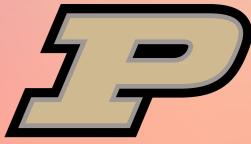
An introduction to home-brew library prep methods for Illumina sequencing and critical factors to ensure success: ChIP-seq, ATAC-seq, Hi-C, and CRISPR screening

Lauren Robinson

Sequencing and Informatics Specialist OH/MI/IN/KY







Agenda

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

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



screening#:~:text=CRISPR%20screening%20is%20used%20to,drug%20resistance%20and%20drug%20sensitivity.

4 https://haematologica.org/article/view/8891

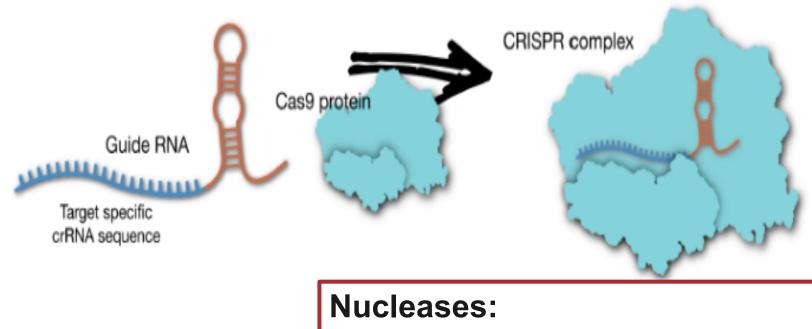


https://www.idtdna.com/pages/education/decoded/article/overview-what-is-crispr-

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..

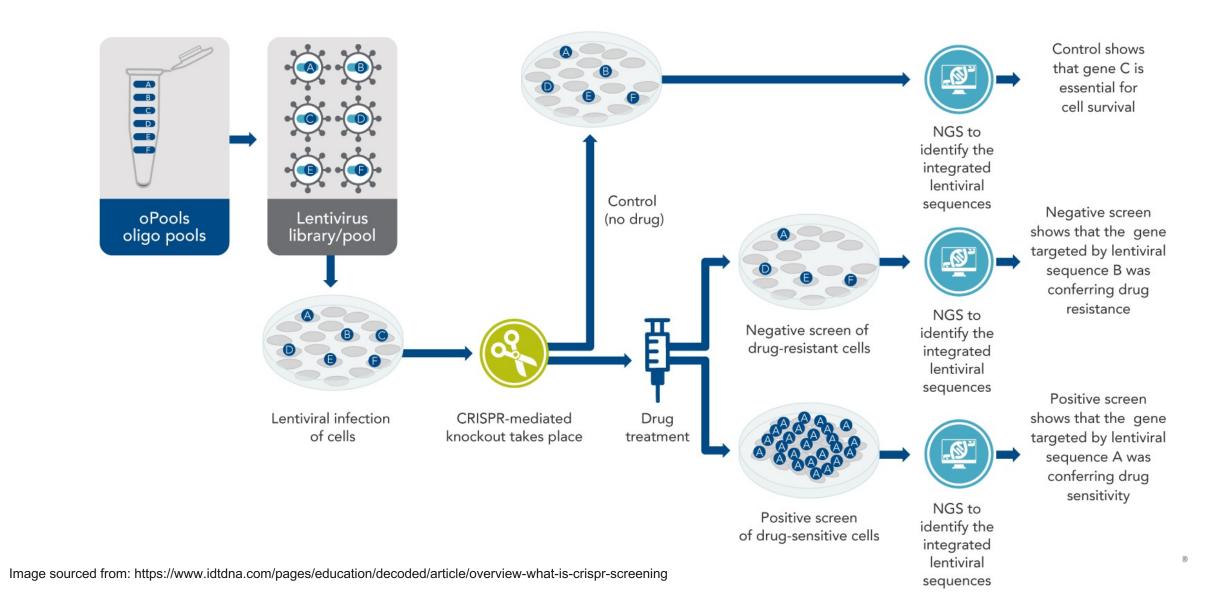


Proteins that bind to gRNAs and are directed by the gRNAs to cut particular DNA sequences. **Cas9** is a very commonly used CRISPR nuclease



