



deepTools

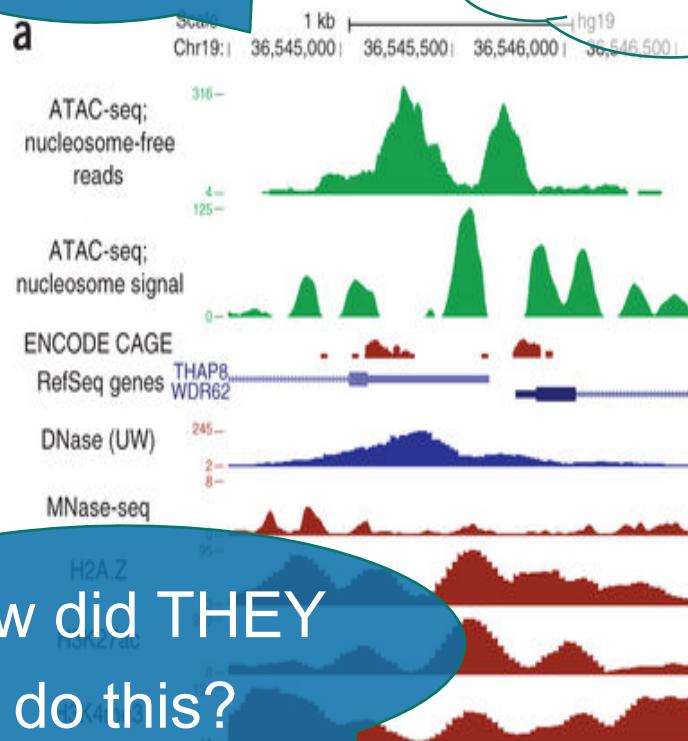
a flexible platform for exploring deep-sequencing data

Sarah Diehl - 2nd Swiss Galaxy Workshop



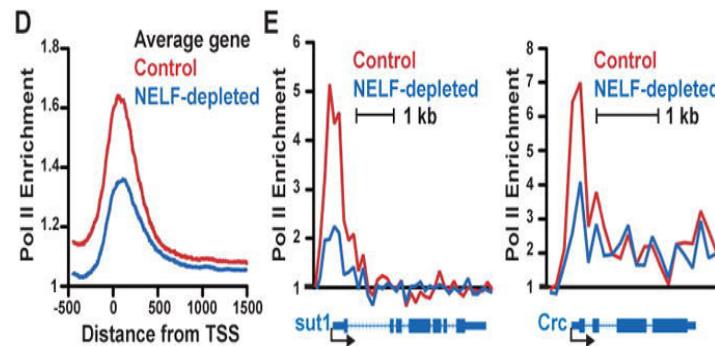
How can I
do this?

Awww, that just
takes too long...

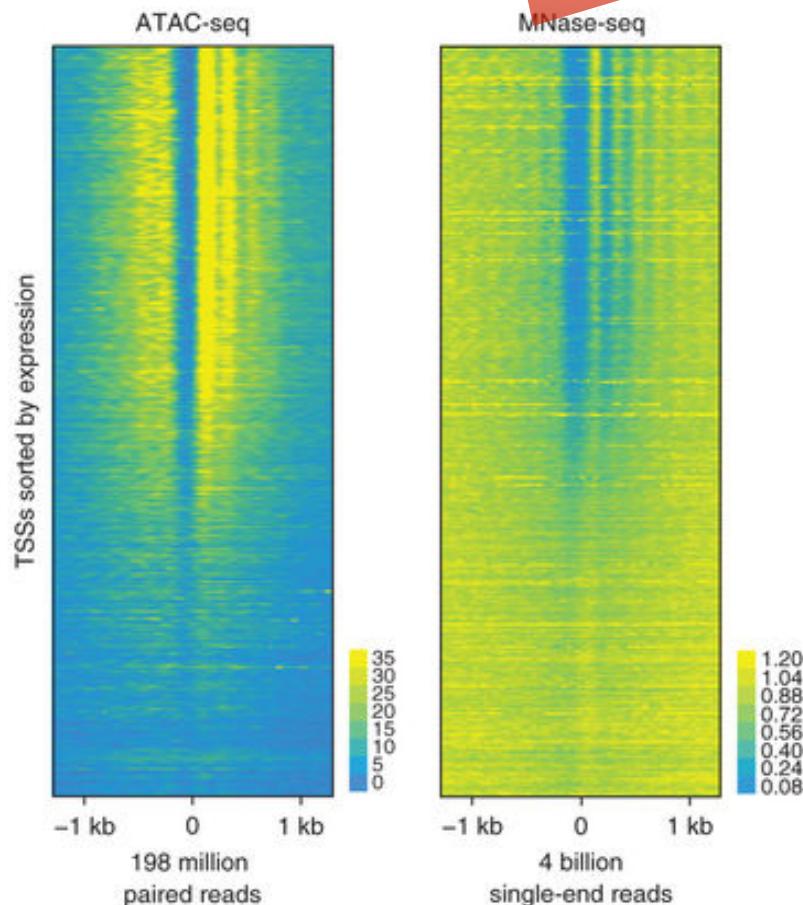


How did THEY
do this?

