

ChIP-Seq

Wednesday (20 September 2023)

New methods and quantitative ChIP-seq (Simon Elsässer and Carmen Navarro Luzón)

09:00 - 10:00 Recap previous day (*online session*)

10:00 - 11:00 ChIP-seq methods (*online session*)

11:00 - 12:00 ChIP-seq alternatives (*online session*)

12:00 - 13:00 lunch (*offline*)

13:00 - 13:30 Introduction to exercises (*online session*)

13:30 - 16:00 [Principles of ChIP-seq](#) and [Advanced ChIP Methods](#) (*online support*)

16:00 - 17:00 Exercises (*offline*)

16:00 - 17:00 Daily challenge

Simon Elsässer, KI/SciLifeLab

simon.elsasser@scilifelab.se

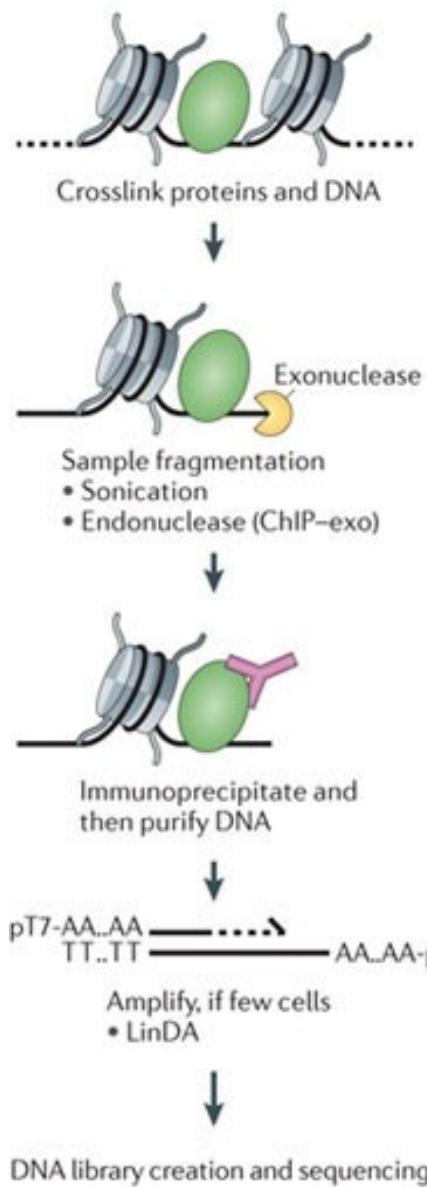
Excercises – big thanks to Carmen



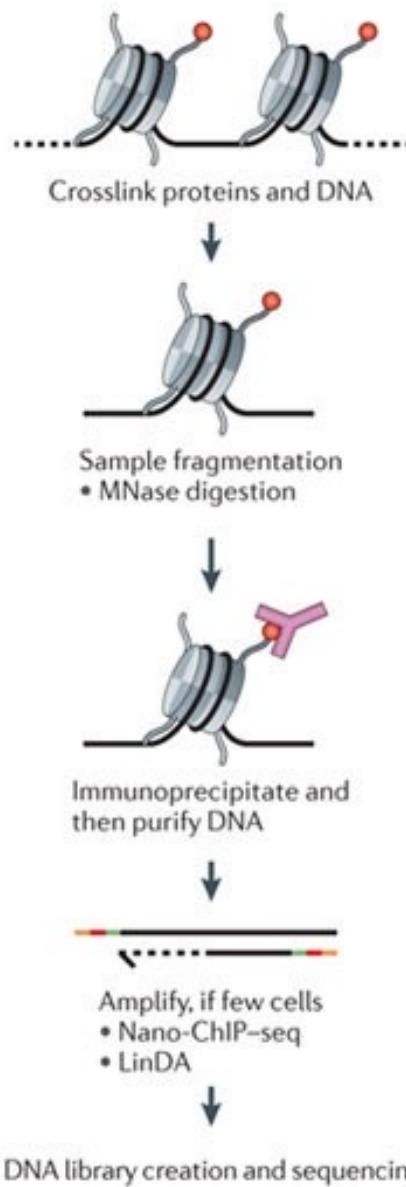
Carmen Navarro

Experimental considerations

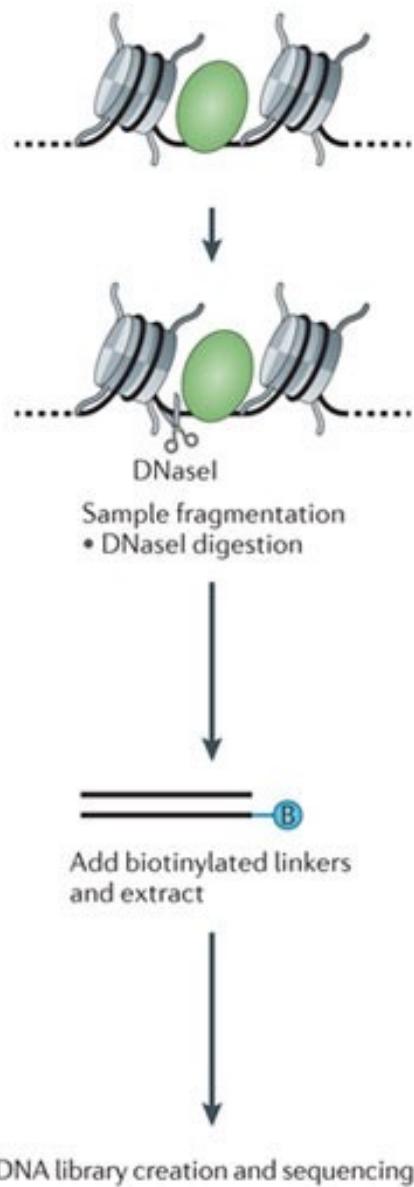
a DNA-binding protein ChIP-seq



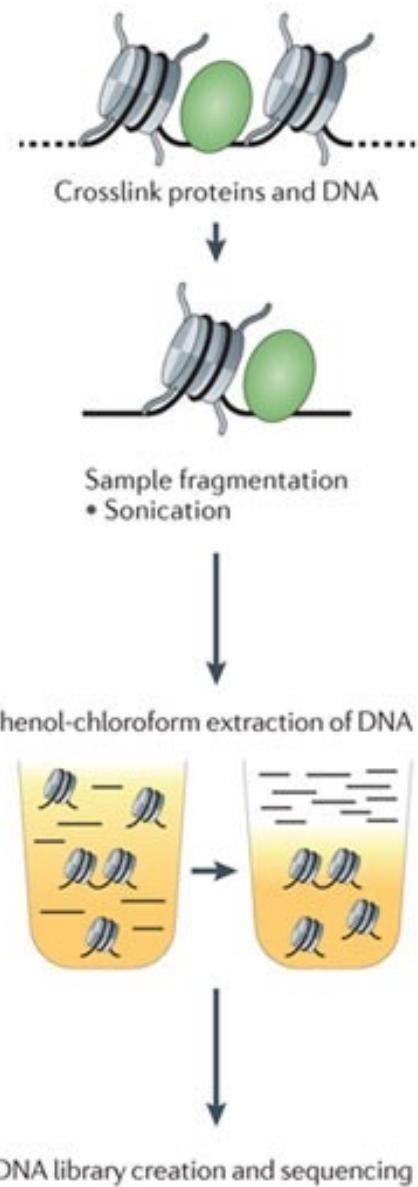
b Histone modification ChIP-seq



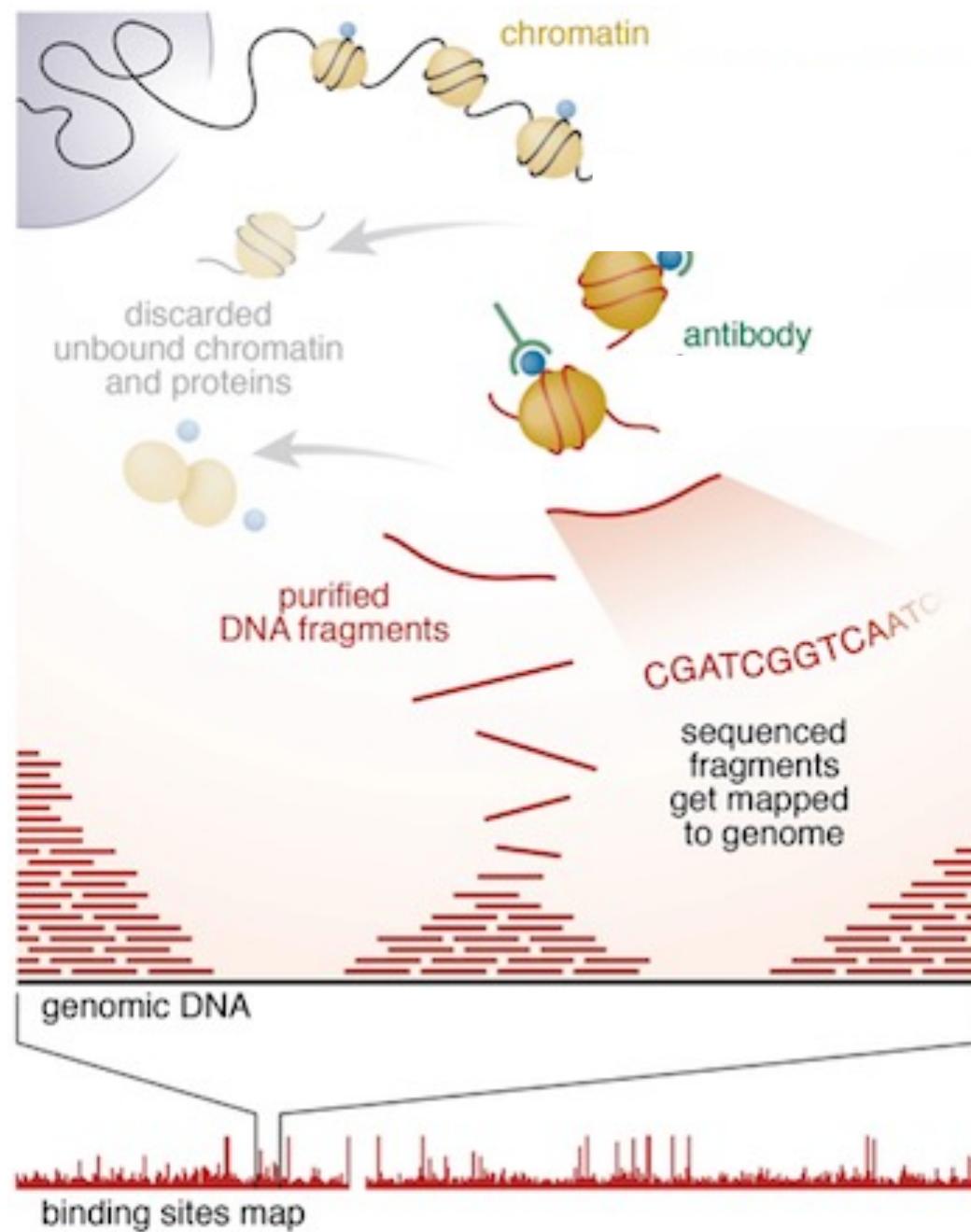
c DNase-seq



d FAIRE-seq



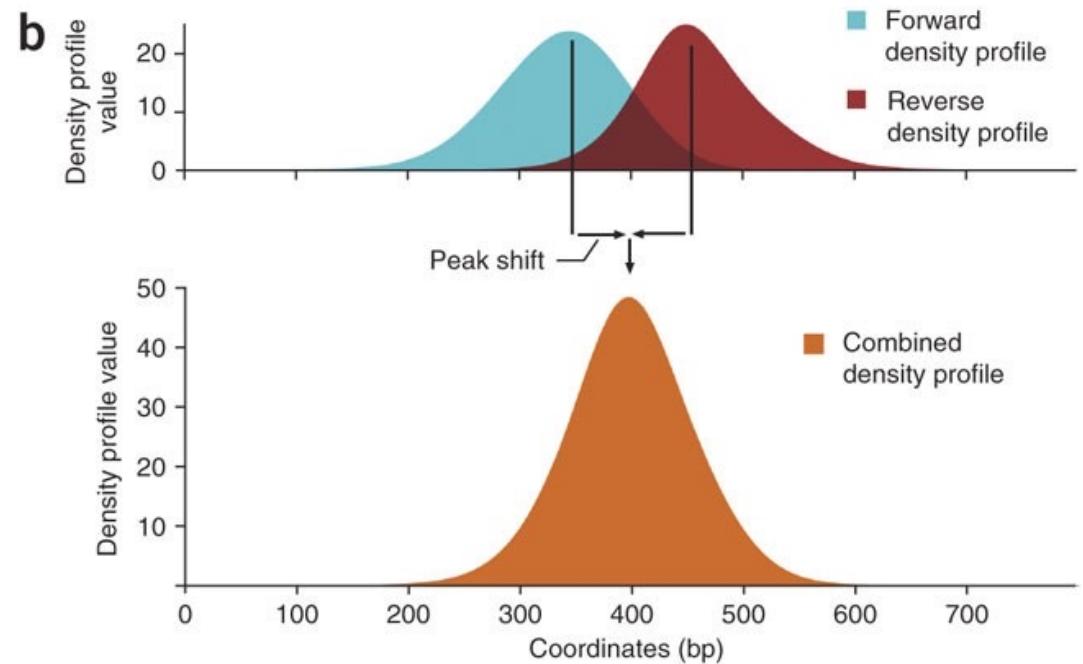
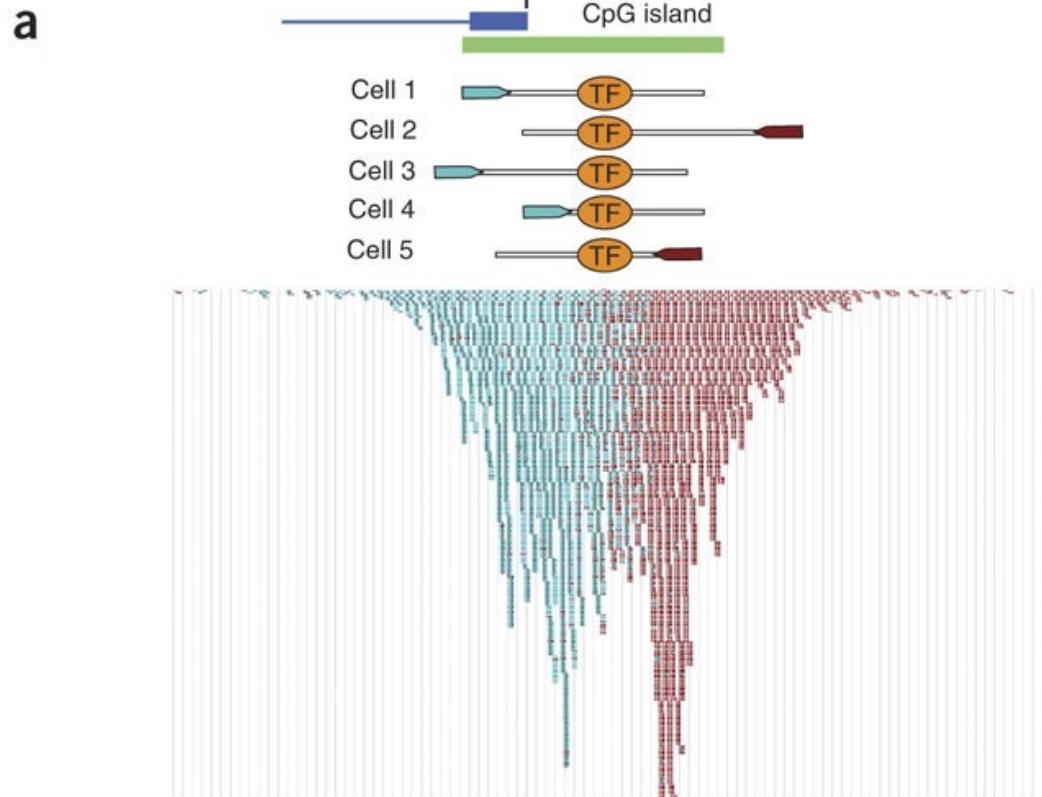
ChIP-seq Principles



ChIP-seq Principles

Mapping transcription factor binding:

- TF-DNA interactions typically don't survive lysis and immunoprecipitation conditions --> we must crosslink the TF to the DNA beforehand

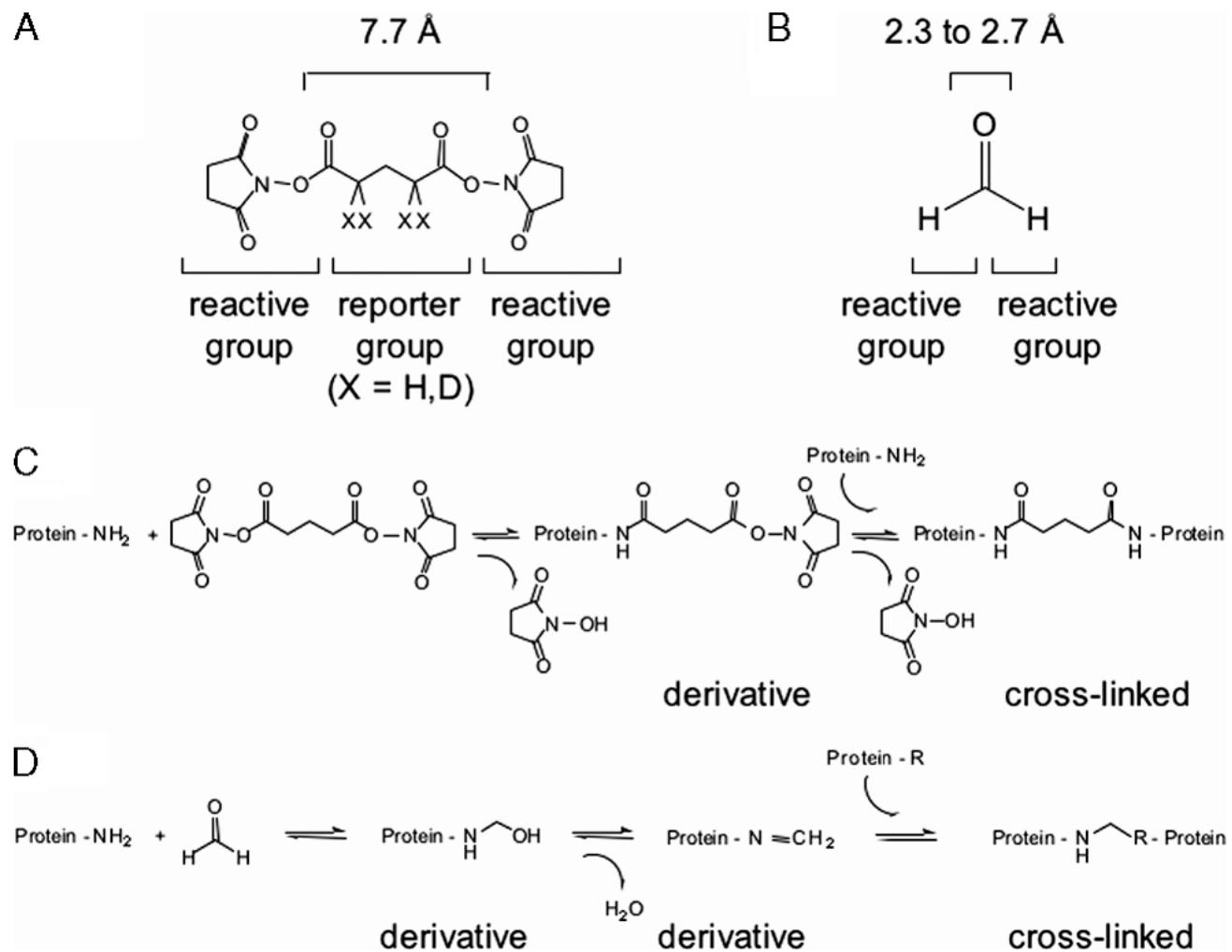


ChIP-seq Principles

Mapping transcription factor binding:

- TF-DNA interactions typically don't survive lysis and immunoprecipitation conditions --> we must crosslink the TF to the DNA beforehand A