Pooled CRISPR Screen Overview

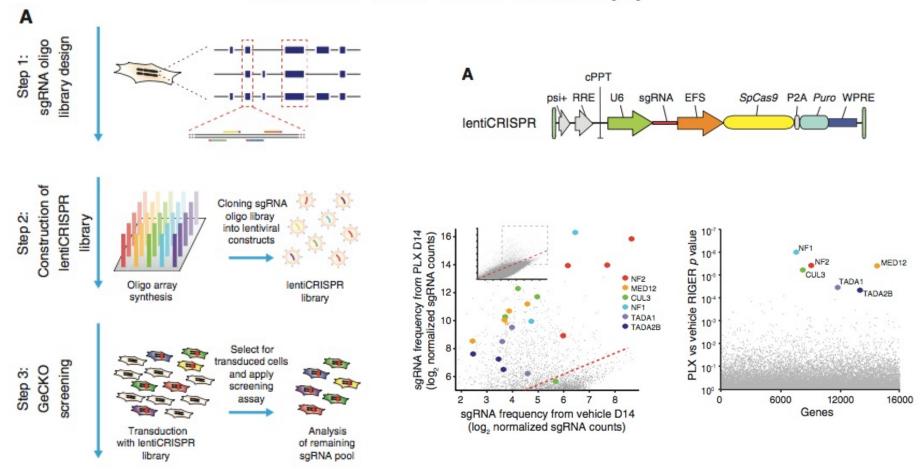
6



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

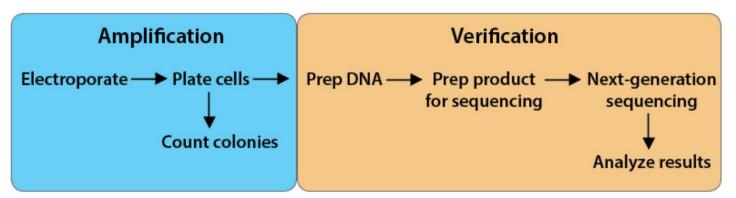
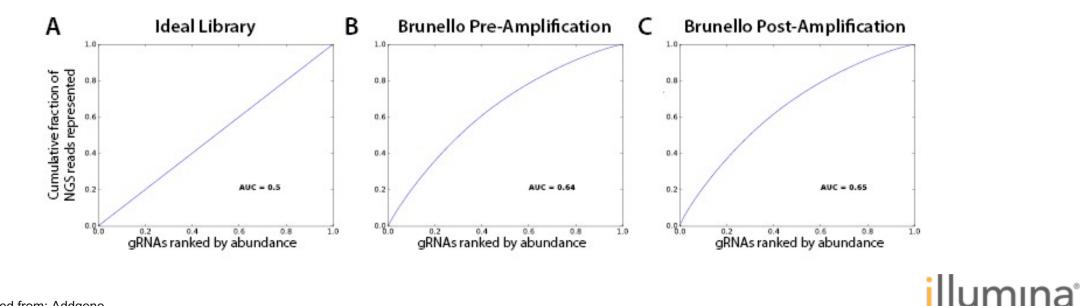
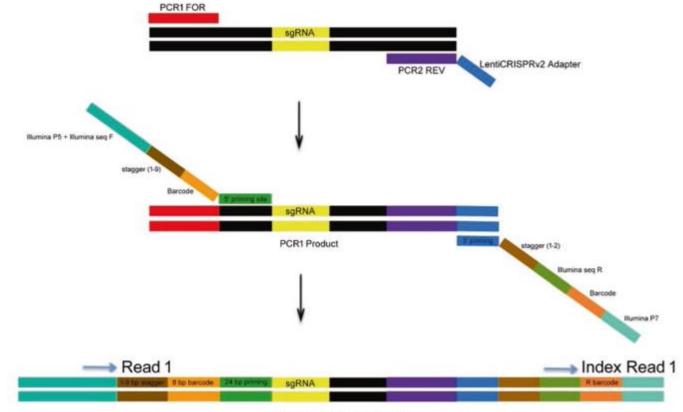


Figure 1. An overview of the pooled library amplification/verification process.



Hit identification using Illumina Sequencing



PCR2 Product (amplicon ~330bp)

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.



Combining CRISPR and Single Cells

Cell

Resource

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

Authors

Aviv Regev

In Brief

scale possible

Correspondence

aregev@broadinstitute.org

Atray Dixit, Oren Parnas, Biyu Li, ...,

Jonathan S. Weissman, Nir Friedman,

A technology combining single-cell RNA

perturbations termed Perturb-seg makes

analyzing complex phenotypes at a large

sequencing with CRISPR-based

Graphical Abstract Pooled CRISPR screens Single cell RNA-Seq of barcoded droplets (3) Next-Gen sequencing of library -88 8 Y = X BRegularized linear mode number of cells Regulatory inference Epistasis Power from perturbations decompositio analysis

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

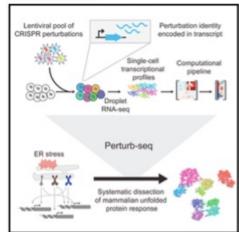




Cell

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

Resource

Correspondence

jonathan.weissman@ucsf.edu

In Brief

A strategy for barcoding CRISPRmediated 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

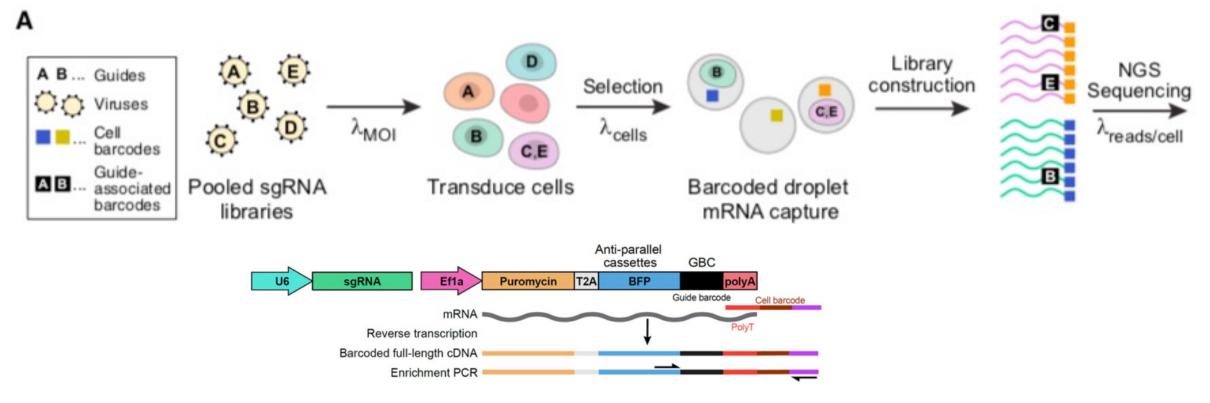
Adamson et al., 2016, Cell 167, 1867–1882 molitei December 15, 2016 © 2016 Elsevier Inc. http://dx.doi.org/10.1016/j.cell.2016.11.048





