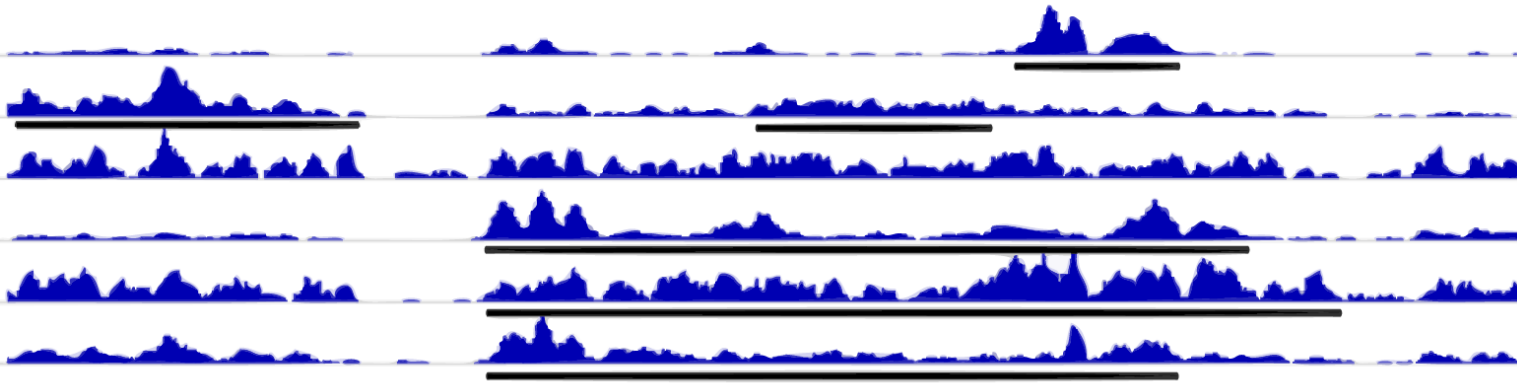


Multiscale Enrichment Calling with MUSIC

MUSIC



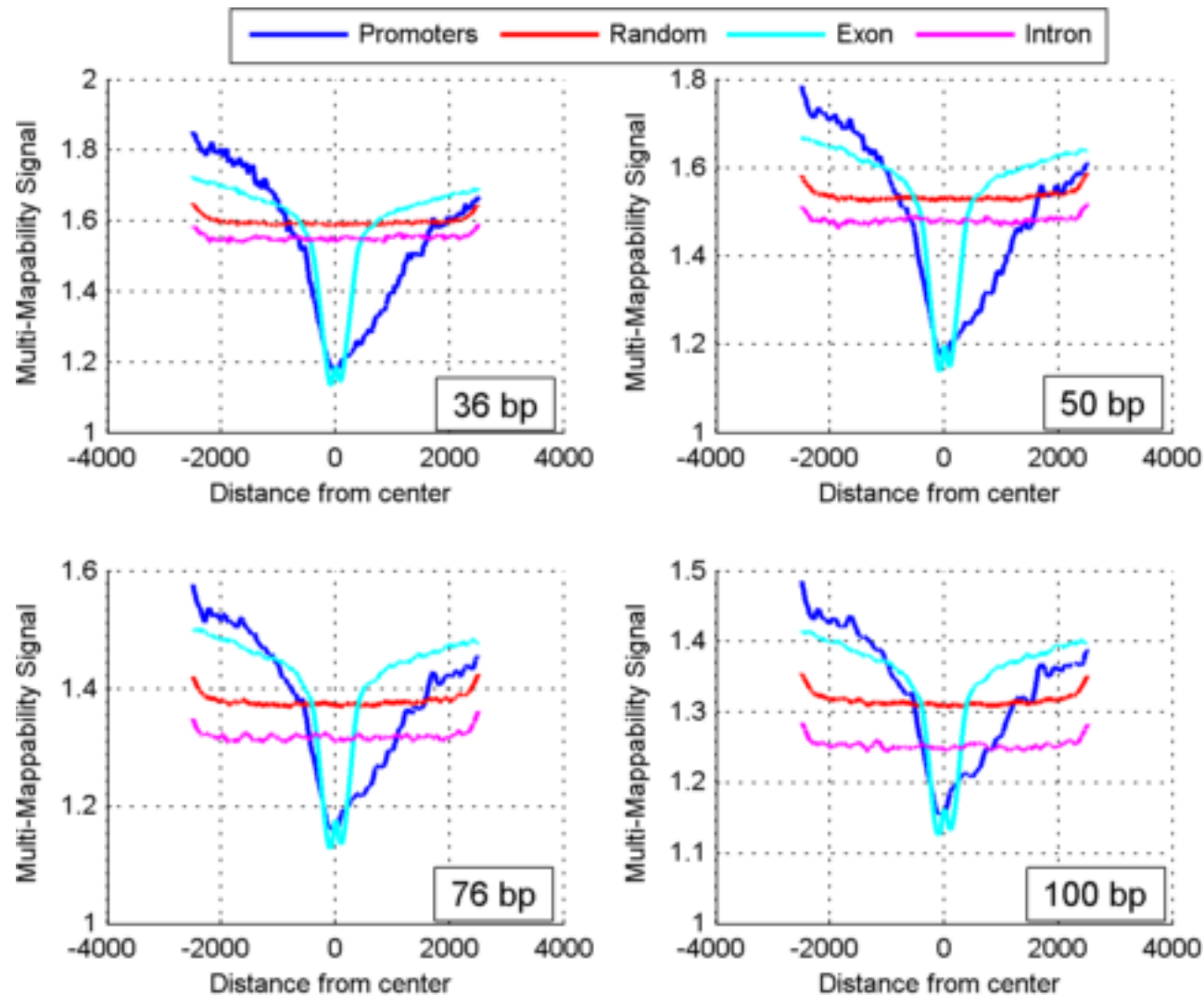
MULTiScale Enrichment Calling

Arif Harmanci
Yale University

Motivation

- Broad enriched regions (ERs) are more tricky to process than punctate ERs
 - Mappability is not uniform over the genome
 - A large spectrum of length scales for the enriched regions in different experimental assays