Summary plots



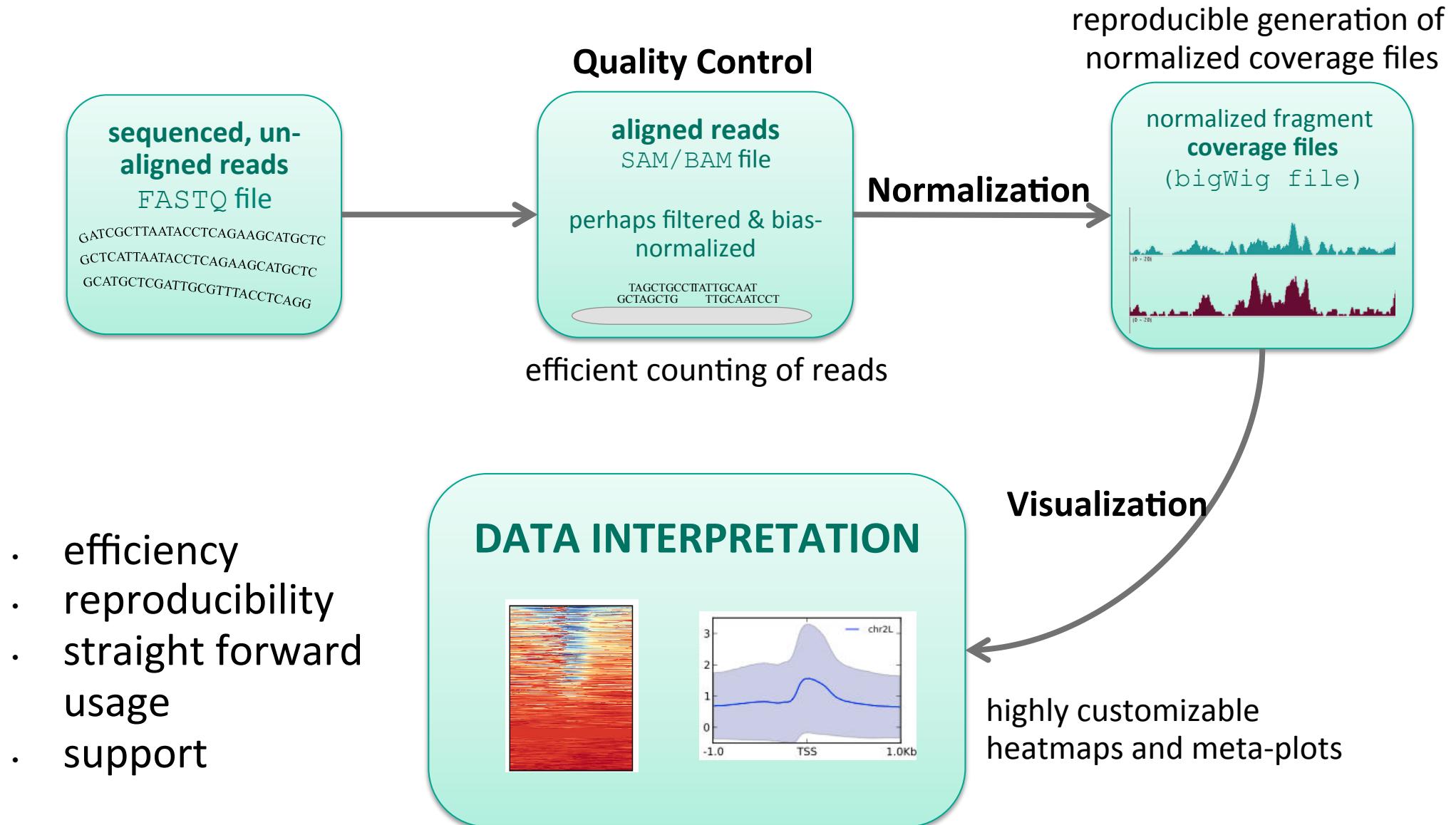
Heatmaps





- NGS analyses often characterized by
 - Specialized and custom-made scripts
 - Hidden filtering strategies
 - Lack of standardization and reproducibility
- Existing tools:
 - Require experience with command-line
 - No graphical output to guide interpretation of results
- Several biases in NGS data
- Deters biologists from taking part in analysis of their data

