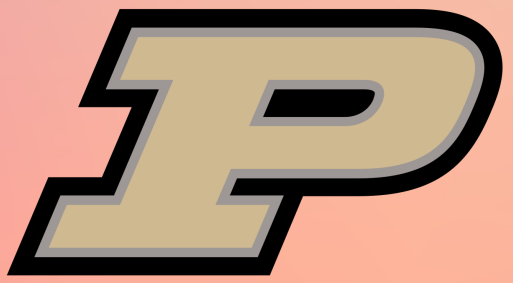
An introduction to home-brew library prep methods for Illumina sequencing and critical factors to ensure success: ChIP-seq, ATAC-seq, Hi-C, and CRISPR screening



Lauren Robinson
Sequencing and Informatics Specialist
OH/MI/IN/KY

illumina®

Agenda

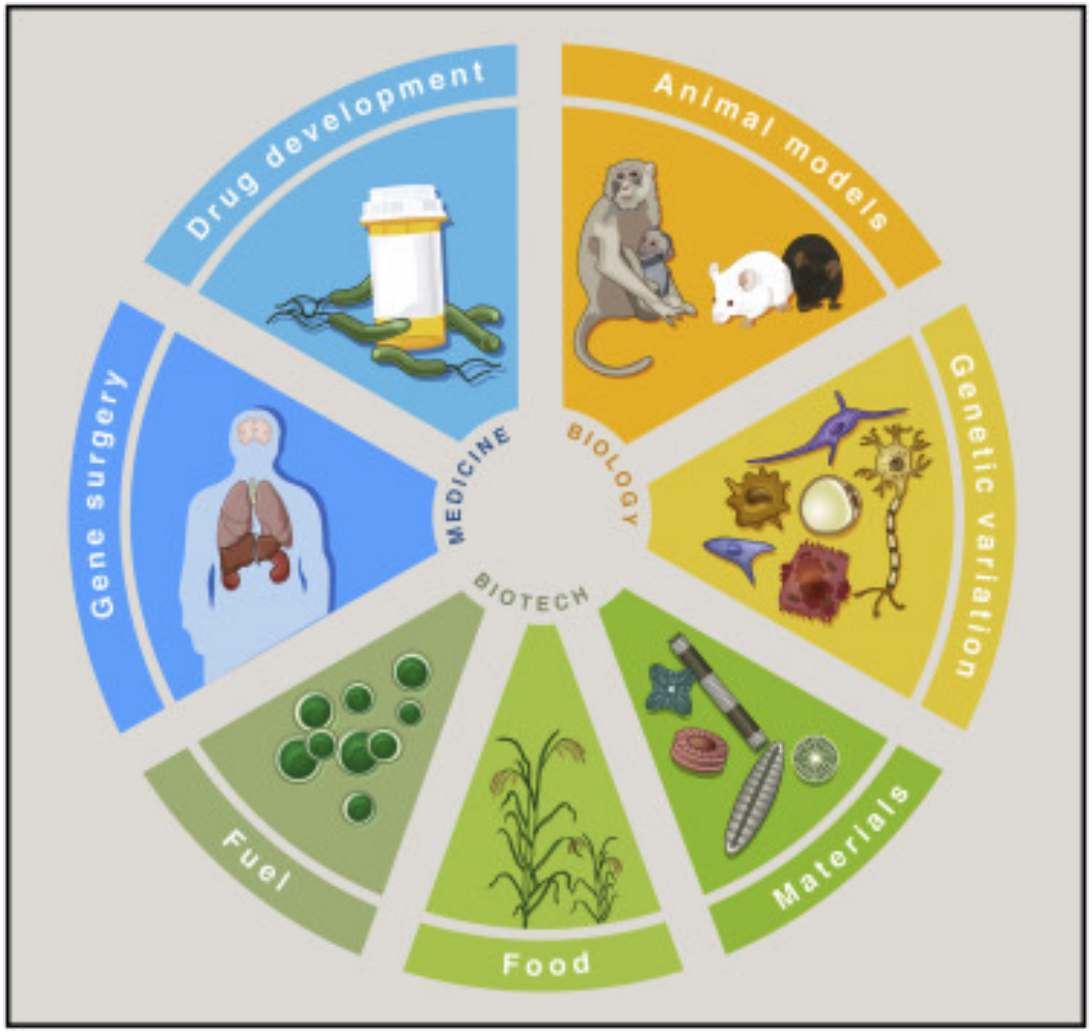
- CRISPR Screening
- ATAC-Seq
- ChIP-Seq, Cut&Run, and Cut&Tag
- HiC/Capture-C
- Questions

illumina

CRISPR Screening

CRISPR Screening

- Identify genes or DNA sequences causing cells to be either resistant or sensitive to a drug
- Identify genes of DNA sequences affecting susceptibility to environmental toxins
- Identify components of a cellular pathway
- Identify genes or DNA sequences leading to a particular disease state

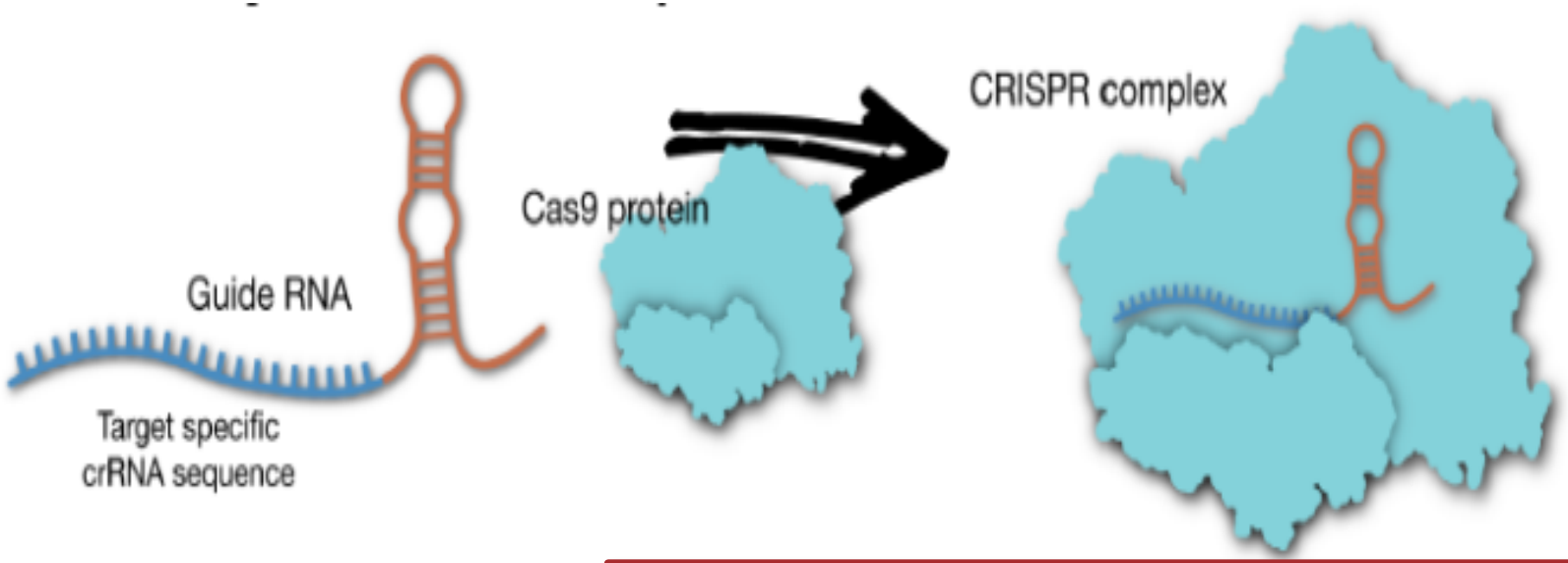


https://www.idtdna.com/pages/education/decoded/article/overview-what-is-crispr-screening#:~:text=CRISPR%20screening%20is%20used%20to,drug%20resistance%20and%20drug%20sensitivity.
https://haematologica.org/article/view/8891

illumina

Natural CRISPR Systems Have 2 Major Components

Repeat Regions –
These are used to create **gRNAs** - sequences of RNA that direct the CRISPR system to cut other DNA sequences..



Nucleases:
Proteins that bind to gRNAs and are directed by the gRNAs to cut particular DNA sequences. **Cas9** is a very commonly used CRISPR nuclease

Costa, Joana R., et al. "Genome editing using engineered nucleases and their use in genomic screening." *Assay Guidance Manual [Internet]*. Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2017.

illumina

Pooled CRISPR Screen Overview

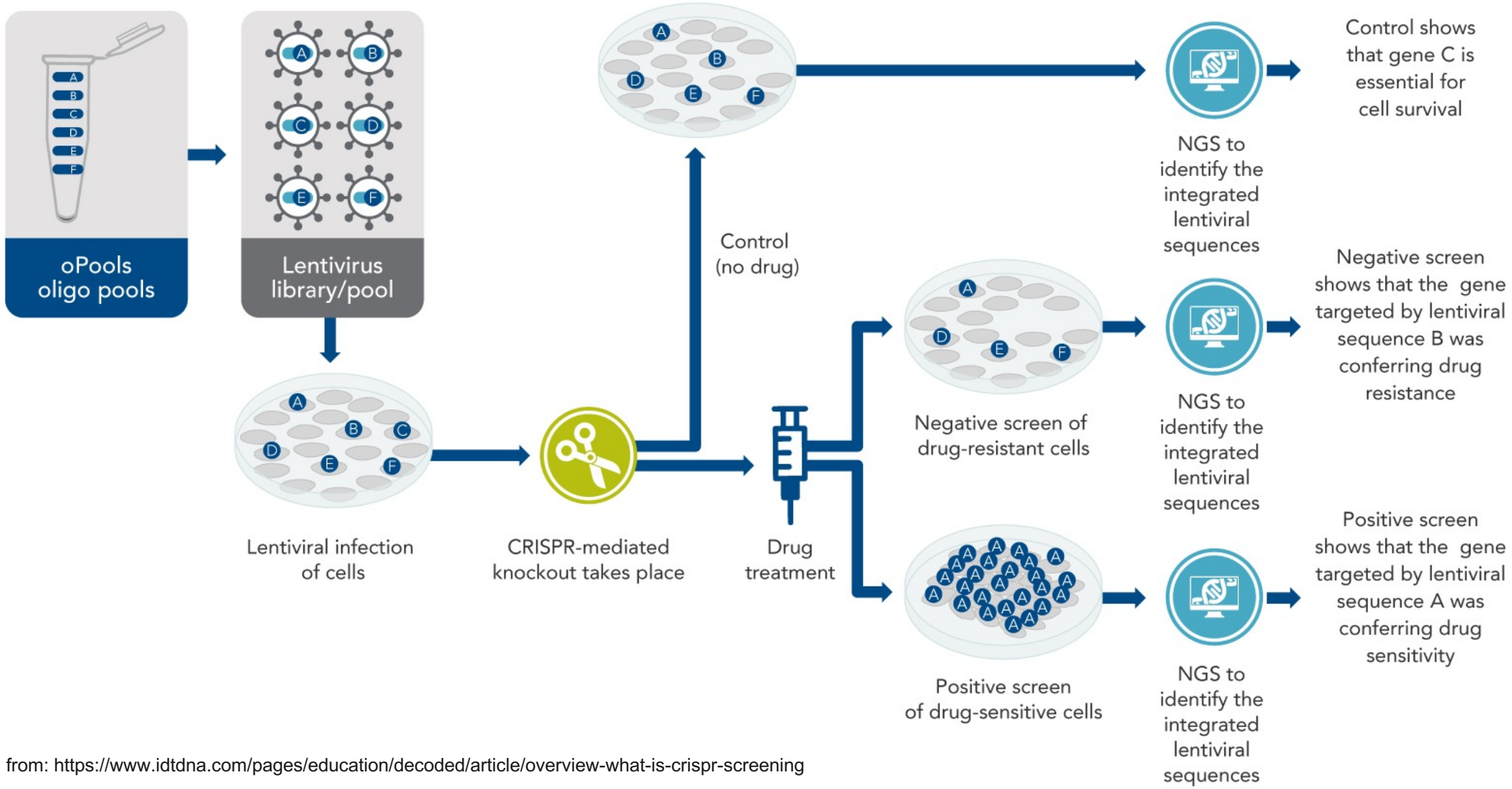
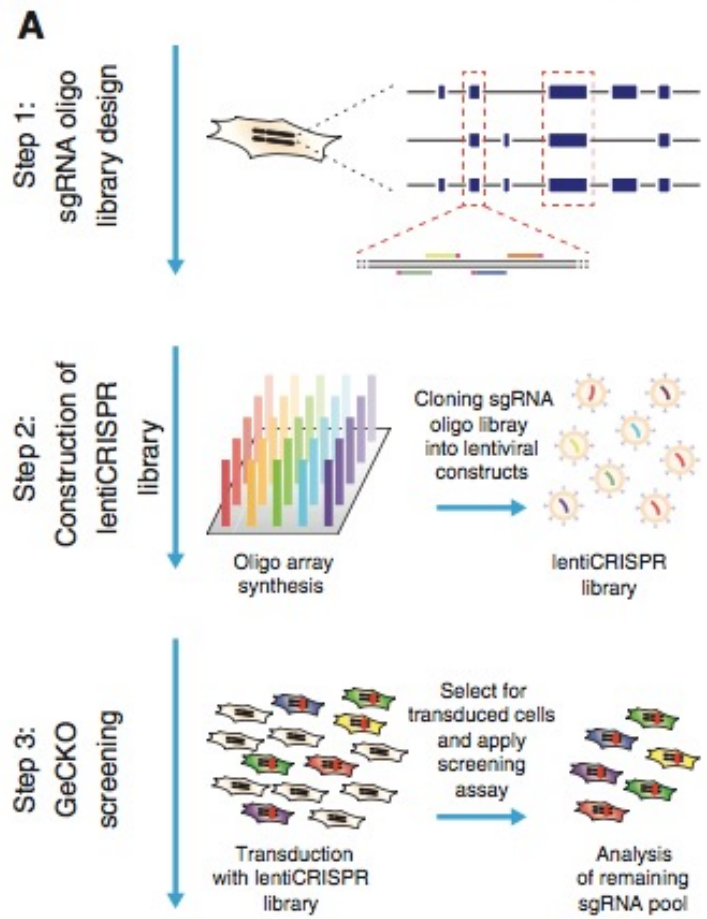