Perturb-Seq:

Genetic Screening with CRISPR



Uniquely Single Cell NGS

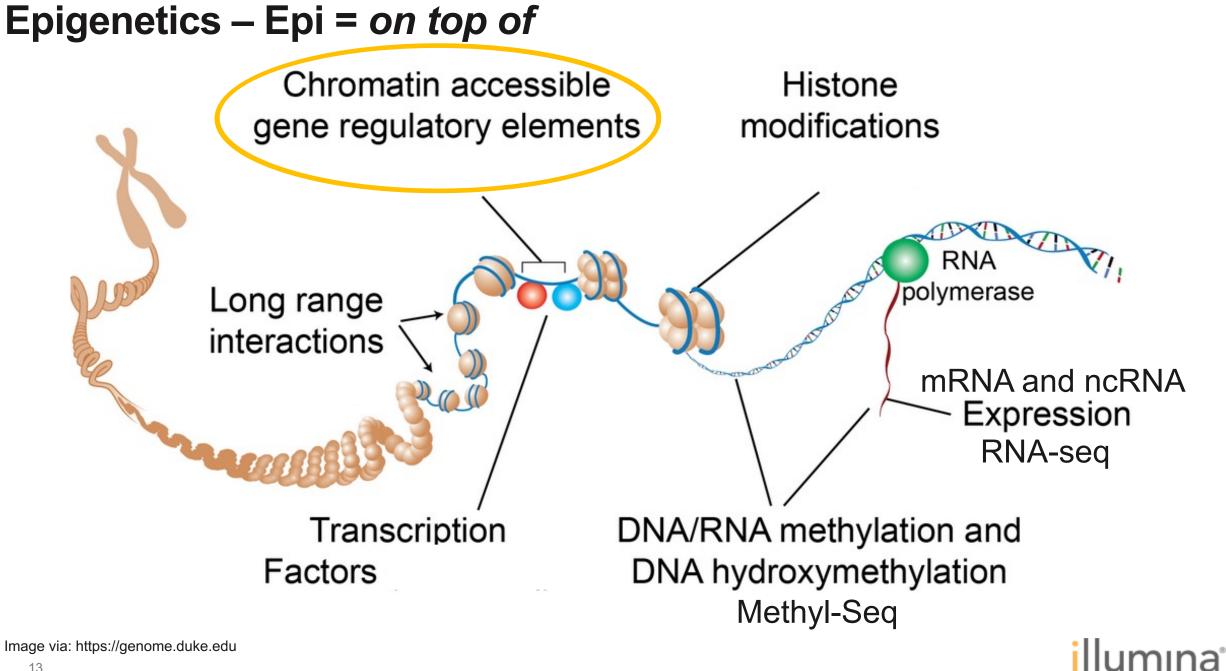
Combines ability to screen many knockout targets simultaneously, both as single gene and combinatorial knockouts paired with rich digital read out of the exact guide RNA and Cell ID as well as the effect on gene expression phenotype.



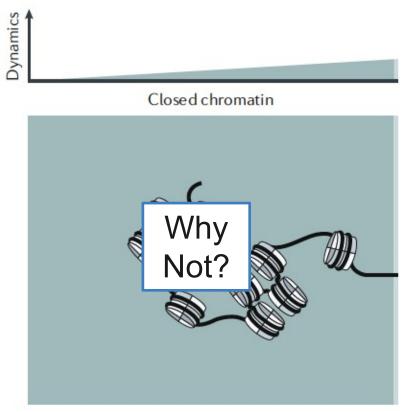
illumina

11 Dixit et al 2016 Cell 167, 1853-1866

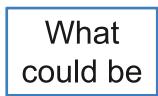
ATAC-seq or sequencing "openness"



ATAC-Seq – Why study chromatin accessibility?



Heterochromatin ~97%



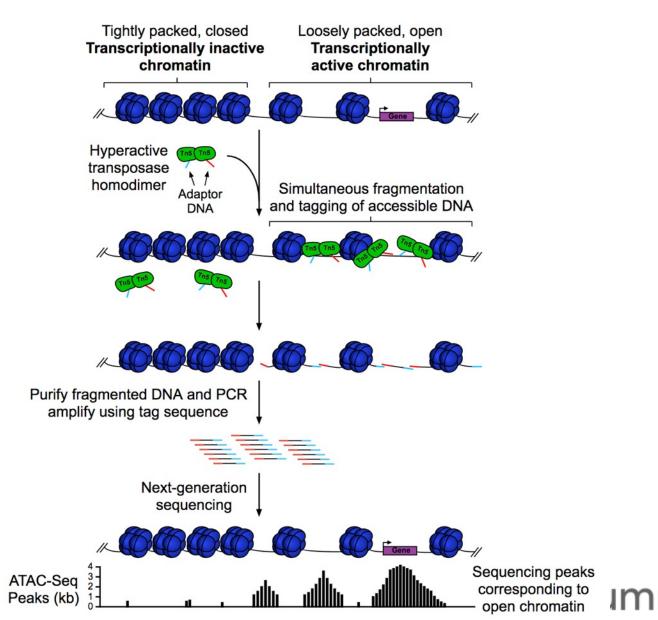




DOI:10.1038/s41576-018-0089-8

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



¹⁵ Image sourced from Wikimedia

For Research Use Only. Not for use in diagnostic procedures.

