LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Pranzatelli *et al. BMC Genomics* (2018) 19:563
https://doi.org/10.1186/s12864-018-4943-z

BMC Genomics

METHODOLOGY ARTICLE Open Access

# ATAC2GRN: optimized ATAC-seq and DNase1-seq pipelines ~~for rapid and accurate genome regulatory network inference~~

Thomas J. F. Pranzatelli, Drew G. Michael and John A. Chiorini*

08.11.19, Journal Club, Evi Berchtold

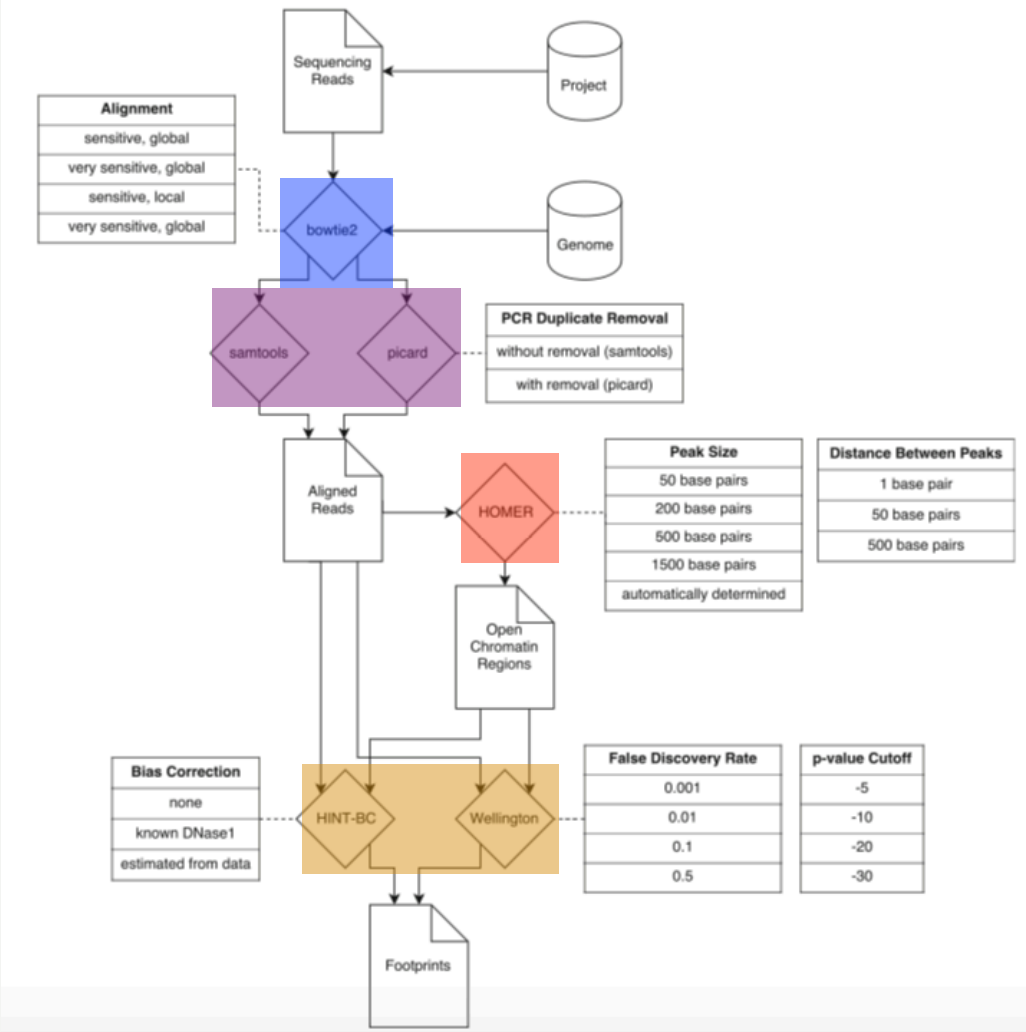
# What it really is

Pipeline for ATAC/DNase-seq data

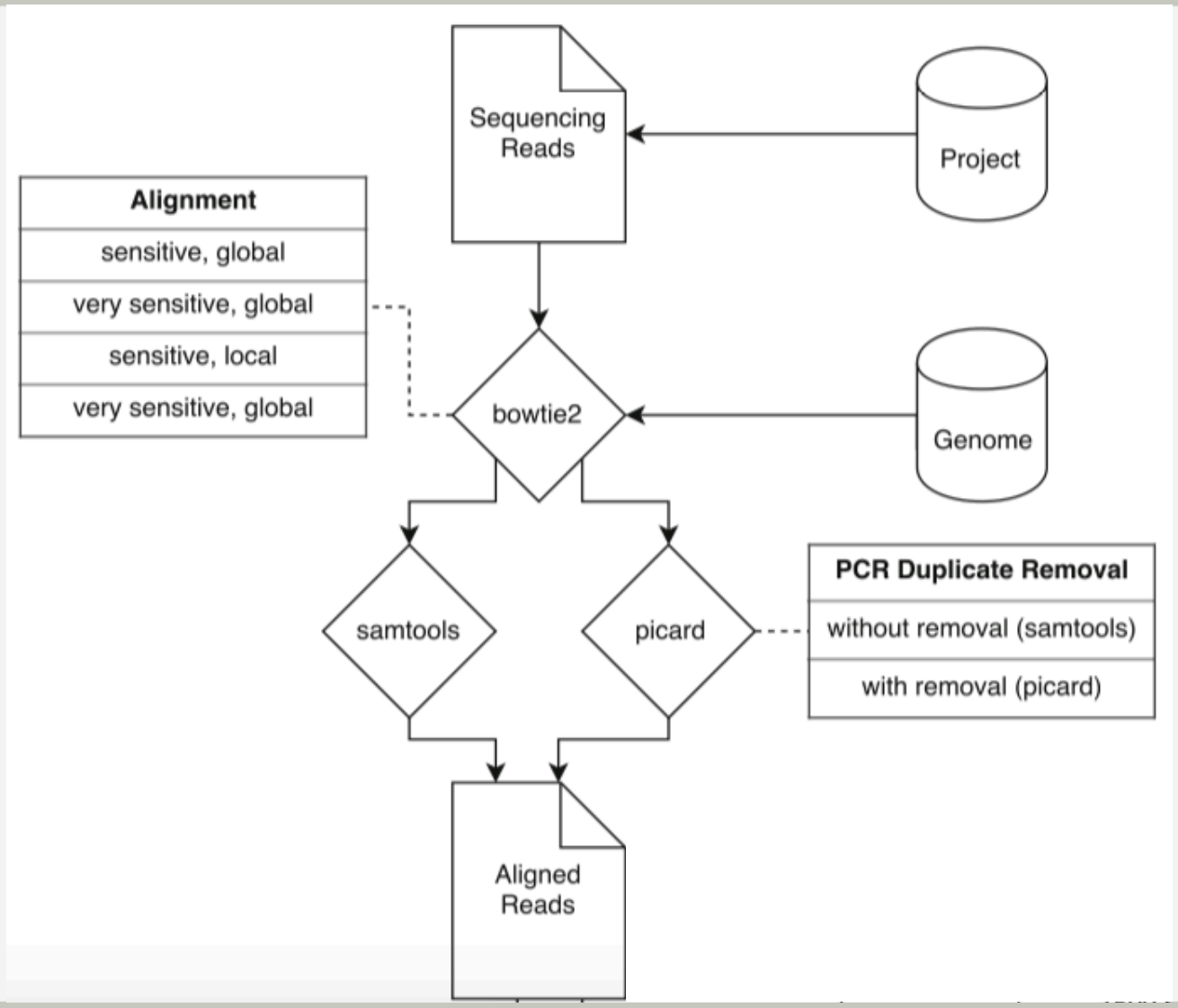
- mapping
- PCR duplicate removal
- peak calling
- footprint calling

parameter search to optimize

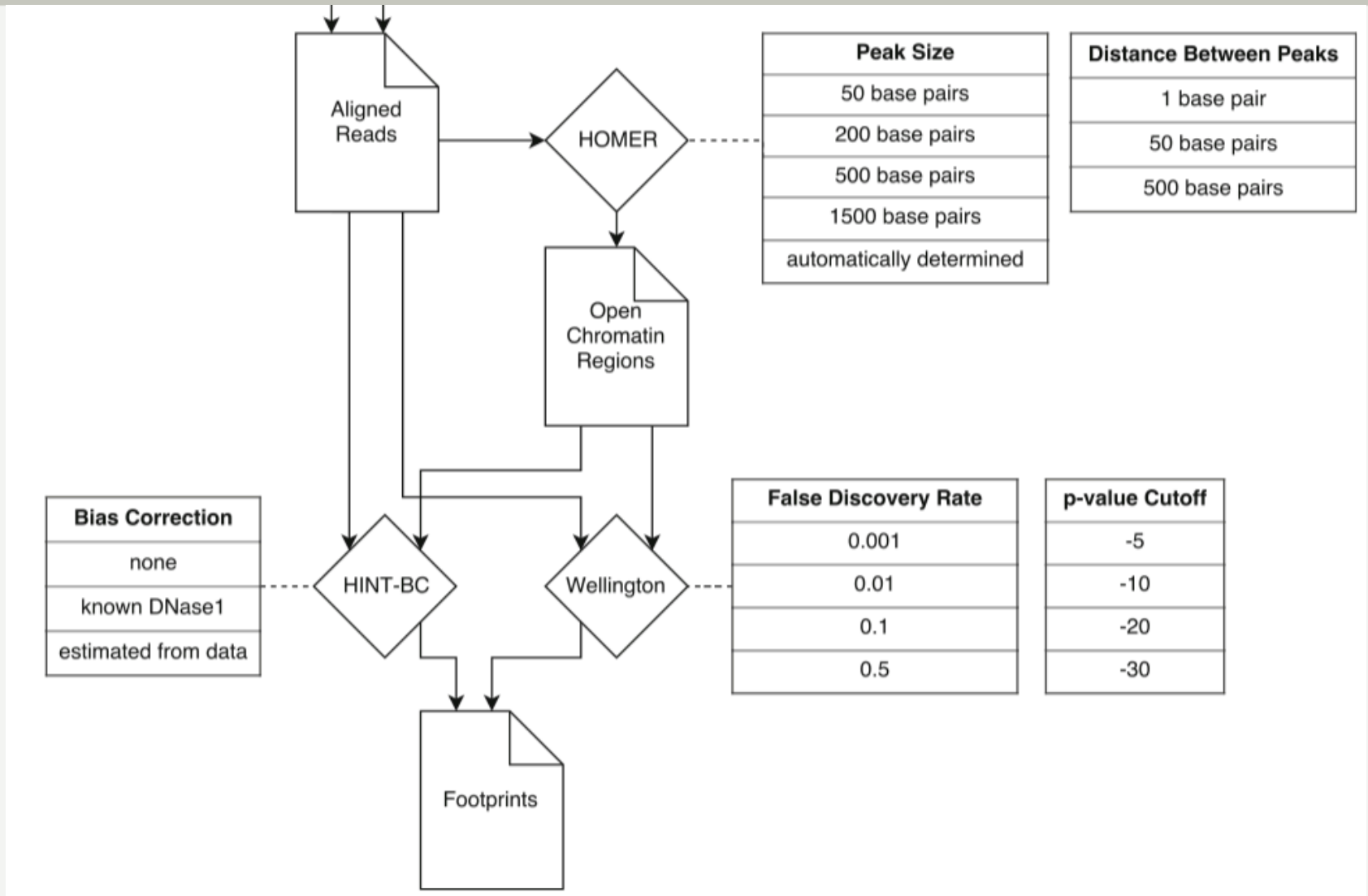
- replicate consistency
- ChIPseq peak overlap