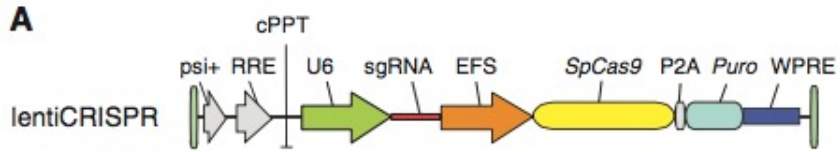
 Image sourced from: https://www.idtdna.com/pages/education/decoded/article/overview-what-is-crispr-screening

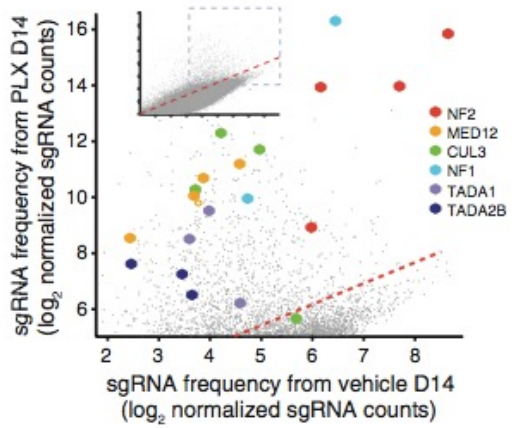
Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells

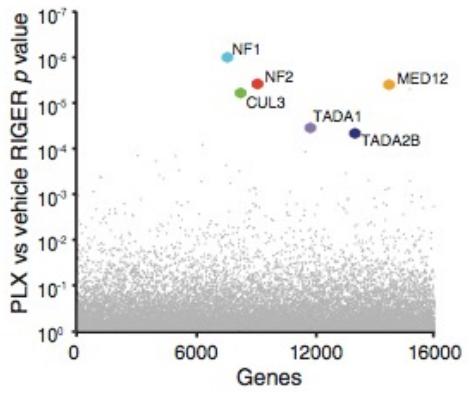
Ophir Shalem,[1,2]* Neville E. Sanjana,[1,2]* Ella Hartenian,[1] Xi Shi,[1,3]
David A. Scott,[1,2] Tarjei S. Mikkelsen,[1] Dirk Heckl,[4] Benjamin L. Ebert,[4] David E. Root,[1]
John G. Doench,[1] Feng Zhang[1,2]†

3 JANUARY 2014 VOL 343 **SCIENCE** www.sciencemag.org









illumina®

NGS Quality Control for Pooled Libraries

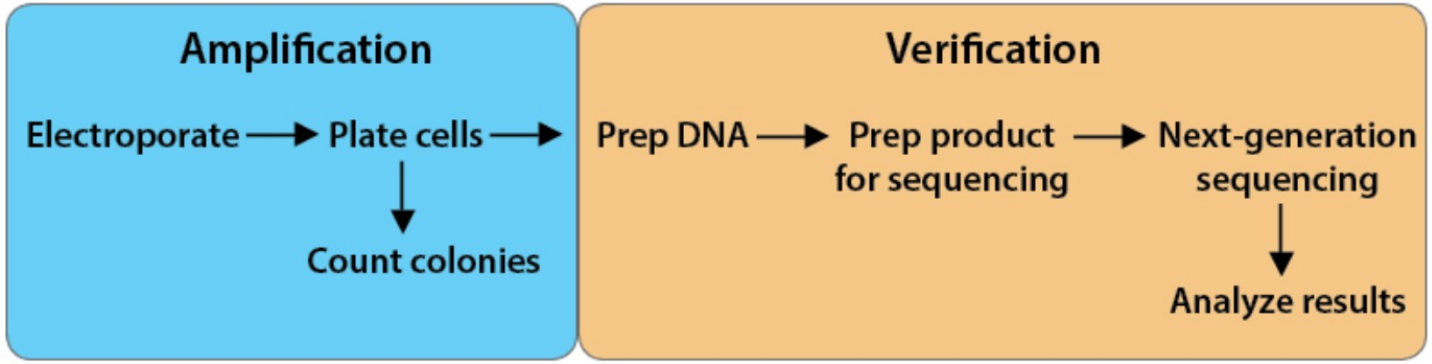


Figure 1. An overview of the pooled library amplification/verification process.

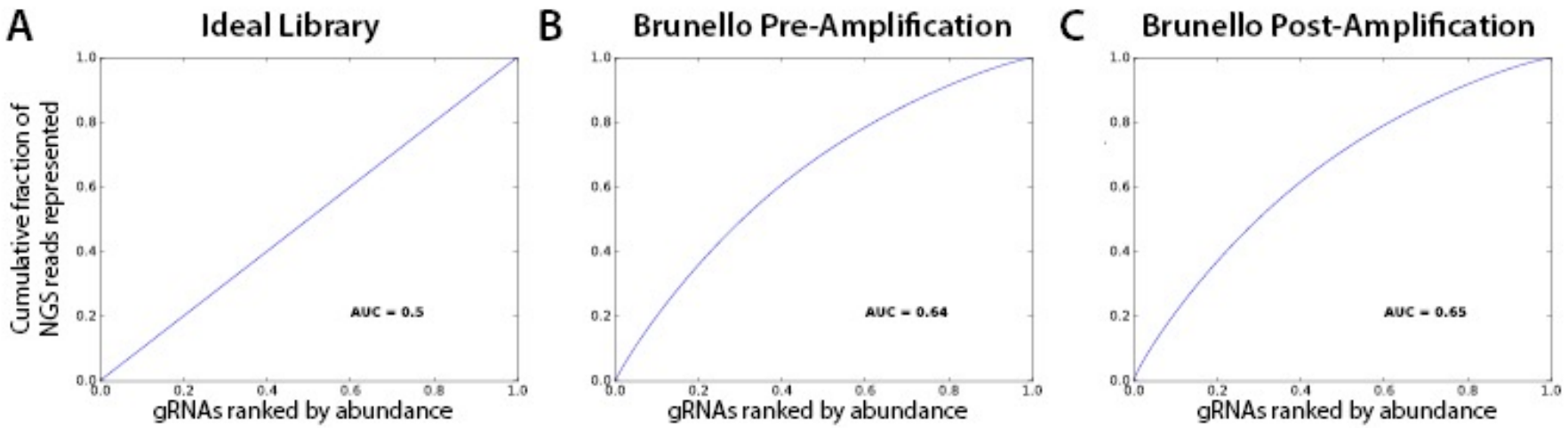
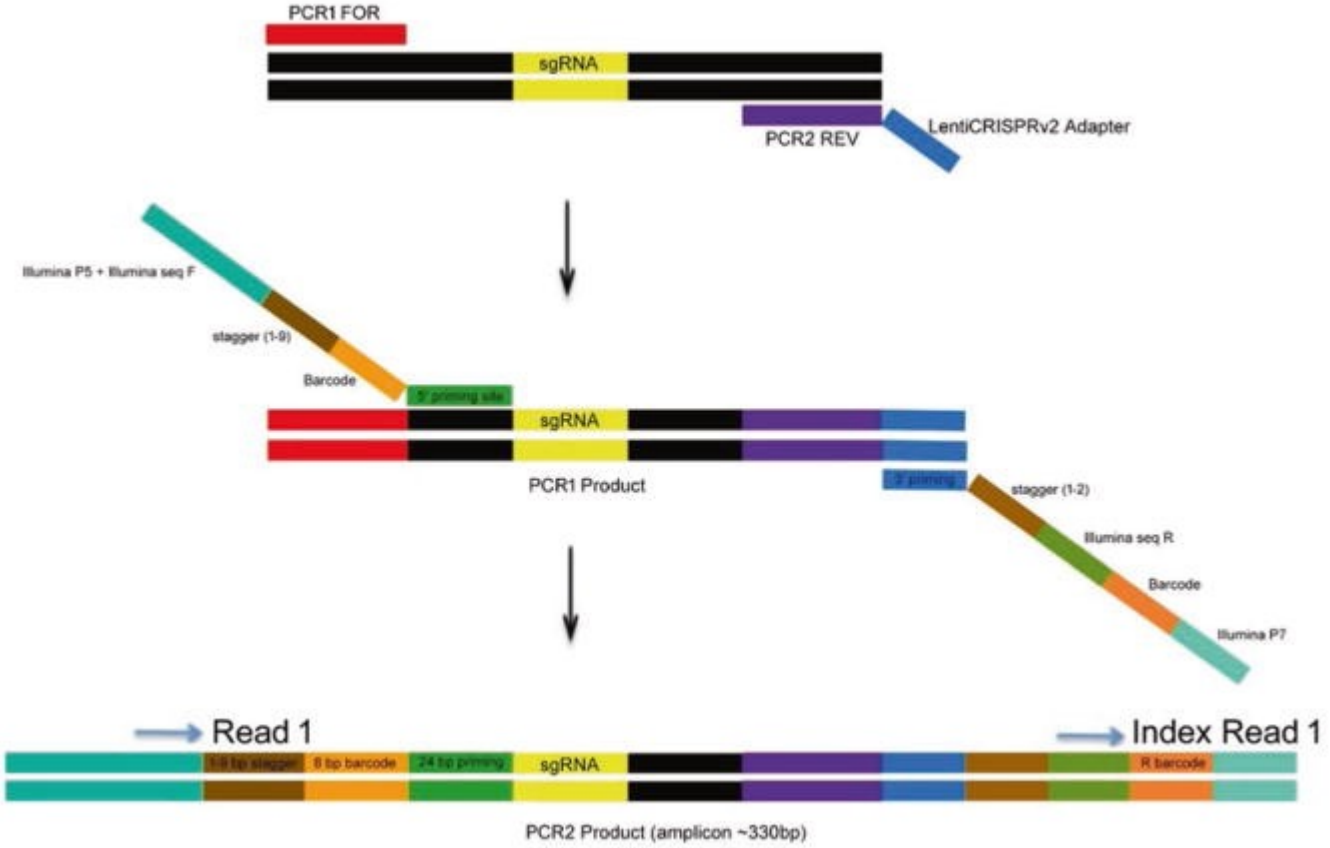


Image sourced from: Addgene

illumina®

Hit identification using Illumina Sequencing



Yau, Edwin H., and Tariq M. Rana. "Next-generation sequencing of genome-wide CRISPR Screens." Next Generation Sequencing. Humana Press, New York, NY, 2018. 203-216.