- Formaldehyde
 - ESG
 - DSG



Crosslinking versus native ChIP

Crosslinking

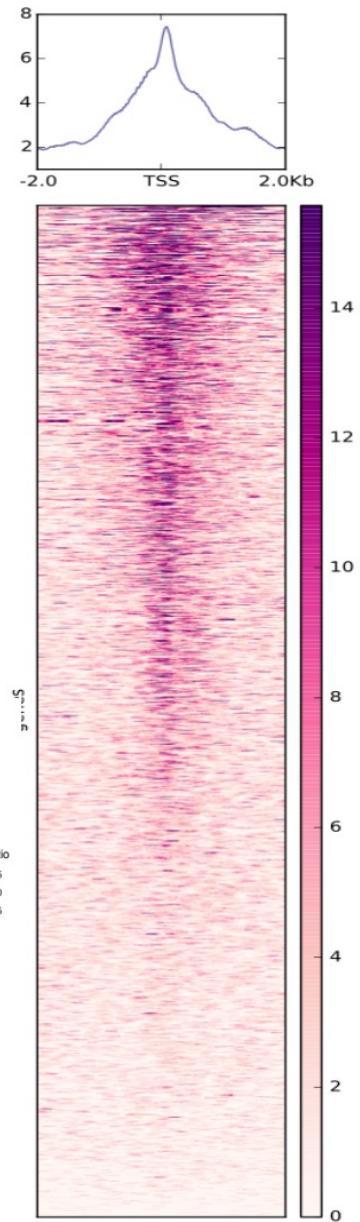
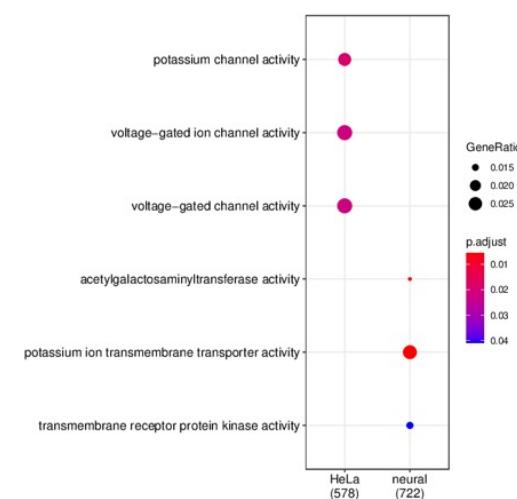
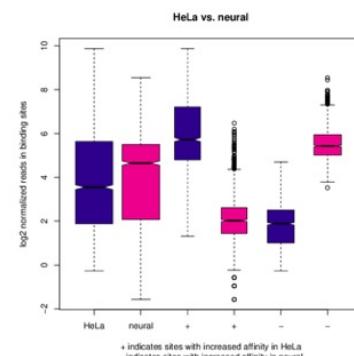
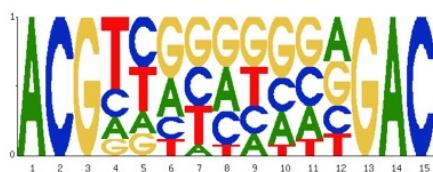
- Formaldehyde fixation
- Shearing (sonications) or MNase
- Increasing signal for weak/transient histone/DNA- interacting proteins
- fragment ends not informative with sonication

Native

- Lower salt/detergent
- Mnase fragmentation
- Better signal-to-noise for strong chromatin interactors, histones
- Fragment ends demarcate footprint, e.g. nucleosome position

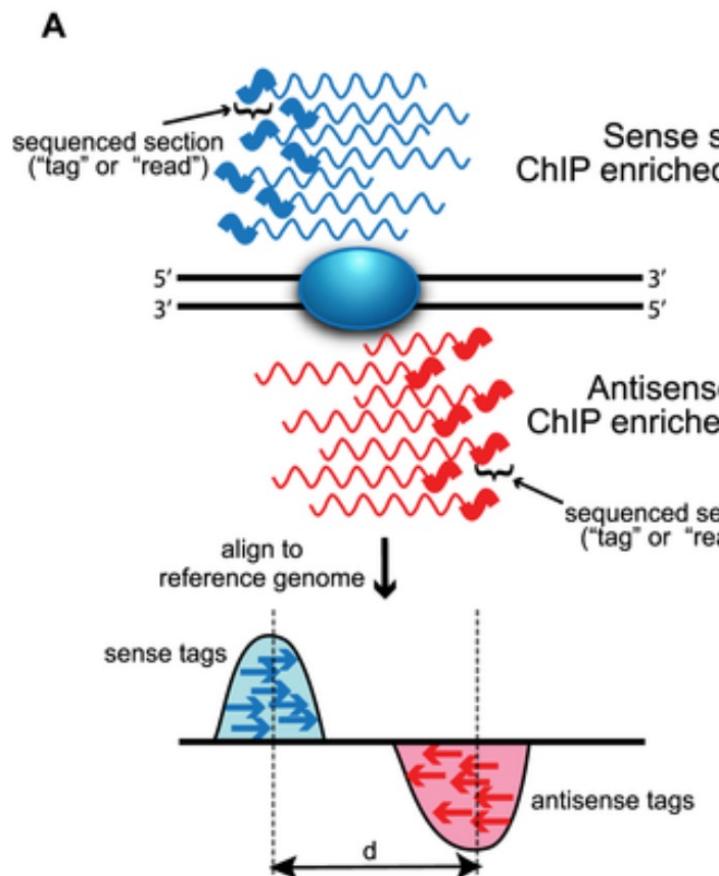
ChIP-seq peak calling downstream analyses

- Validation (wet lab)
- Downstream analysis
 - Motif discovery
 - Annotation
 - Integration of binding and expression data
 - Integration of various binding datasets
 - Differential binding

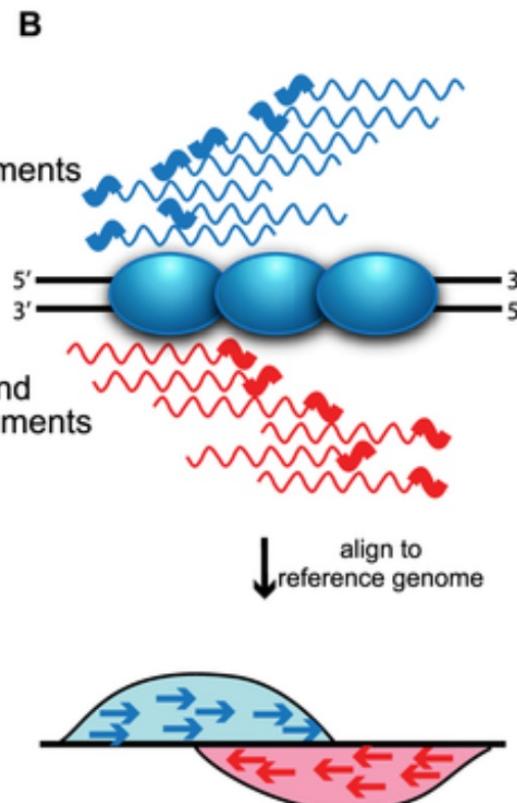


Peak detection

Sequence-specific binding (TFs)

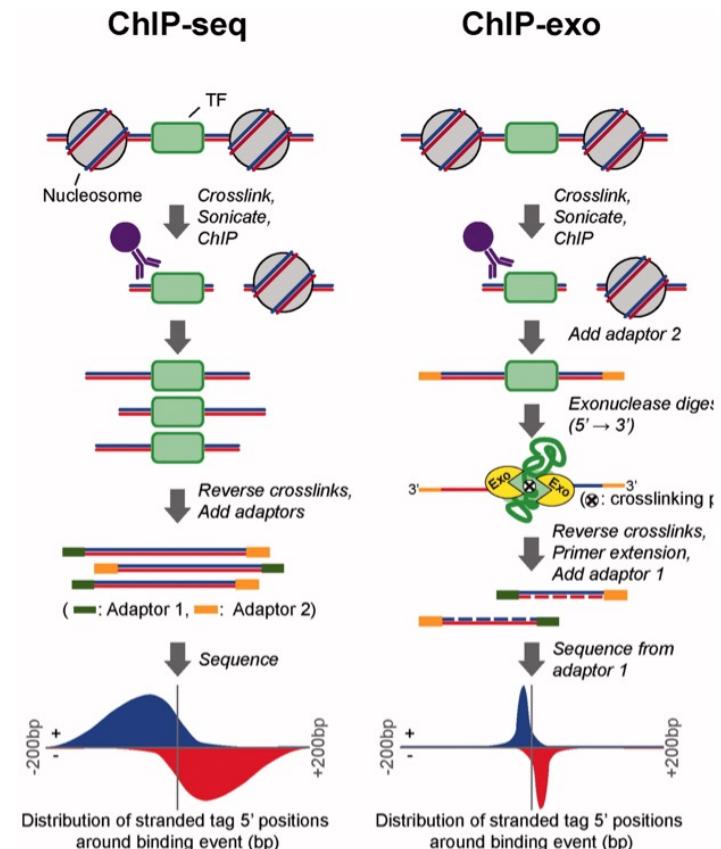
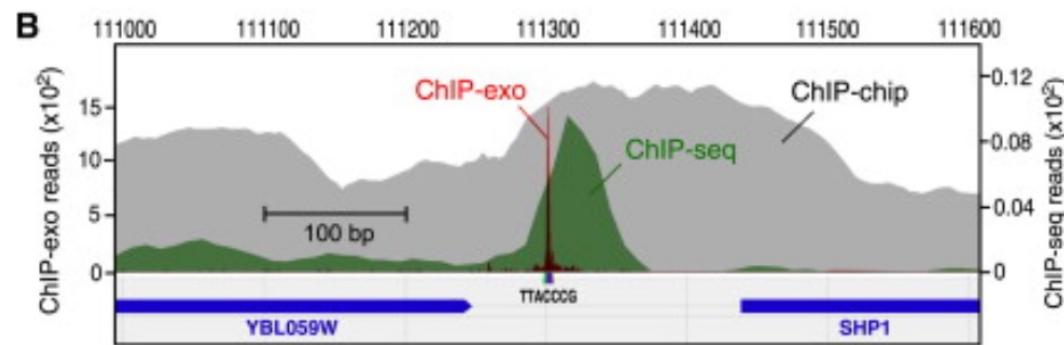
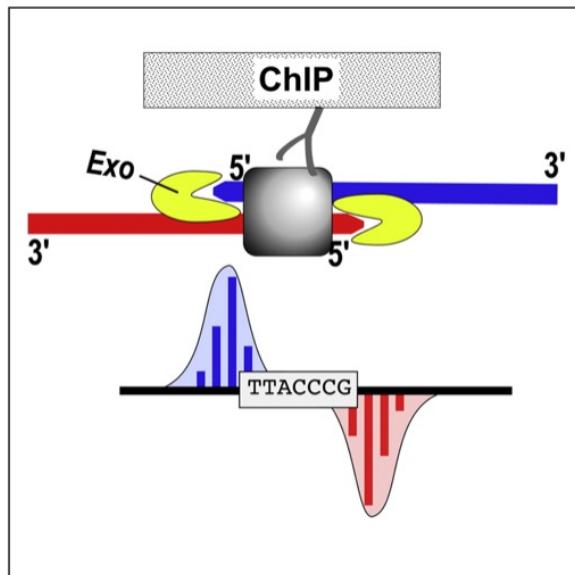


Distributed binding (histones, RNAPol2)



Wilbanks 2010

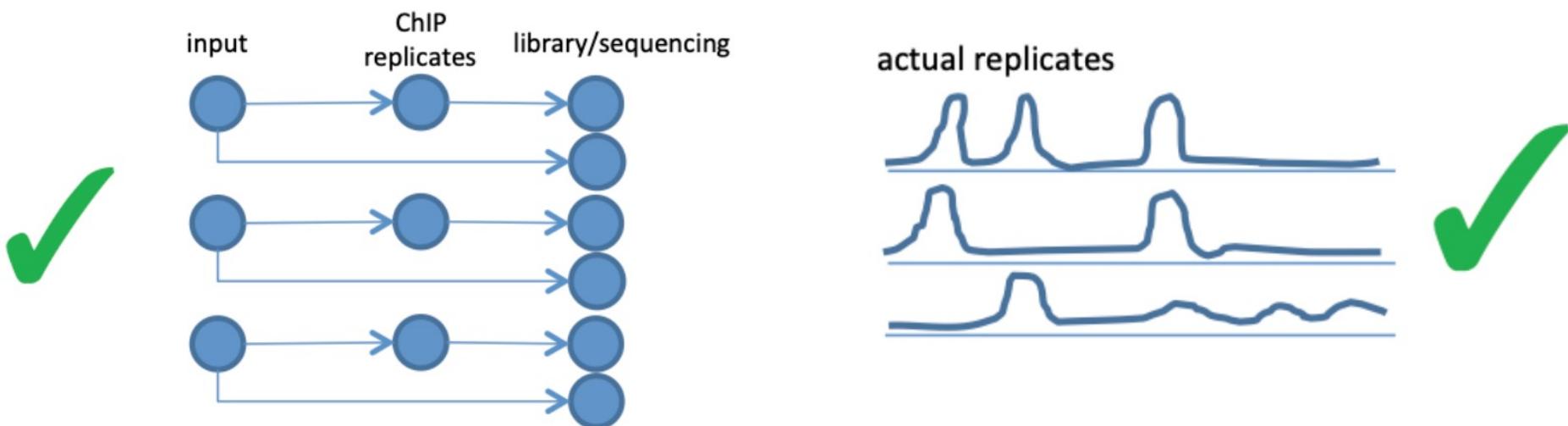
ChIP-exo: improvement in binding site identification



Pugh 2015
Rhee and Pugh, Cell 2011

Experiment design

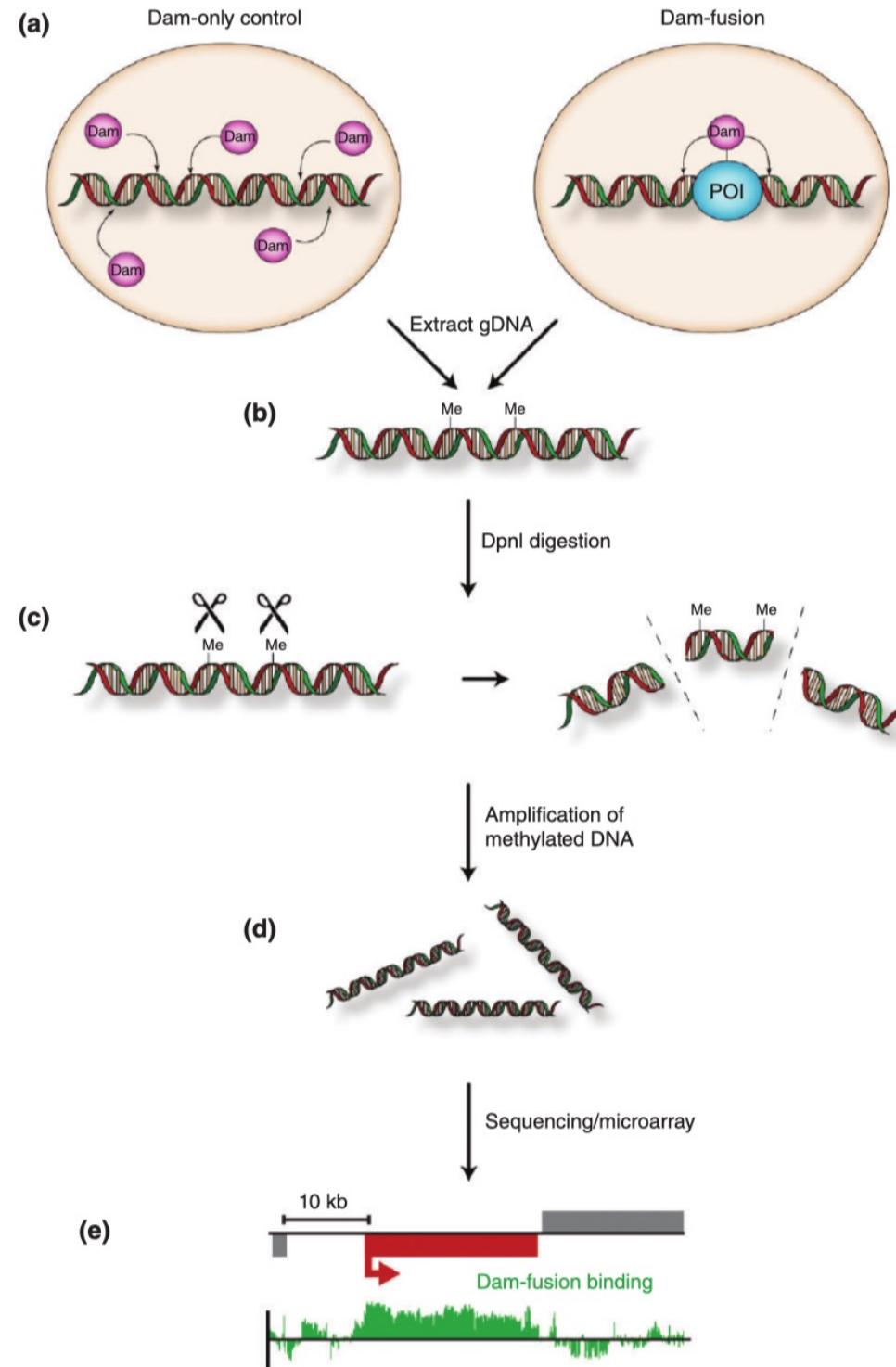
- Sound experimental design: replication, randomisation, control and blocking (R.A. Fisher, 1935)
- In the absence of a proper design, it is essentially impossible to partition biological variation from technical variation
- Please visit section *Experimental Design and Data Management* on the course website for more information