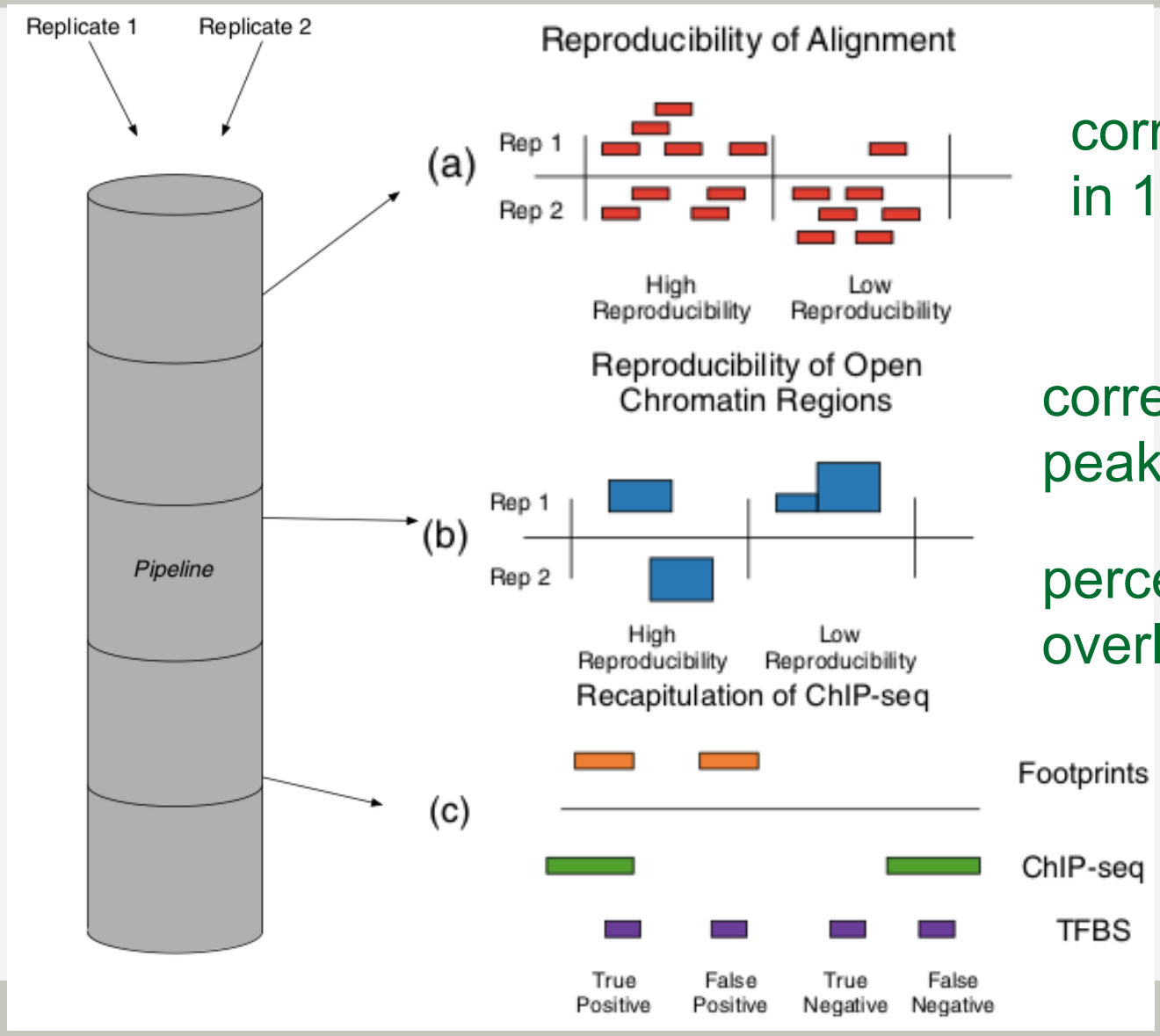
# Mapping + PCR duplicates



# Peak + FP calling



# Validation metrics




correlation bt #reads in 10kb bin

correlation bt avrg peak in 10kb bin