illumina®

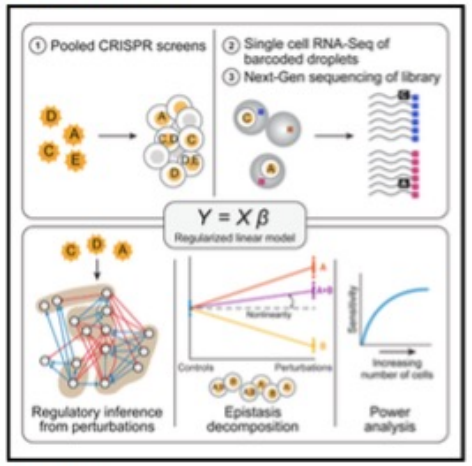
Combining CRISPR and Single Cells

Resource

Cell

Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens

Graphical Abstract



Authors

Atray Dixit, Oren Parnas, Biyu Li, ..., Jonathan S. Weissman, Nir Friedman, Aviv Regev

Correspondence

aregev@broadinstitute.org

In Brief

A technology combining single-cell RNA sequencing with CRISPR-based perturbations termed Perturb-seq makes analyzing complex phenotypes at a large scale possible

Highlights

- Pooled CRISPR screen with scRNA-seq readout
- Integrated model of perturbations, single cell phenotypes, and epistatic interactions
- Effect of TFs on genes, programs, and states in LPS response in immune cells
- Downsampling assessment of feasibility of genome-wide or combinatorial screens

Dixit et al., 2016, Cell 167, 1853–1866
December 15, 2016 © 2016 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2016.11.038

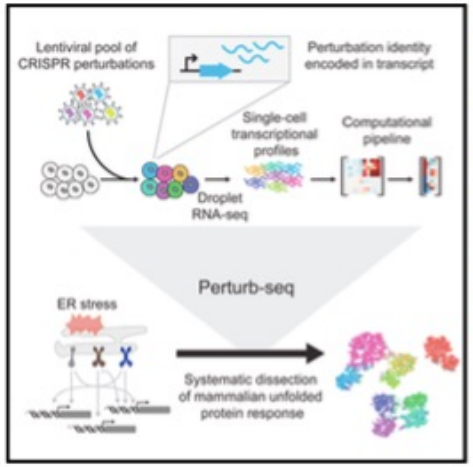
CellPress

Resource

Cell

A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response

Graphical Abstract



Authors

Britt Adamson, Thomas M. Norman, Marco Jost, ..., Oren Parnas, Aviv Regev, Jonathan S. Weissman

Correspondence

jonathan.weissman@ucsf.edu

In Brief

A strategy for barcoding CRISPR-mediated perturbations allows pooled expression profiling via single-cell RNA sequencing. Application to the mammalian unfolded protein response then enabled systematic delineation of the transcriptional arms of the response and functional clustering of genes affecting ER homeostasis.

Highlights

- Perturb-seq allows parallel screening with rich phenotypic output from single cells
- Simultaneous delivery and identification of up to three CRISPR perturbations
- Genome-scale screens dissect the mammalian unfolded protein response
- Analytical methods separate perturbation responses from confounding effects

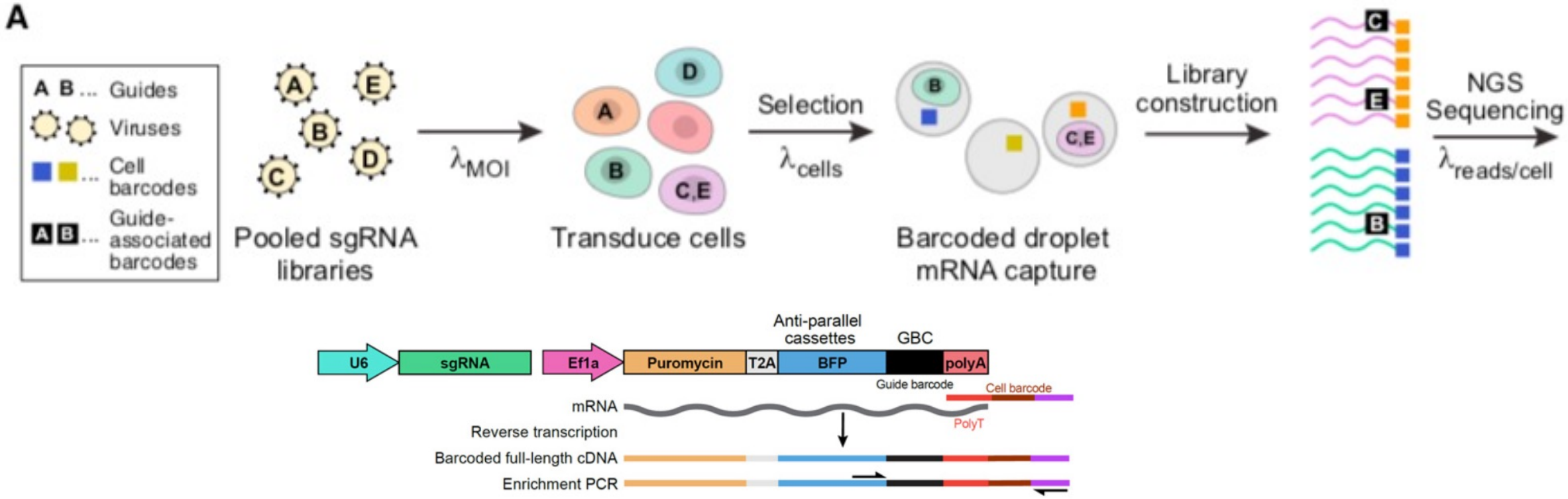
Adamson et al., 2016, Cell 167, 1867–1882
December 15, 2016 © 2016 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2016.11.048

CellPress

illumina

Perturb-Seq:

Genetic Screening with CRISPR



Uniquely Single Cell NGS

Combines ability to screen many knockout targets simultaneously, both as single gene and combinatorial knockouts paired with rich digital read out of the exact guide RNA and Cell ID as well as the effect on gene expression phenotype.

Dixit et al 2016 Cell 167, 1853-1866

illumina

ATAC-seq or sequencing “openness”

Epigenetics – Epi = *on top of*

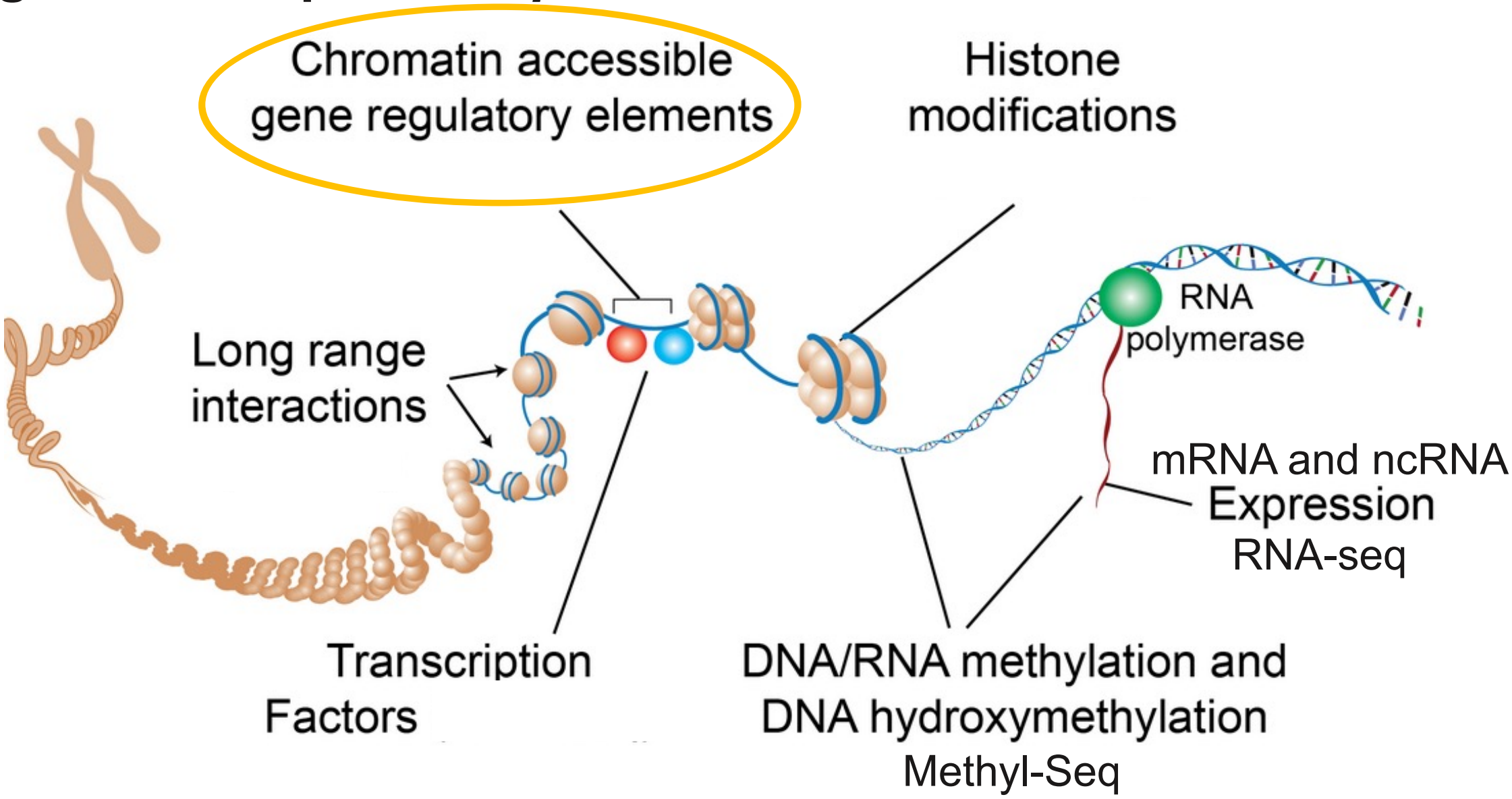
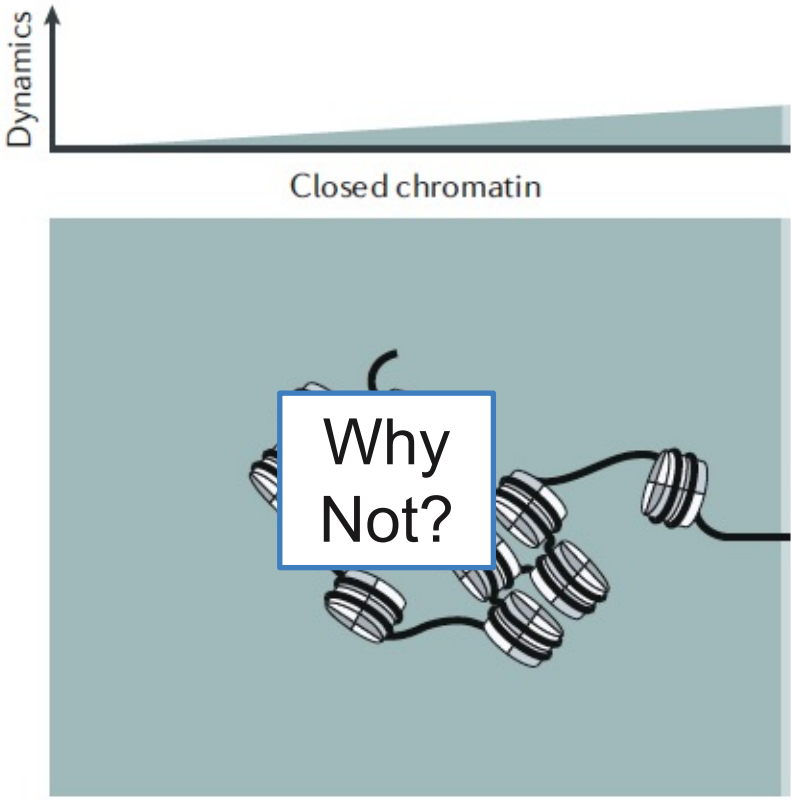


Image via: https://genome.duke.edu

illumina

ATAC-Seq – Why study chromatin accessibility?



Heterochromatin ~97%

What could be

Why?