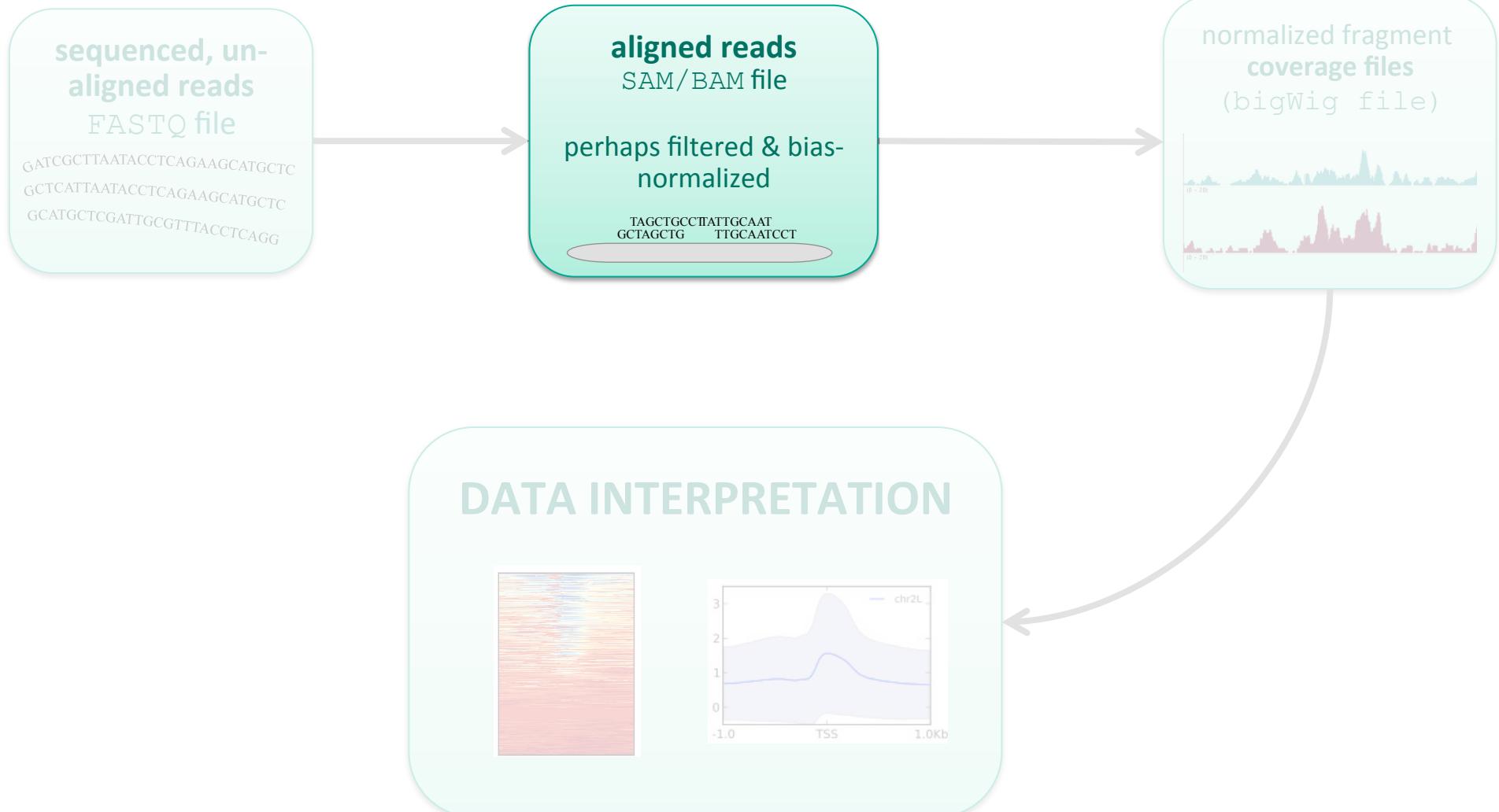
deepTools data processing workflow



- efficiency
- reproducibility
- straight forward usage
- support

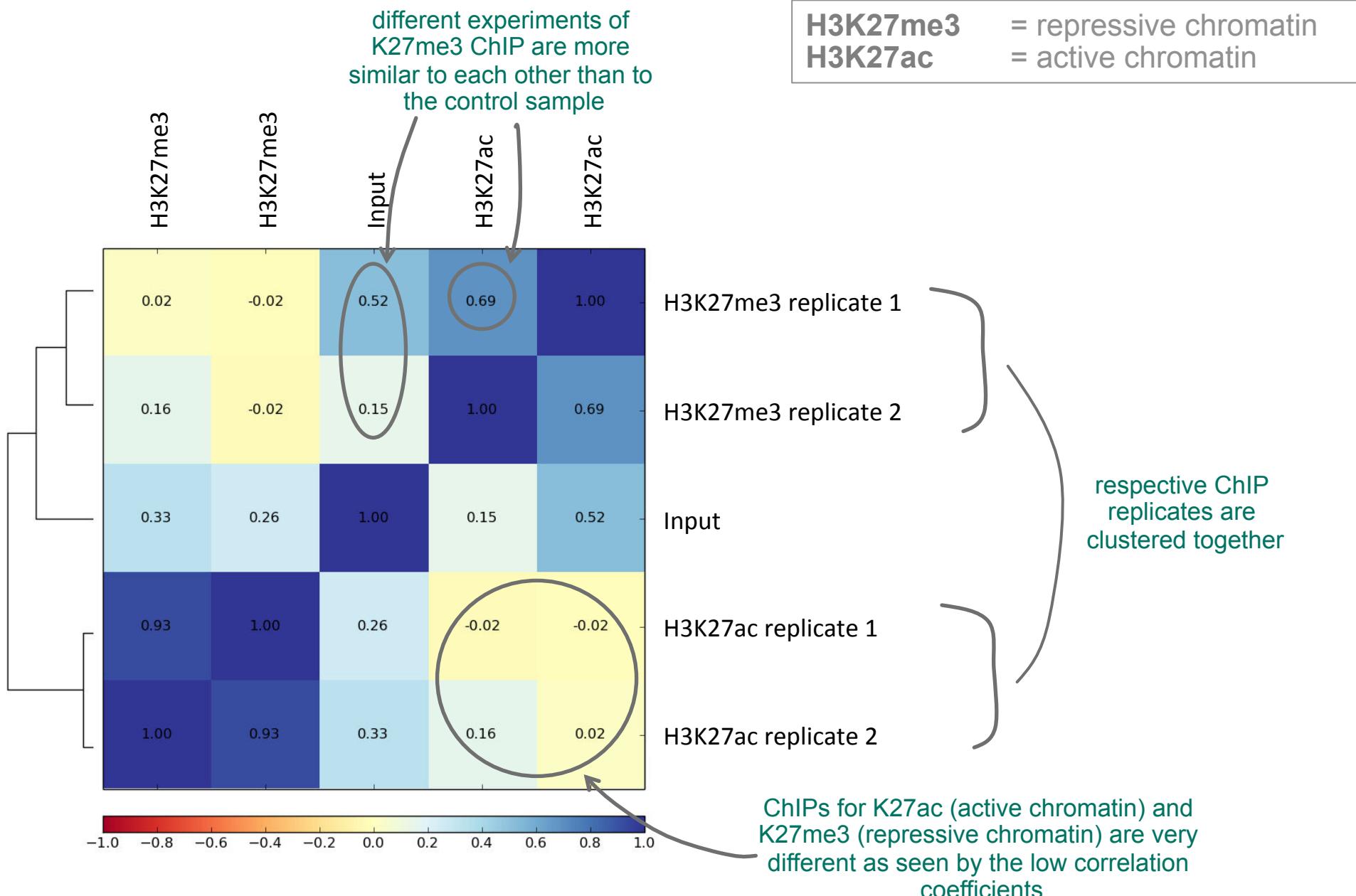


Quality Control





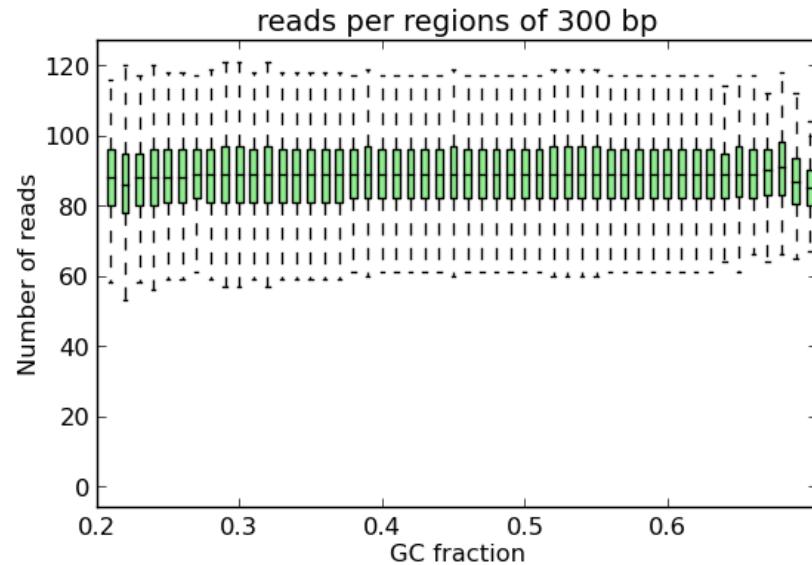
QC I: bamCorrelate



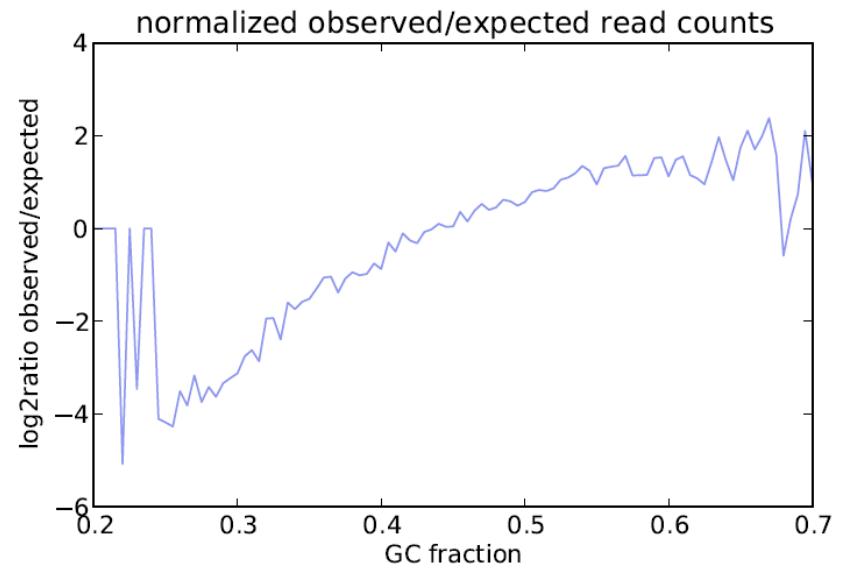
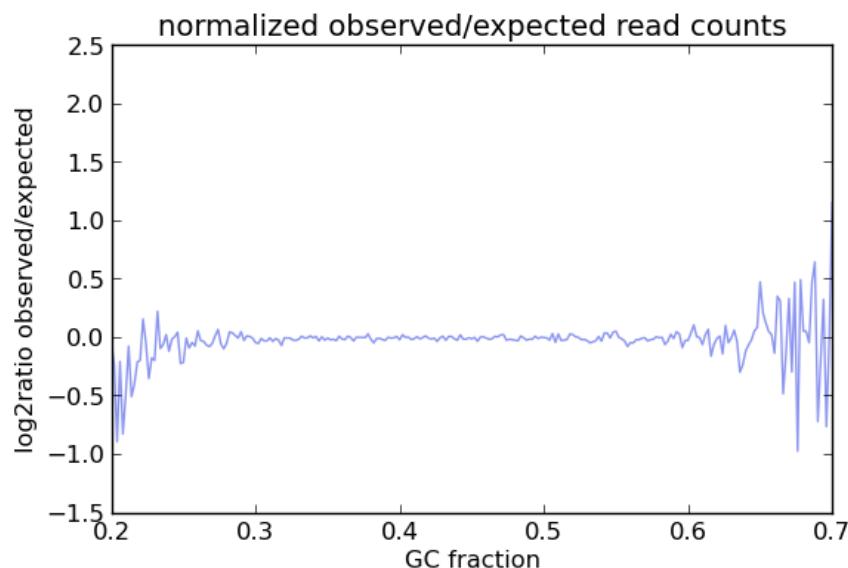
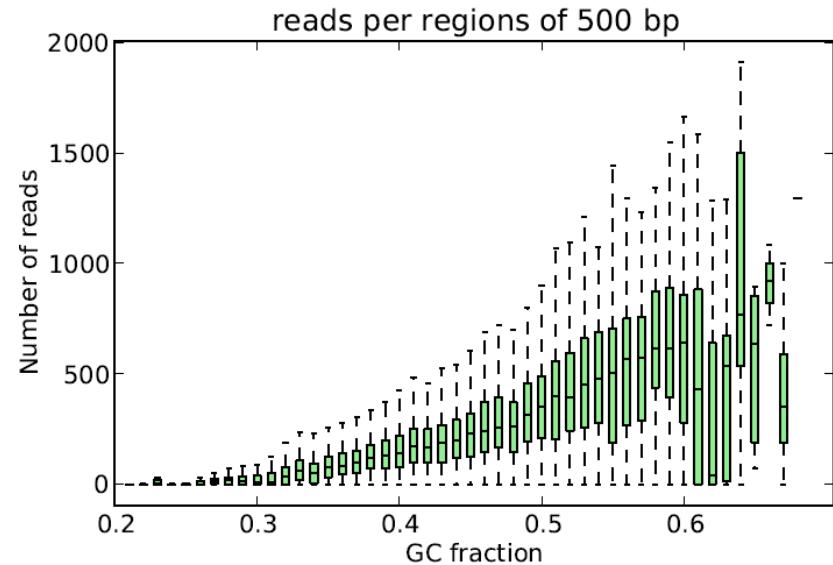
QC II: computeGCbias