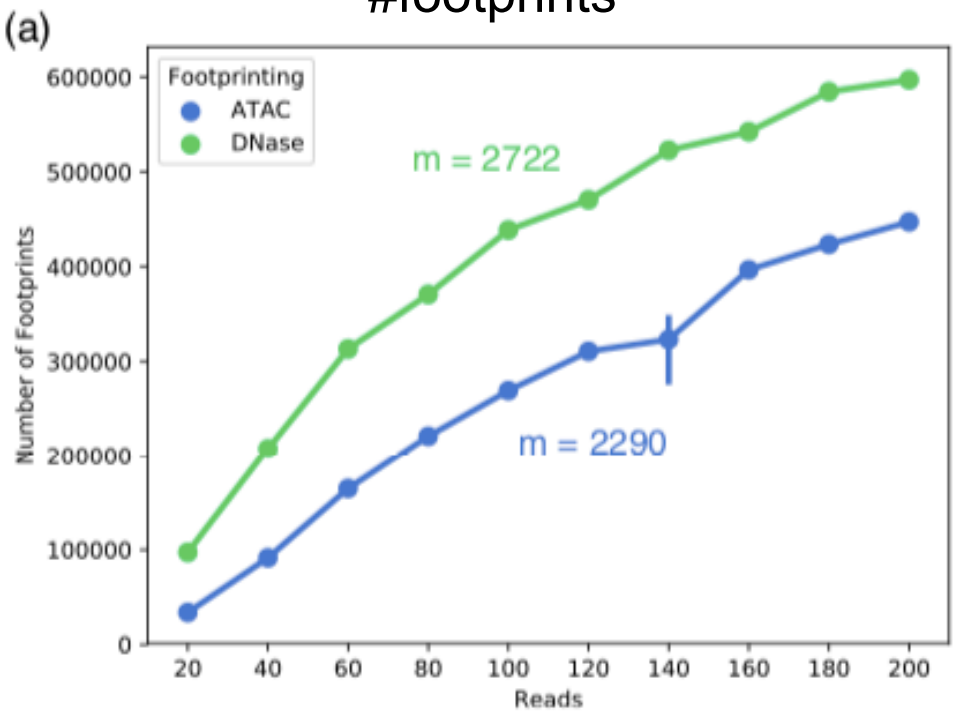
percentage of overlapping footprints

AUC

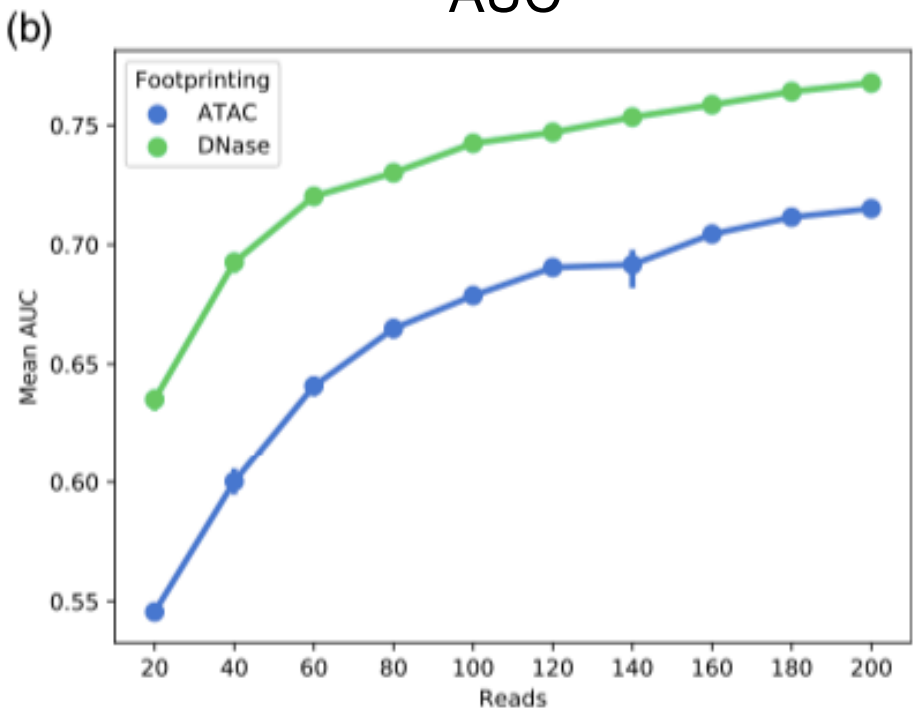
# Sequencing depth

## subsample #reads using seqtk

#footprints