DOI:10.1038/s41576-018-0089-8

illumina

ATAC-Seq: Studying gene regulation through chromatin accessibility

- ATAC-Seq makes libraries from open/accessible chromatin
  - Nextera-based library preparation
- Simple two step protocol with an input of 500 to 50,000 cells

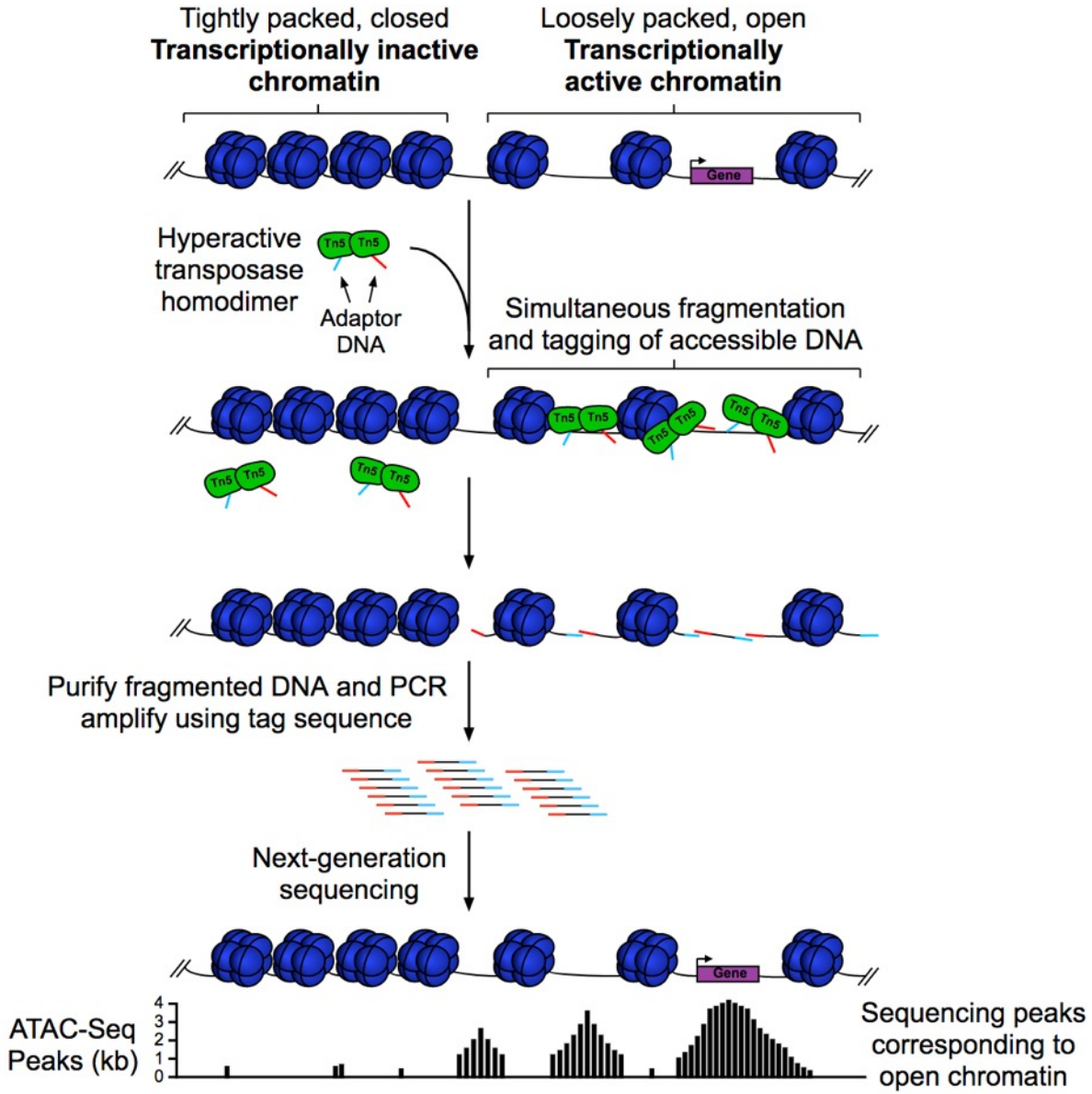


Image sourced from Wikimedia

Bulk ATAC-seq has been improved: Omni-ATAC

- **Fewer Cells, Frozen tissue compatible**
  - Open to more samples
- **Better signal, cleaner data**

2018


An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues

M Ryan Corces[1,2], Alexandro E Trevino[1,3,4], Emily G Hamilton[5], Peyton G Greenside[3,6], Nicholas A Sinnott-Armstrong[3], Sam Vesuna[4], Ansuman T Satpathy[1,7], Adam J Rubin[2], Kathleen S Montine[7], Beijing Wu[3], Arwa Kathiria[3], Seung Woo Cho[1,2], Maxwell R Mumbach[1,3], Ava C Carter[1,2], Maya Kasowski[1,7], Lisa A Orloff[8], Viviana I Risca[3], Anshul Kundaje[3,9], Paul A Khavari[2], Thomas J Montine[7], William J Greenleaf[1,3,10] & Howard Y Chang[1,2]

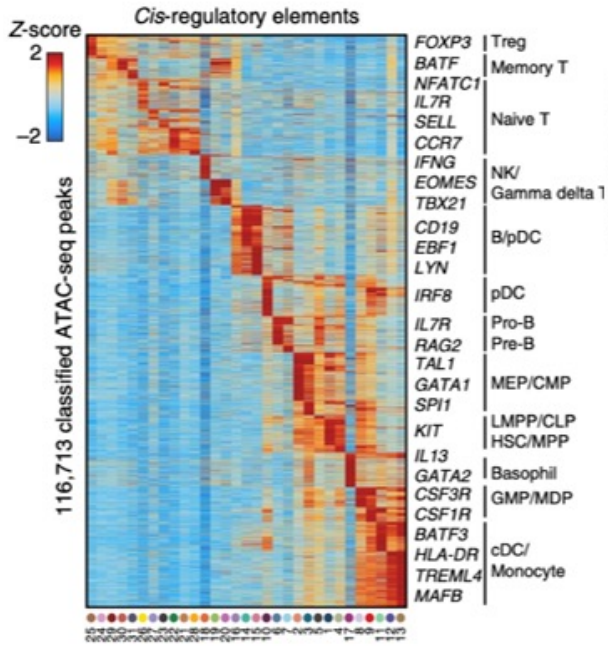
https://www.nature.com/articles/nmeth.4396

illumina

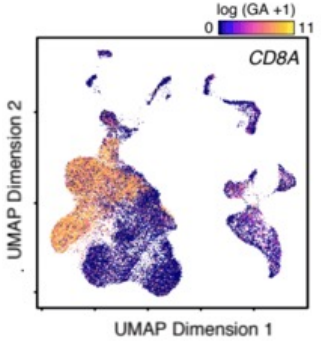
Single Cell ATAC seq

One quick protocol – multiple readouts

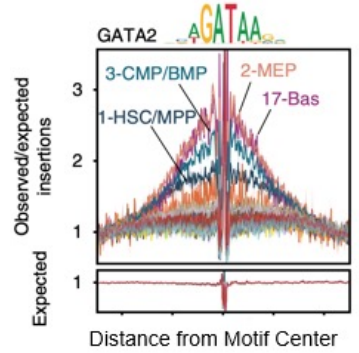
Peaks for open chromatin



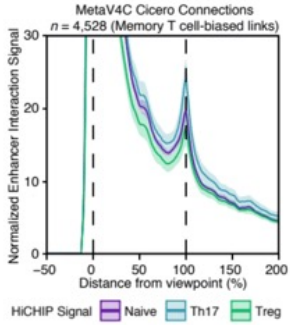
Cell state classification/clustering



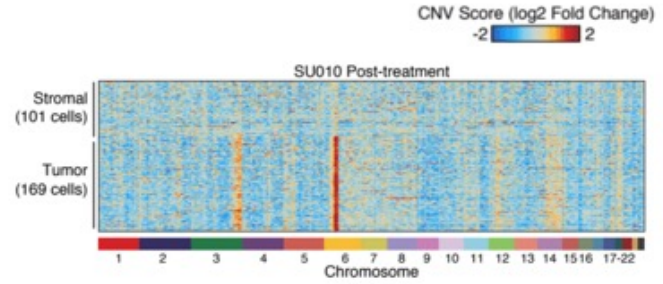
DNA binding motifs



Inferred gene expression



Copy # variations



Satpathy*, Granja*, et al. Nature Biotechnology (2019).

illumina®

Coverage Recommendations for ATAC-Seq

The minimum required sequencing coverage for ATAC-Seq varies according to research objectives.

Paired-end reads are recommended as they offer:

- Higher unique alignment rates
- Removal of PCR duplicates
- More complete information about accessible sequences
- Ability to categorize reads as nucleosome-free, mono-nucleosomal, or di-nucleosomal

Research Goal	Recommended Depth
Identification of open chromatin differences in human samples	>50M paired-end reads
Transcription factor foot printing to construct gene regulatory networks	>200M paired-end reads
Single-cell analysis	50K paired-end reads per nucleus/cell

*read depth may vary depending on experimental conditions

https://emea.illumina.com/techniques/popular-applications/epigenetics/atac-seq-chromatin-accessibility.html

illumina

Mapping Protein DNA Interactions with ChIP-Seq

Epigenetics – Epi = *on top of*

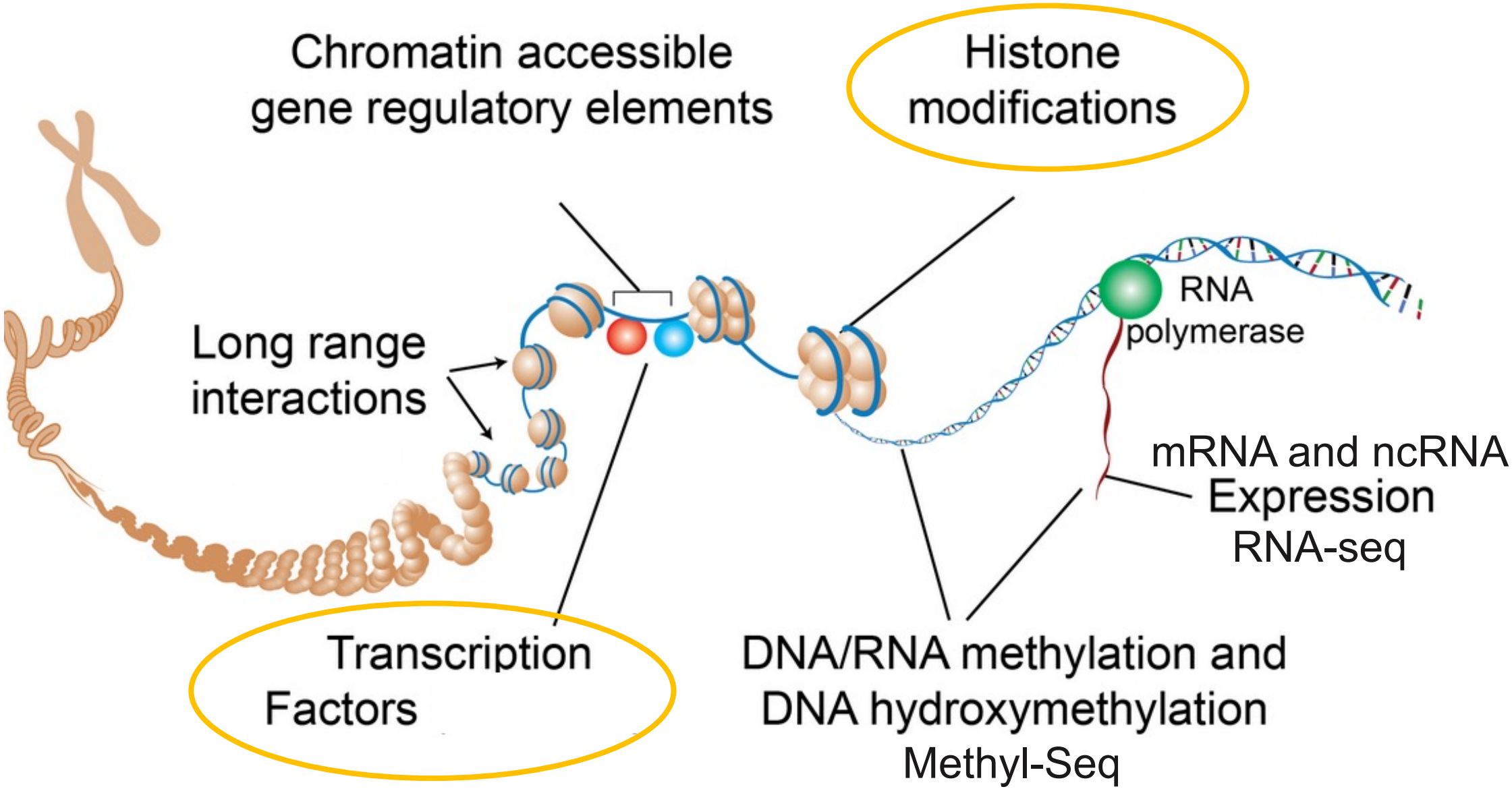


Image via: https://genome.duke.edu

illumina

ChIP (Chromatin Immunoprecipitation)

- Identifies the genome-wide binding sites of DNA-associated proteins.