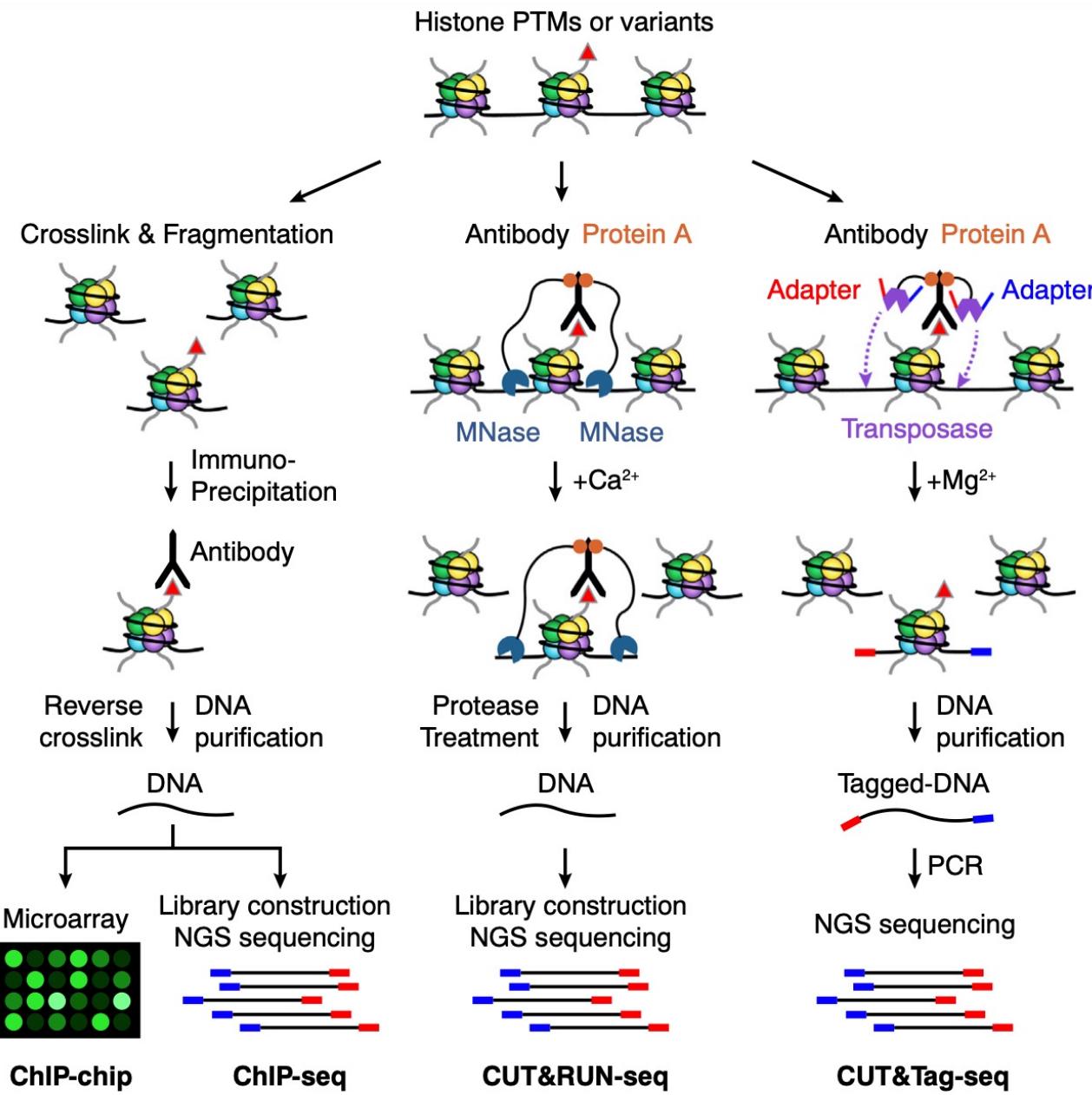
ChIP-Seq alternatives

- Dam-ID
- CUT&RUN
- CUT&Tag

Dam-ID



CUT&Run, CUT&Tag



ChIC, ChEC-Seq CUT&Run

Molecular Cell, Vol. 16, 147–157, October 8, 2004, Copyright ©2004 by Cell Press

ChIC and ChEC: Genomic Mapping of Chromatin Proteins

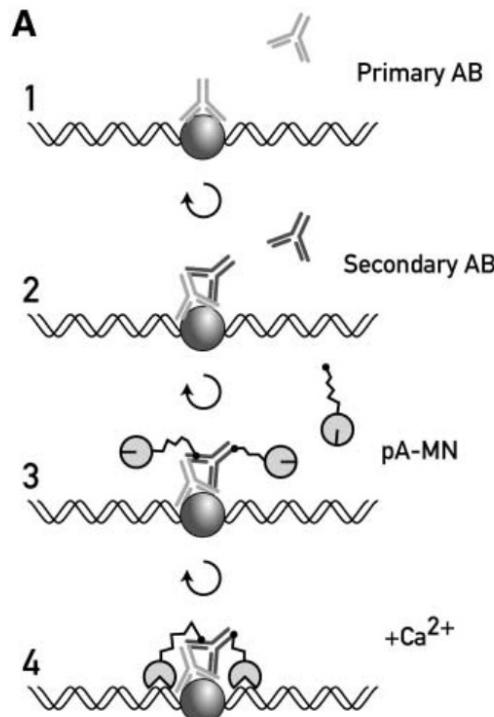
Manfred Schmid, Thérèse Durussel,
and Ulrich K. Laemmli*
Departments of Biochemistry and Molecular Biology
NCCR Frontiers in Genetics
University of Geneva
30, Quai Ernest-Ansermet
CH1211, Geneva 4
Switzerland

Technique

ChIC: Chromatin-immuno cleavage

ble, and significant amounts are lost into the pellet during centrifugation.

While ChIP is highly successful when applied to soluble proteins, such as transcription regulatory proteins, unpublished experiments with insoluble-type proteins (such as scaffolding components) in this laboratory appeared less promising. ChIP analyses with such insoluble-type proteins appear afflicted with increased background.



HOME MAGAZINE INNOVATION

ABOUT COMMUNITY SUBMIT MY RESEARCH LOG IN/REGISTER



Chromosomes and Gene Expression



An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites



Peter J Skene, Steven Henikoff*

Howard Hughes Medical Institute, Basic Sciences Division, Fred Hutchinson Cancer Research Center, United States

Brief Communication | Published: 28 March 2019

Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification

Wai Lim Ku, Kosuke Nakamura, Weiwu Gao, Kairong Cui, Gangqing Hu, Qingsong Tang, Bing Ni✉ & Keji Zhao✉

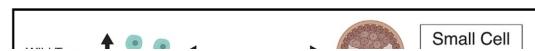
Nature Methods 16, 323–325(2019) | Cite this article

Resource

Cell

Profiling of Pluripotency Factors in Single Cells and Early Embryos

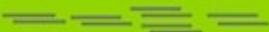
Graphical Abstract



Authors

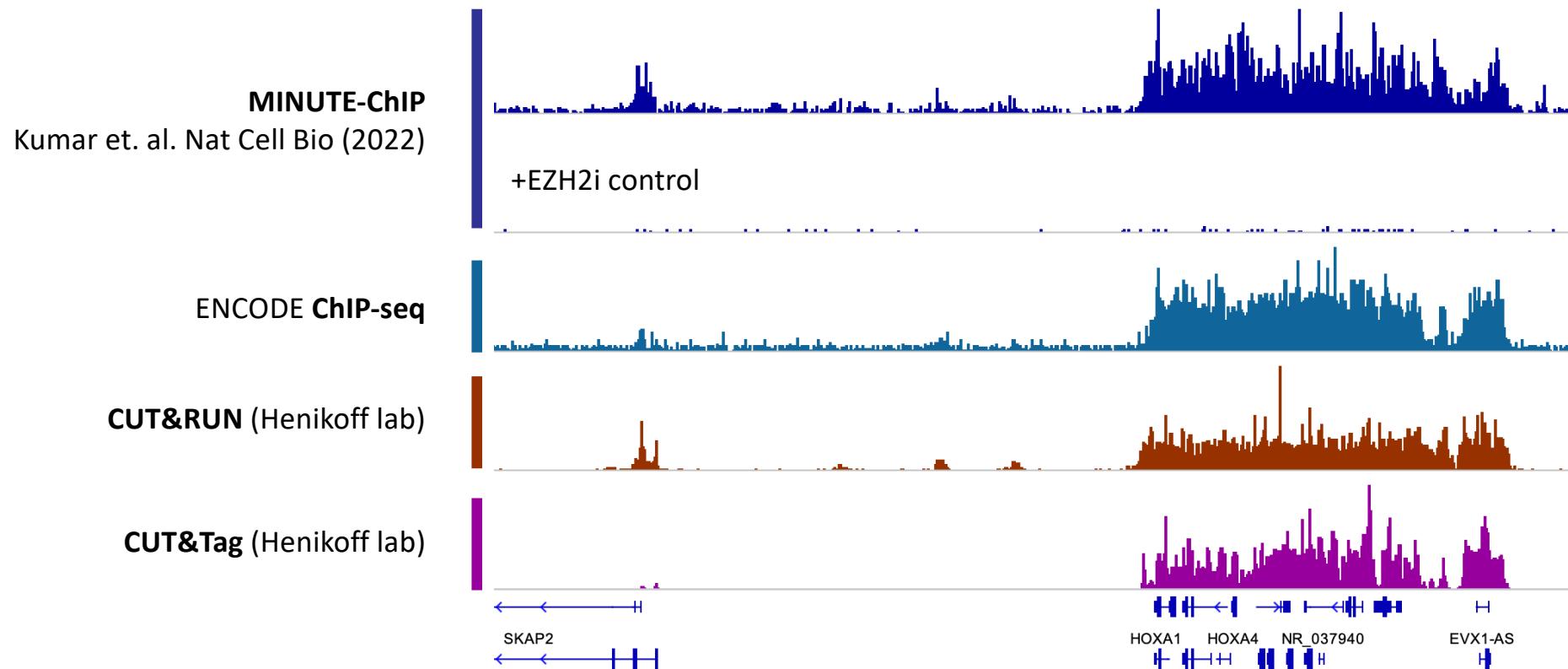
Sarah J. Hainer, Ana Bošković,
Kurtis N. McCannell, Oliver J. Rando,

CUT&Run, CUT&Tag

Features	ChIP-Seq	CUTANA™ CUT&RUN	CUTANA™ CUT&Tag
Sample Input	Sheared Chromatin 	Cells OR nuclei 	Nuclei (recommended) 
Typical Required Cell #	> 1 Million	500K	100K
Ideal Targets	Histone PTMs & chromatin-interacting proteins	Histone PTMs & chromatin-interacting proteins, including remodelers	Histone PTMs & select validated targets
Secondary Antibody	No	No	Yes
Library Preparation	Yes	Yes	No (Direct to PCR)
Protocol Time (Cells → NGS libraries)	~ 1 week	2 days (can be automated)	2 days (can be automated)
Sequencing Depth	> 30 million	3-5 million	3-5 million
Signal : Noise	Low	High	High
Experimental Throughput	Low	High	High

Excercise – CUT&RUN vs CUT&Tag vs ChIP

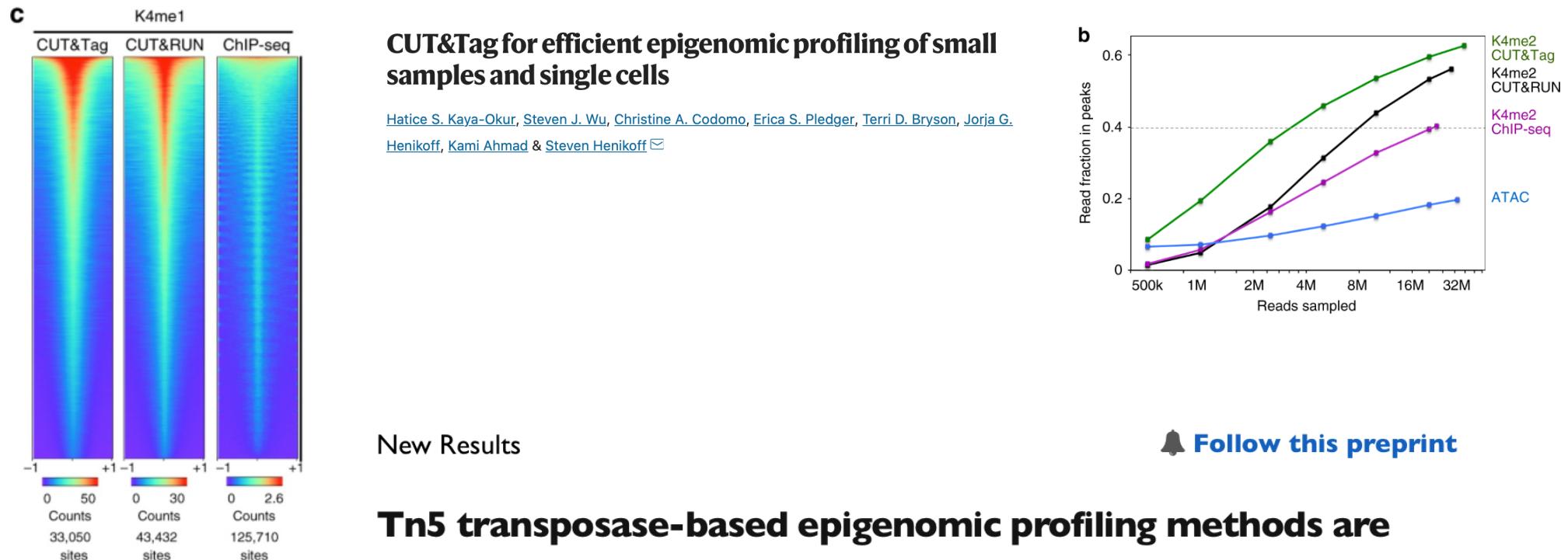
Comparison: H3K27me3 signal in human embryonic stem cells across different methods



<https://nbis-workshop-epigenomics.readthedocs.io/en/latest/content/tutorials/quantitativeChip/cut-and-tag-data.html>

Weighing pros and cons

- Cell number needed
- Time aspect
- Signal-to-noise
- Background (think technical versus biological!)
- Reproducibility
- QCability (e.g. no input for CUT&Tag, CUT&RUN)



New Results

🔔 [Follow this preprint](#)

Tn5 transposase-based epigenomic profiling methods are prone to open chromatin bias

Meng Wang, Yi Zhang

doi: <https://doi.org/10.1101/2021.07.09.451758>

New Results

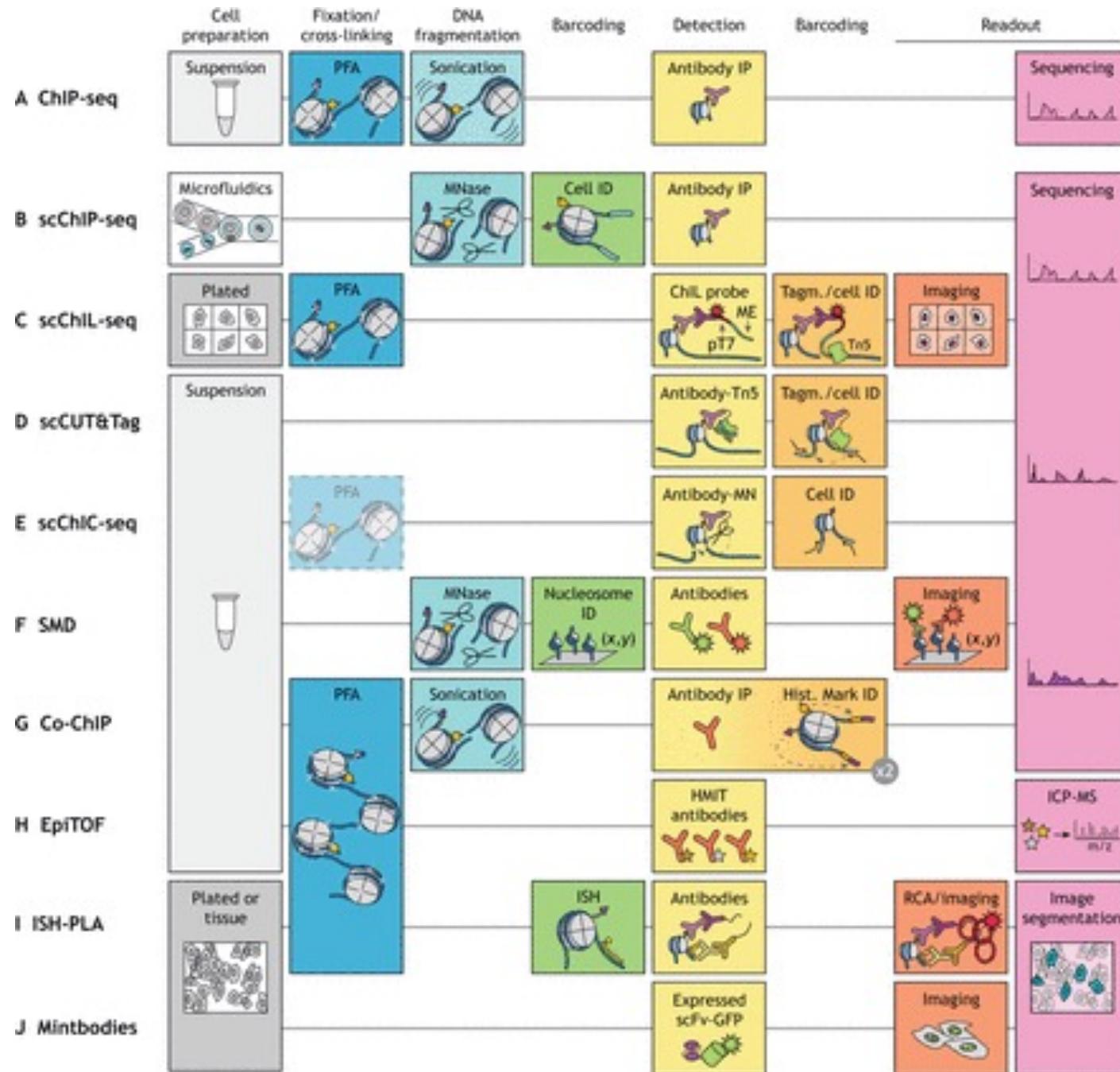
🔔 [Follow this preprint](#)

CUT&Tag recovers up to half of ENCODE ChIP-seq peaks

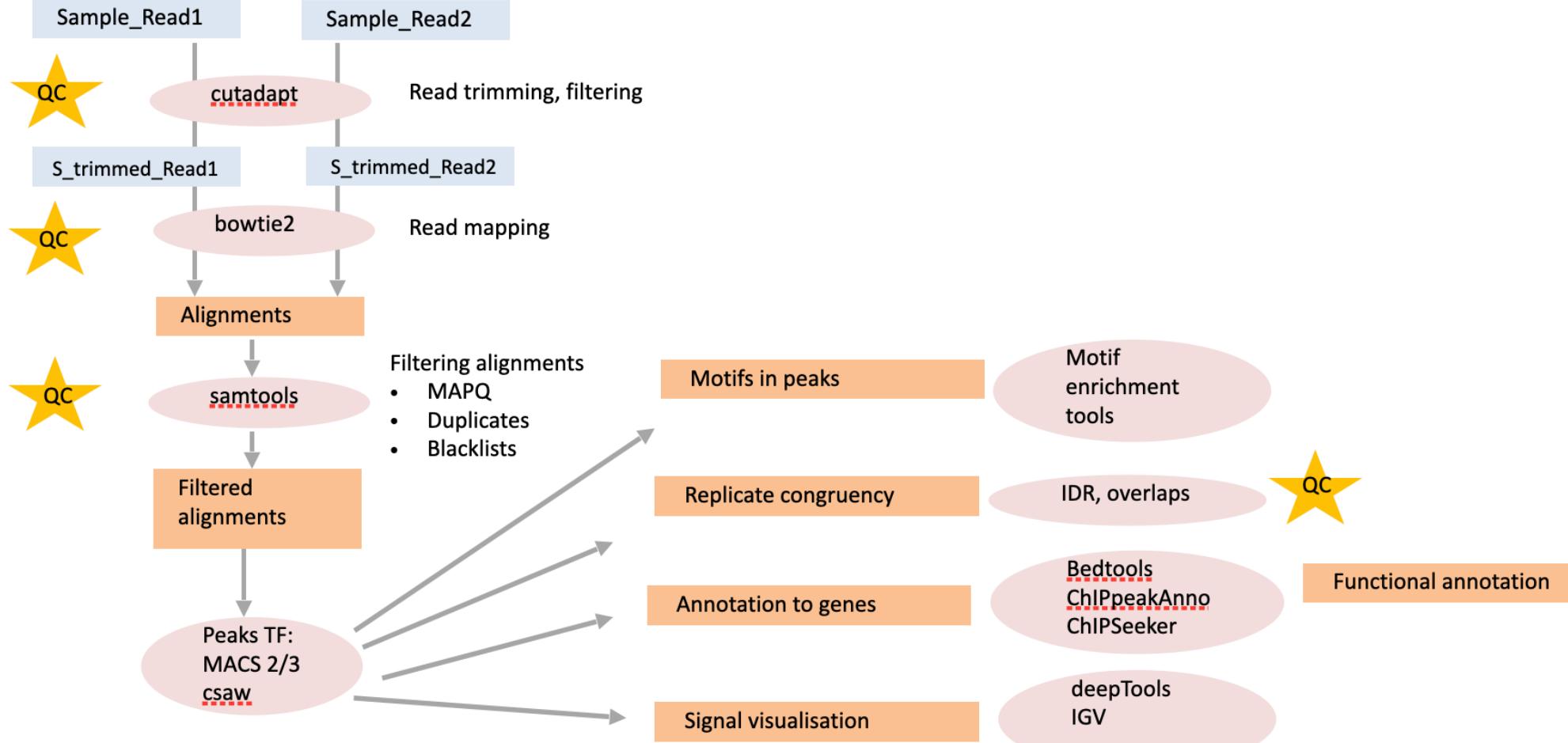
Di Hu, Leyla Abbasova, Brian M Schilder, Alexi Nott, Nathan G Skene, Sarah J Marzi

doi: <https://doi.org/10.1101/2022.03.30.486382>

Single-cell revolution



Analysis



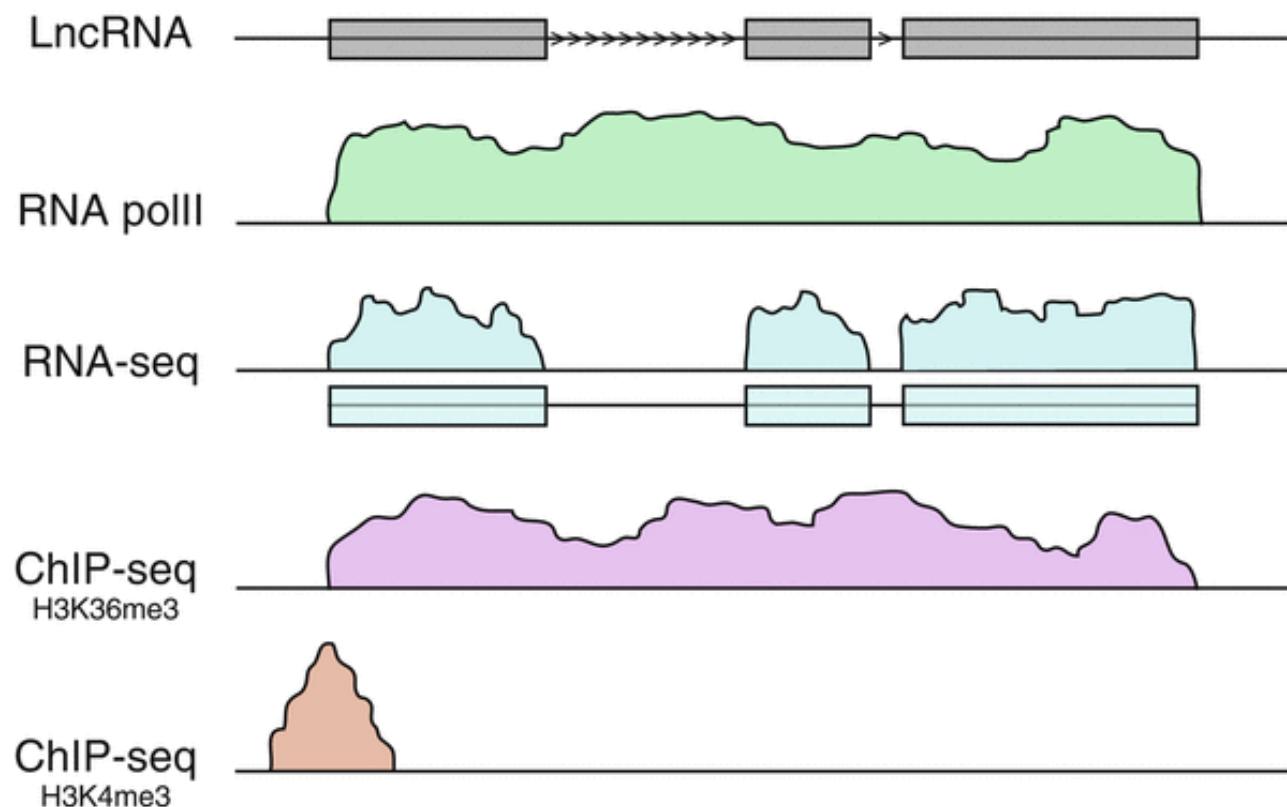
Applications of ChIP-Seq and related methods