AUC

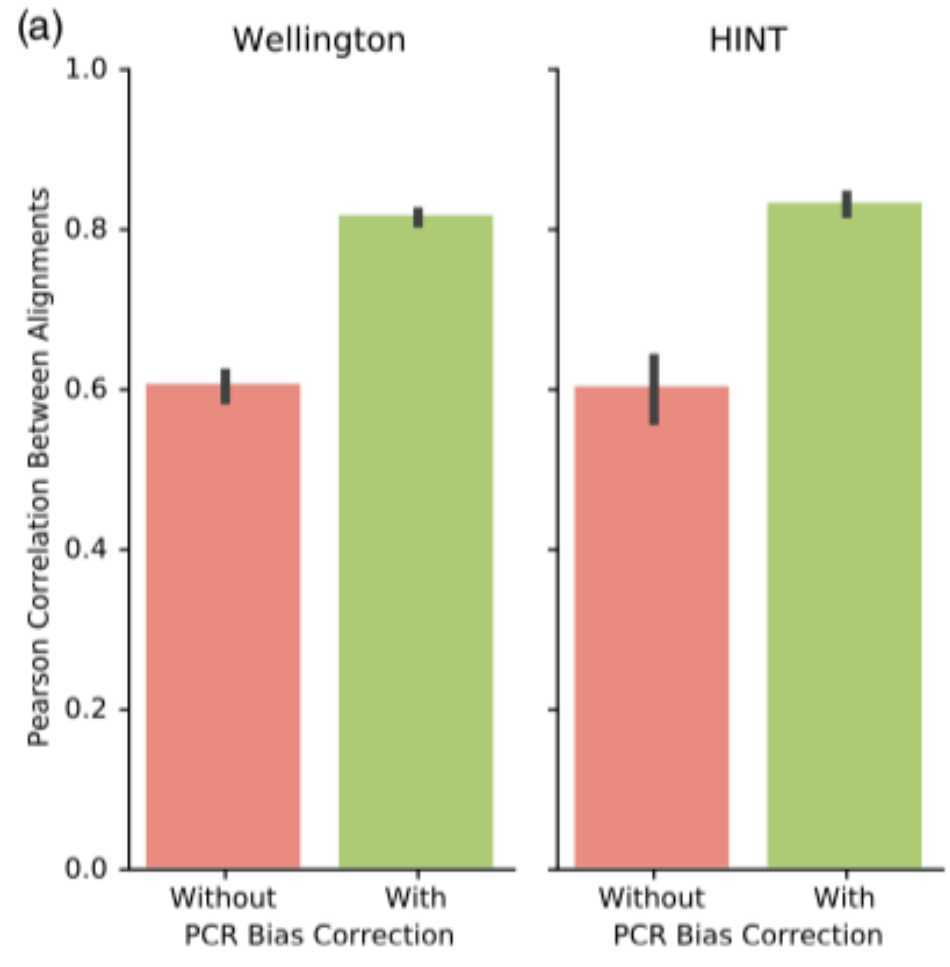



ATAC-seq: 2.85 billion reads for AUC=0.95

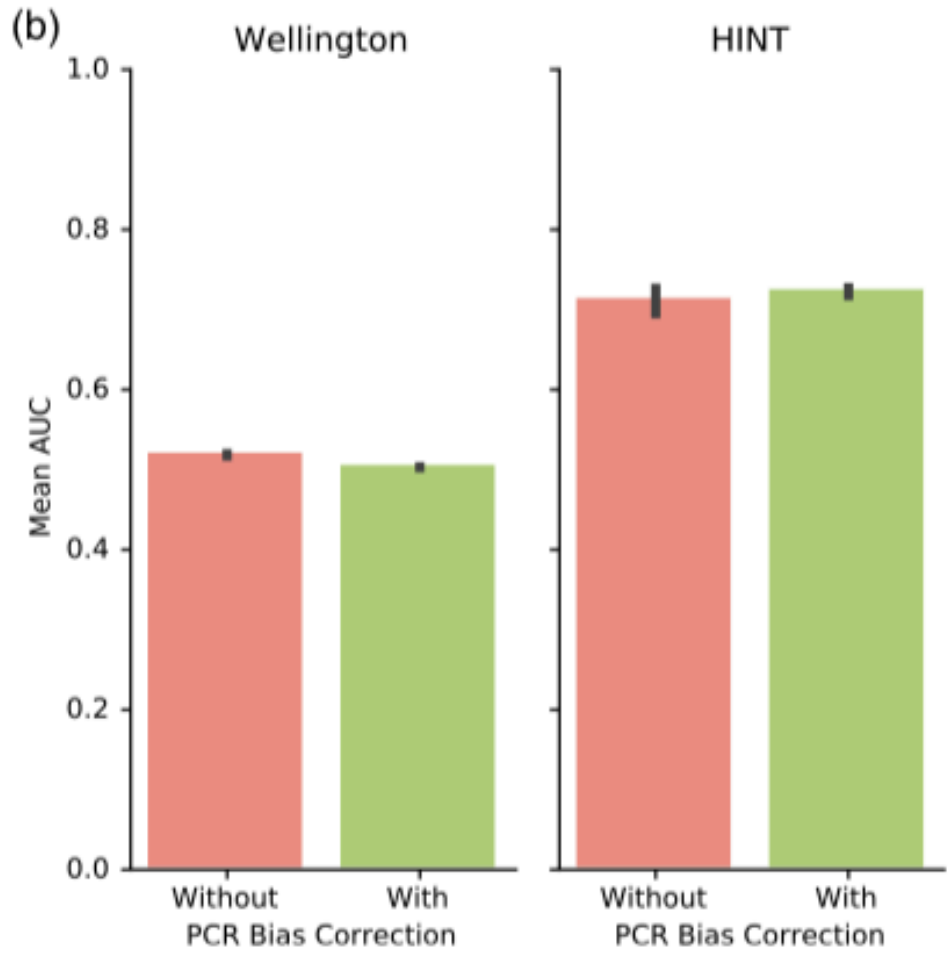
100 mio reads “defensable”, 160 mio “optimal” read depth