- Widely used to map histone modifications and transcription factor binding on a genome-wide level

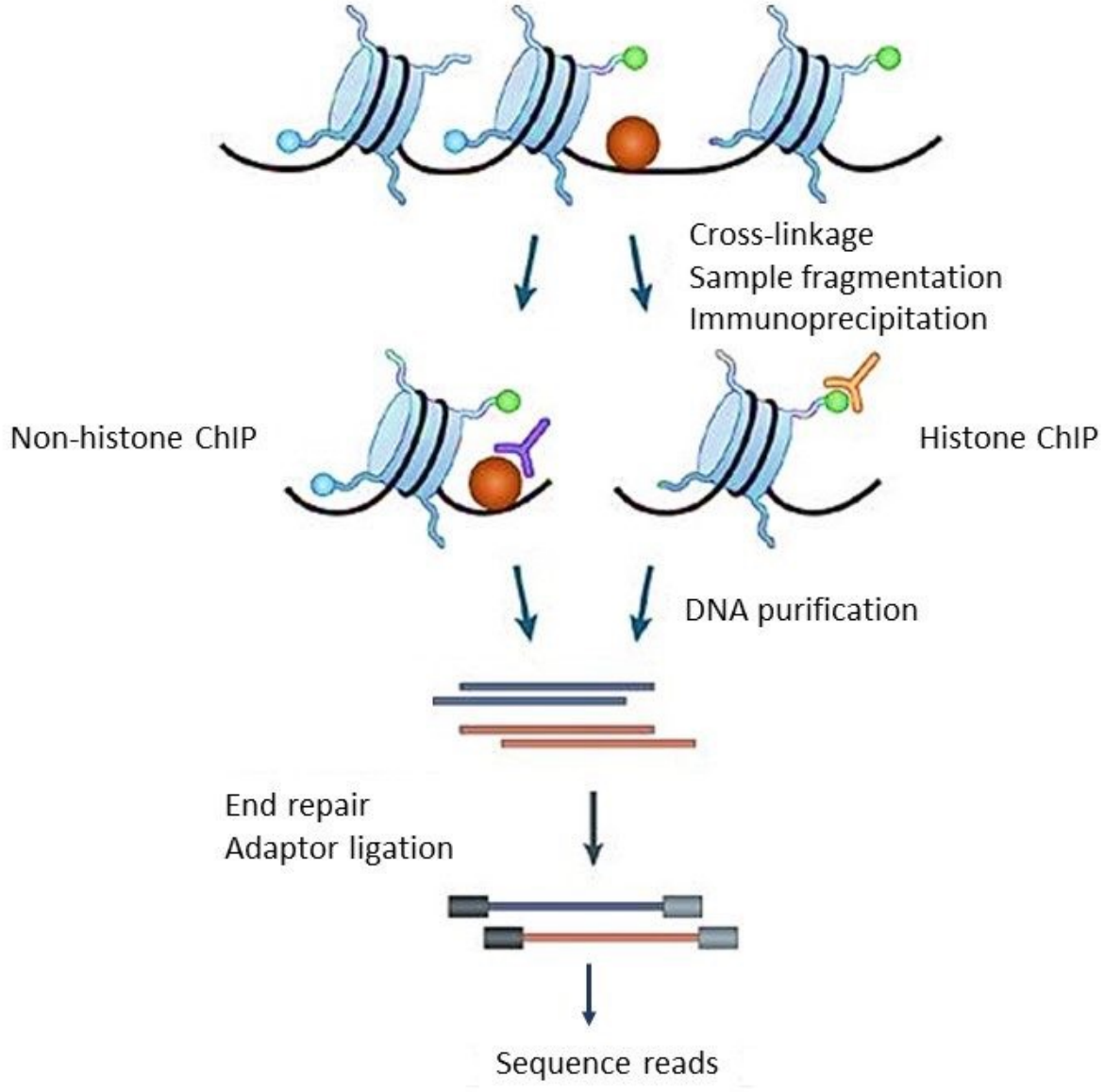


Image sourced from: https://www.france-genomique.org/technological-expertises/regulome/chip-sec

ChIP (Chromatin Immunoprecipitation)

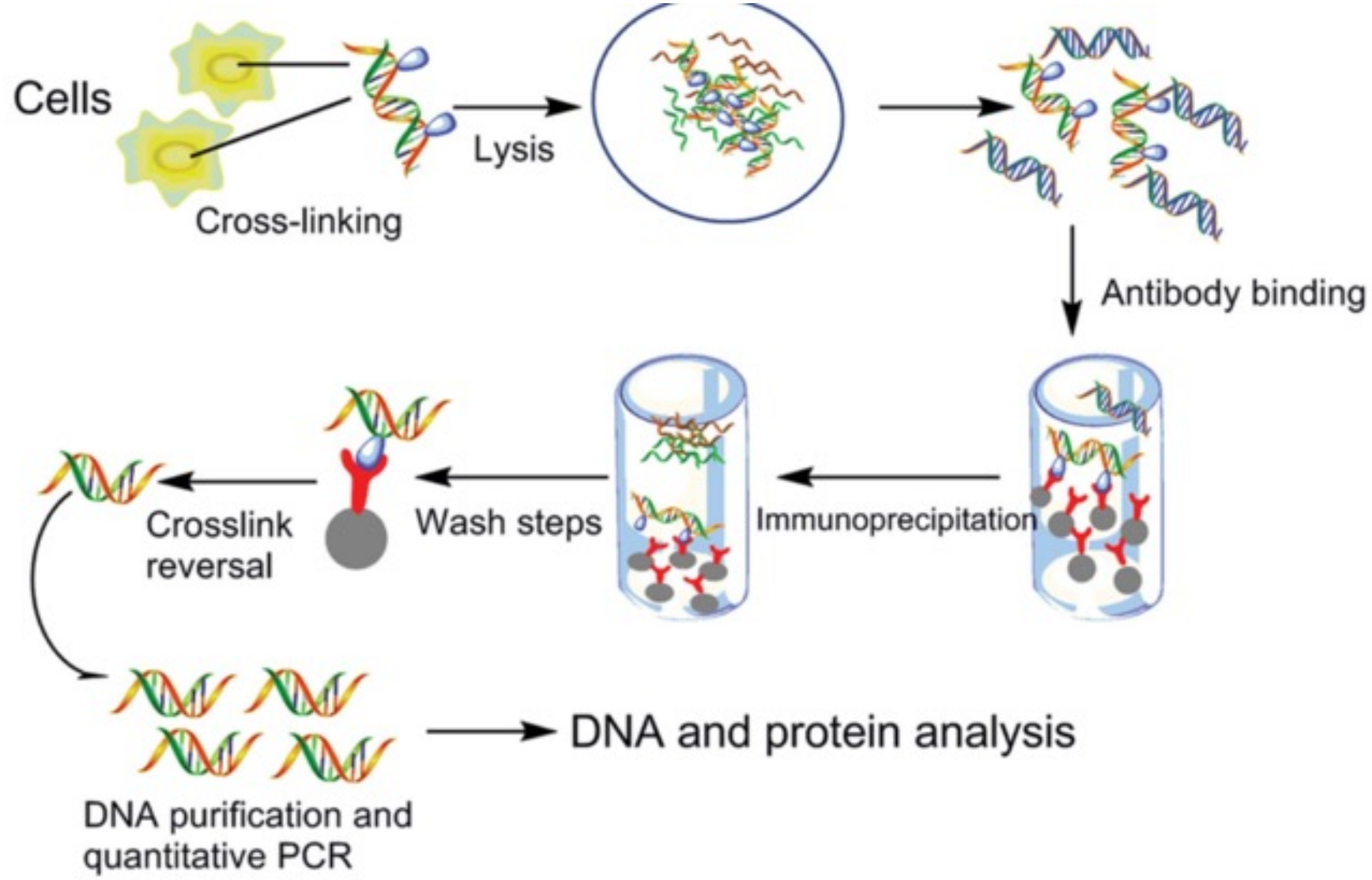


Image sourced from: https://www.sigmaaldrich.com/technical-documents/articles/biology/chromatin-immunoprecipitation-chip.html

illumina®

ChIP-Seq Recommended Read Depths

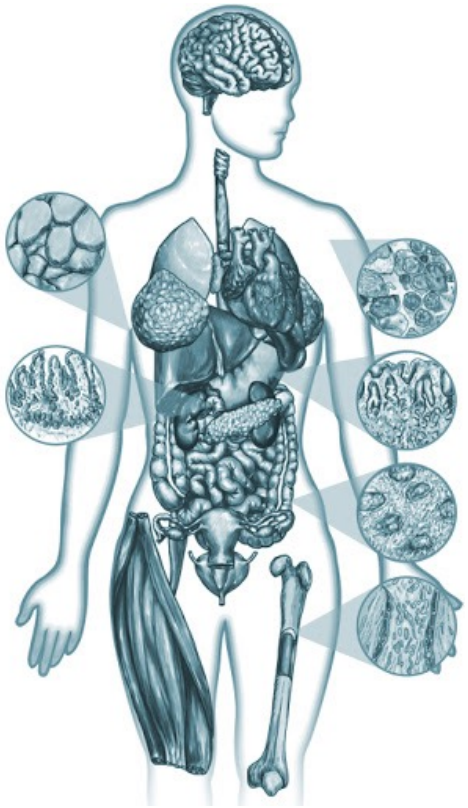
Transcription Factor ChIP-Seq:

- Each replicate should have 20 million reads

Histone ChIP-Seq:

- narrow-peak histone experiments: each replicate should have 20 million reads.
  - Ie: H3F3A, H3K27me3, H3K36me3, etc.

- broad-peak histone experiments: each replicate should have 45 million reads.
  - Ie: H2AFZ, H3ac, H3K27ac, etc.

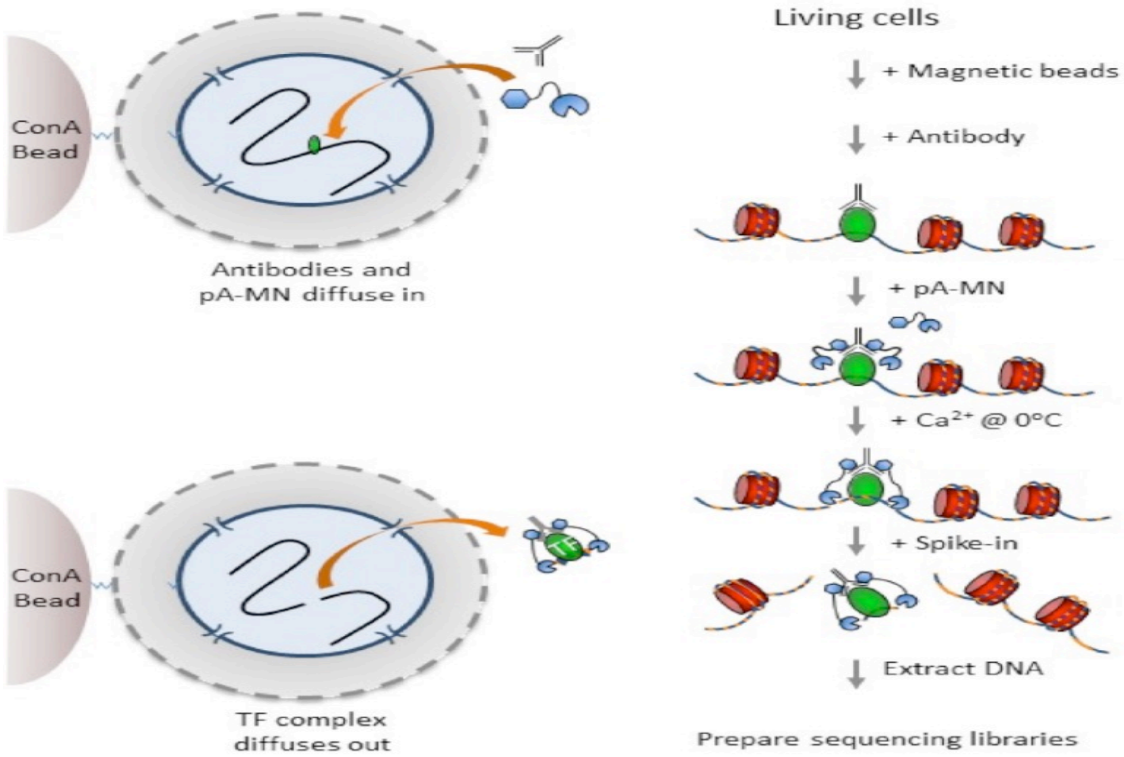


https://www.encodeproject.org/chip-seq/transcription_factor/
http://twww.roadmapepigenomics.org/files/protocols/data/histone-modification/REMC_ChIP-seqStandardsFINAL.pdf

illumina

CUT&RUN (Cleavage Under Targets and Release Using Nuclease)