- Map features to genome (--> knowing where a feature is *may imply* function)
- Specific versus genome-wide feature
- Discover genome-wide correlations (--> generate experimental hypothesis --> test to establish causation)
 - Think about the meaning: “repressive chromatin”, “activating mark”, “silencing factor”
 - A histone PTM “recruits” a factor
 - A factor ‘protects’ a gene from spurious transcription
- Compare conditions, e.g. how does a knockout, inhibitor, external change in condition affect epigenome?

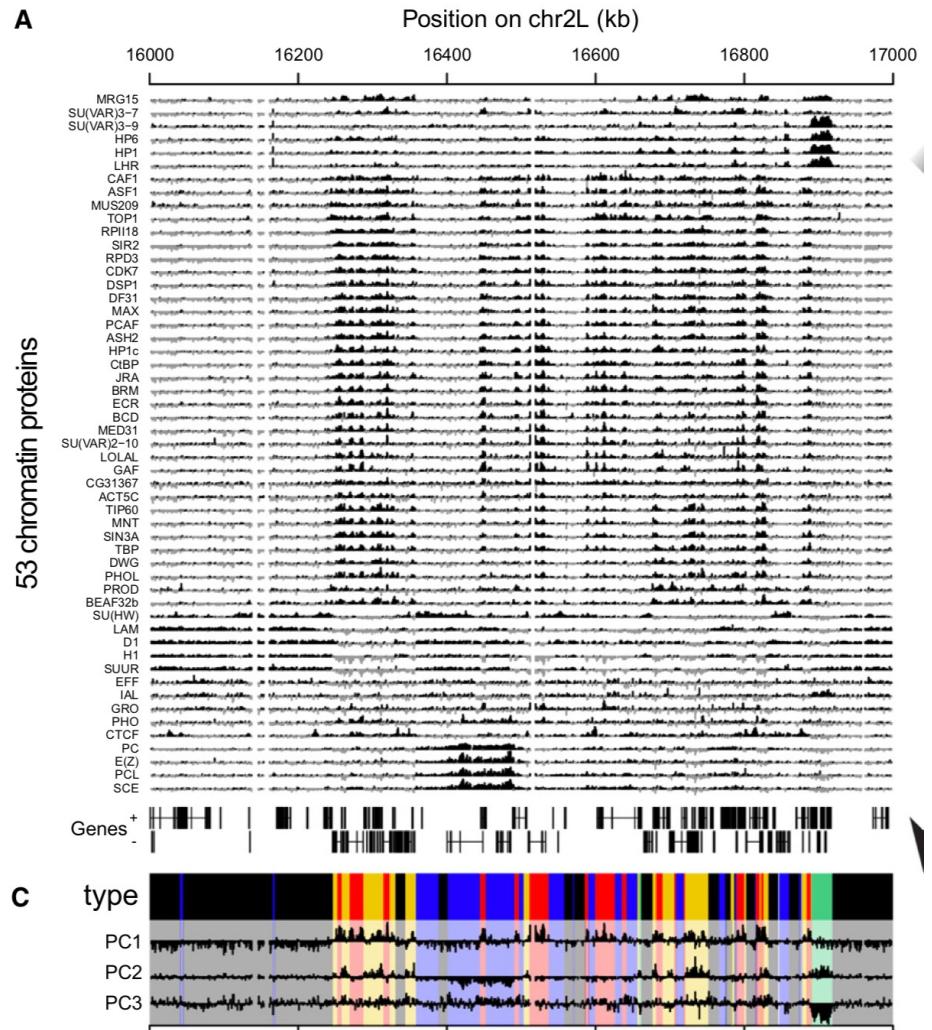
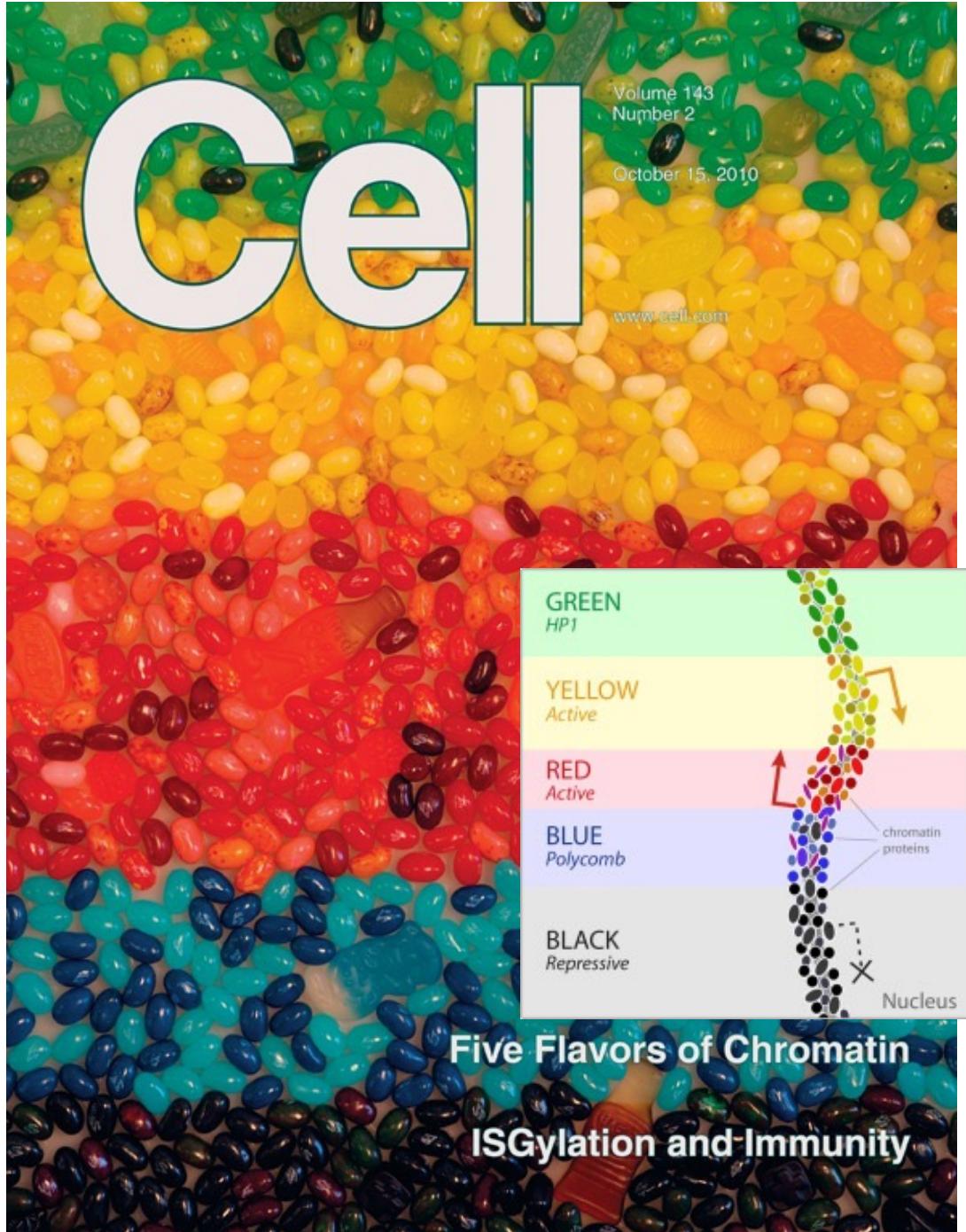
Using histone modifications to predict functional regions in the genome

Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals

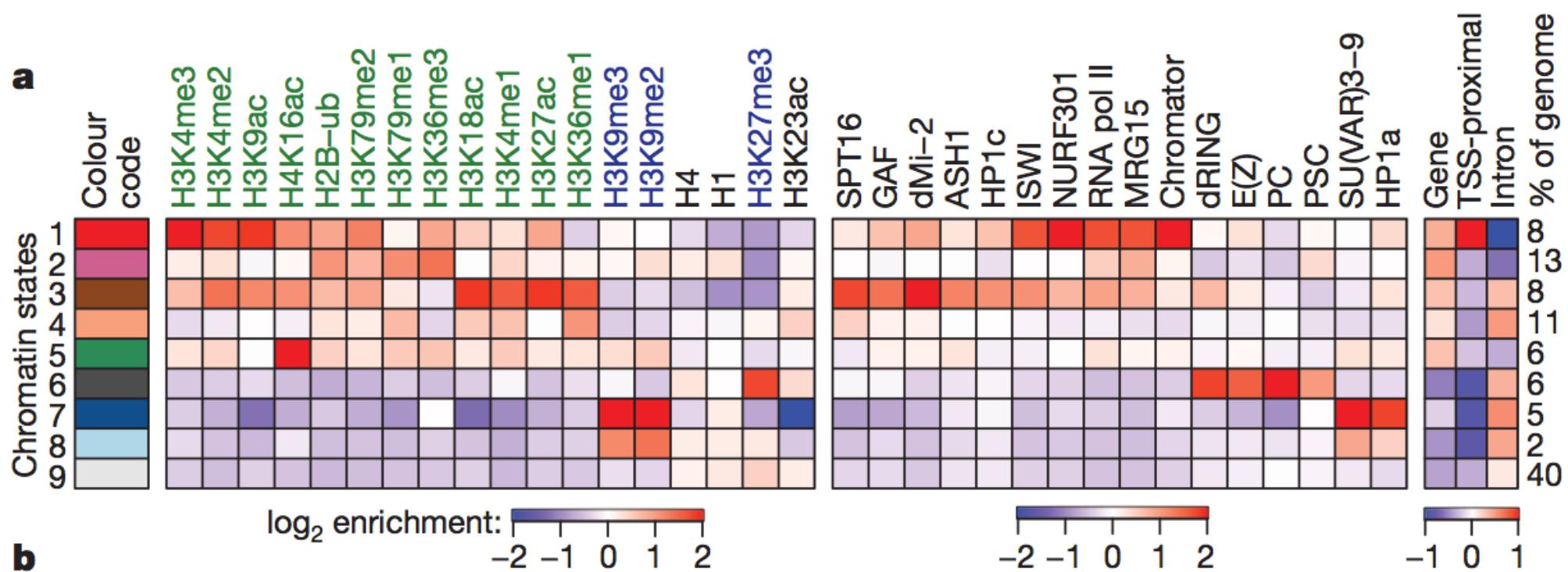
Mitchell Guttman^{1,2}, Ido Amit¹, Manuel Garber¹, Courtney French¹, Michael F. Lin¹, David Feldser³, Maite Huarte^{1,6},



Chromatin 'states' define functional regions

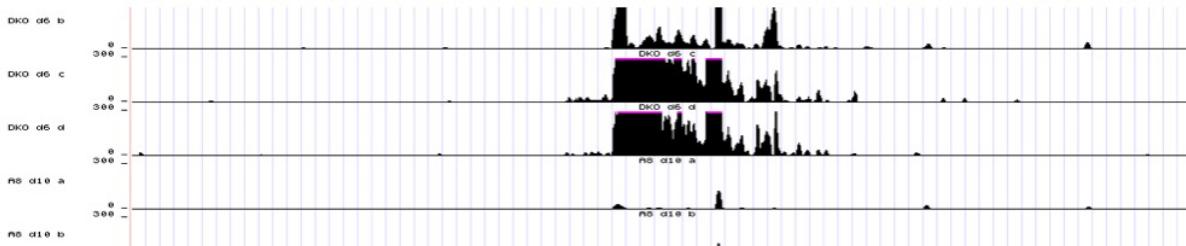


9 chromatin states defined by a combinatorial pattern of enrichment and depletion for specific chromatin marks

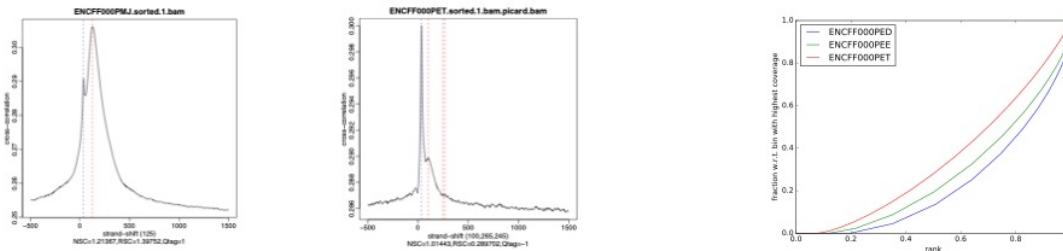


ChIP-seq QC: did the ChIP work?

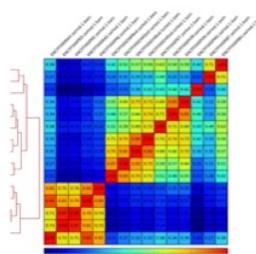
- 1. Inspect the signal (mapped reads, coverage profiles) in genome browser



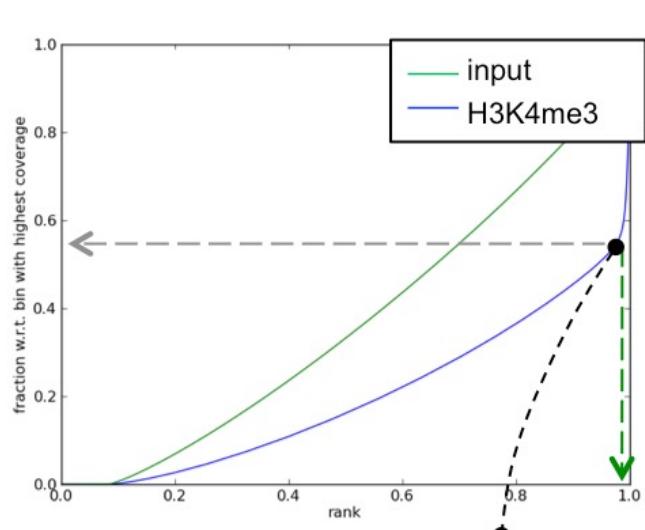
- 2. Compute peak-independent quality metrics (cross correlation, cumulative enrichment)



- 3. Assess replicate consistency (correlations between replicates of the same condition)

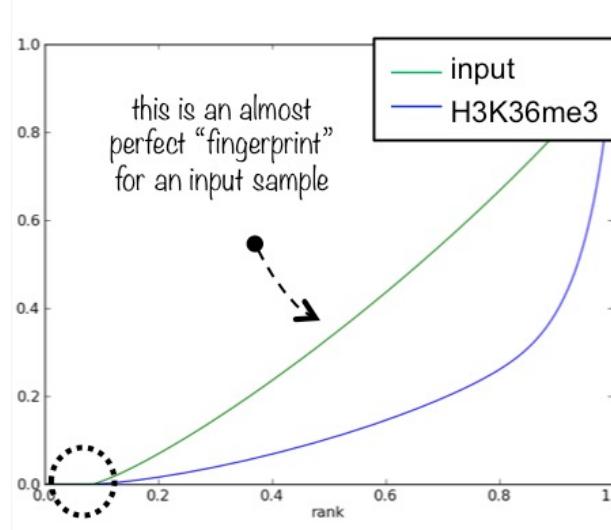


Fingerprint plot (deepTools)

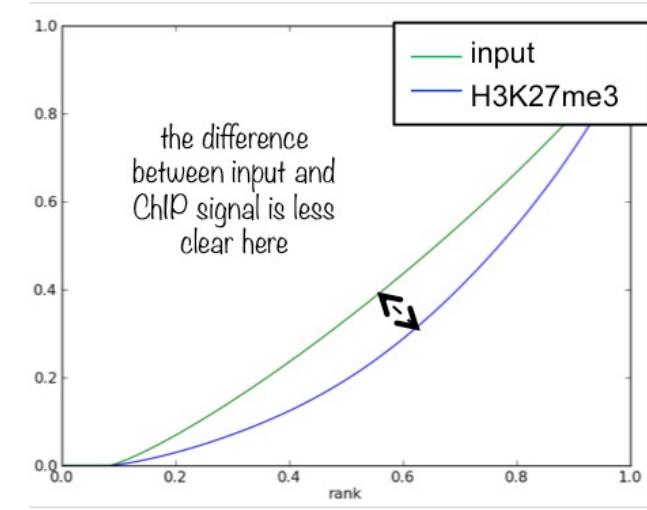


when counting the reads contained in **97%** of all genomic bins, only ca. **55%** of the maximum number of reads are reached, i.e. 3% of the genome contain a very large fraction of reads!