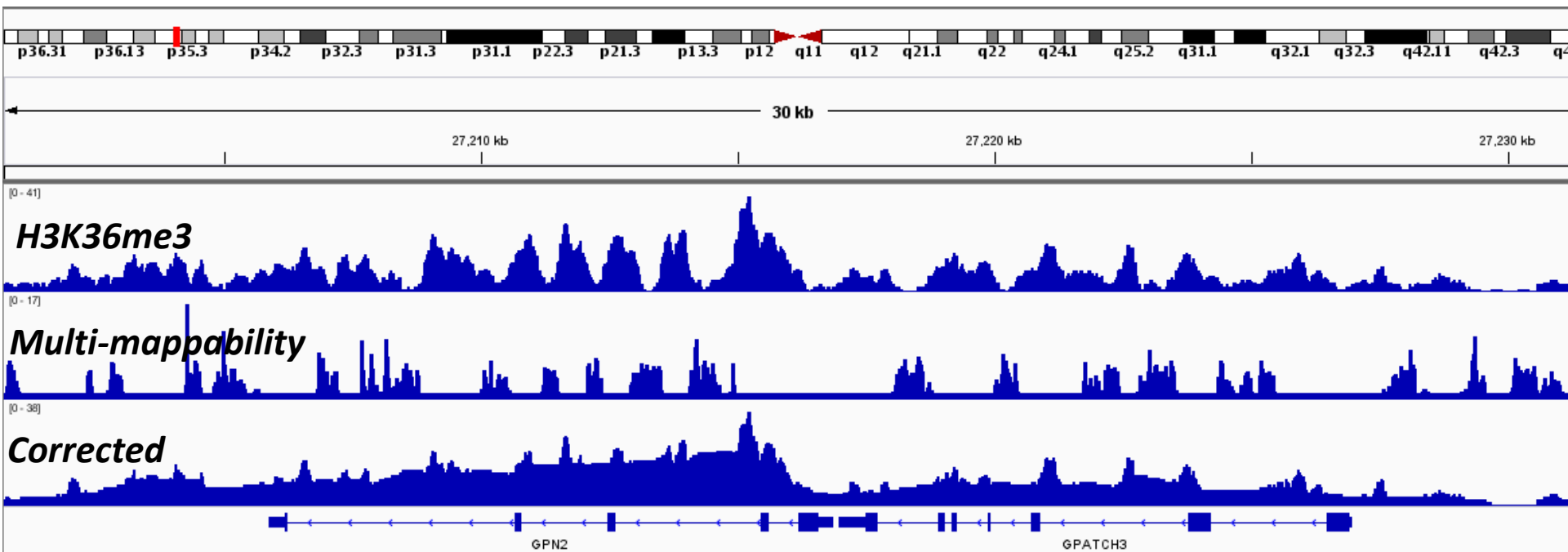
Multi-mappability Profile



- “Number of reads mapping at a position from uniformly sampled genome”

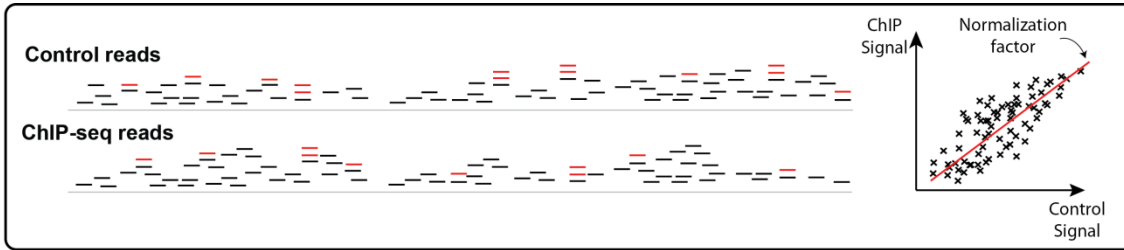
Multi-Mappability Correction

- At each position, use the surrounding highly mappable positions to update and smooth the signal

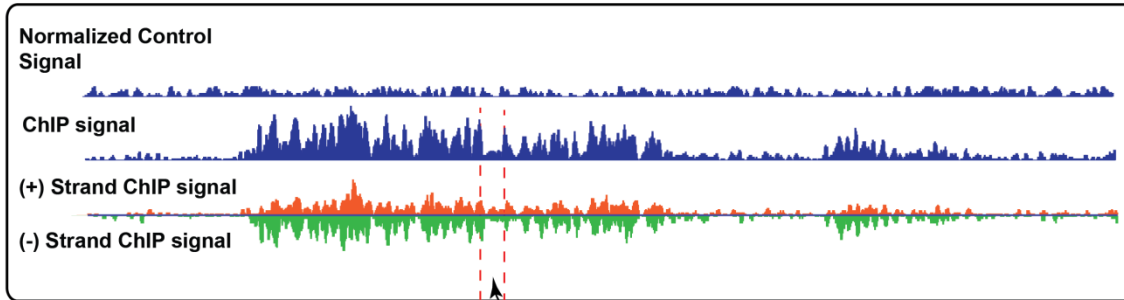


MUSIC Algorithm

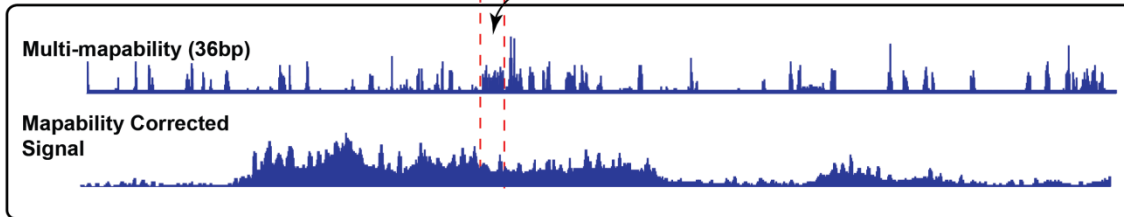
1 Remove PCR duplicate reads, Normalize Control Profile