simulated reads



real-life sample

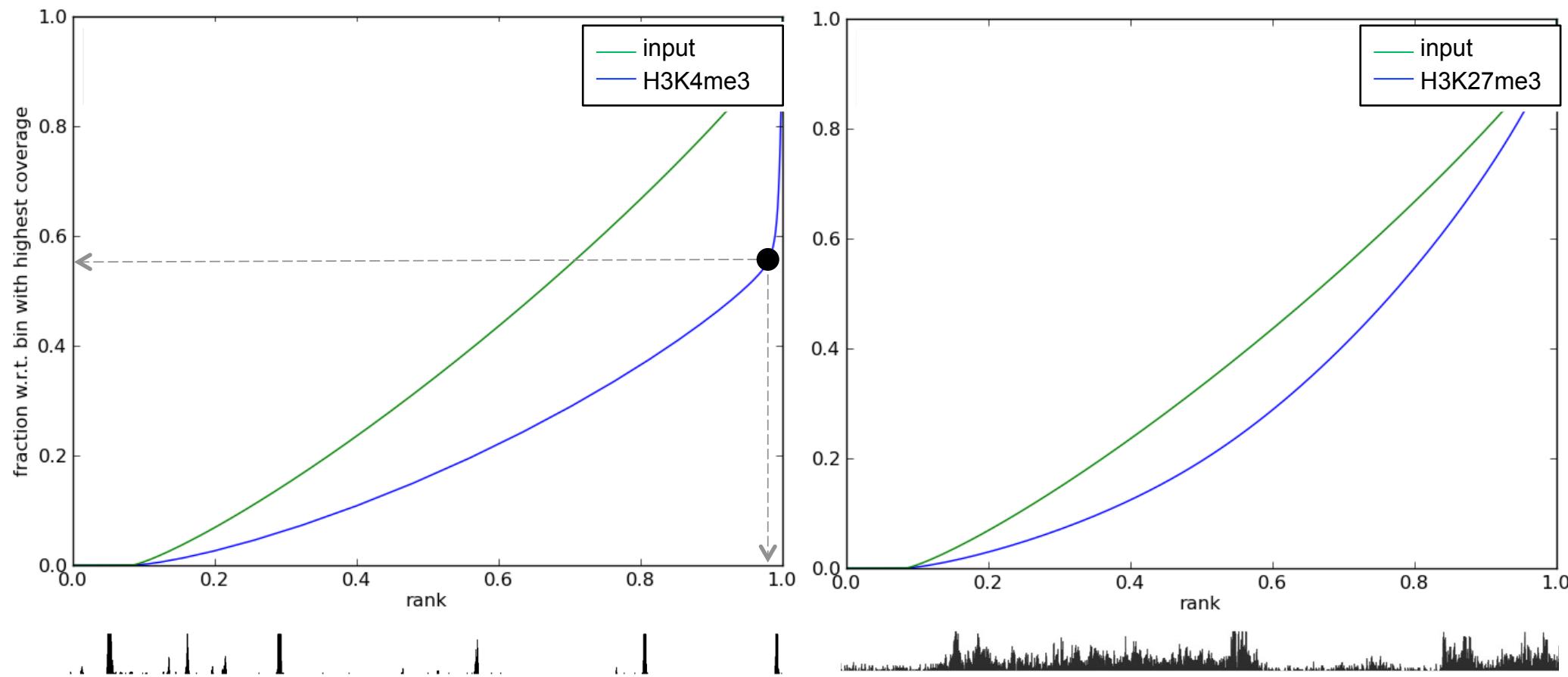


Benjamini & Speed. Summarizing and correcting the GC content bias in high-throughput sequencing. NAR (2012)