Bulk ATAC-seq has been improved: Omni-ATAC

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

An improved ATACseq protocol reduces background and enables interrogation of frozen tissues

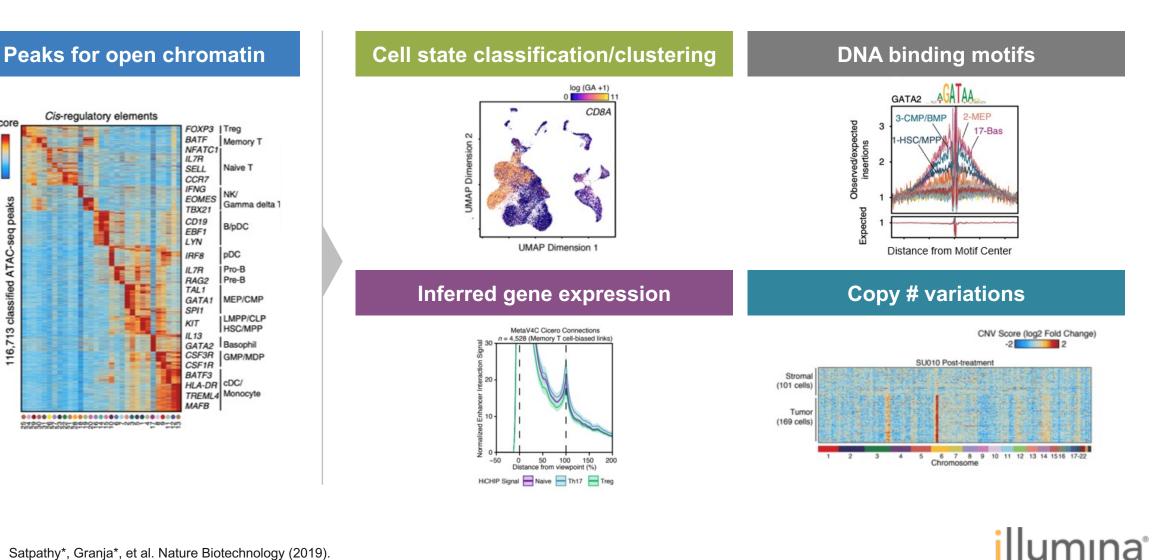
2018

M Ryan Corces^{1,2}, Alexandro E Trevino^{1,3,4}, Emily G Hamilton⁵, Peyton G Greenside^{3,6}, Nicholas A Sinnott-Armstrong³, Sam Vesuna⁴, Ansuman T Satpathy^{1,7}, Adam J Rubin², Kathleen S Montine⁷, Beijing Wu³, Arwa Kathiria³, Seung Woo Cho^{1,2}, Maxwell R Mumbach^{1,3}, Ava C Carter^{1,2}, Maya Kasowski^{1,7}, Lisa A Orloff⁸, Viviana I Risca³, Anshul Kundaje^{3,9}, Paul A Khavari², Thomas J Montine⁷, William J Greenleaf^{1,3,10} & Howard Y Chang^{1,2}

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



Single Cell ATAC seq One quick protocol – multiple readouts



Z-score

-2

peaks

116,713 classified ATAC-seq

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

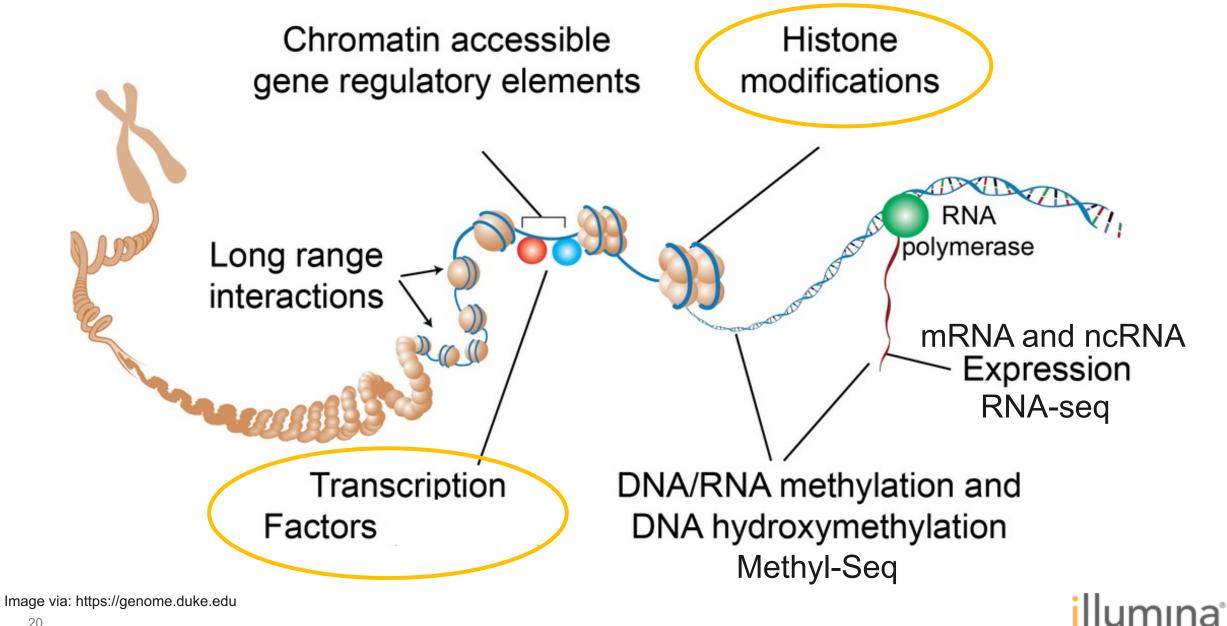
Research Goal	Recommended Depth
Identification of open chromatin differences in human samples	>50M paired-end reads
Transcription factor foot printing to construct gene regulatory networks	>200M paired-end reads
Single-cell analysis	50K paired-end reads per nucleus/cell