→ this indicates very localized, very strong enrichments!
(as every biologist hopes for in a ChIP for H3K4me3)



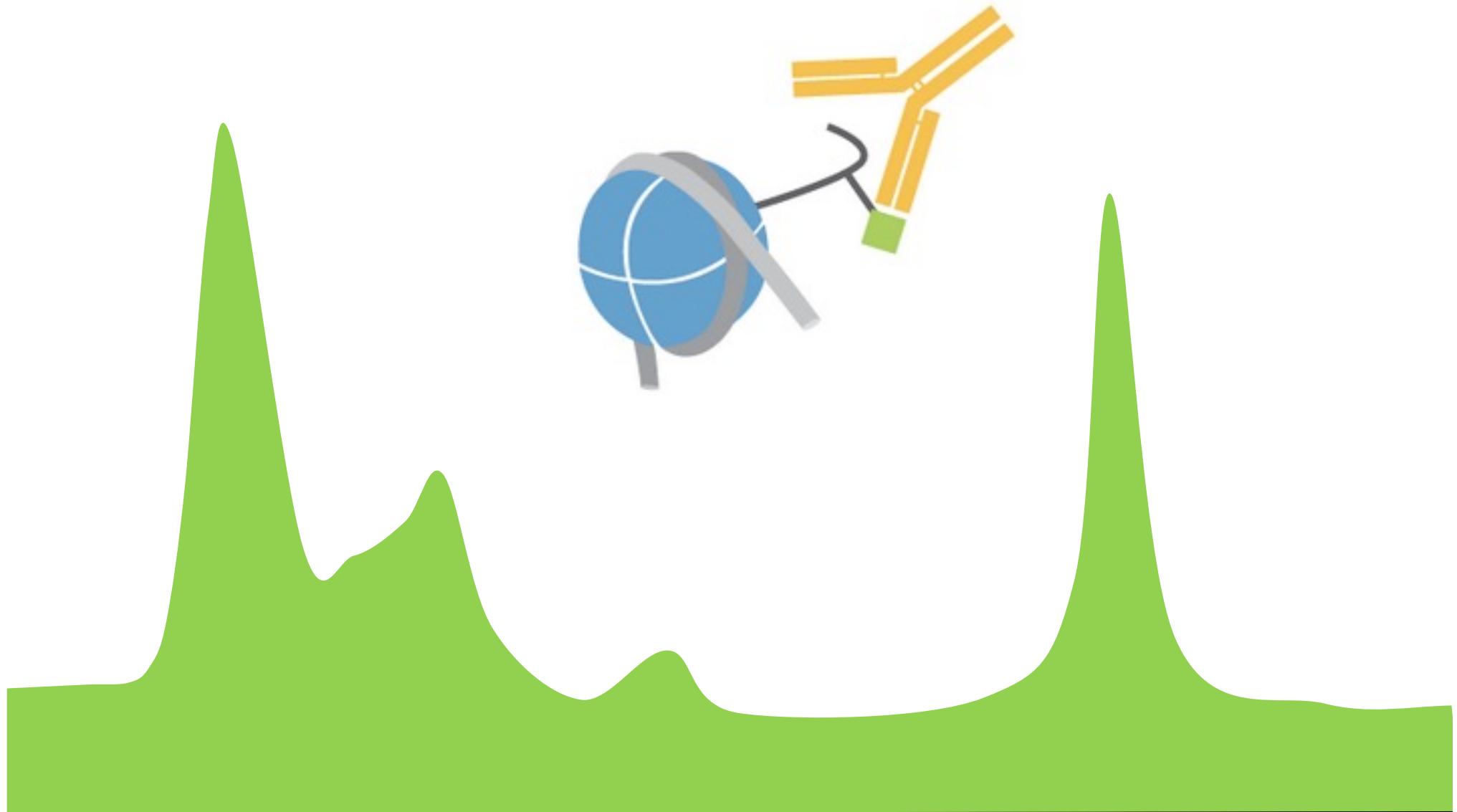
this is an almost perfect “fingerprint” for an input sample
pay attention to where the curves start to rise – this already gives you an assessment of how much of the genome you have not sequenced at all (i.e. bins containing zero reads – for this example, ca. 10% of the entire genome do not have any read)



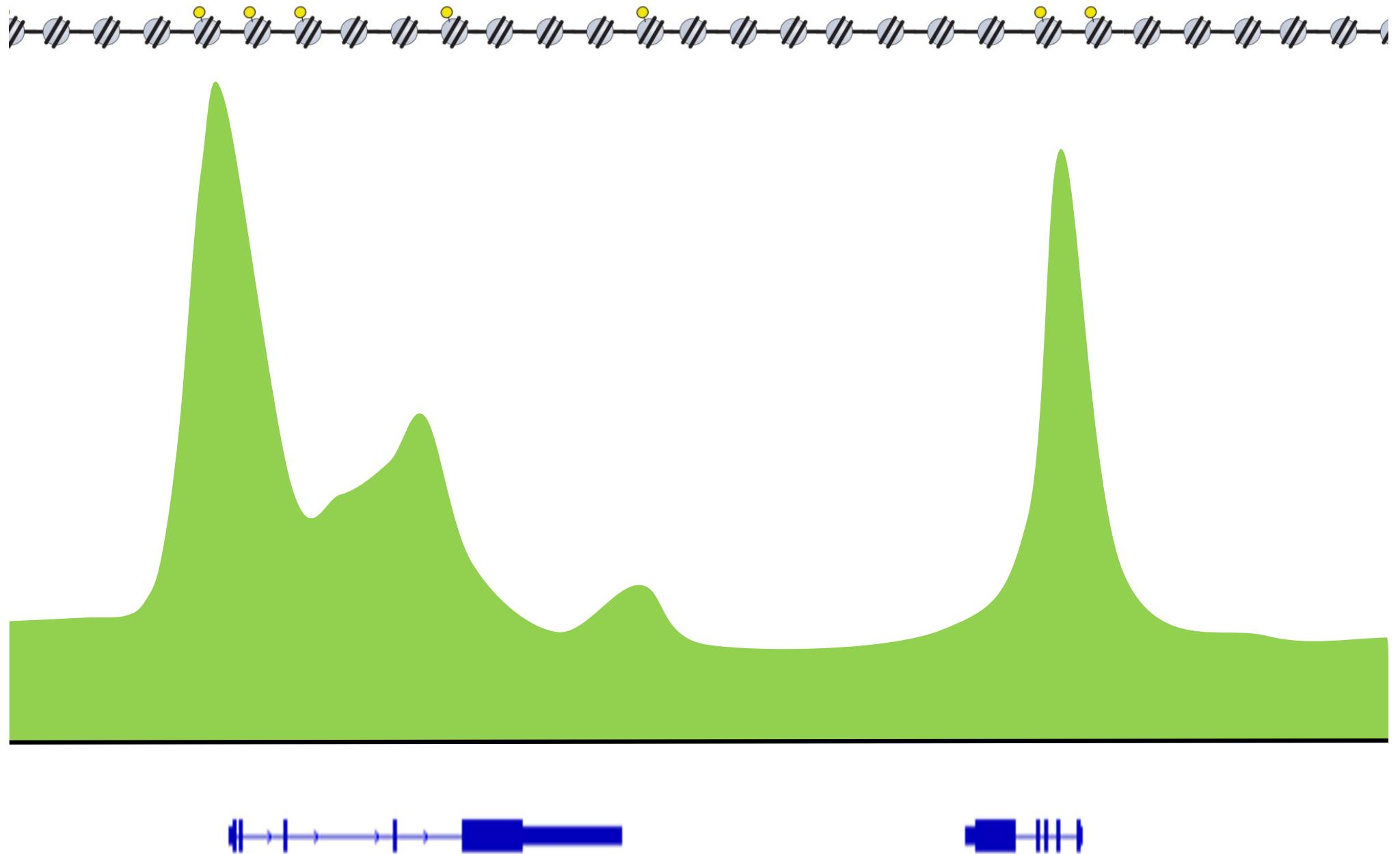
the difference between input and ChIP signal is less clear here
H3K27me3 is a mark that yields broad domains instead of narrow peaks

it is more difficult to distinguish input and ChIP, it does not mean, however, that this particular ChIP experiment failed

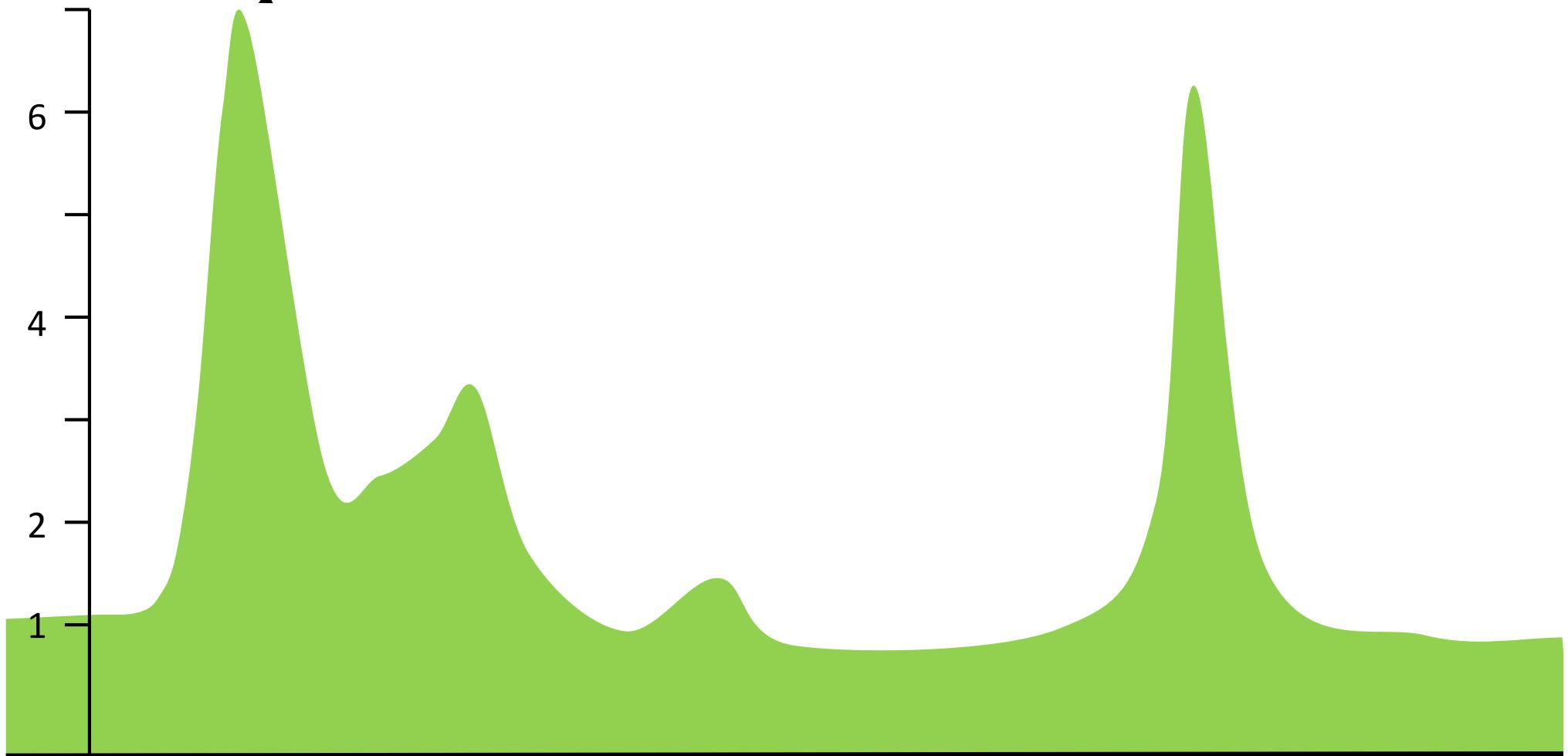
ChIP-Seq signal (histone PTM or Binding Protein occupancy)

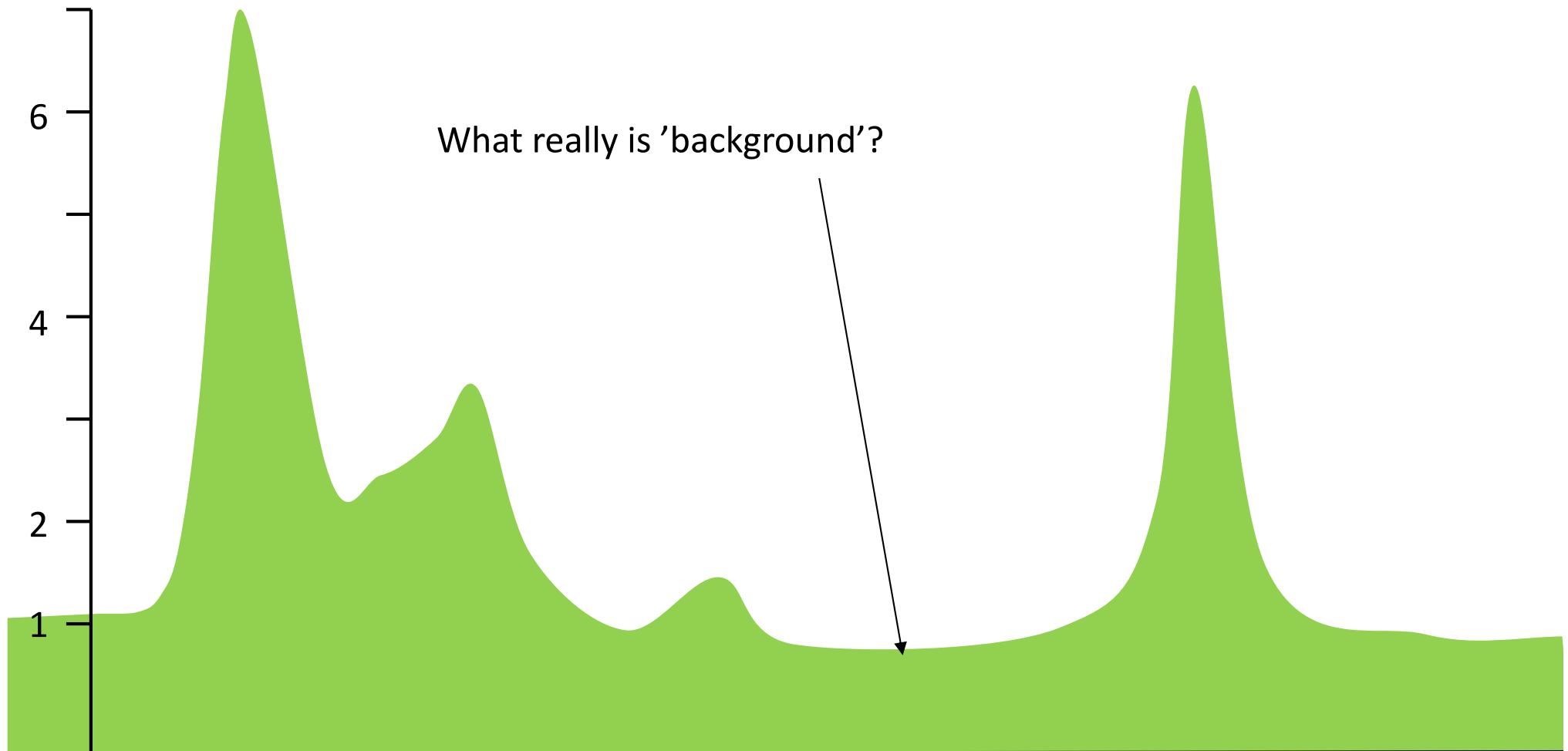


Interpreting ChIP-Seq signal

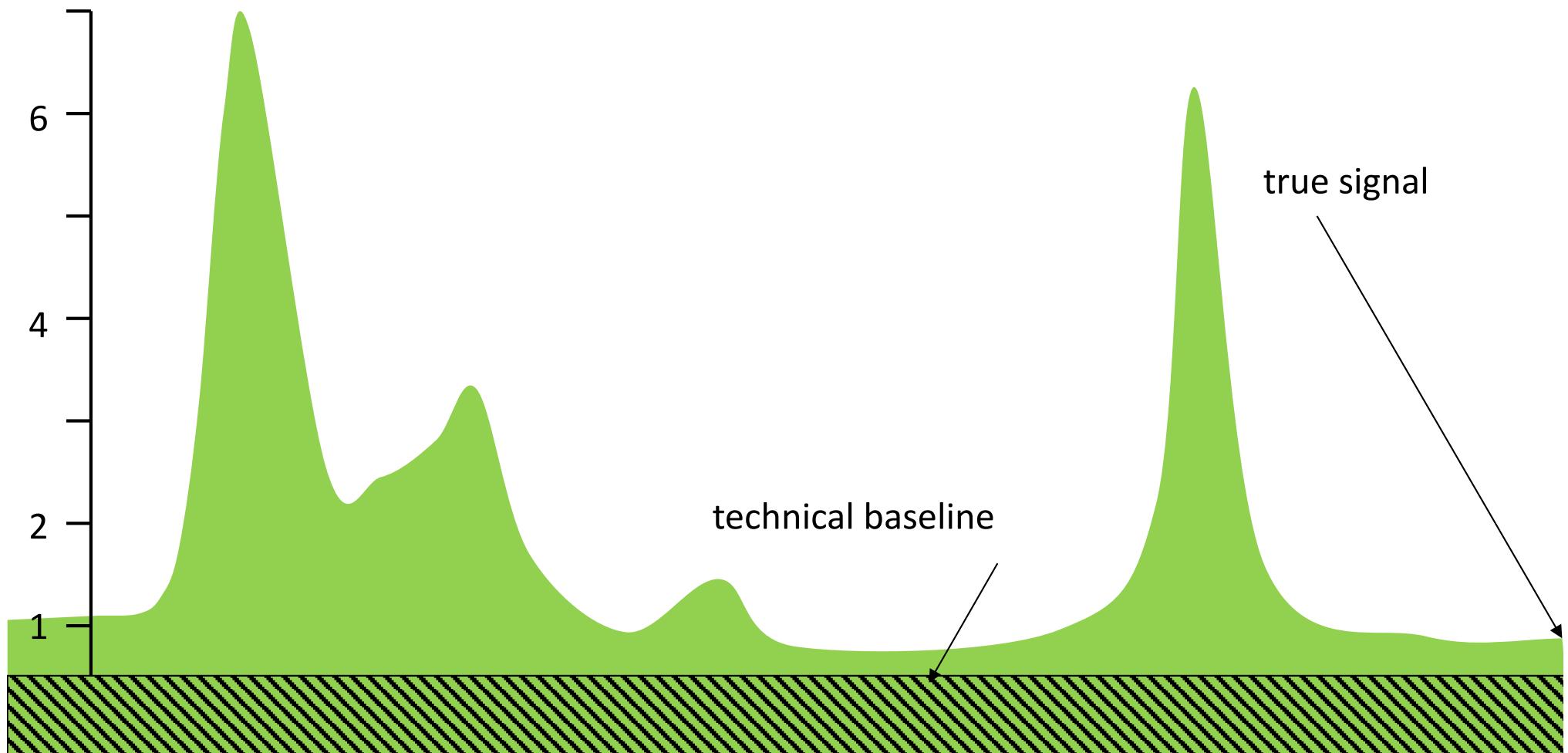


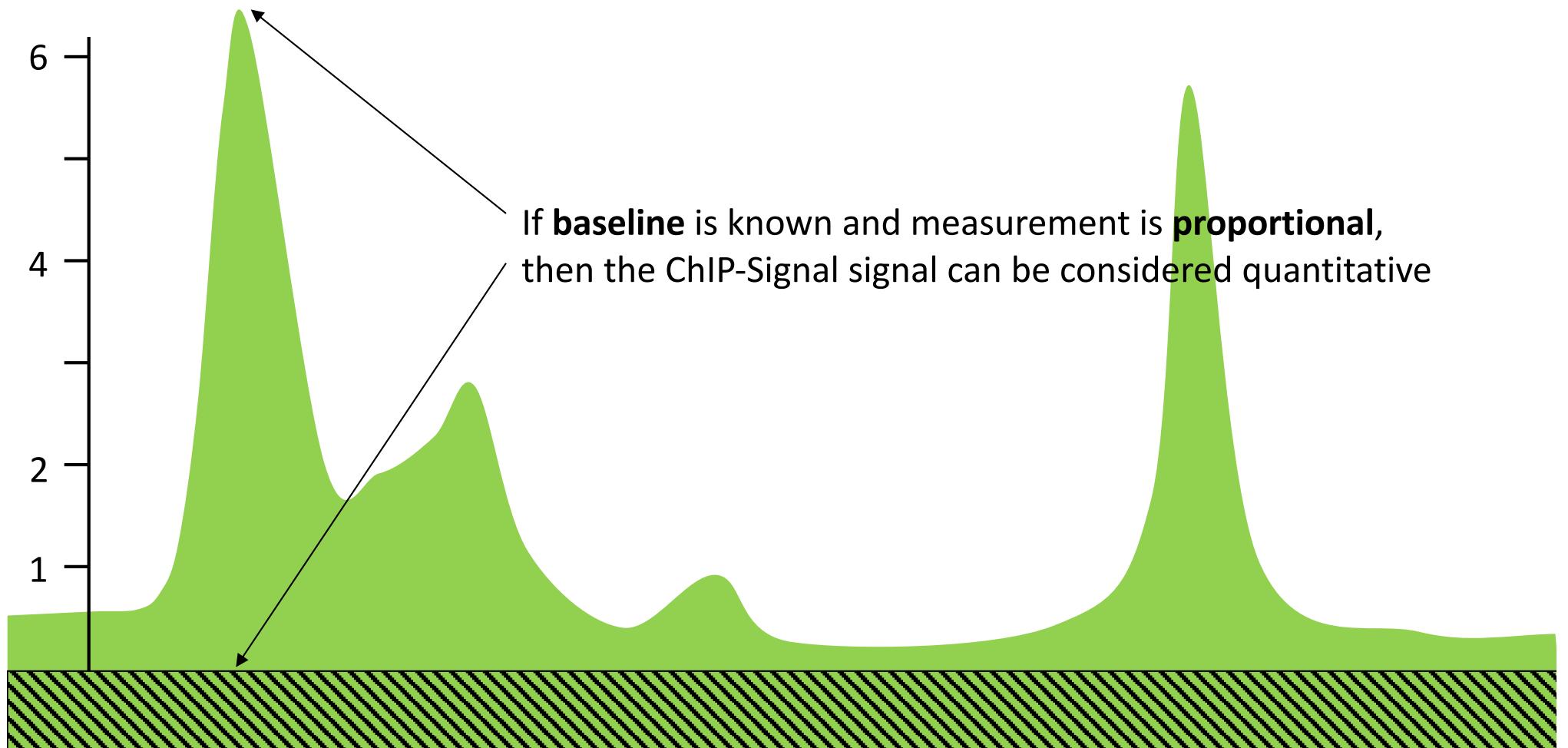
What does the peak height mean?