QC III: bamFingerprint



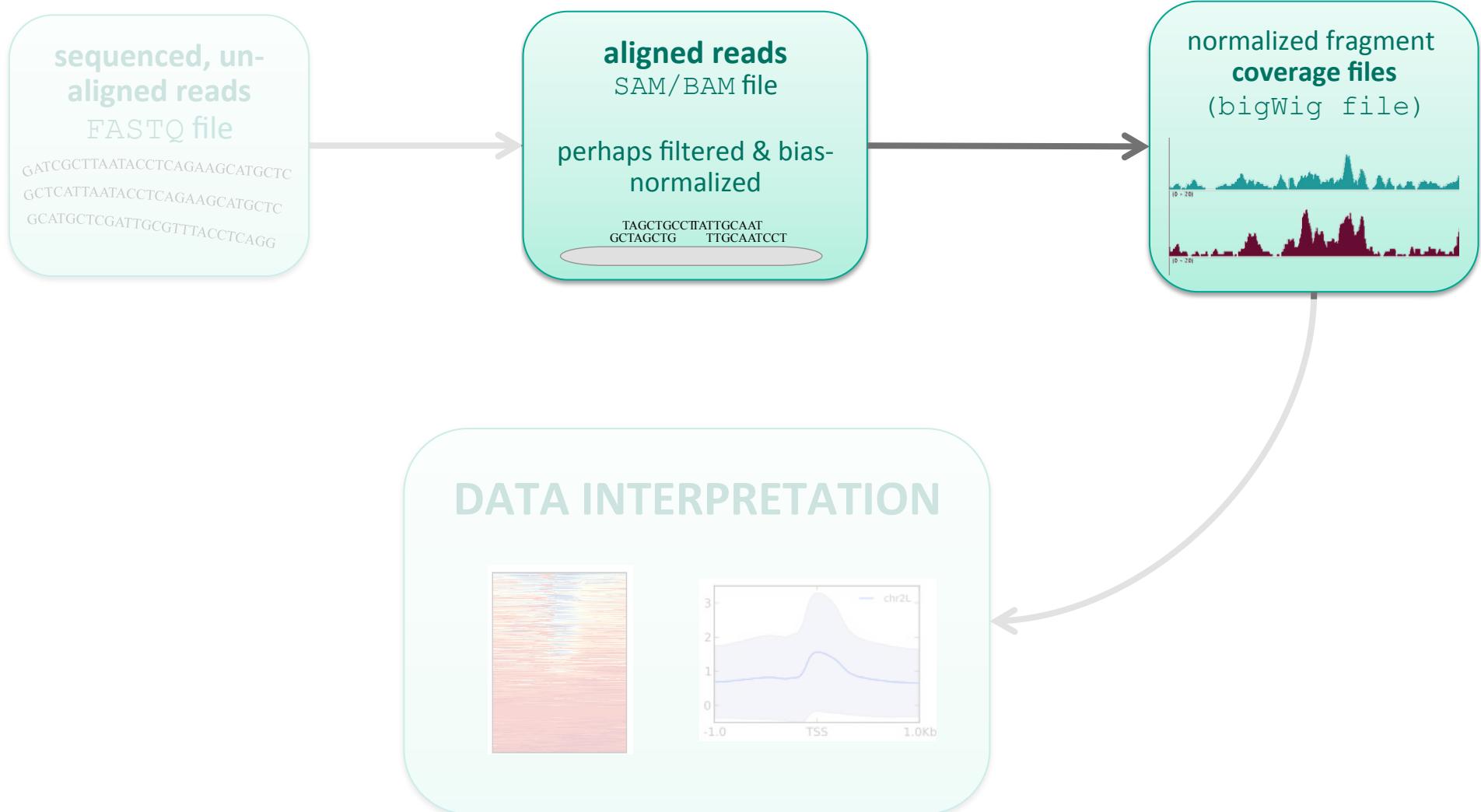
- How well can the signal in the ChIP-seq sample be differentiated from the background?



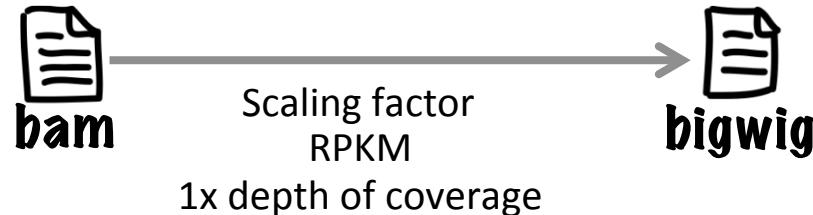
Diaz A., Park K., Lim D.A., Song J.S. Normalization, bias correction, and peak calling for ChIP-seq. Stat. Appl. Genet. Mol. Biol. 2012;11:9.