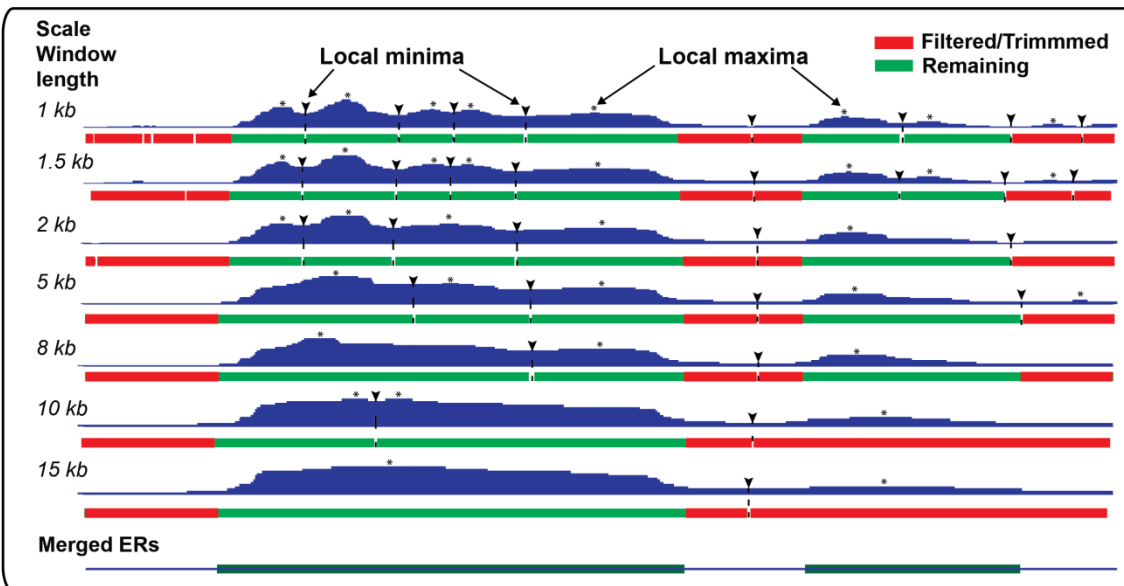
2 Preprocess reads, Generate Signal Profiles, Normalize Control Profile