- Epigenomic profiling strategy in which antibody-targeted controlled cleavage by micrococcal nuclease releases specific protein-DNA complexes into the supernatant



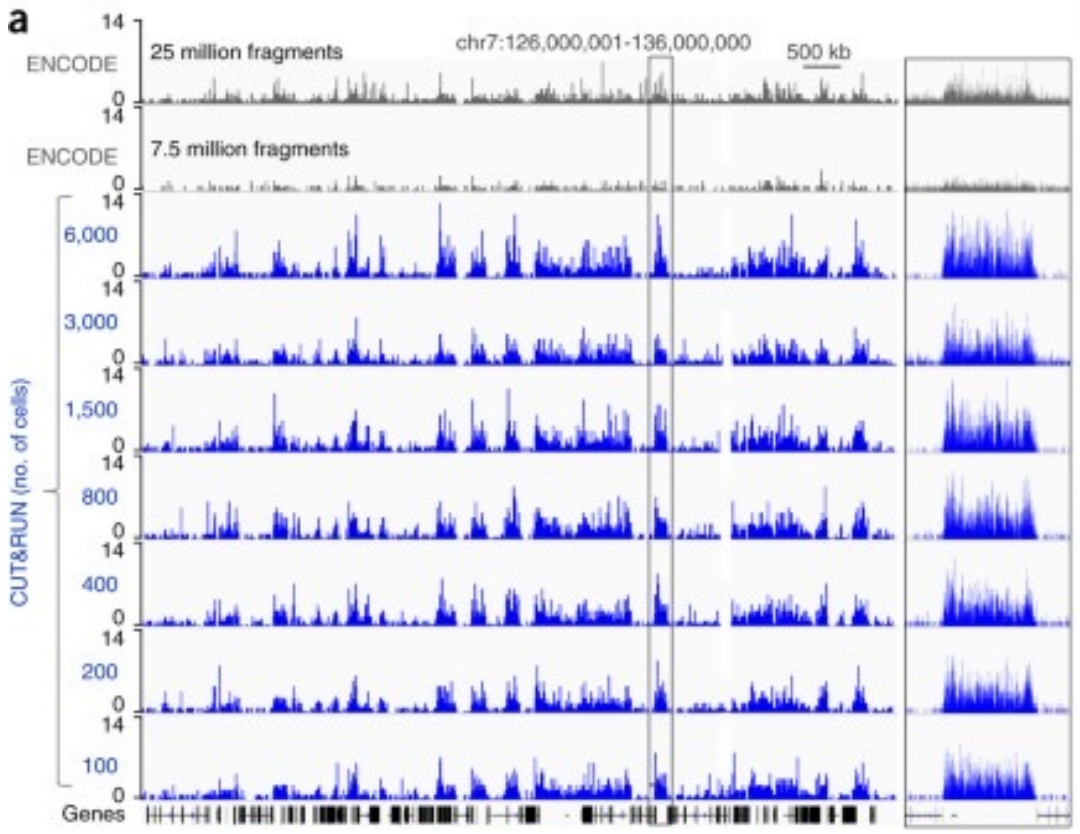
- Only the targeted fragments enter into solution, and the vast majority of DNA is left behind
- Exceptionally low background levels.
- Does not require isolation of nuclei
- 100 cells for histone modification / 1000 cells for TF binding

Skene, Peter J., Jorja G. Henikoff, and Steven Henikoff. "Targeted in situ genome-wide profiling with high efficiency for low cell numbers." Nature protocols 13.5 (2018): 1006.

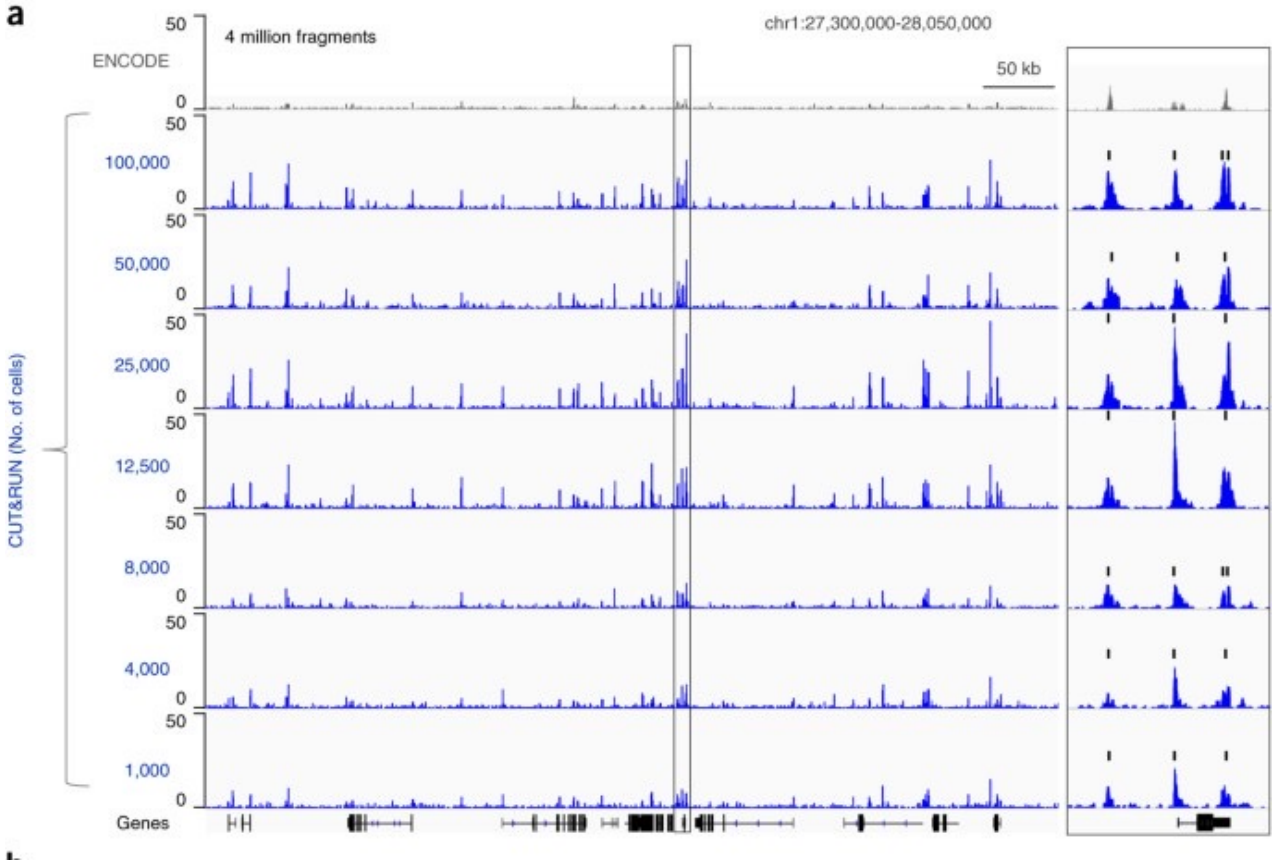
illumina

CUT&RUN Transcription Occupancy Profiles

H3K27me3 requires only 100



CTCF peaks requires only 1000 cells



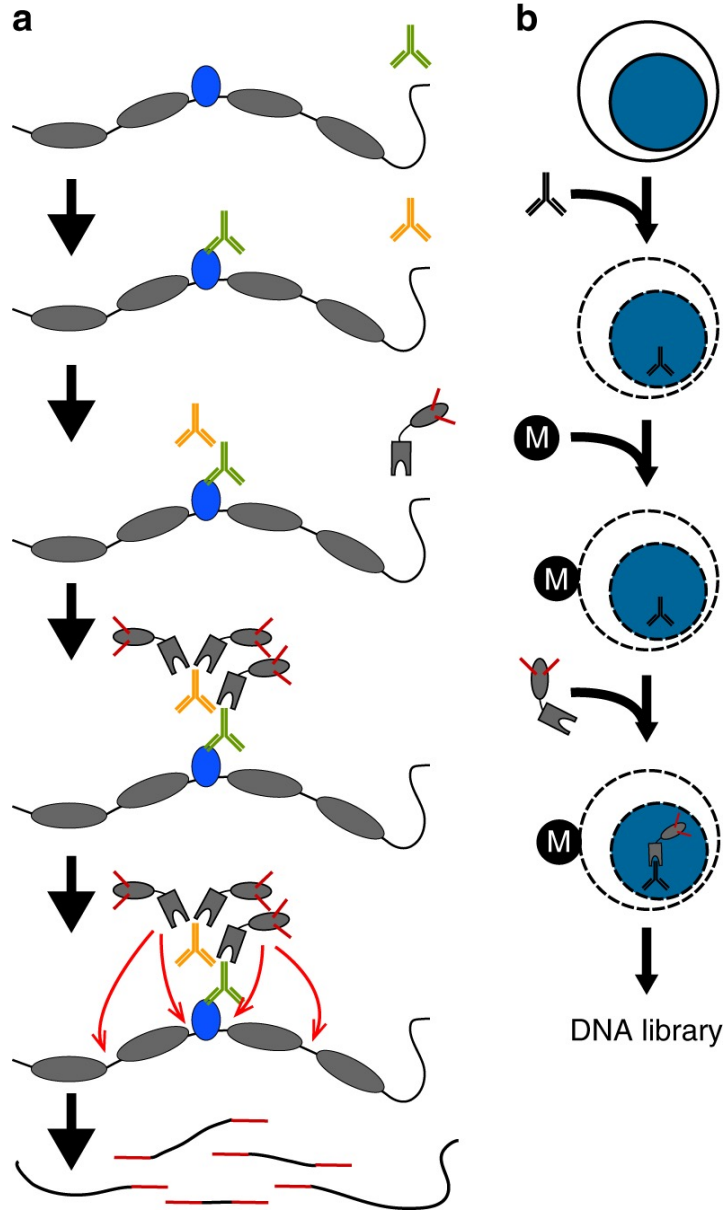
Skene, Peter J., Jorja G. Henikoff, and Steven Henikoff. "Targeted in situ genome-wide profiling with high efficiency for low cell numbers." Nature protocols 13.5 (2018): 1006.

illumina

CUT&TAG:

Cleavage Under Targets and Tagmentation

- Utilizes antibodies to bind to specific chromatin proteins *in situ*
- Lower background signals when compared to CHiP-seq libraries
- Protein A(pA)-Tn5 transposase is utilized in place of pA-Mnase from CUT&RUN-seq allowing for single cell resolution



Kaya-Okur et al., Nature Communications 2019

illumina

But Standard ChIP-Seq Still has its place…

- Transcription factors that are…
  - not abundantly expressed
  - weakly or transiently bind to DNA
  - indirectly bind to chromatin
- Most ChIP-validated antibodies are validated to work with fixed samples
- The Tn5 transposase used in Cut&Tag has a high affinity to open-chromatin regions

illumina

illumina

Mapping Long Range DNA Interactions with Hi-C

Why is Chromatin Conformation important? Expression is regulated by *interaction*

- **Promoters**
  - for Expression to start
- **Enhancers**
  - Distal non-coding elements that interact and amplify expression
- **Silencers and insulators**
  - Adjacent sequences that help “temper” expression signals
- **Transcription Factors**
  - DNA binding proteins that can assist in turning on and off gene expression
- **Repressors**
  - Proteins that inhibit expression through silencer binding