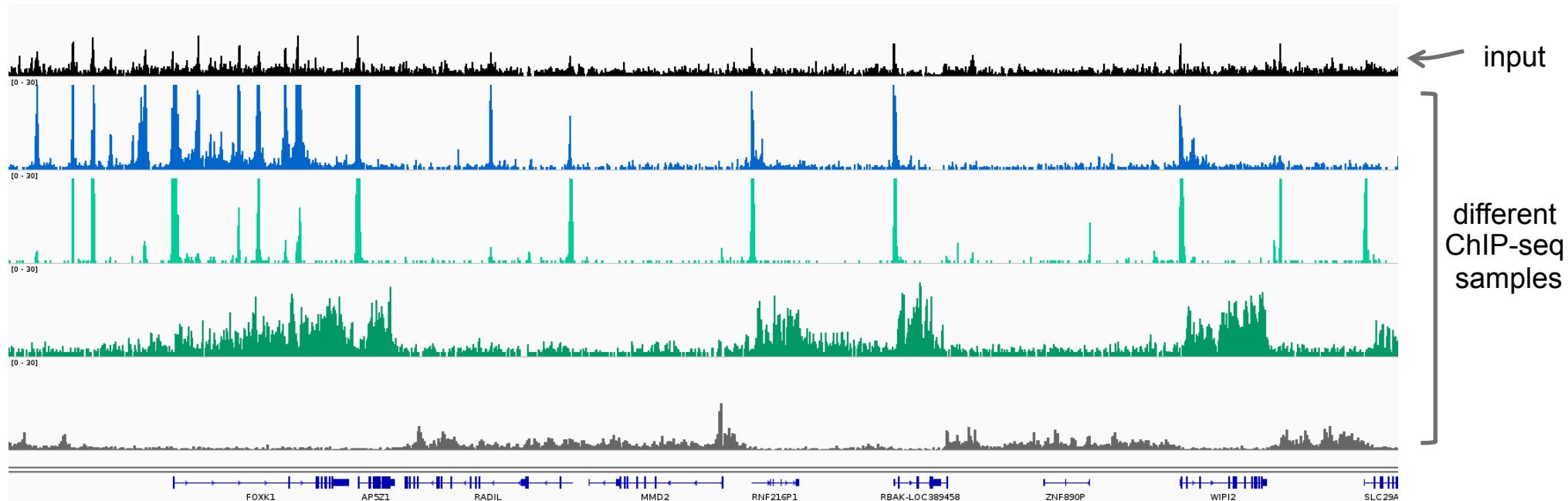
Normalization



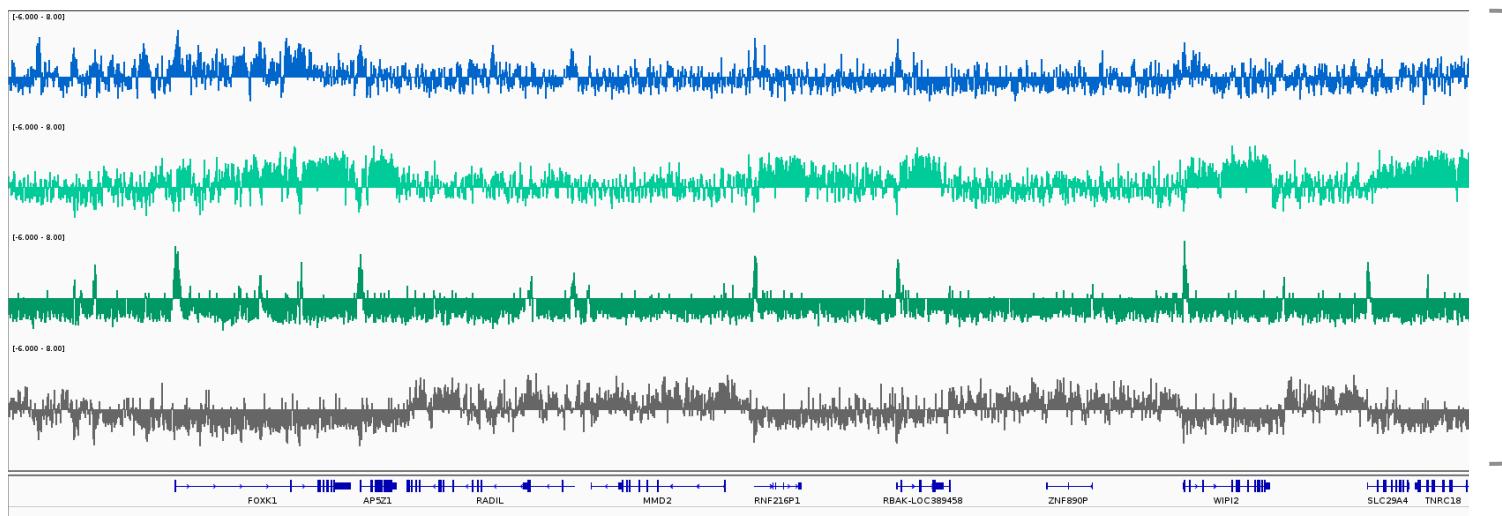
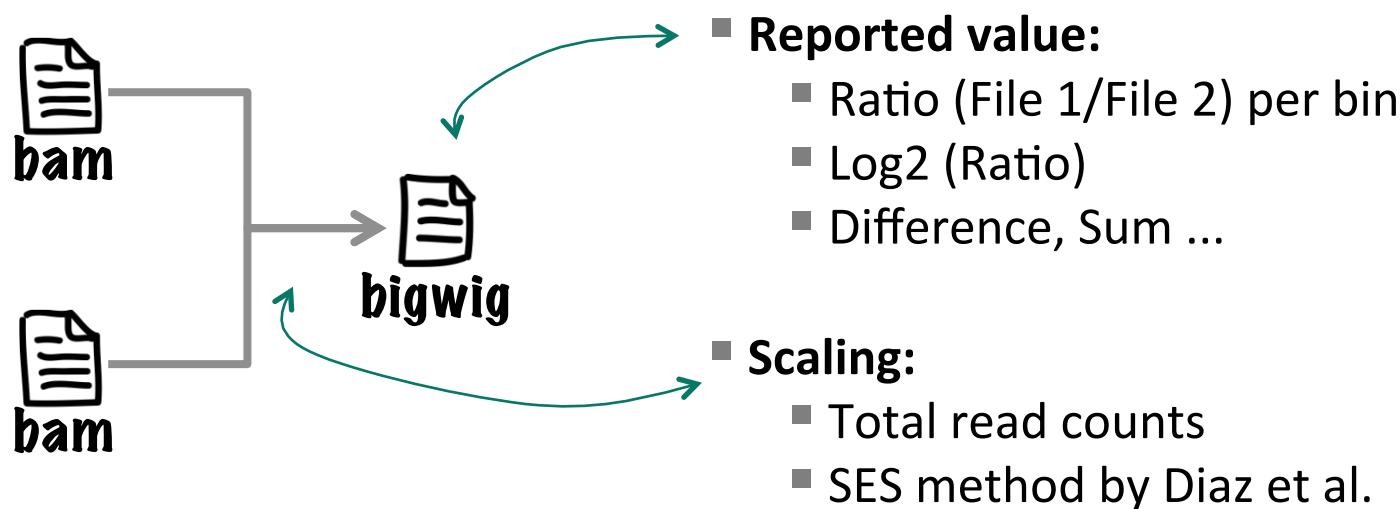
Normalization I: bamCoverage



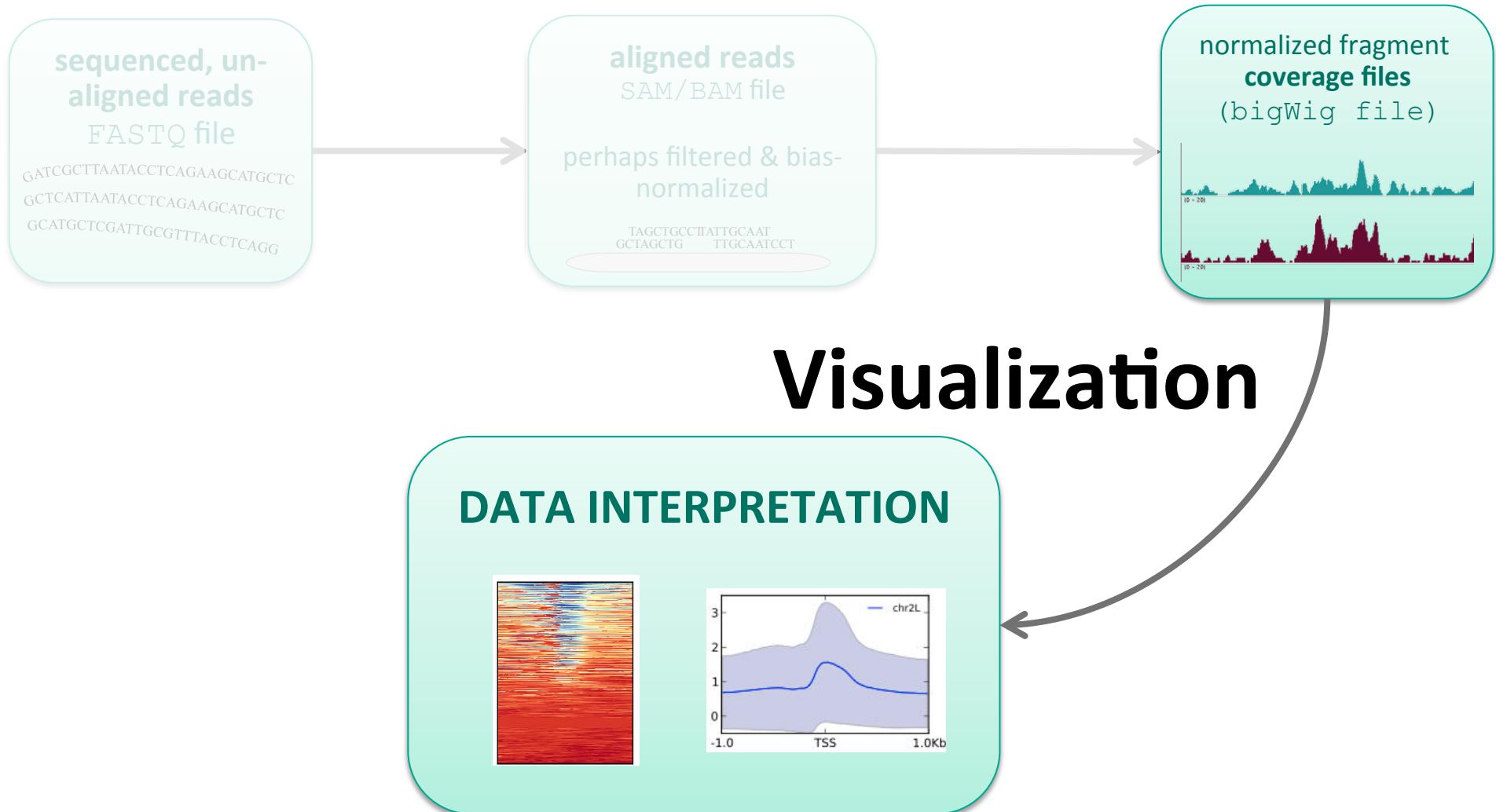
- Data sharing & storage
- Visualization in Genome Browsers (e.g. IGV)
- More efficient downstream analyses