What really is 'background'?





Good ChIP requires good Antibody

Potential problems with antibodies:

- specificity modified histone (crossreactivity for other histone sites/modifications)
- specificity for target protein
- cross-reactivity to other epitopes
- Affinity/Avidity of the interaction and stability against harsh wash conditions
- sensitivity to formaldehyde modification of the target protein

2010

Cell

ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner

Martin J. Law,^{1,8} Karen M. Lower,^{1,8} Hsiao P.J. Voon,¹ Jim R. Hughes,¹ David Garrick,¹ Vip Viprakasit,³ Matthew Mitson,¹ Marco De Gobbi,¹ Marco Marra,⁷ Andrew Morris,⁴ Aaron Abbott,⁴ Steven P. Wilder,⁵ Stephen Taylor,² Guilherme M. Santos,⁶ Joe Cross,¹ Helena Ayyub,¹ Steven Jones,⁷ Jiannis Ragoussis,⁴ Daniela Rhodes,⁶ Ian Dunham,⁵ Douglas R. Higgs,¹ and Richard J. Gibbons^{1,*}

2013

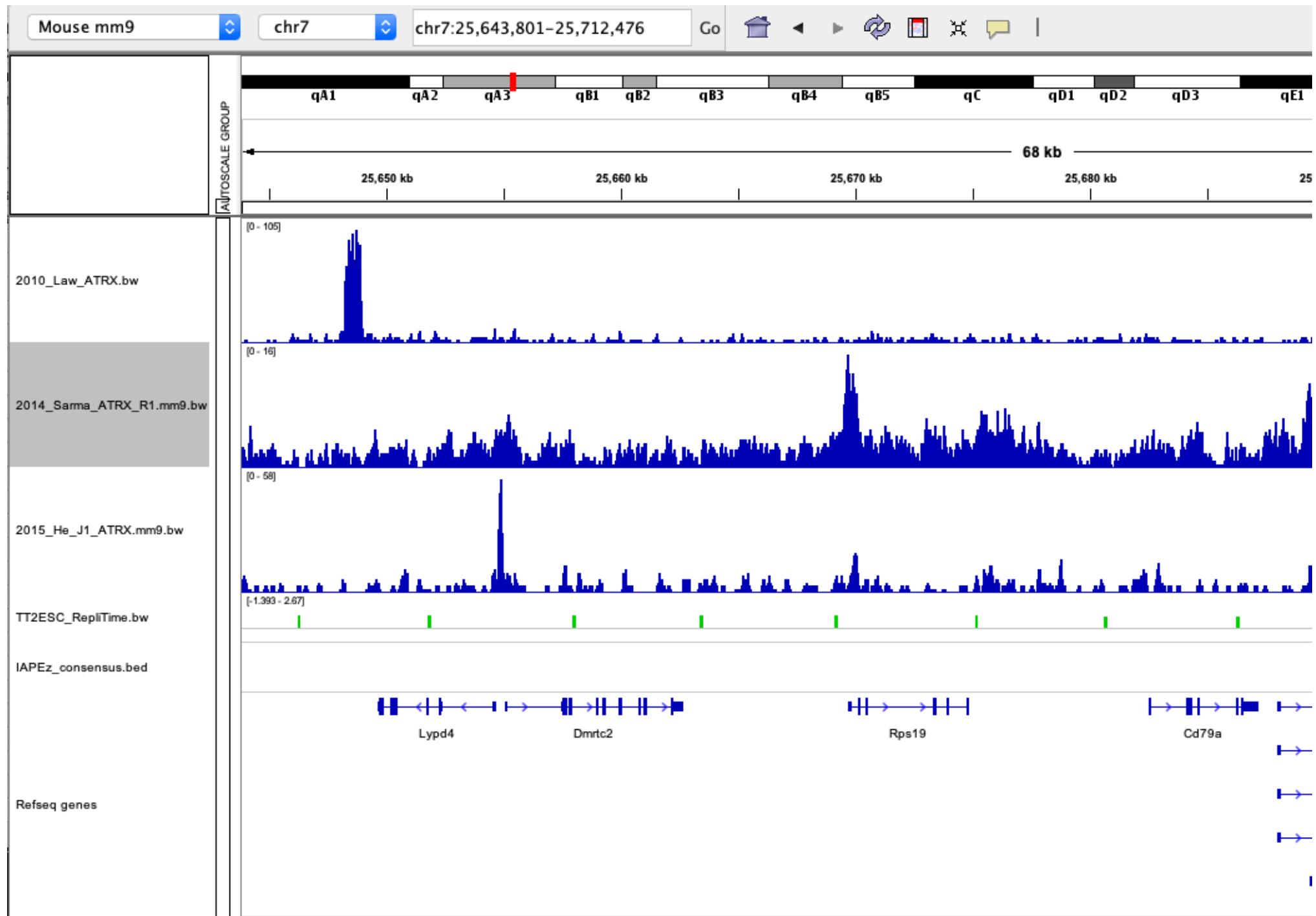
Cell

Article

ATRX Directs Binding of PRC2 to Xist RNA and Polycomb Targets

Kavitha Sarma,^{1,2,3} Catherine Cifuentes-Rojas,^{1,2,3} Ayla Ergun,^{2,3} Amanda del Rosario,⁵ Yesu Jeon,^{1,2,3} Forest White,⁵ Ruslan Sadreyev,^{2,3,4} and Jeannie T. Lee^{1,2,3,4,*}



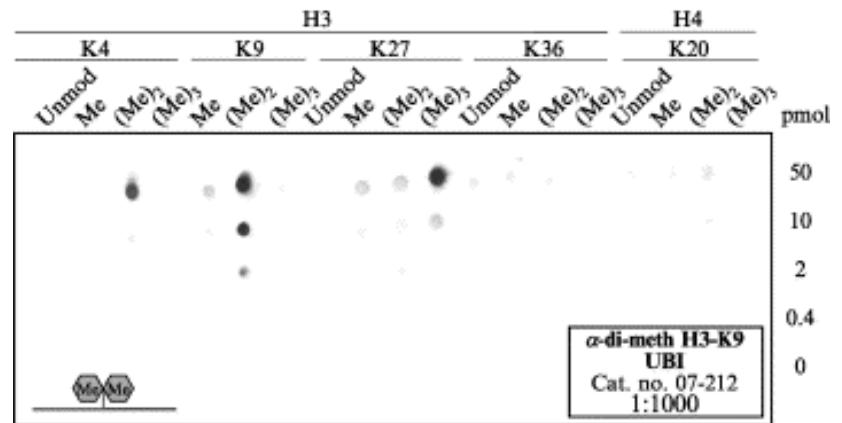


Good ChIP requires good Antibody

Potential problems with antibodies:

background affinity for unmodified peptide epitope proximal to the site of modification

- specificity for seq. context



Good ChIP requires good Antibody

The worse the antibody, the more ChIP will look like input

And normalizing for uneven input is tricky! Options

- ratio ChIP versus background
- background subtraction



Cistrome Data Browser

ⓘ Tips

- Check what factors regulate your gene of interest, what factors bind in your interval or have a significant binding overlap with your peak set. Have a try at [CistromeDB Toolkit](#).
- If you have a Transcription Factor ChIP-seq (and TF perturbed expression) data, [Cistrome-GO](#) help you predict the function of this TF.
- Please help us curate the samples which has incorrect meta-data annotation by clicking the button on the inspector page. Thank you!

Containing word(s):

Search

Options ▾

Species

All

Homo sapiens

Mus musculus

Biological Sources

All

1-cell pronuclei

1015c

10326

1064Sk

106A

Factors

All

AATF

ABCC9

ACSS2

ACTB

ADNP

Results

Batch	Species	Biological Source	Factor	Publication	Quality Control
<input type="checkbox"/>	Homo sapiens	HeLa; Epithelium; Cervix	BTAF1	Johannes F, et al. Bioinformatics 2010	
<input type="checkbox"/>	Homo sapiens	HeLa; Epithelium; Cervix	GAPDH	Johannes F, et al. Bioinformatics 2010	

<http://cistrome.org/db/#/>

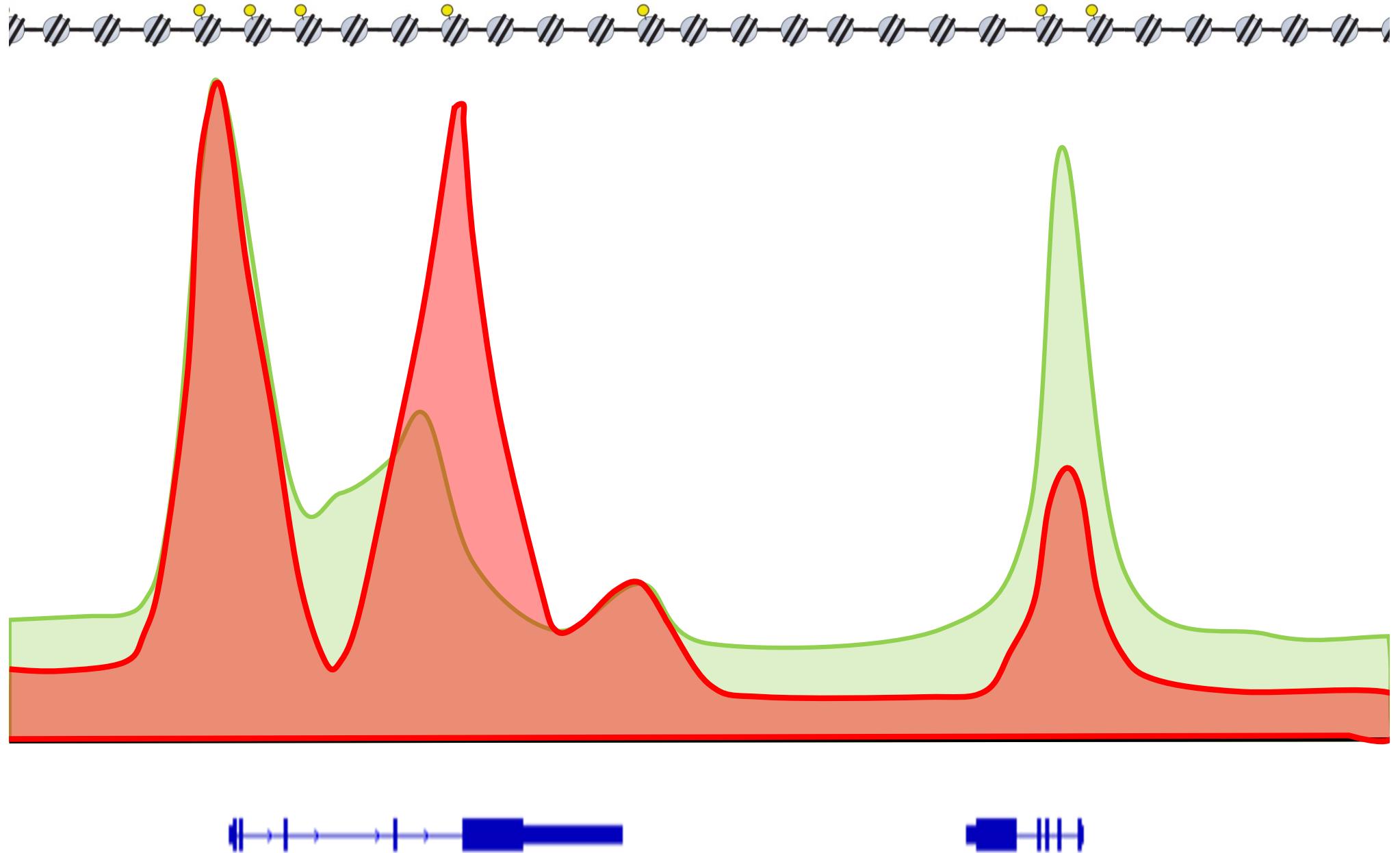


Applications of ChIP-Seq and related methods

- Map features to genome (--> knowing where a feature is *may imply* function)
- Specific versus genome-wide feature
- Discover genome-wide correlations (--> generate experimental hypothesis --> test to establish causation)
 - Think about the meaning: “repressive chromatin”, “activating mark”, “silencing factor”
 - A histone PTM “recruits” a factor
 - A factor ‘protects’ a gene from spurious transcription
- Compare conditions, e.g. how does a knockout, inhibitor, external change in condition affect epigenome?

These applications imply that you are interested in quantitatively comparing occupancies or levels of PTMs!

Interpreting ChIP-Seq signal



Relative quantitation Epigenomic profiles can be compared

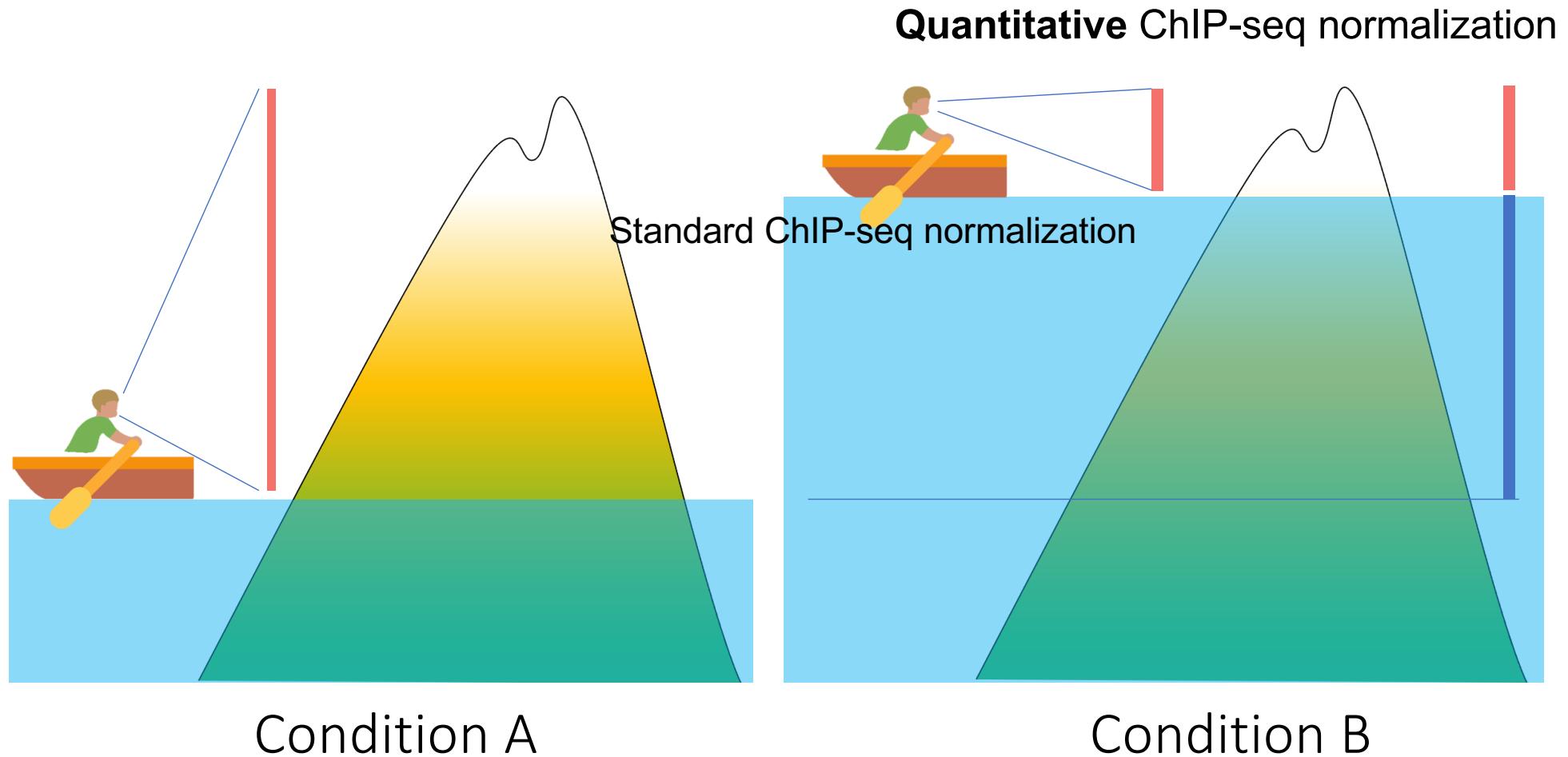
Before comparing, we need to ensure that samples are normalized. Traditional normalization brings all samples to the same effective sequencing depth, and it is assumed that then the samples can be compared quantitatively.

Normalization methods used

- RPKM/FPKM (Reads/Fragments Per Kilobase Million)
- RPGC (Reads Per Genome Coverage, “1x normalization”)

Assumes that relative signal changes but global levels and background do not change!

Capturing quantitative epigenome landscapes



Quantitative ChIP-Seq

- Measured signal (=read density) scales linear (proportional) with 'true' signal
- Signal is comparable quantitatively between samples
- Technical or batch does not influence the quantitative answer

Absolute quantification further requires that the output can be understood in a real-world unit (e.g. binding occupancy or PTM density in fraction/percent)

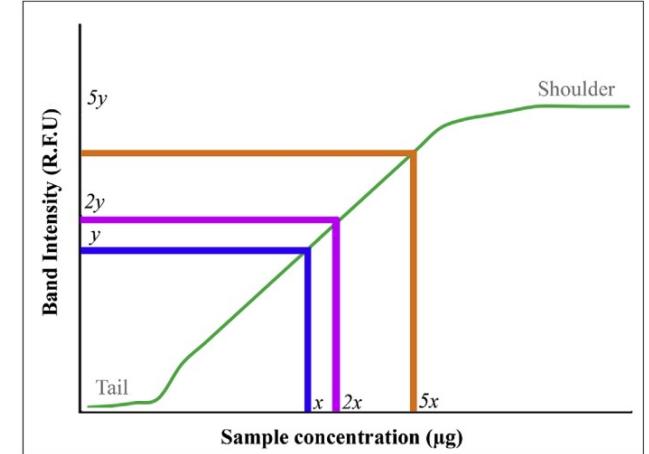


Fig. 2. Linear relationship between sample concentration and band intensity. The signal derived from the protein bands on a Western blot varies with the amount of sample extract loaded onto the protein gel. The illustrated graph depicts a linear and proportional relationship between amount of sample loaded (x , $2x$, $5x$) and the relative fluorescence units (R.F.U) captured from the target bands (y , $2y$, $5y$). Tail and shoulder end of the data curve capture noise and saturated signal, respectively. µg, micrograms; R.F.U, Relative fluorescence units.