*read depth may vary depending on experimental conditions

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

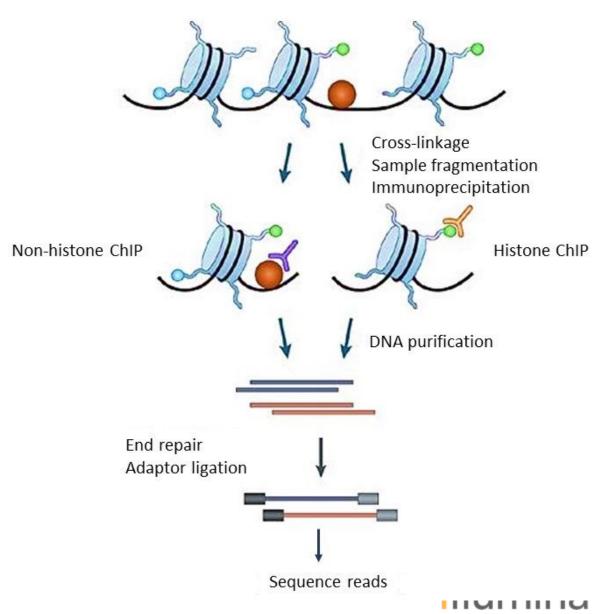
Mapping Protein DNA Interactions with ChIP-Seq

Epigenetics – Epi = on top of



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



ChIP (Chromatin Immunoprecipitation)

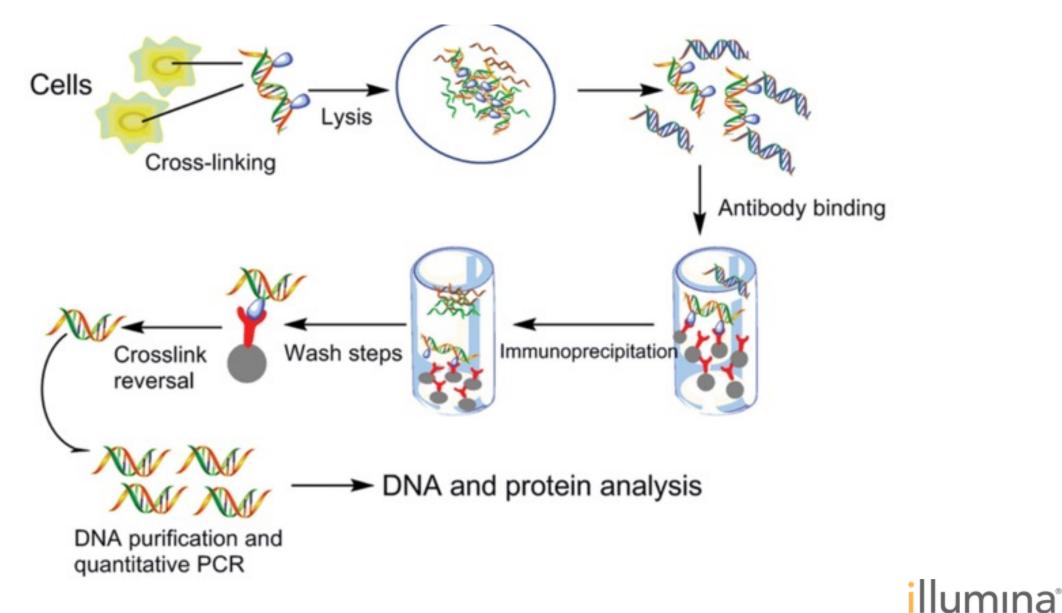


Image sourced from: https://www.sigmaaldrich.com/technical-documents/articles/biology/chromatin-immunoprecipitation-chip.html

ChIP-Seq Recommended Read Depths

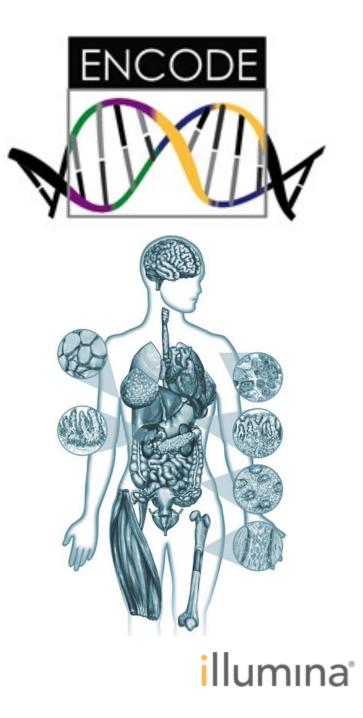
Transcription Factor ChIP-Seq:

• Each replicate should have 20 million reads

Histone ChIP-Seq:

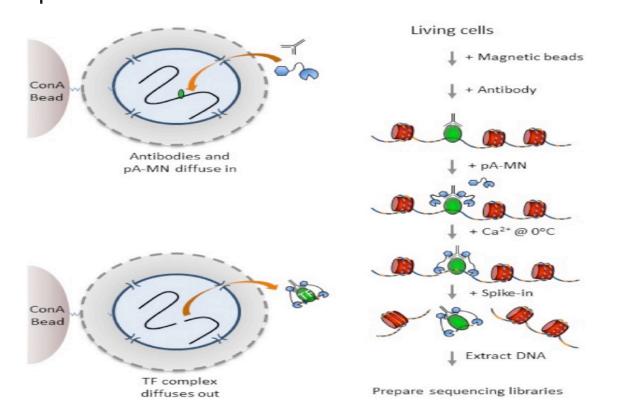
- <u>narrow-peak histone experiments</u>: each replicate should have 20 million reads.
 - Ie: H3F3A, H3K27me3, H3K36me3, etc.
- <u>broad-peak histone experiments</u>: each replicate should have 45 million reads.
 - Ie: H2AFZ, H3ac, H3K27ac, etc.

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



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

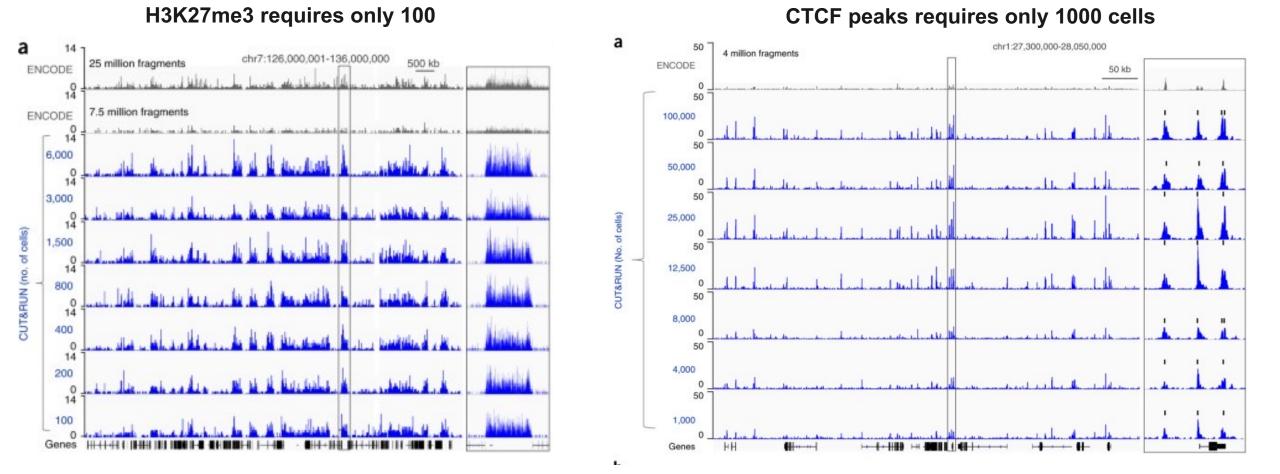


- Only the targeted fragments enter into solution, and the vast majority of DNA is left behind
- Exceptionally low background levels.
- Does not require isolation of nuclei
- 100 cells for histone modification
 / 1000 cells for TF binding

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.



CUT&RUN Transcription Occupancy Profiles



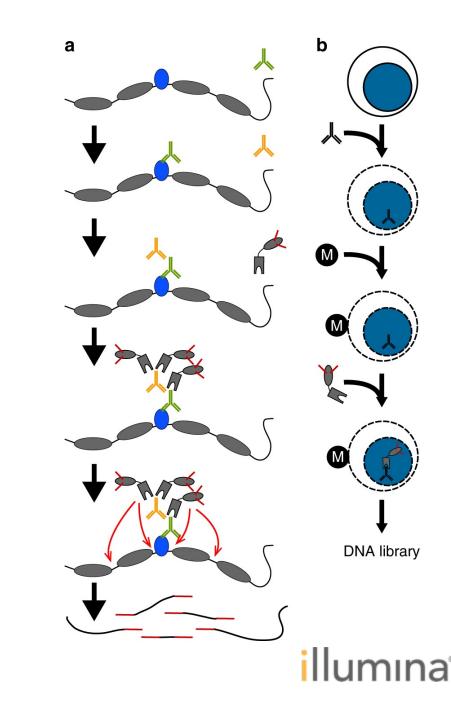
Skene, Peter J., Jorja G. Henikoff, and Steven Henikoff. "Targeted in situ genome-wide profiling with high efficiency for low cell numbers." Nature protocols 13.5 (2018): 1006.

illumina

CUT&TAG:

<u>Cleavage</u> <u>Under</u> <u>Targets</u> and <u>Tag</u>mentation

- 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



But Standard ChIP-Seq Still has its place...

- Transcription factors that are...
 - not abundantly expressed
 - weakly or transiently bind to DNA
 - o 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