Normalization II: bamCompare



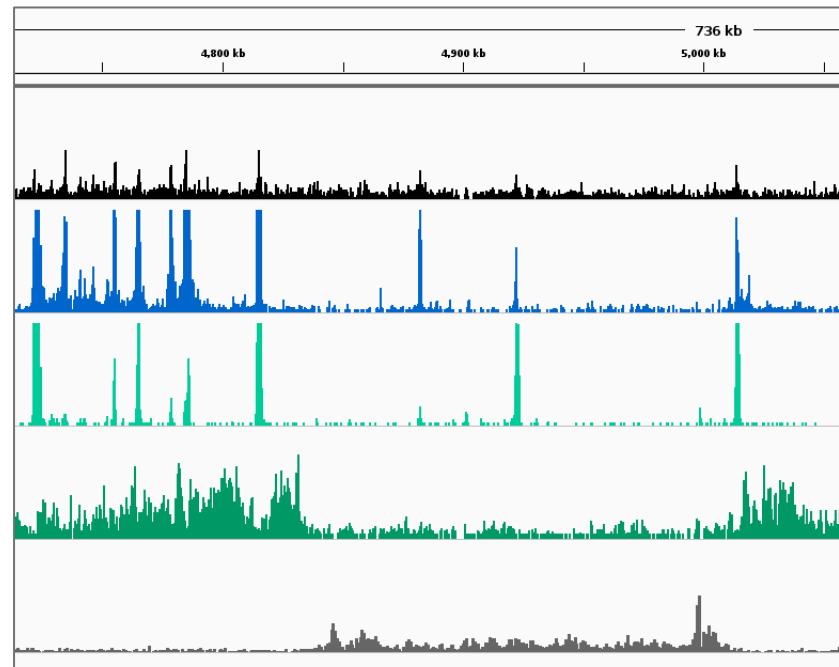
Diaz A., Park K., Lim D.A., Song J.S. Normalization, bias correction, and peak calling for ChIP-seq. Stat. Appl. Genet. Mol. Biol. 2012;11:9.



Visualization: local to global



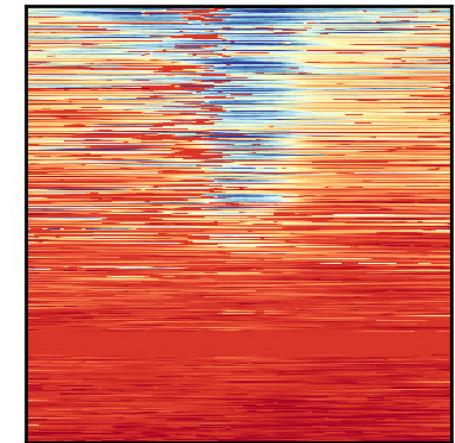

bigwig



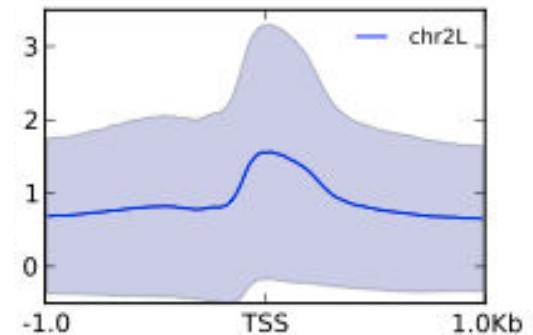
computeMatrix



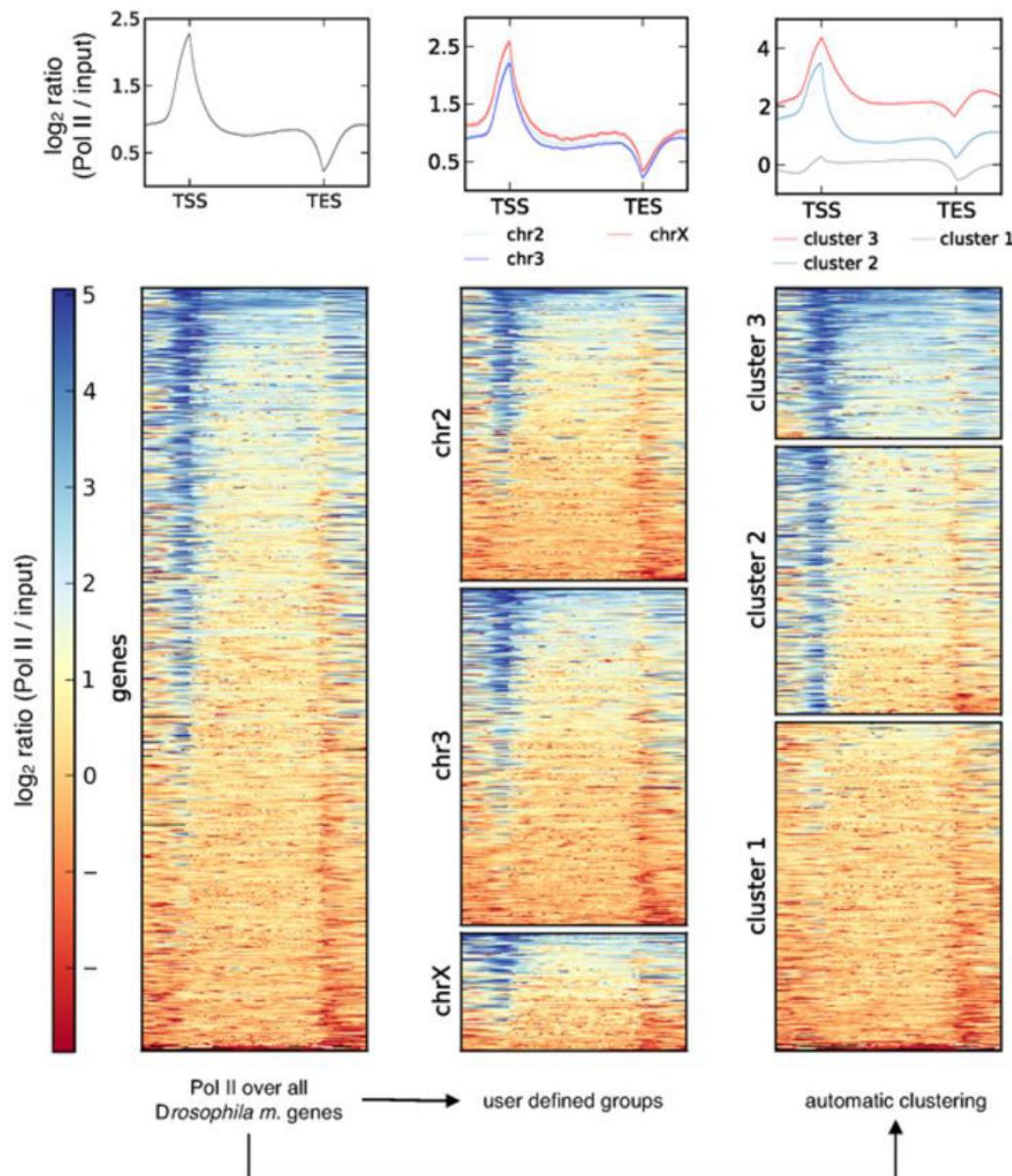
heatmapper



profiler

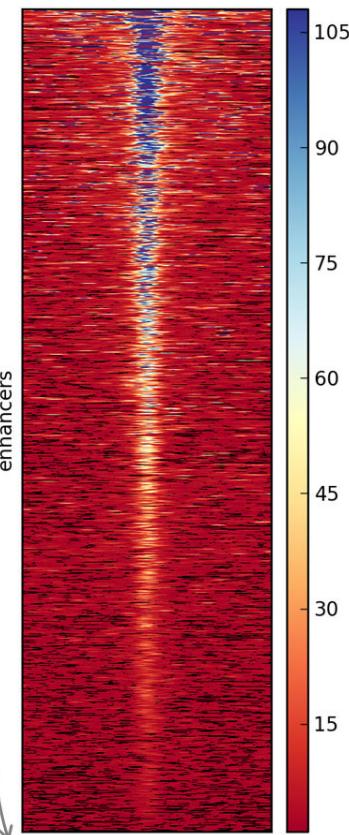
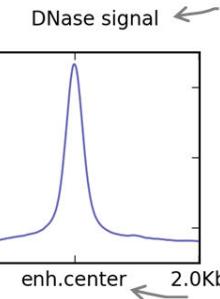


