ENCODE enhancers

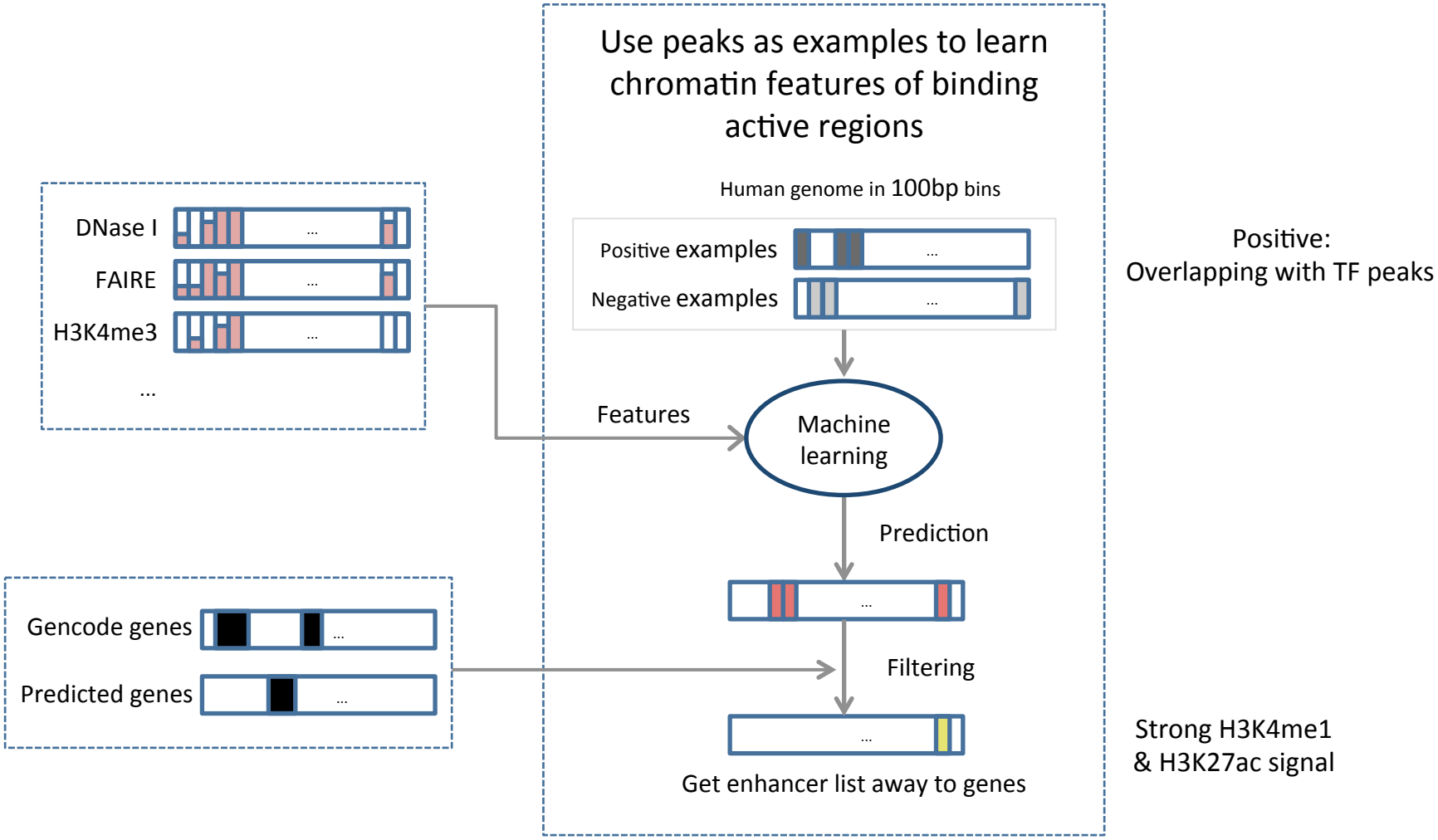
12/13/2013

Yao Fu

Gerstein lab

‘Supervised’ enhancer prediction

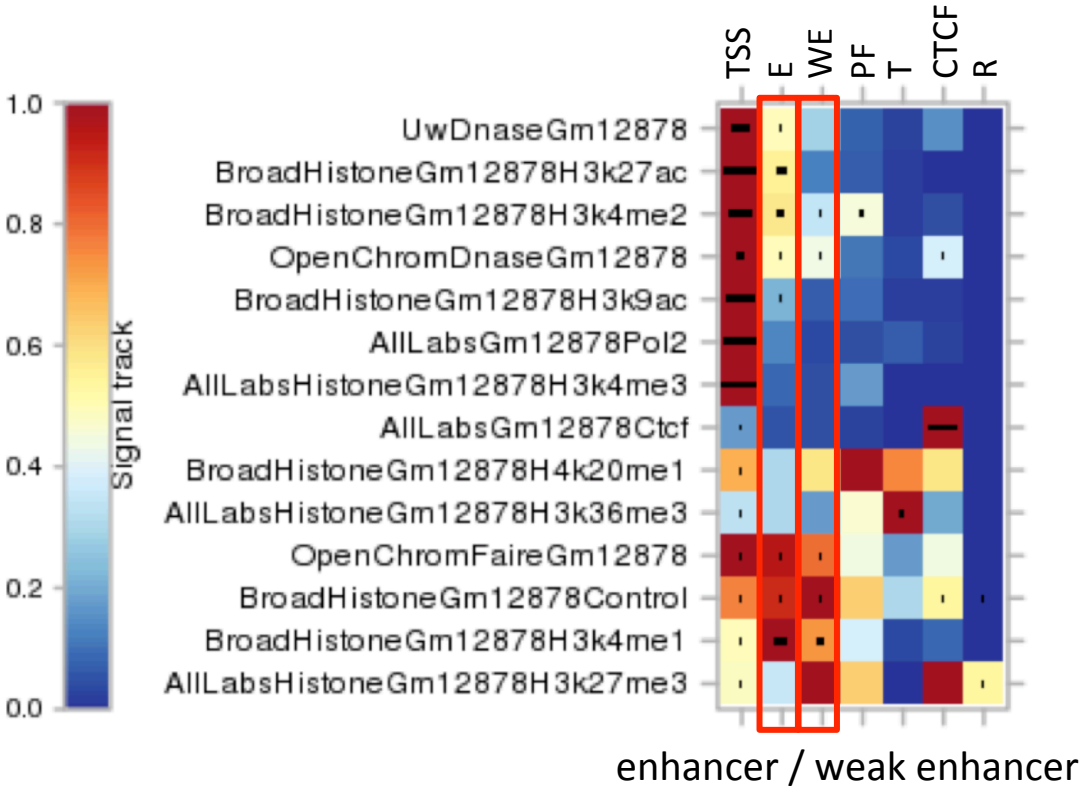
- Identifying Potential Enhancer-like Elements from Discriminative Model



Yip et al., *Genome Biology (2012)*

“Unsupervised” Segway/chromHMM

- Enhancer “states” from unsupervised segmentations (Hoffman et al. & Ernst et al.)



