regulatory network
  - Correlation between binding and gene expression

Networks

- Co-association vs. regulation hierarchy (figure)
  - Hierarchy level vs. co-association

Networks

- Regulation at non-promoter regions (figure)
  - Method: For each of five selected cell lines,
    - Obtain TF binding peaks…
      - Not in blacklist regions
      - In binding active regions
      - Not in predicted promoters
      - Not intersecting exons
    - Identify target genes of regulatory regions
      - Correlating gene expression with specific features (H3K4me1, H3K27me3, H3K9me2/3) across all cell lines with data
      - Distance
      - Blocking CTCF (?)
      - Long-range interaction (?)
    - For each (regulatory region, target gene) pair, find cell lines with active regulation (by feature values)
    - Associate TFs with peaks at those regions with the regulations

Materials for Elements Tracks Companion Paper

(With Chao, Joel, Nitin, Anshul Kundaje, Ben Brown and Peter Bickel)

Outline

- Definitions/production methods of tracks
  - Promoters
  - Enhancers
  - Silencers
  - High TF occupancy (HOT) regions
  - Low TF occupancy (LOT) regions
- Overlap with related tracks
- Experimental validation results (enhancers and silencers)

Outline

- Feature comparisons
  - TF binding, with vs. without motifs, co-association
  - Open chromatin
  - Histone modifications
  - Gene expression
- Cell-line specificity