3 Mapability Correction



4 Multiscale Decomposition and Enrichment Feature Identification

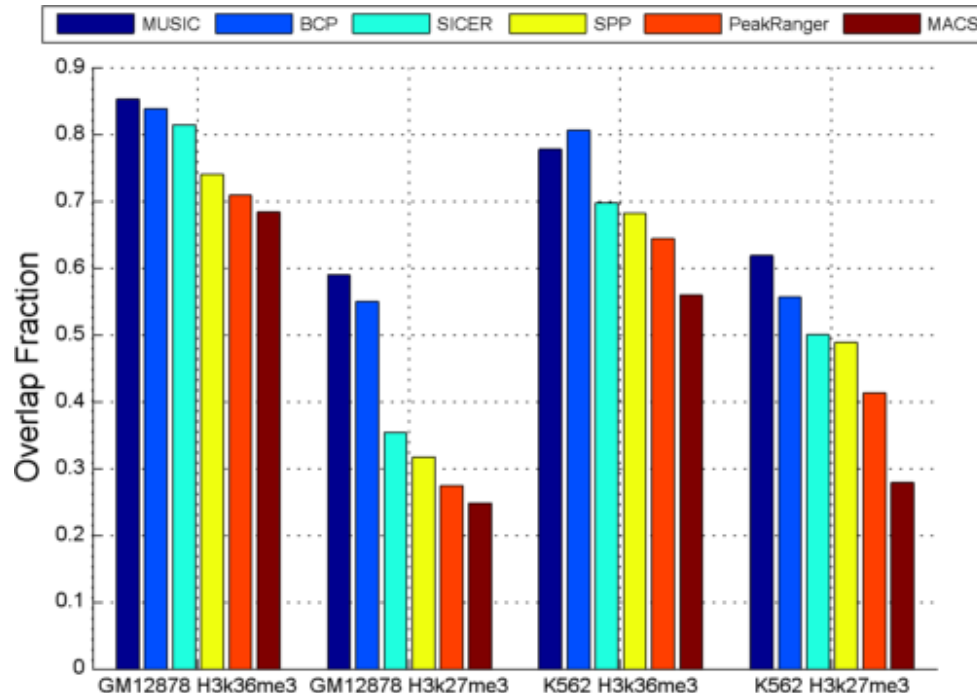
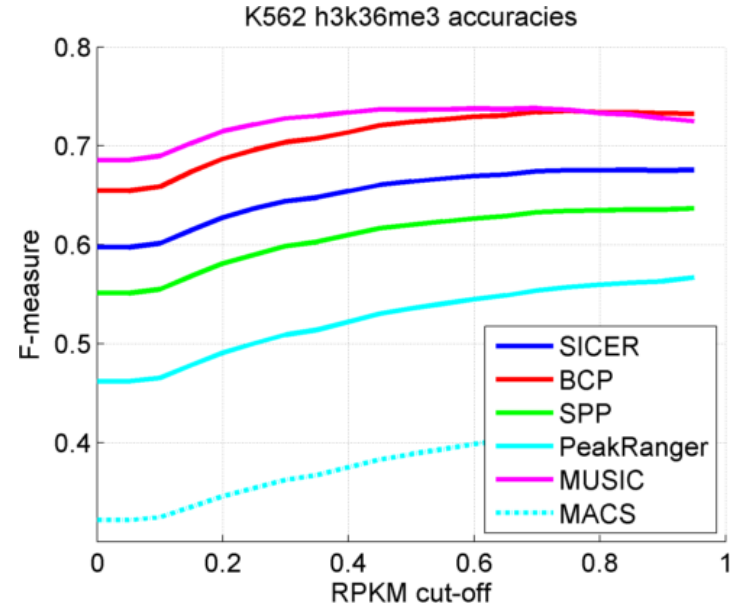
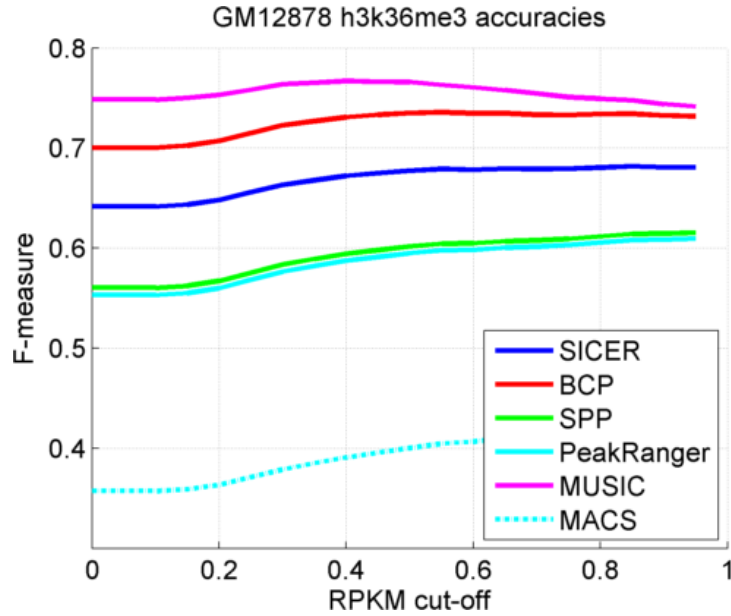


5 Significant Enrichment Feature Filtering

6 Feature Merging P-value assignment

Binomial p-value computation
Poisson thresholding
Per strand ChIP Signal

Comparison with Other Methods



Feature Pileup Signal: Another way to look at multiscale features

- Rather than merging the enrichment features, pileup the features and count the number of features at each nucleotide
- Small scale pileup signal: Punctate features
- Large scale pileup signal: Broad features