# PCR duplicate removal

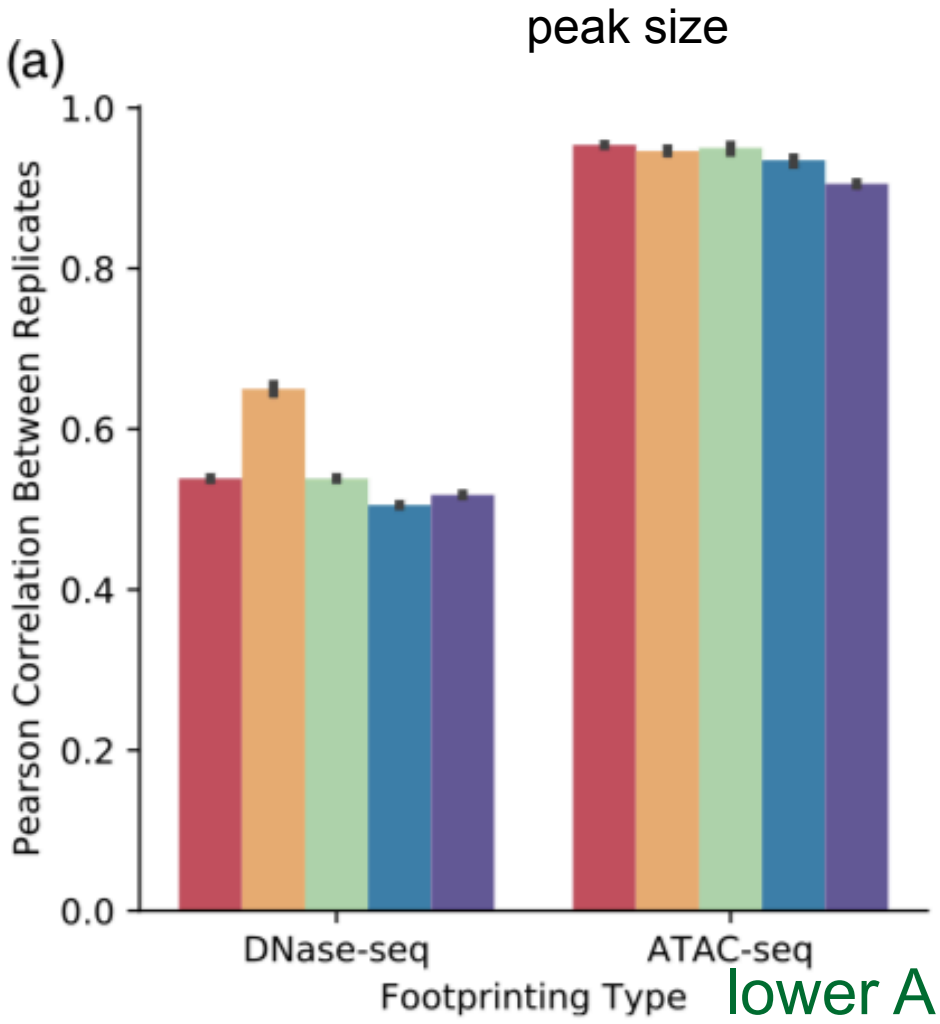
Correlation of #reads in bins

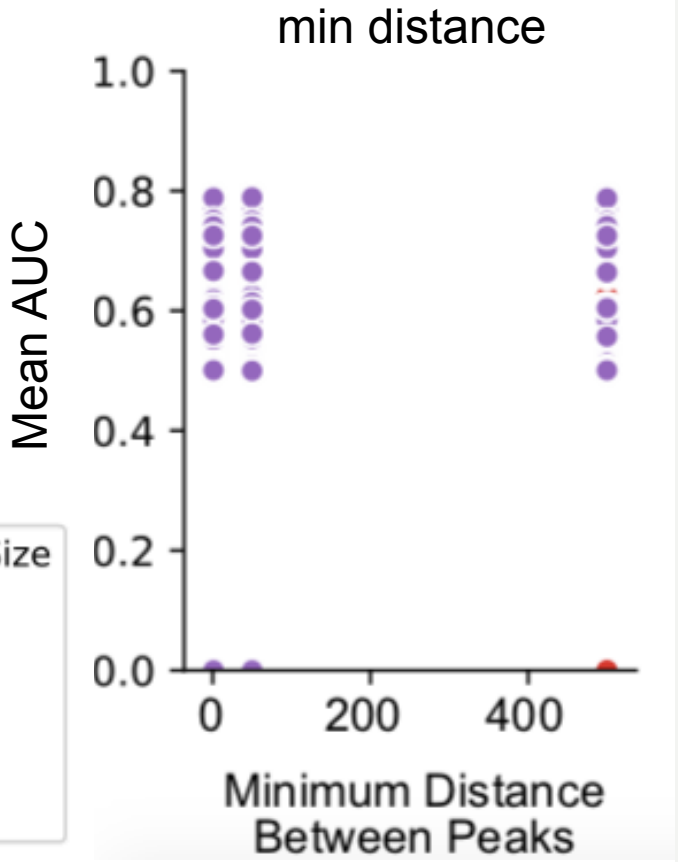



AUC