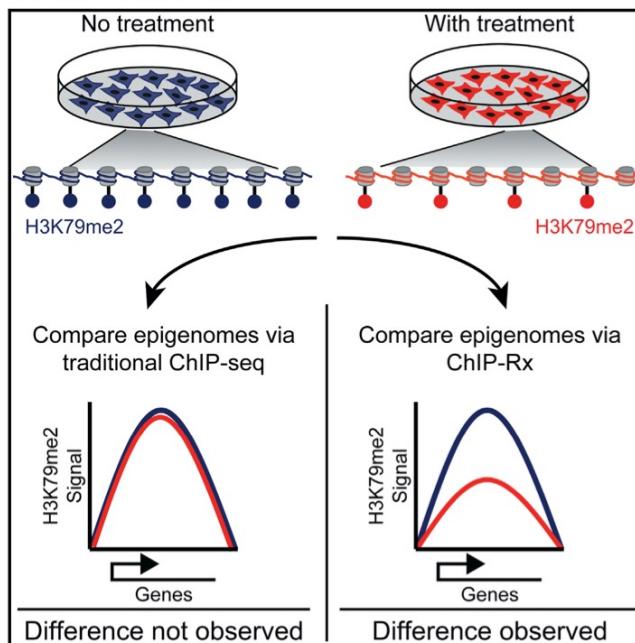
Quantitative ChIP with *Drosophila* Spike-in

Cell Reports

Resource

Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome

Graphical Abstract



Authors

David A. Orlando, Mei Wei Chen, ..., James E. Bradner, Matthew G. Guenther

Correspondence

dorlando@syros.com (D.A.O.),
mguenther@syros.com (M.G.G.)

In Brief

The lack of an empirical methodology to enable normalization among chromatin immunoprecipitation coupled with massively parallel DNA sequencing (ChIP-seq) experiments has limited the precision and comparative utility of this technique. Orlando et al. describe a method, called ChIP with reference exogenous genome (ChIP-Rx), that allows one to perform genome-wide quantitative comparisons of histone modification status across cell populations using defined quantities of a reference epigenome. They use the method to detect disease-relevant epigenomic changes following drug treatment.

Accession Numbers

GSE60104

Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq

Bin Hu, Naomi Petela, Alexander Kurze, Kok-Lung Chan, Christophe Chapard, Kim Nasmyth

Nucleic Acids Research, Volume 43, Issue 20, 16 November 2015, Page e132,
<https://doi.org/10.1093/nar/gkv670>

Published: 30 June 2015 Article history

ChIPSeqSpike: An R/Bioconductor package for ChIP-Seq data scaling according to spike-in control

Descostes N, Tsirigos A, Reinberg D

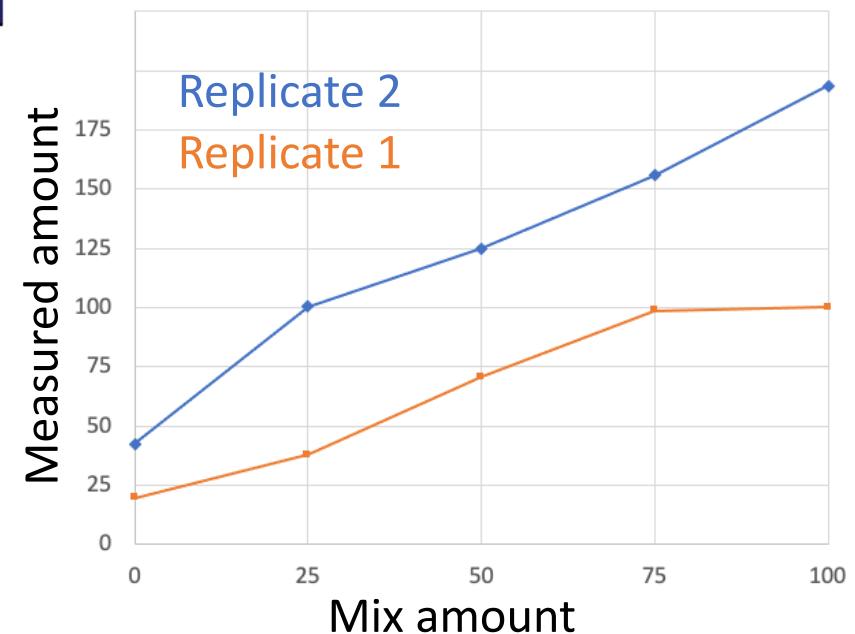
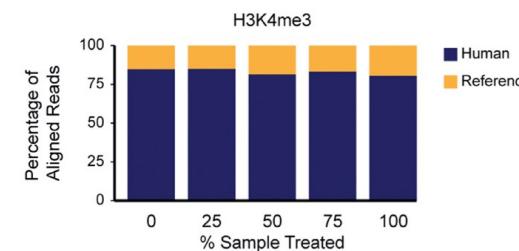
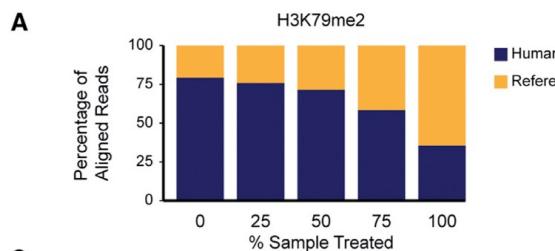
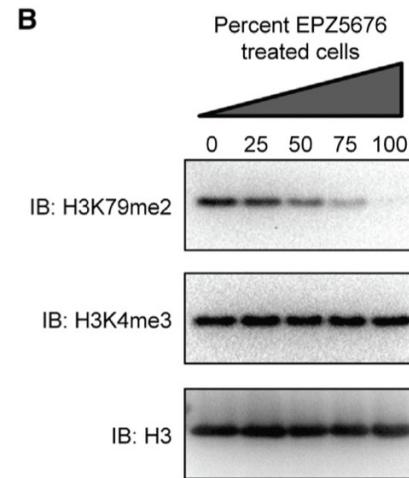
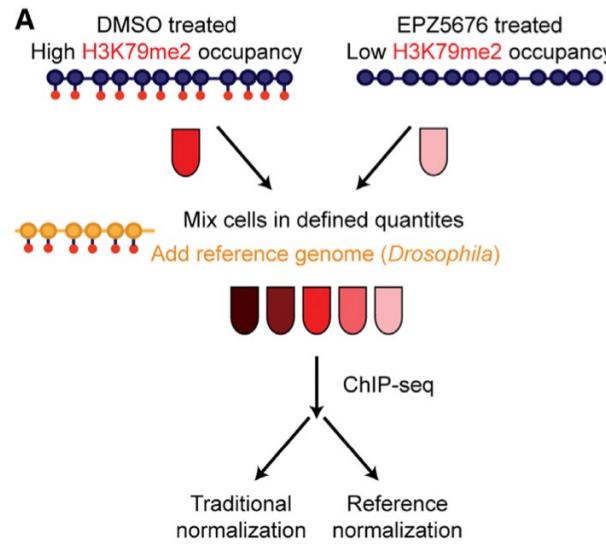
Preprint from bioRxiv, 22 Feb 2018

Highlights

ChIP-seq is a prevailing methodology to investigate and compare epigenomic states

Lack of an empirical normalization strategy has limited the usefulness of ChIP-seq

ChIP-Rx allows genome-wide quantitative comparisons of histone modification status



Multiple replicates are essential for correct spike-in normalization!

Practical solutions to quantitative scaling

- BAM-based, then carry along scale factor determined by BAM read counts for scaling on-the-fly
- BigWig-based, scaled bigwig file can be used in any downstream analysis

Limitations of Spike-in ChIP

- Spike-in amount has to be accurate in relation to chromatin amount (think about error in cell counting, protein assay or pipetting)
- Alternatively, fraction Drosophila reads spiked-in needs to be determined experimentally by sequencing input. Some confusion in the field exists if or not input is taken into account
- Some confusion exists if and when a background normalization using input can be done while also using spike-in normalization (how to normalize input? divide or subtract?)
- Antibody must crossreact with the spike-in species.

Excercise – reanalysis of Orlando data using Bioconductor package

ChIP-seq with exogenous chromatin spike

This tutorial is included from [previous workshop](#). Thanks!

- Requirements
 - Uppmax
 - Local
- Data
- Data preparation
- Fingerprint plots
- Disclaimer
- Using [ChIPSeqSpike](#) for ChIPseq signal scaling
 - Files and directories
 - Scaling of signal to exogenous chromatin spike
 - Data visualization
 - Visualization with gene meta-profiles
 - Visualization with Boxplots
 - Correlation plots
- What to do next

ChIPSeqSpike: A R/Bioconductor package for ChIP-Seq data scaling according to spike-in control

Descotes N, Tsirigos A, Reinberg D

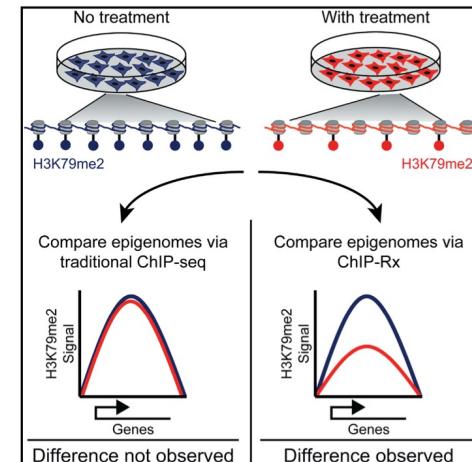
Preprint from [bioRxiv](#), 22 Feb 2018

Resource

Cell Reports

Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome

Graphical Abstract



Authors

David A. Orlando, Mei Wei Chen, ..., James E. Bradner, Matthew G. Guenther

Correspondence

dorlando@syros.com (D.A.O.), mguenther@syros.com (M.G.G.)

In Brief

The lack of an empirical methodology to enable normalization among chromatin immunoprecipitation coupled with massively parallel DNA sequencing (ChIP-seq) experiments has limited the precision and comparative utility of this technique. Orlando et al. describe a method, called ChIP with reference exogenous genome (ChIP-Rx), that allows one to perform genome-wide quantitative comparisons of histone modification status across cell populations using defined quantities of a reference epigenome. They use the method to detect disease-relevant epigenomic changes following drug treatment.

Highlights

ChIP-seq is a prevailing methodology to investigate and compare epigenomic states

Lack of an empirical normalization strategy has limited the usefulness of ChIP-seq

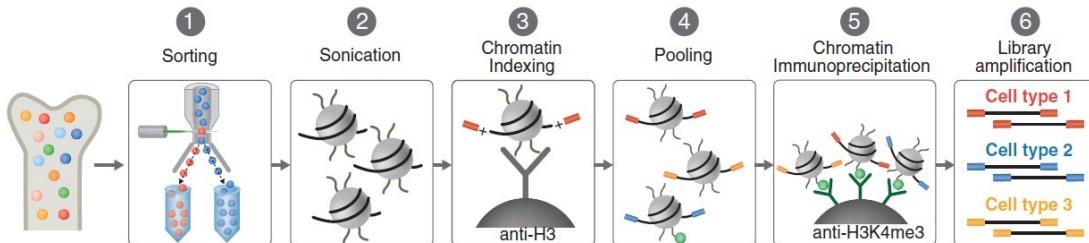
ChIP-Rx allows genome-wide quantitative comparisons of histone modification status

Accession Numbers

GSE60104

Barcode-first methods

I-ChIP: on-beads barcoding



One-pot methods ChIP

A high-throughput ChIP–Seq for large-scale chromatin studies

Christophe D Chabbert, Sophie H Adjalley, Bernd Klaus, Emilie S Fritsch, Ishaan Gupta, Vicent Pelechano, Lars M Steinmetz

Molecular Cell

All Content

Molecular Cell

Explore Online Now Current Issue Archive Journal Information For Authors

< Previous Article Volume 61, Issue 1, p170–180, 7 January 2016

TECHNOLOGY

A Multiplexed System for Quantitative Comparisons of Chromatin Landscapes

Peter van Galen, Aaron D. Viny, Oren Ram, Russell J.H. Ryan, Matthew J. Cotton, Laura Donohue, Cem Sievers, Yotam Drier, Brian B. Laiu, Shawn M. Gillespie, Kaitlin M. Carroll, Michael B. Cross, Ross L. Levine, Bradley E. Bernstein

Published Online: December 10, 2015

Open Archive | PlumX Metrics

DOI: <http://dx.doi.org/10.1016/j.molcel.2015.11.003> | CrossMark

Cell Reports

Quantitative Multiplexed ChIP Reveals Global Alterations that Shape Promoter Bivalency in Ground State Embryonic Stem Cells

Graphical Abstract

Quantitative ChIP with large linear dynamic range

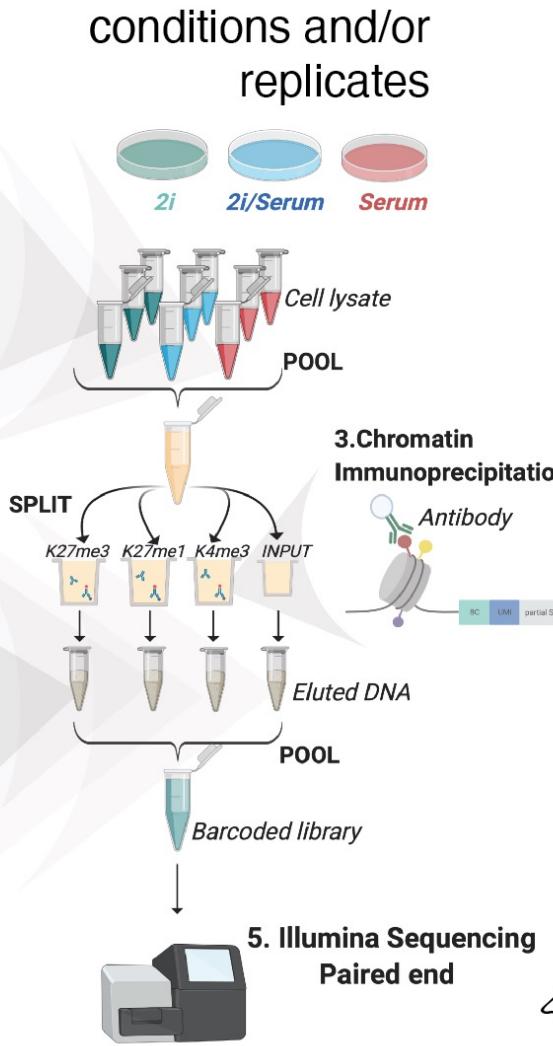
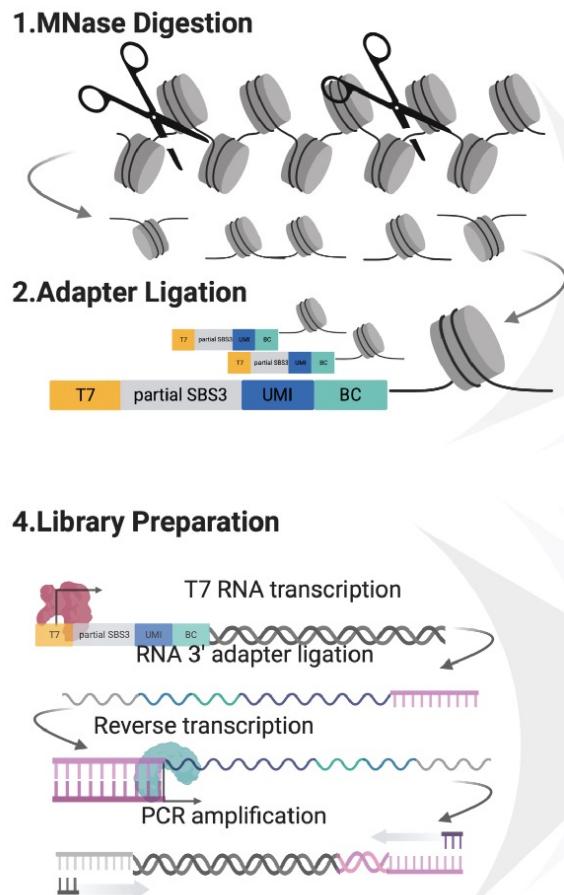
Authors

Banushree Kumar, Simon J. Elsässer

Resource

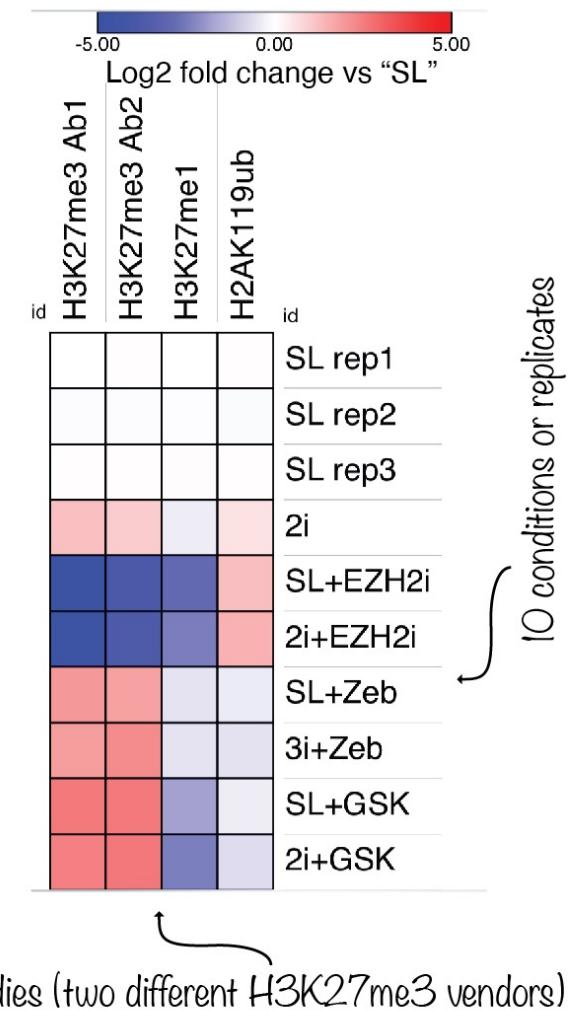
Multiplexed ChIP (MINUTE-ChIP)

A



B

Example multiplexed ChIP

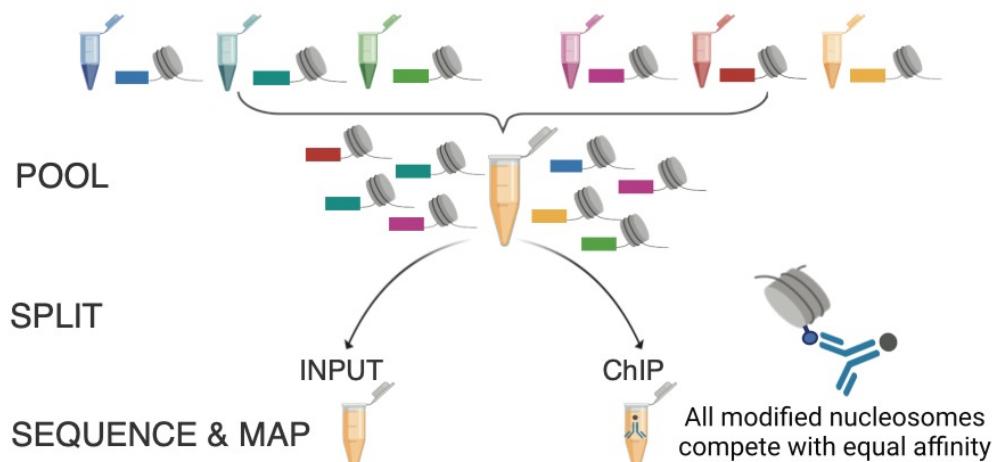


MINUTE-ChIP Quantification

Conditions A and B with different levels of a histone modification 



LYSIS & FRAGMENTATION

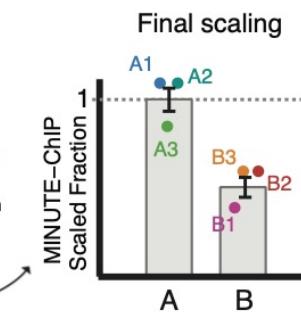


NORMALIZE

$$\text{Input-normalised read count (INR)} = \frac{\text{Reads}_{\text{ChIP, condition}}}{\text{Reads}_{\text{Input, condition}}}$$