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

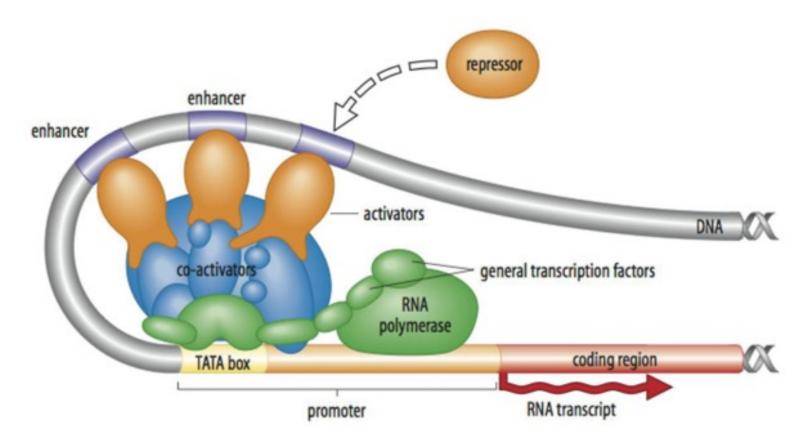
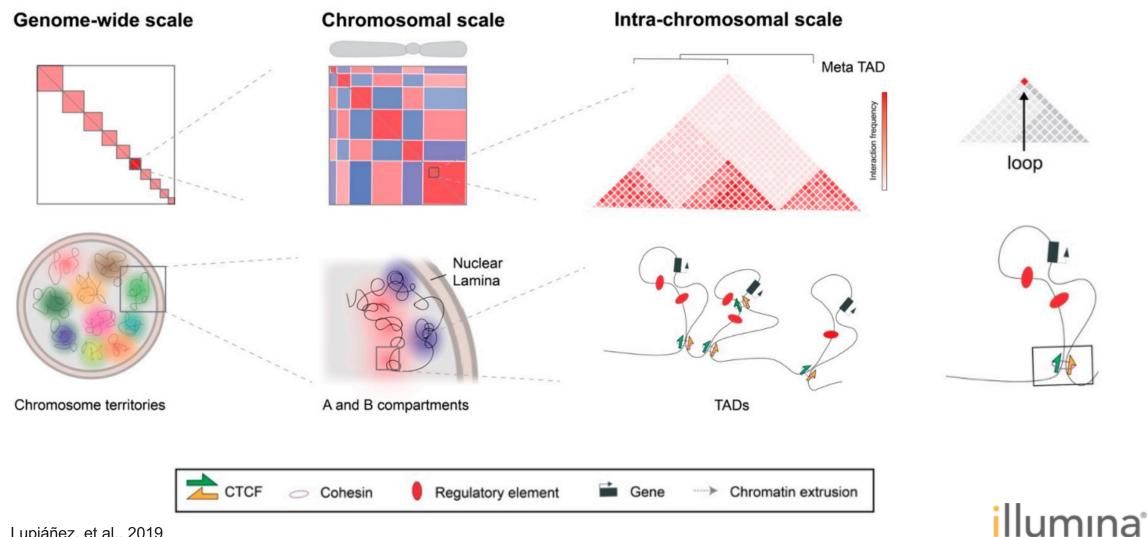


Image source: http://www.mun.ca/biology/desmid/brian/BIOL3530/DEVO_10/devo_10.html

illumina

Prominent Structures of 3D Nuclear Organization



Hi-C

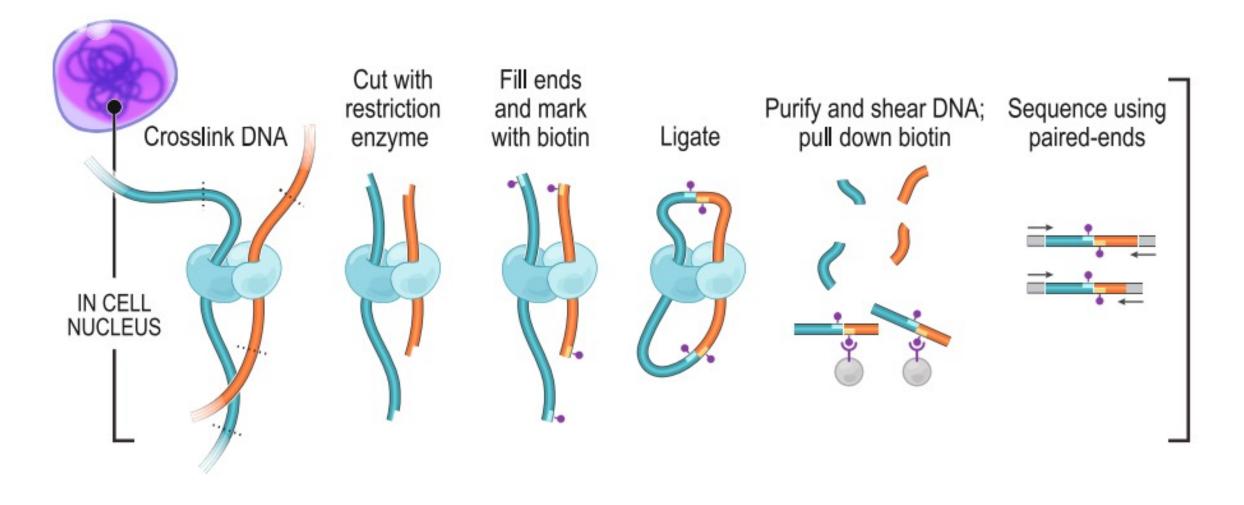
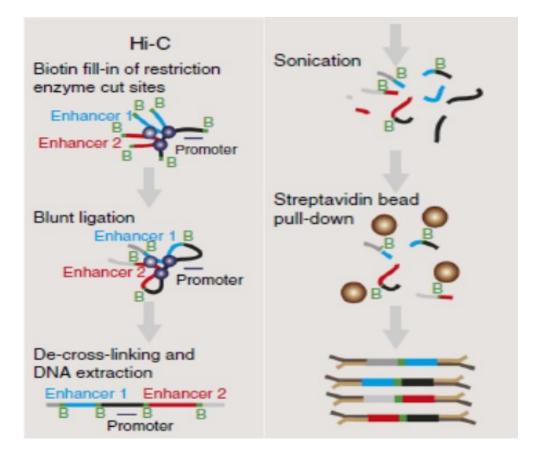