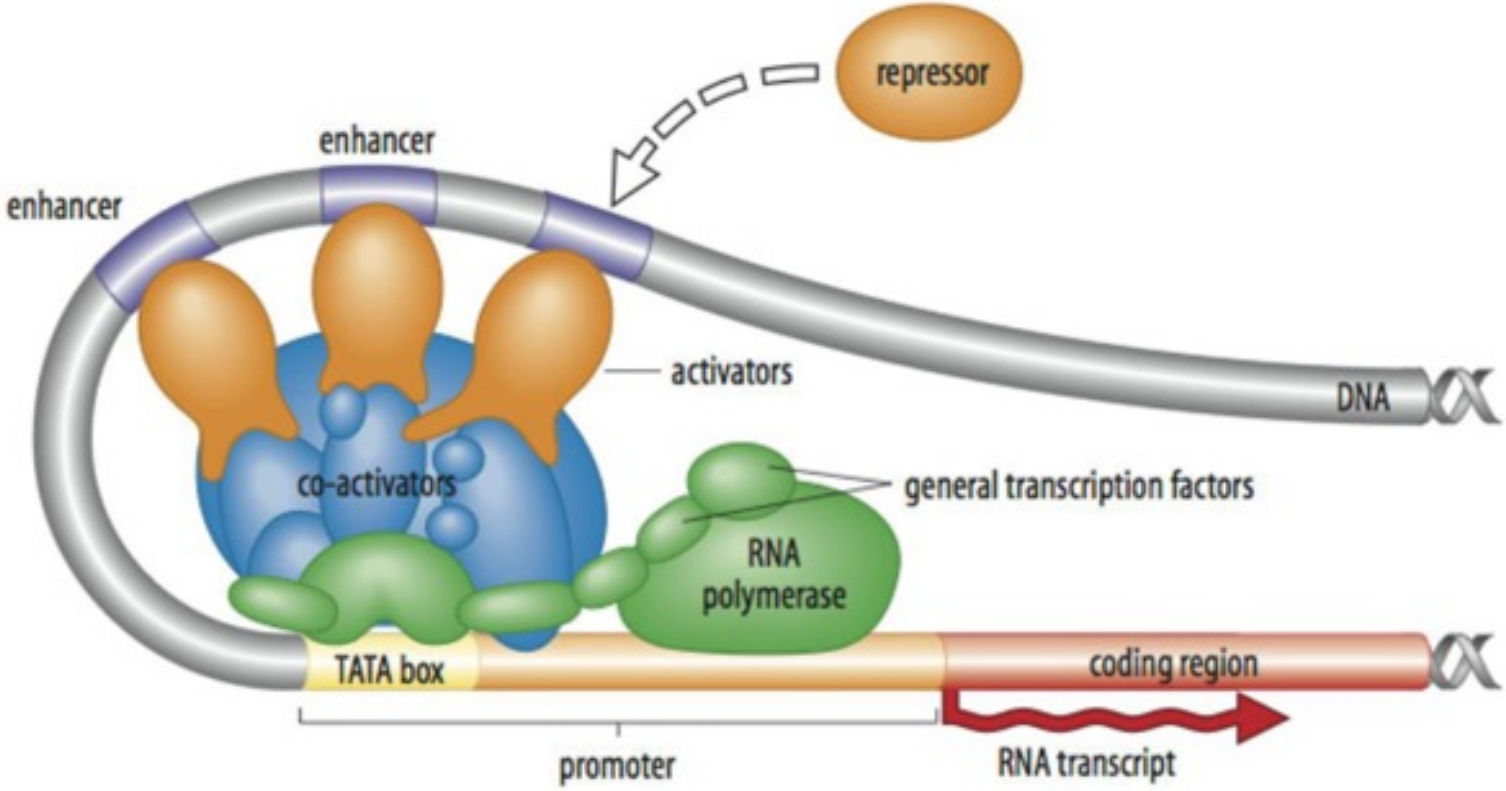
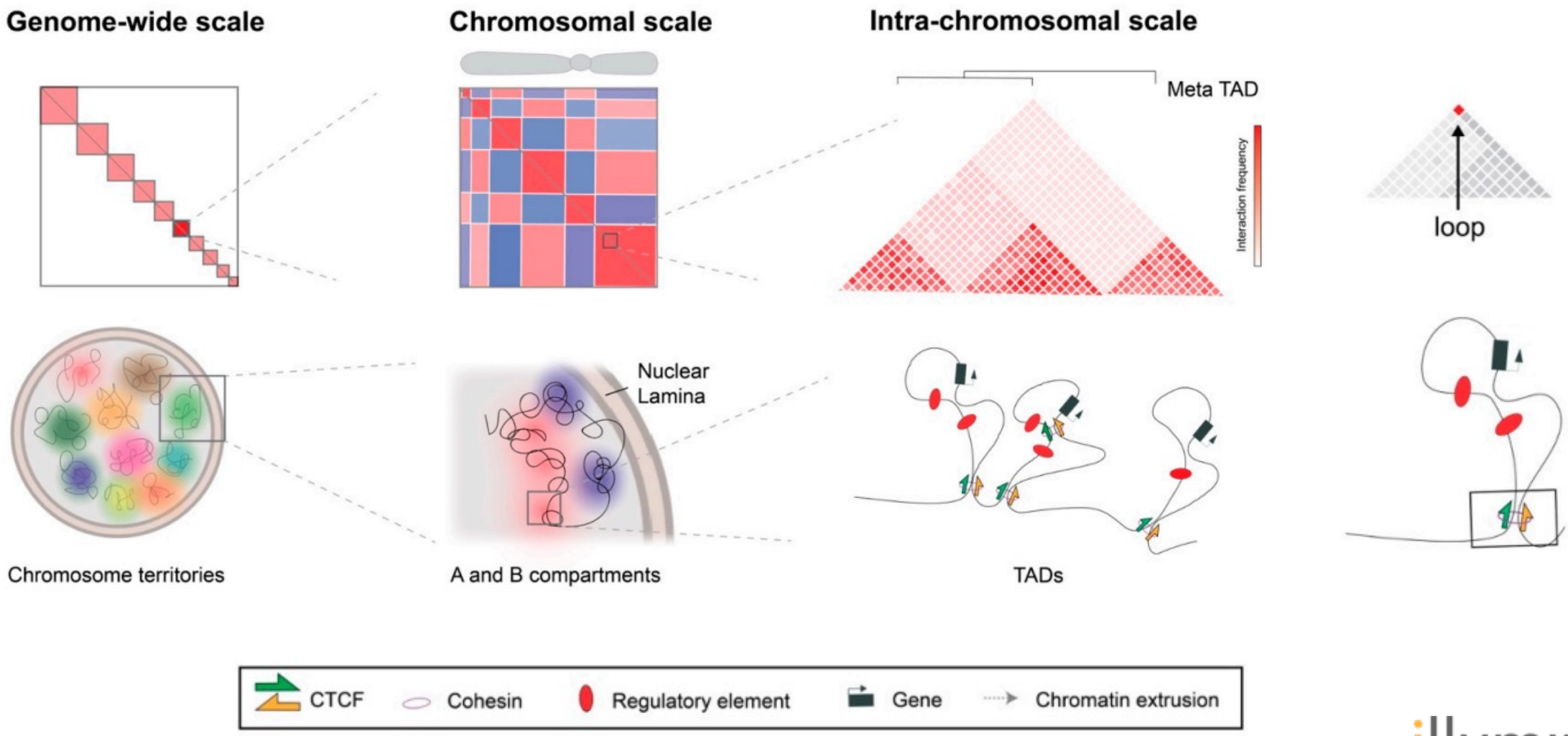


Image source: http://www.mun.ca/biology/desmid/brian/BIOL3530/DEVO_10/devo_10.html

illumina

Prominent Structures of 3D Nuclear Organization



Lupiáñez, et al., 2019

illumina

Hi-C

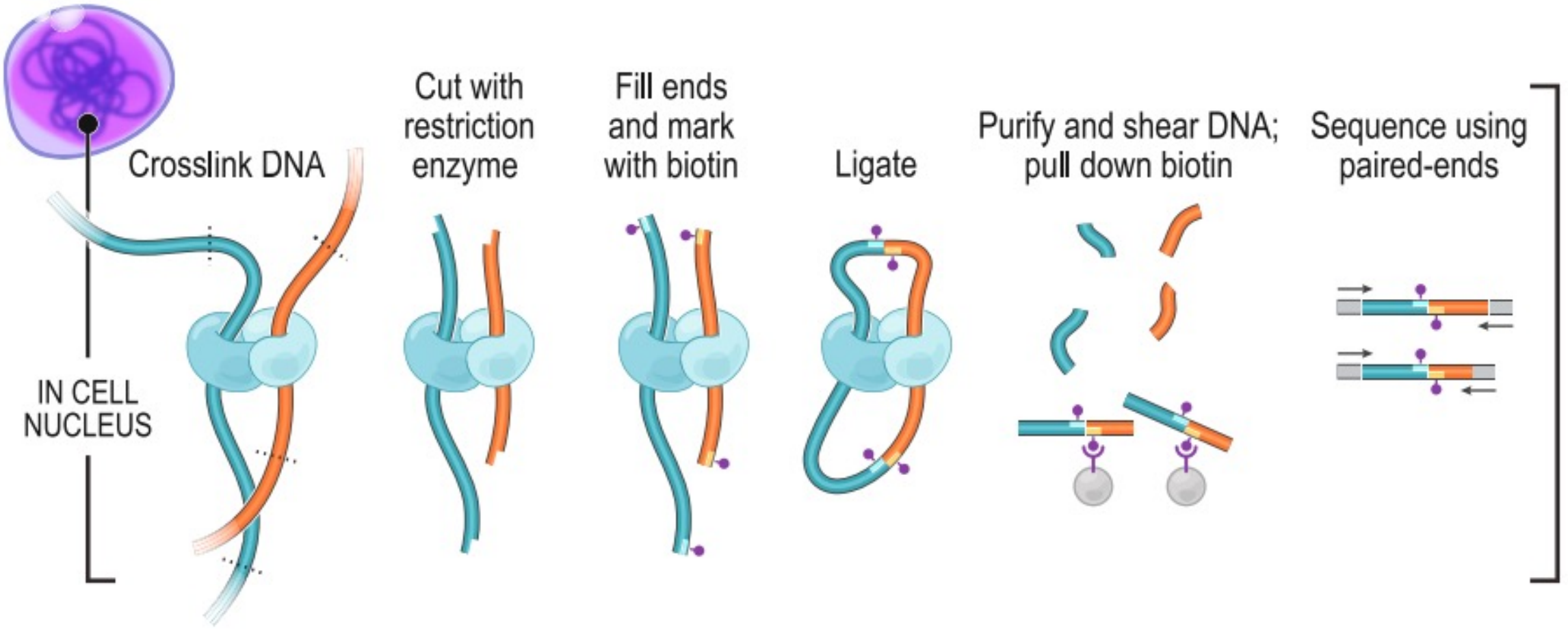
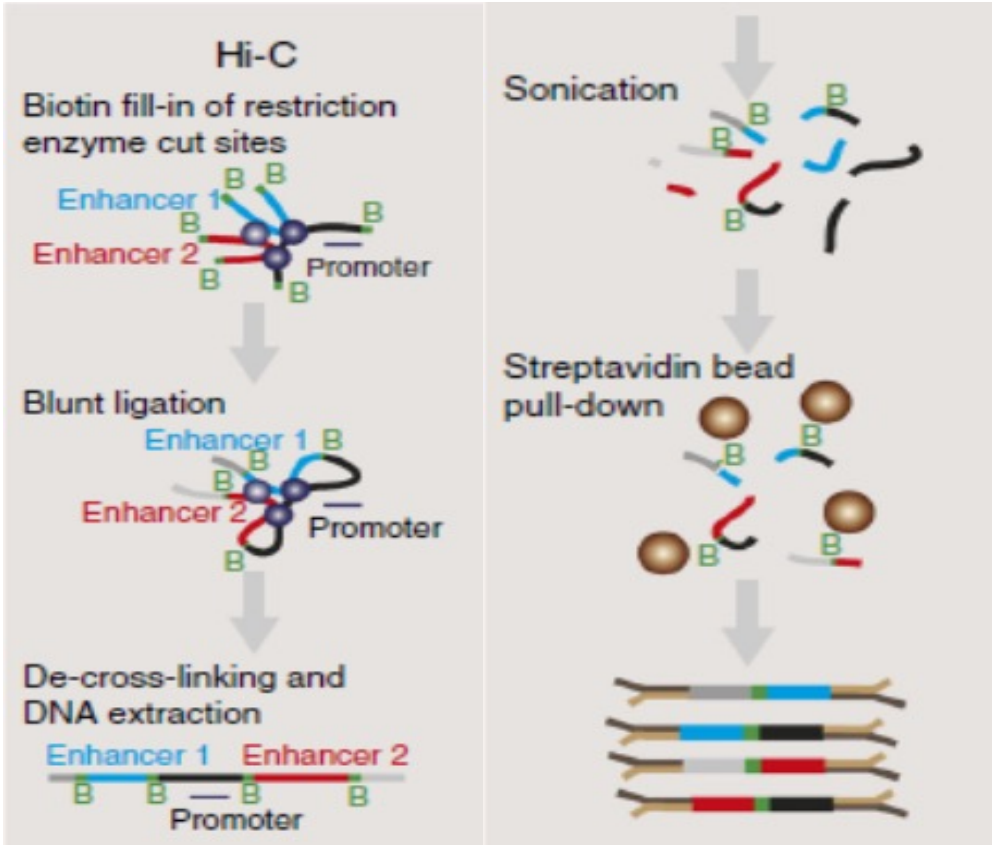


Image sourced from: https://www.jove.com/t/1869/hi-c-a-method-to-study-the-three-dimensional-architecture-of-genomes