# Peak calling parameters

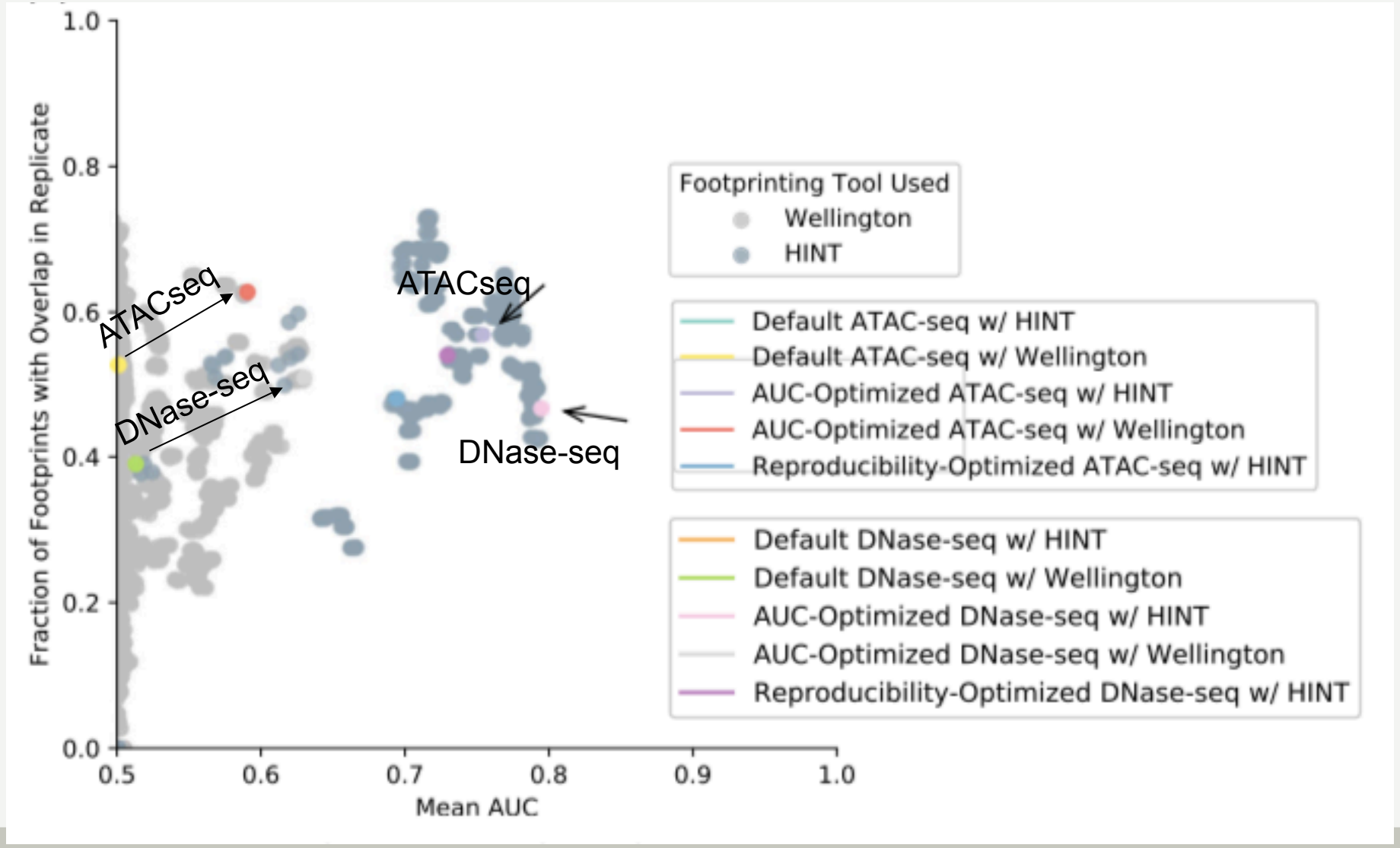


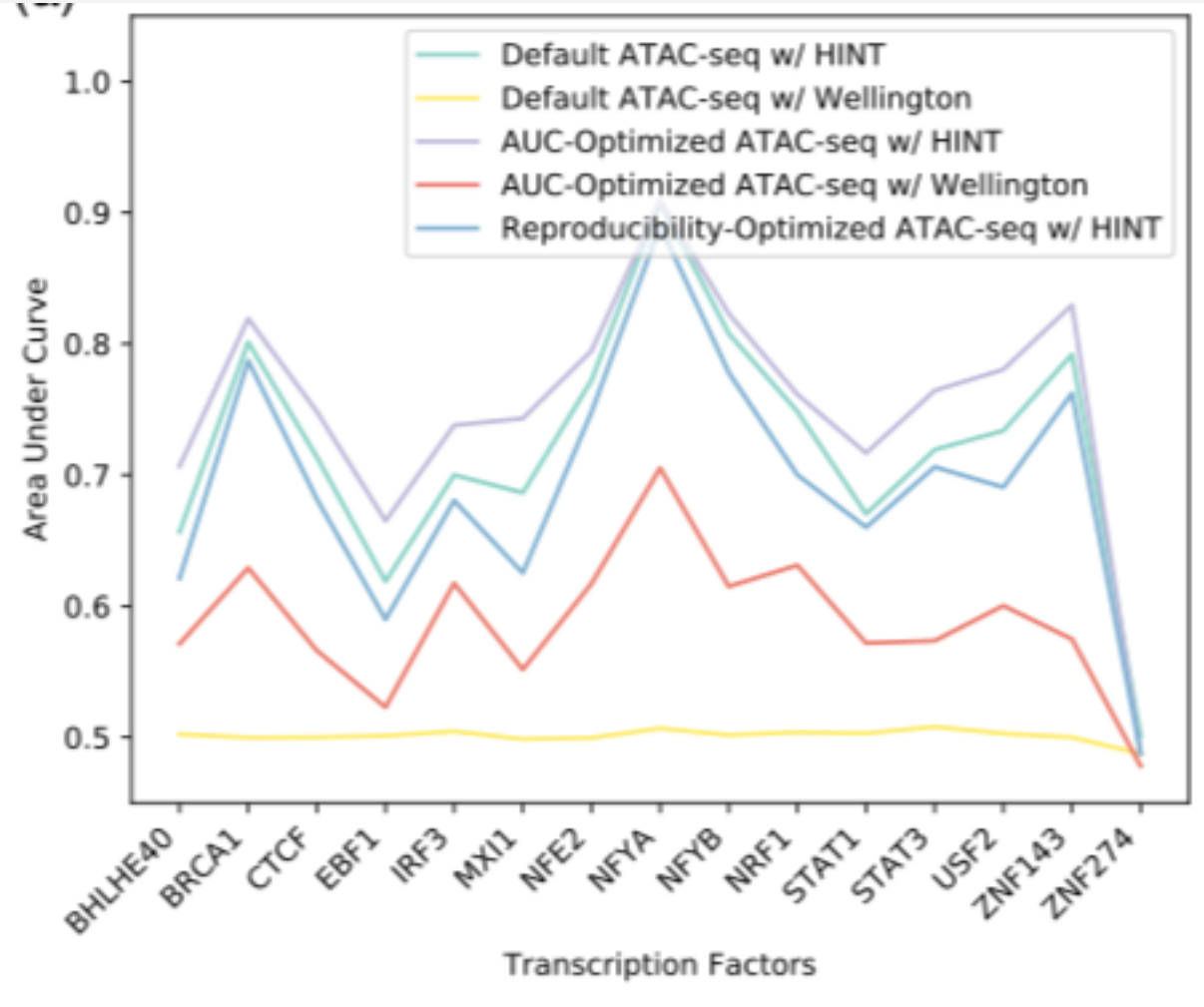

lower AUC at lower peak sizes (no plot)