ENCODE combined segmentation from Segway and chromHMM

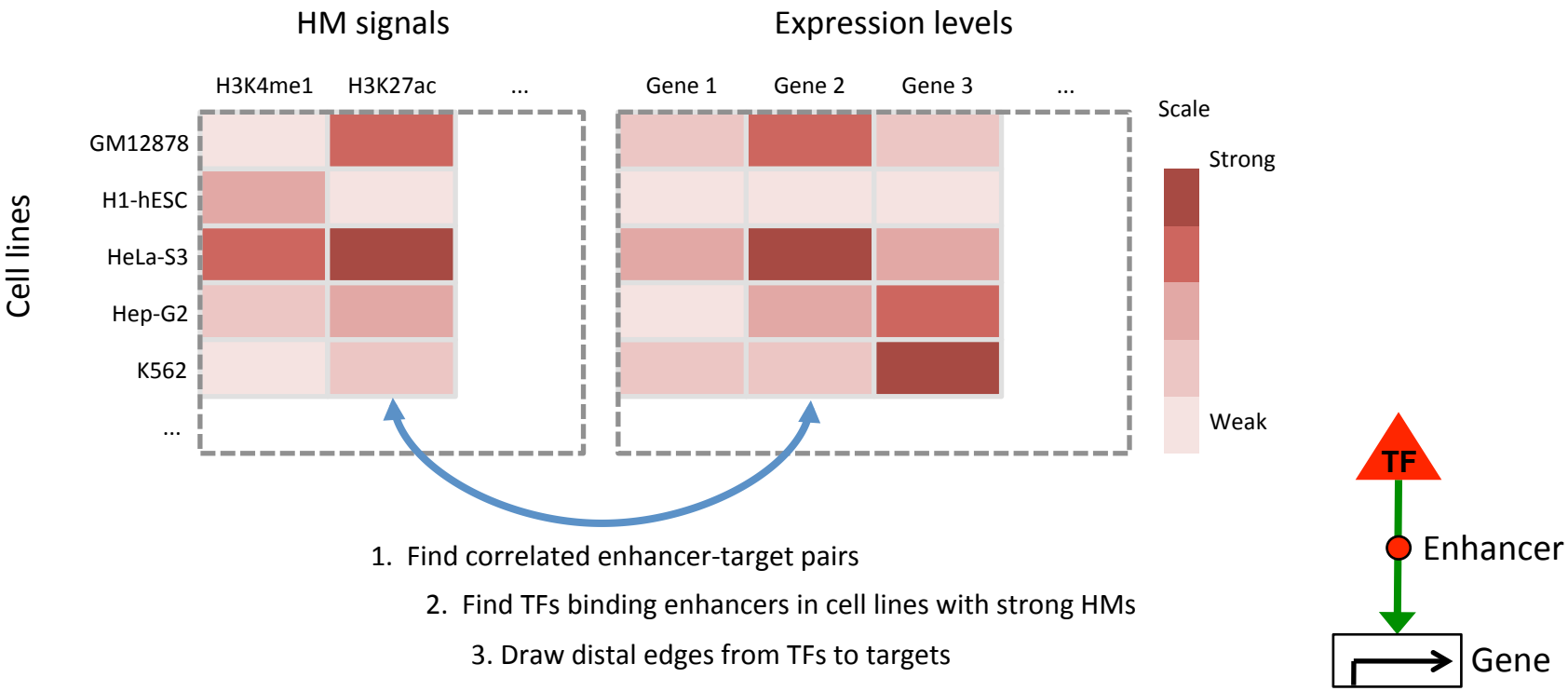
Combine with Segway/chromHMM

- ~130K enhancer-like elements from Yip et al.
- ~ 291K “enhancer state” elements from segmentation.
- Intersection : ~71K