Visualization: heatmapper

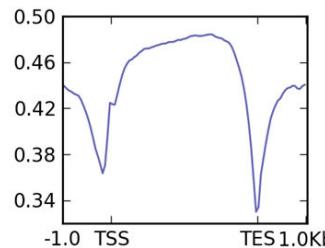




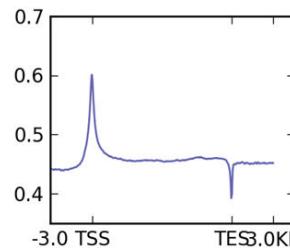
Visualization: heatmapper



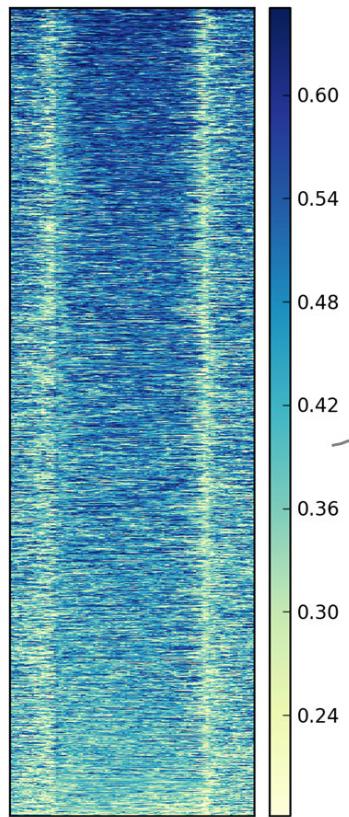
↗ GC content fly



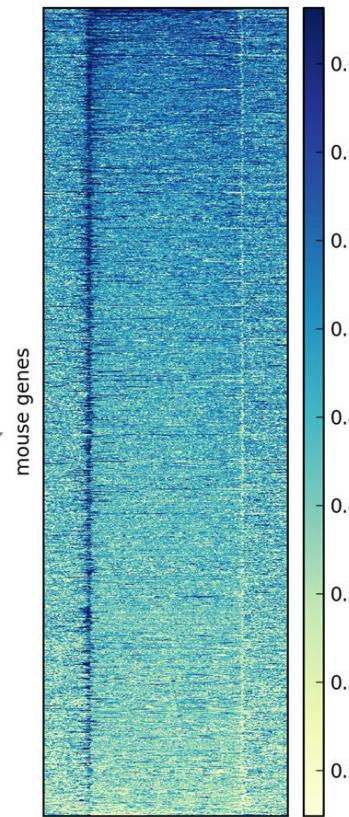
↗ GC content mouse



↗ fly genes

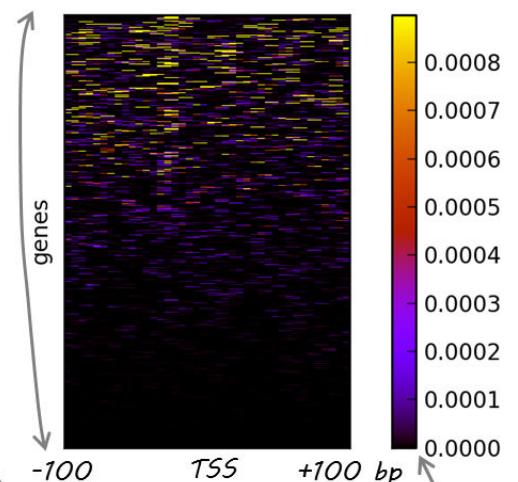


↗ mouse genes



labels size colors
intensity range
file format
what to show ...

↗ TATA motif





analyze more, process less

- Efficiency and highly parallelized processing
- Customized down-stream analyses (store intermediate data)
- Standardized output files
- Modular approach – compatibility, flexibility, scalability

- Automatic installation with dependencies
 - Command-line version: Python Package Index (easy_install, pip)
 - Galaxy: from main toolshed
 - Docker image

- Public Galaxy Server:
deepTools.ie-freiburg.mpg.de

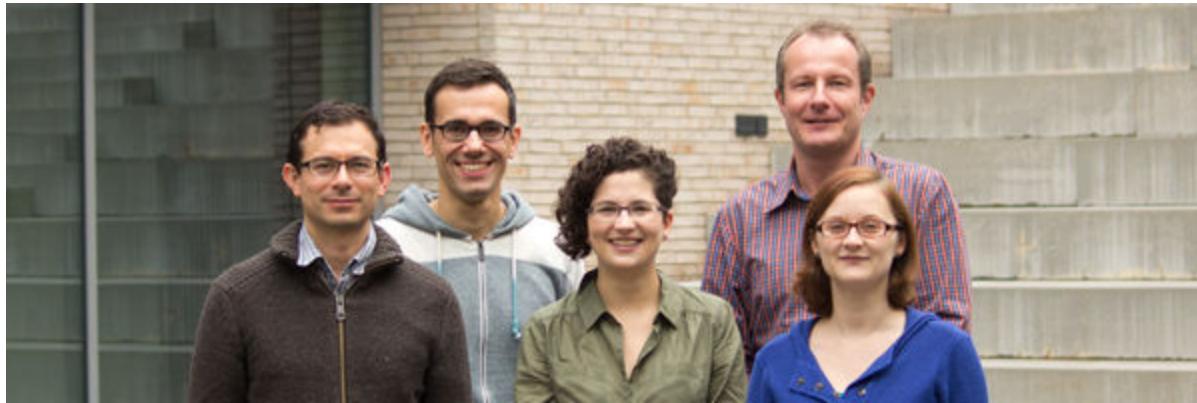


<http://deeptools.github.io>

About deepTools	The Tools	Example Workflows	Technical Documentation	Galaxy
About deepTools	Tools Overview	Example Workflows	Technical Documentation	Galaxy introduction
Introducing deepTools - why we built it and what it does	Detailed information about the individual tools	Step-by-step protocols of typical deepTools applications (for Galaxy and command line usage)	Information on technical requirements and how to install deepTools	General introduction into deepTools Galaxy

+ Google group, FAQ, image gallery

Acknowledgements



MPI-IE

- Fidel Ramírez
- Friedrike Dündar
- Thomas Manke

More information & support:

deepTools.github.io

University of Freiburg

- Björn Grüning

All users of deepTools for
their feedback

```
docker run -d -p 8080:80 bgruening/galaxy-deeptools
```