highest ChIPseq recovery at 200-500

# ChIP-seq AUC

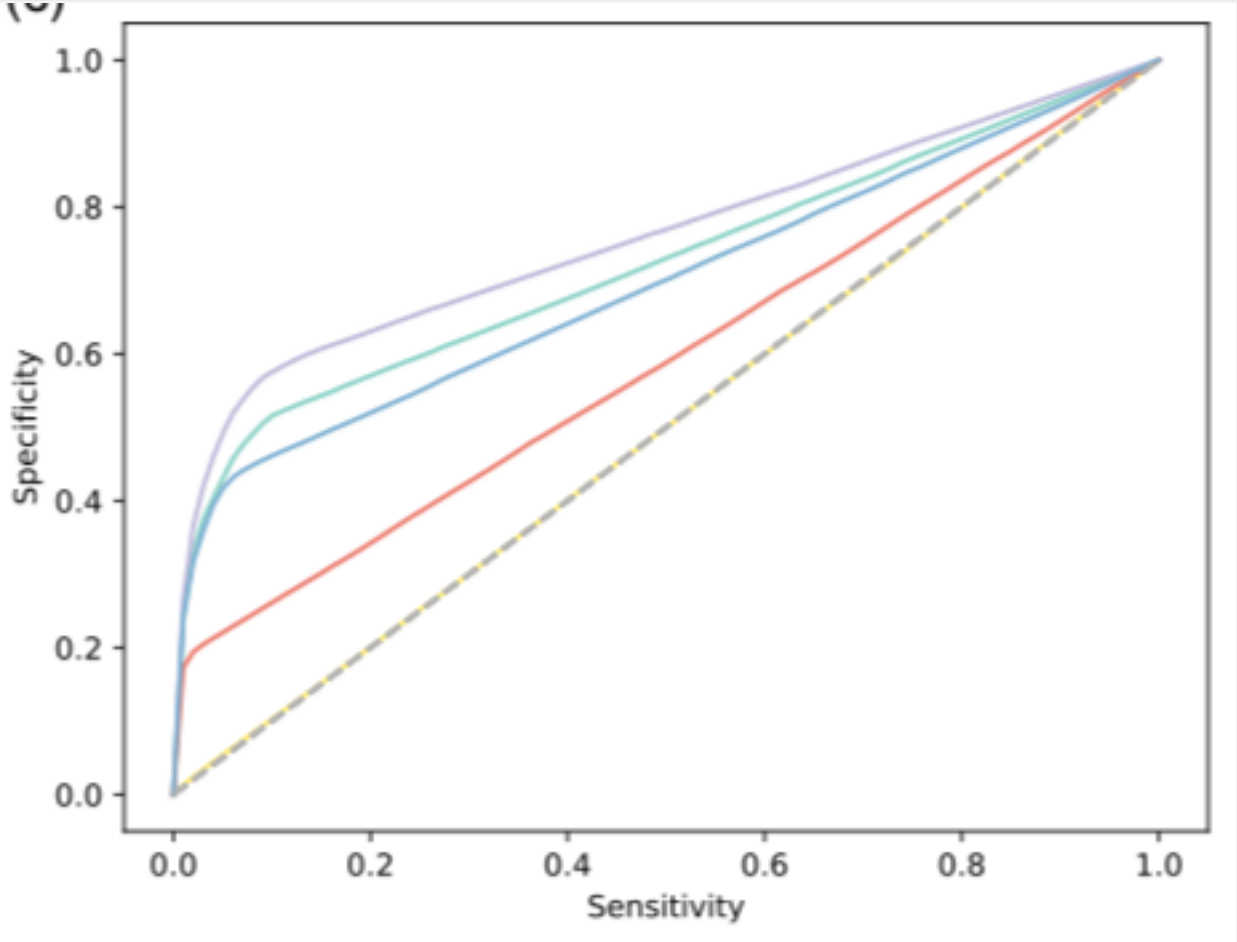

Fraction of Footprints with Overlap in Replicate
Mean AUC
1.0
0.8
0.6
0.4
0.2
0.0
0.5
0.6
0.7
0.8
0.9
1.0
ATACseq
DNase-seq
ATACseq
DNase-seq
Footprinting Tool Used
Wellington
HINT
Default ATAC-seq w/ HINT
Default ATAC-seq w/ Wellington
AUC-Optimized ATAC-seq w/ HINT
AUC-Optimized ATAC-seq w/ Wellington
Reproducibility-Optimized ATAC-seq w/ HINT
Default DNase-seq w/ HINT
Default DNase-seq w/ Wellington
AUC-Optimized DNase-seq w/ HINT
AUC-Optimized DNase-seq w/ Wellington
Reproducibility-Optimized DNase-seq w/ HINT

# TF overview




imbalance of negative and positive TFBS of approx. 10:1

# ROC




imbalance of negative and positive TFBS of approx. 10:1

# Conclusion

- grid search over 4560 parameter combinations
  - but still limited to 1 mapper, 1 peak caller and 2 footprint caller
- claim that they produce a GRN is just wrong
- ChIP recovery plateaus at 160 mio reads
- PCR duplicate removal only affects reproducibility but not ChIP recovery
- smaller DHS peaks good for reproducibility but bad to AUC
- HINT's bias correction negatively affects AUC
- only little information in supplement, some conclusions only described in text without any figures/data available