- http://info.gersteinlab.org/Encode-enhancers

Associate enhancers with target genes

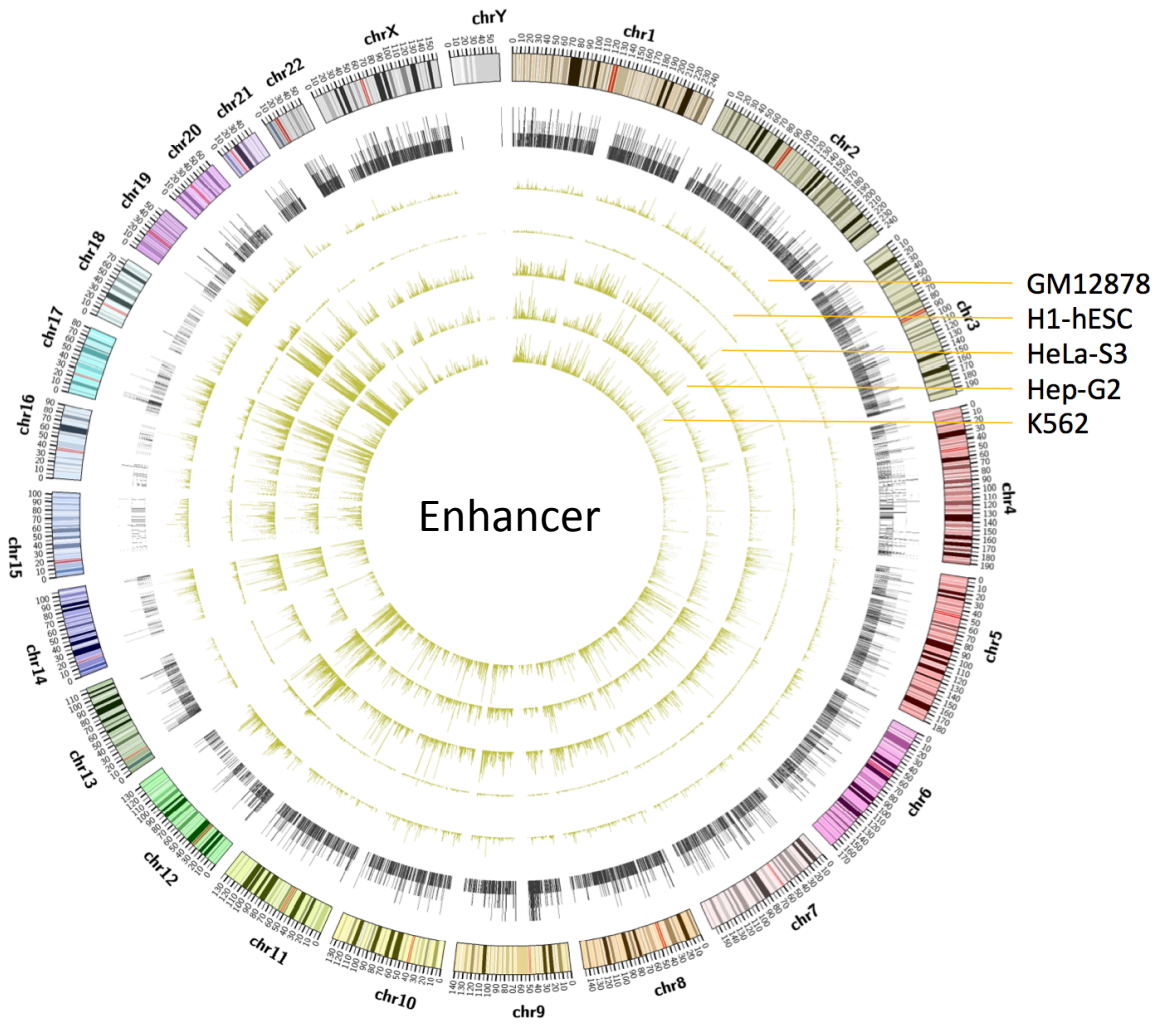
- Idea: Histone modifications to predict gene expression.



- Another direction is to use whole-genome DNA long-range interaction data
- Form distal regulatory networks (~20k distal edges in ENCODE rollout; we extend to edges with ~17k genes)

Yip et al., *Genome Biology (2012)*

Cell-line specific enhancers



ENCODE Work Products (beyond standardized DCC pipelines)

Examples of element lists:

- Enhancers
- DNAse HS sites
  - broad vs cell-type tissue specific sites
- TF targets
  - proximal and distal
  - regulatory networks
- Allelic genes (ASE) and TF binding sites (ASB)
- Fusion (chimeric) transcripts
- Non-coding transcription (contigs or transcripts)
  - classification (e.g. eRNAs)
- High-occupancy (HOT) regions
- Regions of active chromatin
- Chromatin states (e.g. segmentation)
- TF motifs (PWMs & sites)