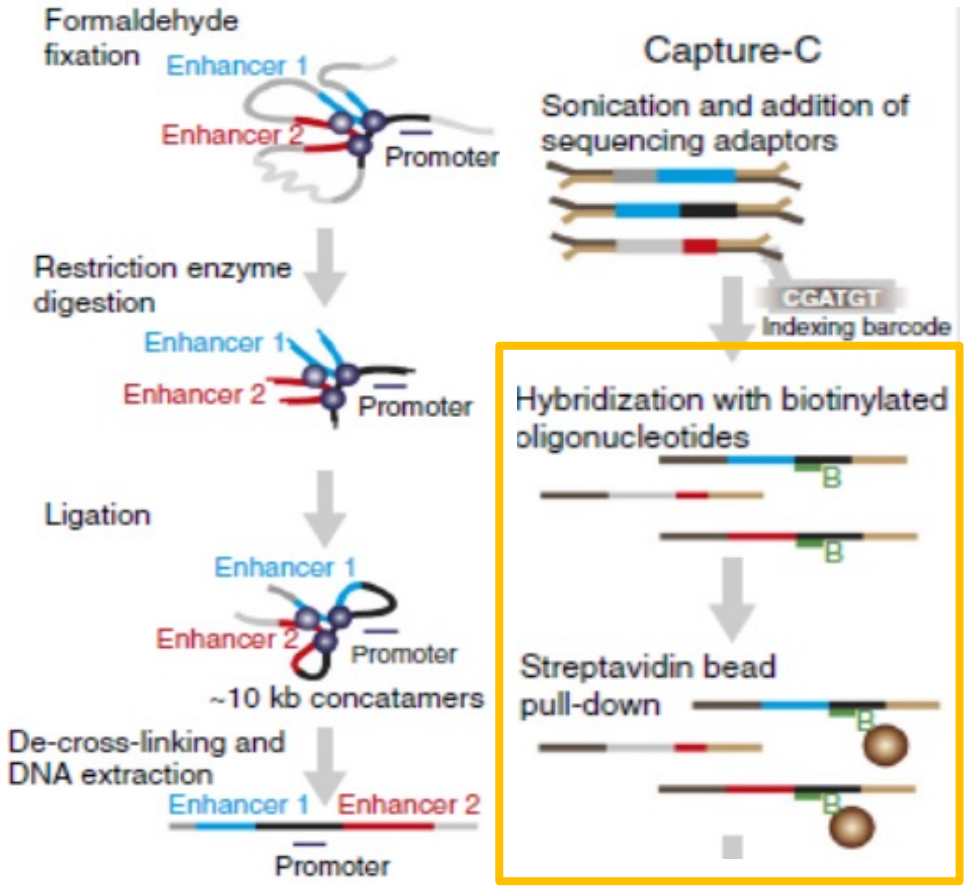
illumina

Hi-C: Whole Contactome Sequencing



- Highest Sequencing requirements
- Best for raw discovery

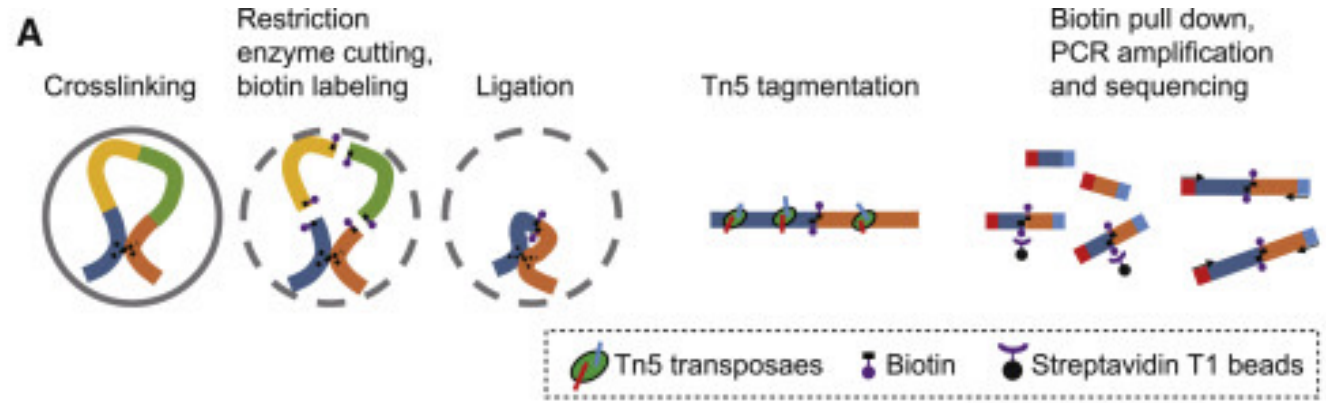
Capture-C: Targeted Contactome Sequencing

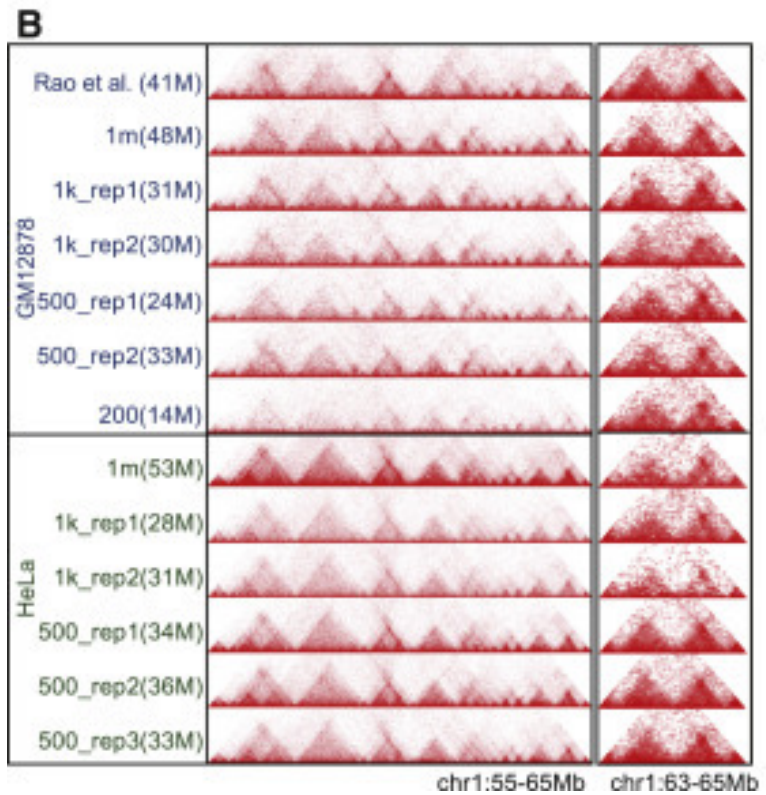


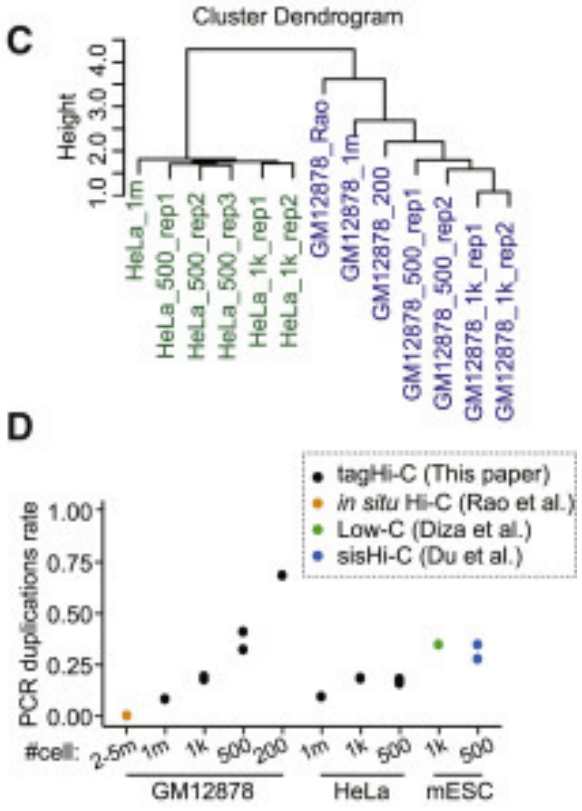
- Highest Resolution, focused
- Less sequencing depth

illumina

tagHi-C





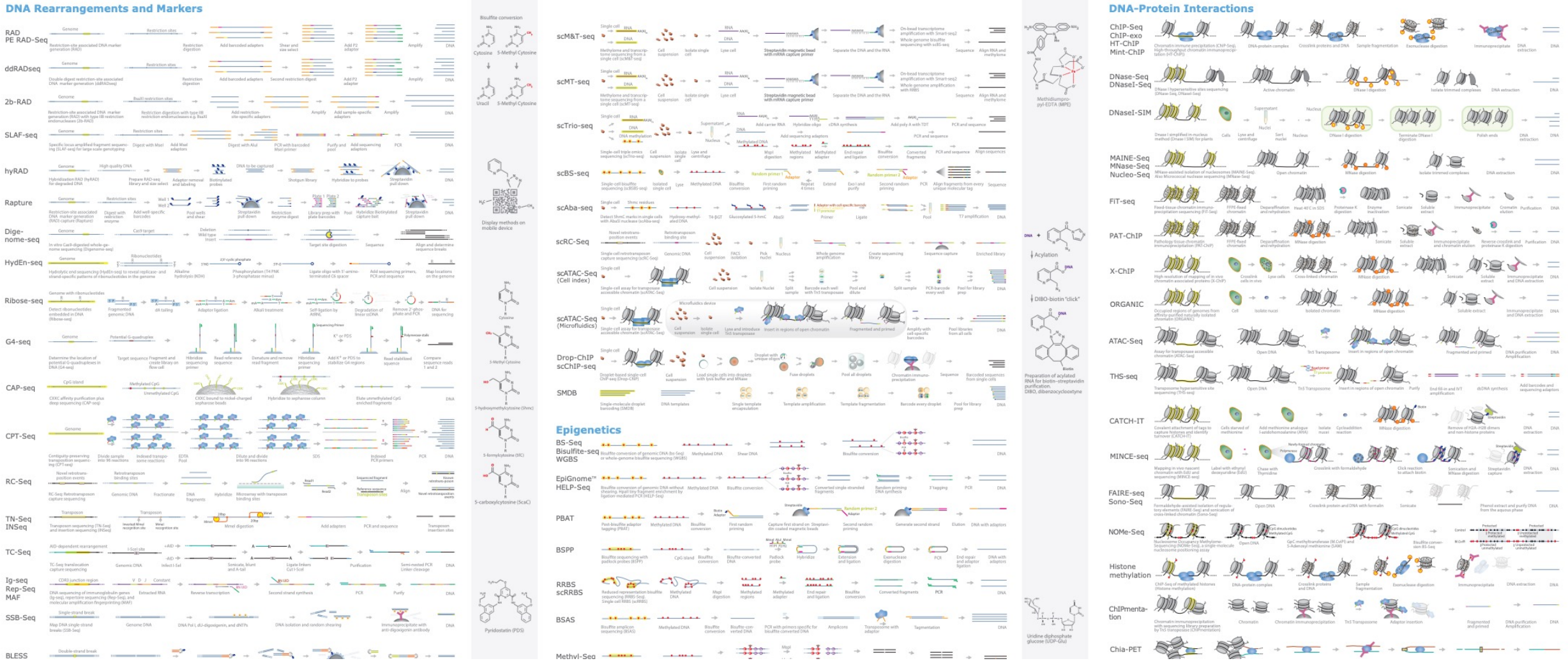


- Low input: 200-1000 cells
- Incorporates tagmentation into the fragmentation step to simultaneously add on anchor adapters
- 2-day workflow

Cheng, et al., 2020

DNA

For all you seq...



We are here for you!

Illumina:

Lauren Robinson
Sequencing and Informatics Specialist
lrobinson@illumina.com

Robin Ball
Territory Account Manager
rball@illumina.com

Illumina Technical Support
+1 800 809 4566
techsupport@illumina.com

Indiana University Center for Genomics & Bioinformatics:

Matthew Hahn, PhD, Director
mwh@Indiana.edu

Bloomington

Indiana School of Medicine Center for Medical Genomics:

Yunlong Liu, PhD, Director
yunliu@iu.edu

Notre Dame Genomics & Bioinformatics Core Facility

Michael Pfrender, PhD, Director
Michael.Pfrender.1@nd.edu

Purdue University Bindley Core for Genomics

Phillip SanMiguel, PhD, Director
pmiguel@purdue.edu

35

illumina

Questions?

lrobinson@illumina.com