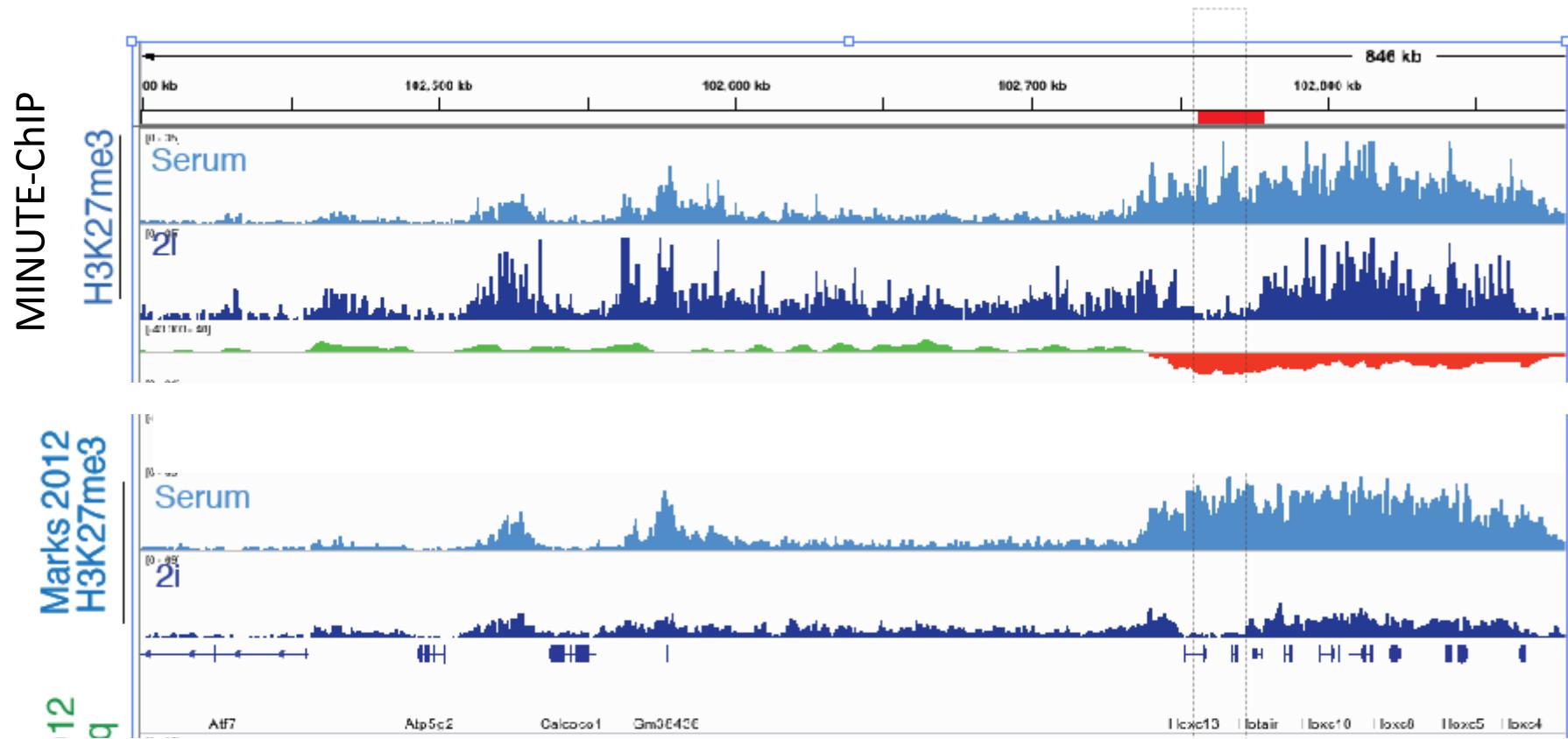
$$\text{INR}_{A1} = \frac{4}{2} \quad \text{INR}_{A2} = \frac{6}{3} \quad \dots \quad \text{INR}_{B3} = \frac{2}{2}$$

SCALE TO REFERENCE

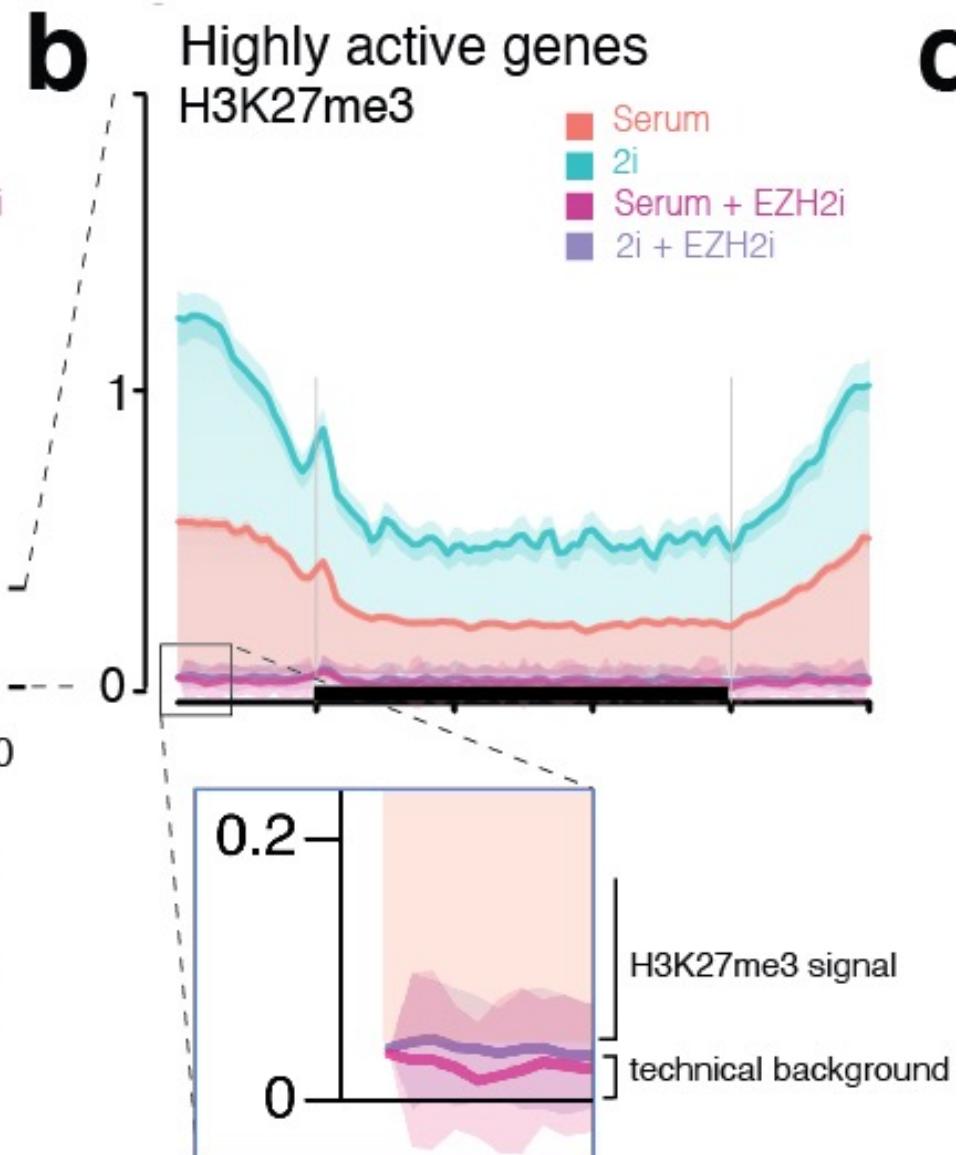
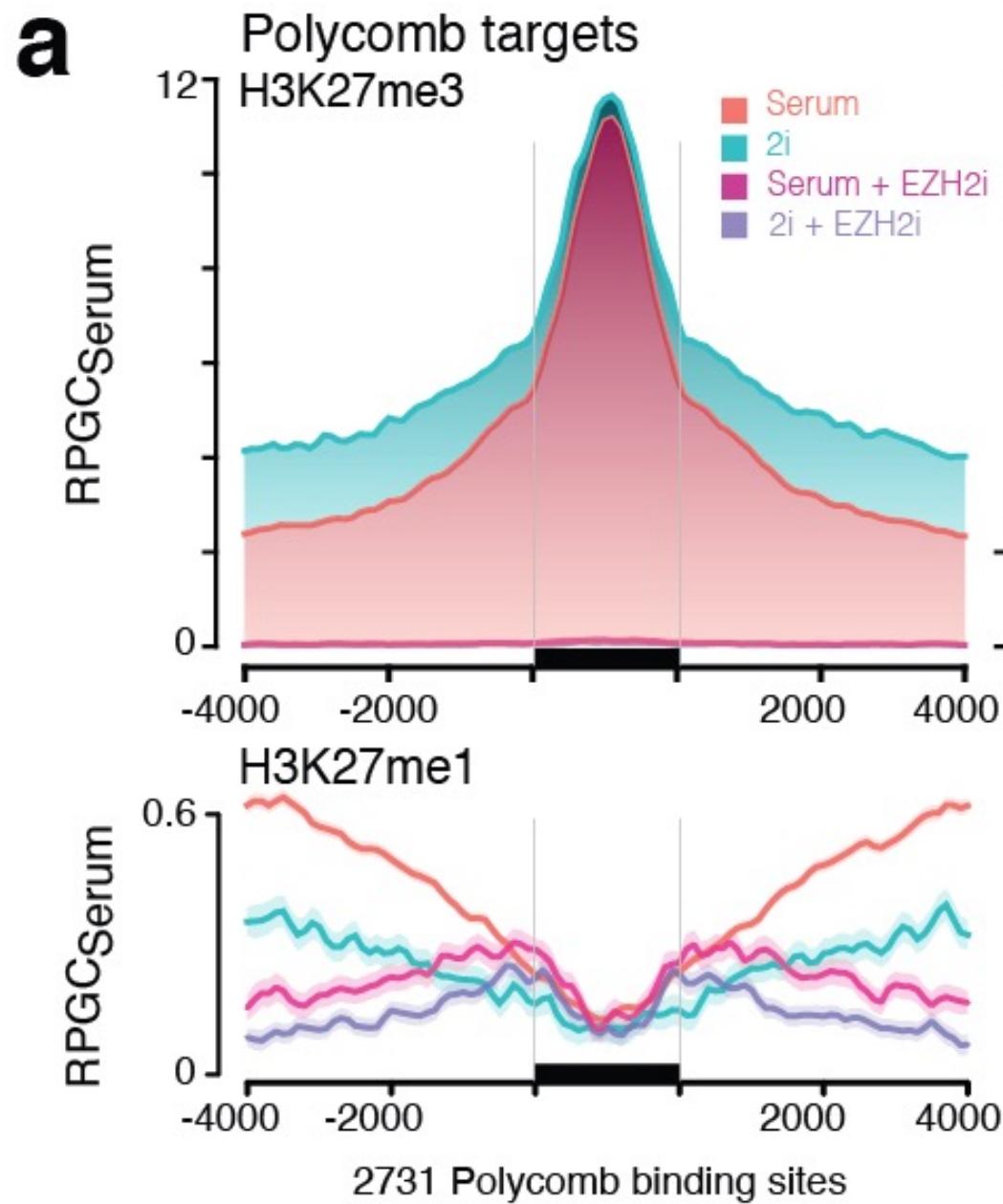


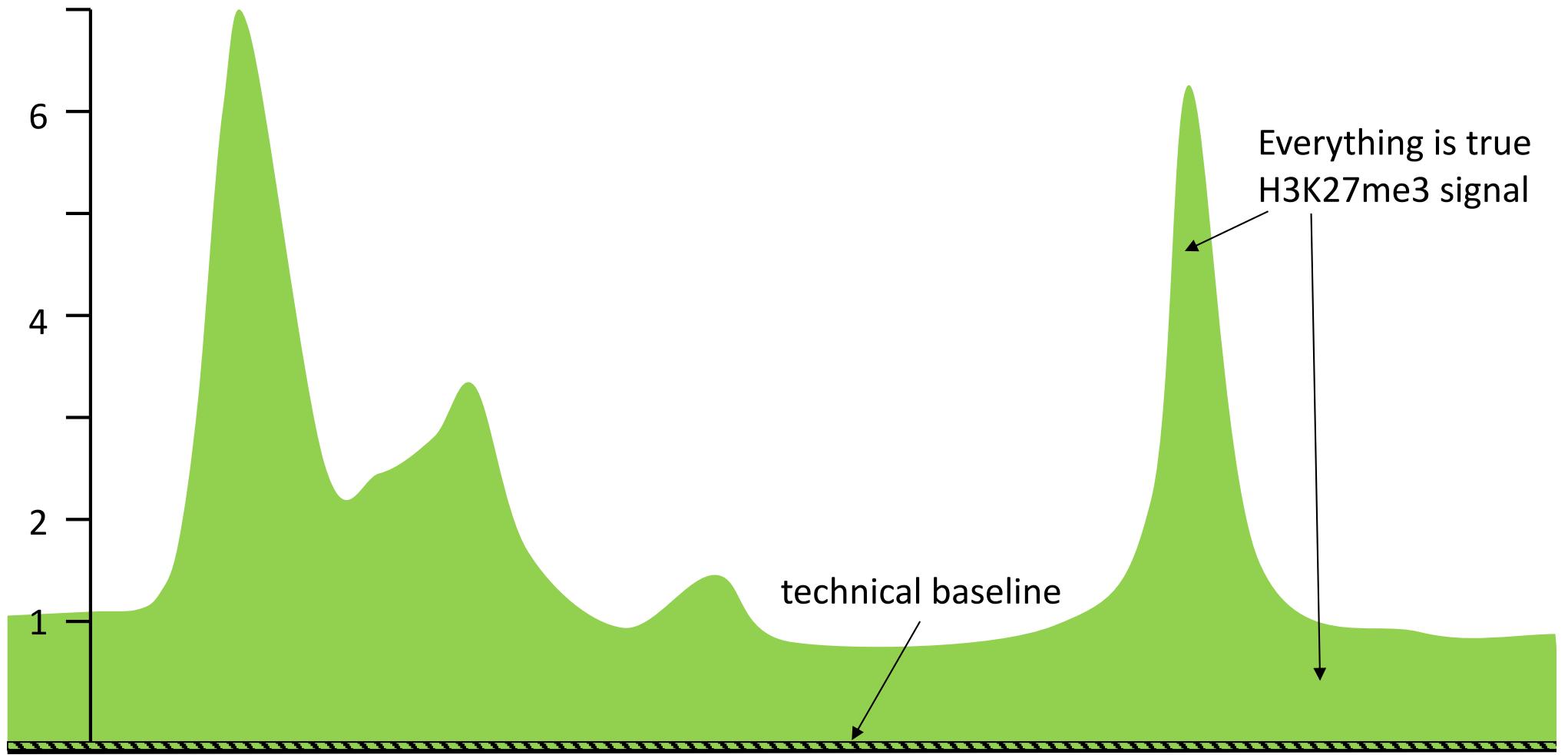
Comparison normal and quantitative ChIP

Comparing naïve and ‘primed’ mouse embryonic stem cells

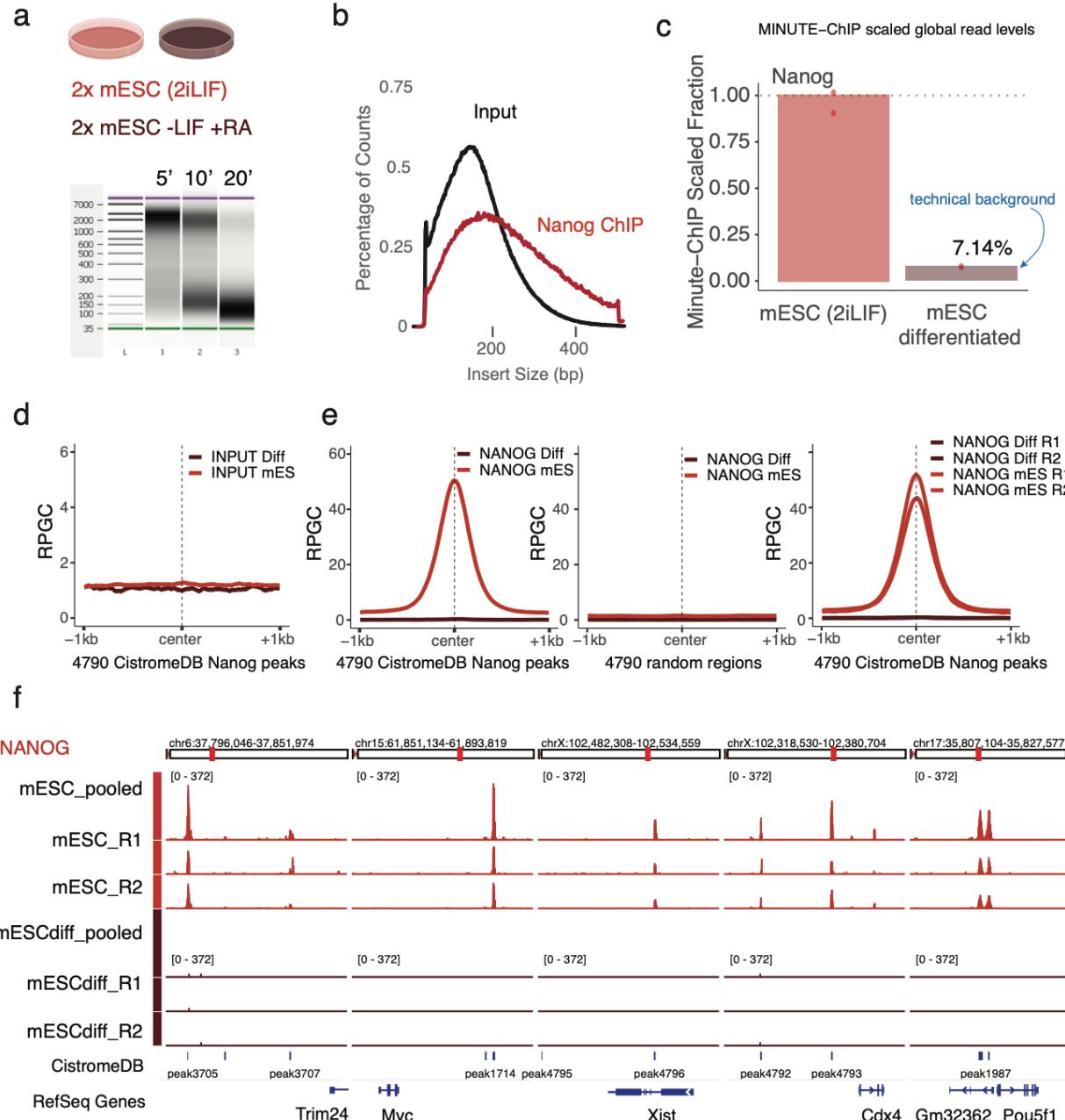


Multiplexed ChIP has very little technical background





MINUTE-ChIP for transcription factors



Excercise – Reanalysis of MINUTE-ChIP data

MINUTE-ChIP

- [Background](#)
- [Primary analysis](#)
 - [Conda environment](#)
 - [Files](#)
 - [Running Minute](#)
 - [Scaling info](#)
 - [IGV tracks](#)
- [Downstream analysis](#)
 - [Files](#)
 - [Looking at bivalent genes](#)
 - [Genome-wide bin distribution](#)

Article | [Open Access](#) | [Published: 30 May 2022](#)

Polycomb repressive complex 2 shields naïve human pluripotent cells from trophectoderm differentiation

[Banushree Kumar](#), [Carmen Navarro](#), [Nerges Winblad](#), [John P. Schell](#), [Cheng Zhao](#), [Jere Weltner](#),
[Laura Baqué-Vidal](#), [Angelo Salazar Mantero](#), [Sophie Petropoulos](#), [Fredrik Lanner](#)✉ & [Simon J. Elsässer](#)✉

Nature Cell Biology **24**, 845–857 (2022) | [Cite this article](#)