Image sourced from: https://www.jove.com/t/1869/hi-c-a-method-to-study-the-three-dimensional-architecture-of-genomes

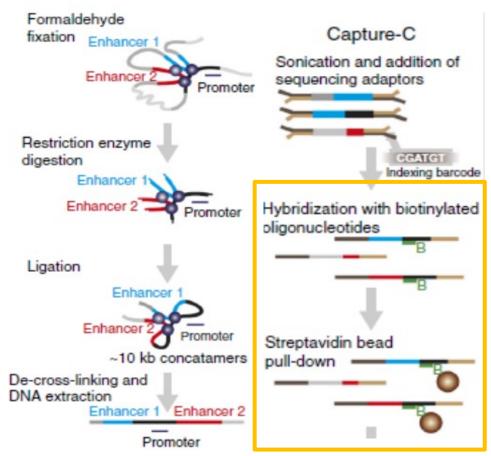


<u>**Hi-C</u>**: Whole Contactome Sequencing</u>



- Highest Sequencing requirements
- Best for raw discovery

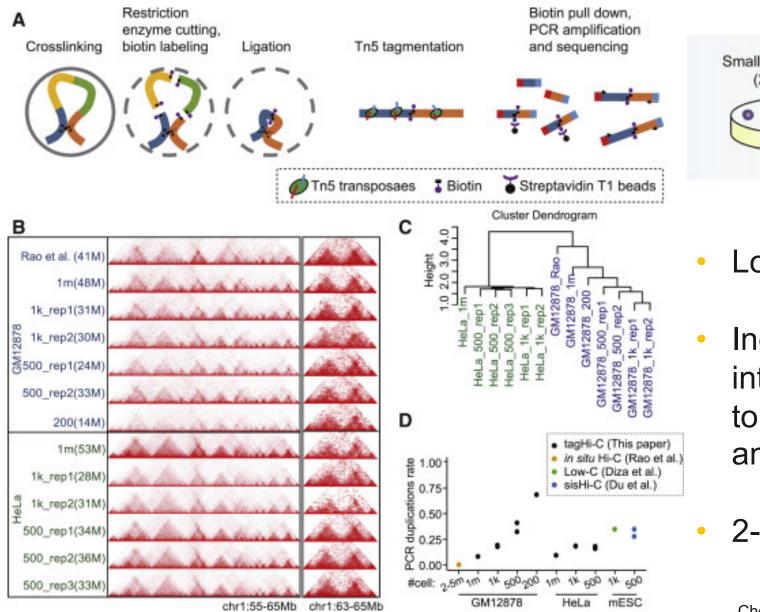
<u>Capture-C</u>: Targeted Contactome Sequencing

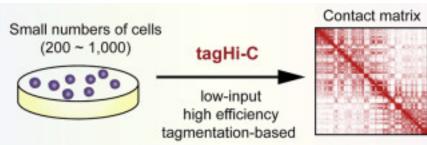


- Highest Resolution, focused
- Less sequencing depth illumina^{*}

32

tagHi-C





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

illumina

Cheng, et al., 2020

For all you seq... Grand D

Doo y a

DNA Rearrangements and Markers

RAD	Genome Austriction sites	10, 10, 10, 10, 10, 10, 10, 10, 10, 10,
PE RAD-Seq	Rutriction-Jin associated DNA-markee Restriction Add barcoded adapters Shear and Add P2 Amplify DNA generation (MAD) digetSen size select adapter	Cytosine S-Methyl Cytos
ddRADseq	General Restriction data Image: Construction data	+ + ,L ,L.
2b-RAD	Genute Institution of a second DNA meter Institutio	Uracii 5-Methyl Cytosi
SLAF-seq	Genome Institution step	
hyRAD	Convert High qualty Disk. Disk bit contrast - Convert	
Rapture	Encoder Instruction Mail Part of the second sec	H ₂ C - CH
Dige- nome-seq	Genome Curl tapet Duktom Agrico di Astrono Balancia di Astrono Bal	mobile device
HydEn-seq	Generer Banchestellity [104,
Ribose-seq	Construction and Induced Information Interview Intervi	
G4-seq	Dentrie & guadaptive + + + + + + + + + + + + + + + + + + +	
CAP-seq	Operation Implementation Implementati	S-Methyl Cytastee
CPT-Seq	Group Control	
RC-Seq	Company parameters in plane stranges indexet tranges indexet transmission indexe	
TN-Seq INSeq	Tangeon Tangeon Tangeon	5-carboxylcytosine (ScaC)
TC-Seq	AD dependent expansioner. AD	
Ig-seq Rep-Seq MAF	Colli) uncline registerize planetoris (Managerize) planetoris (Managerize	y gly y
SSB-Seq	Neg (Pilk Ingle strate) Consistent Constant Cons	Pyridostatin (PDS)
BLESS		

nversion			
lethyl Cytosine	scM&T-seq	Stafford Bask and the	
tethyl Cytosine	scMT-seq	Angle relations and an end service and an end servi	Methidiumpro- pyl-EDTA (MPE)
	scTrio-seq	Stopford III A Add care IBA Hydrolic sign (DA reprint print) ORA enderstation CRA enderstation Notice Noti	
, , , , , , , , , , , , , , , , , , ,	scBS-seq	Struke officialities many Call and the struke is struke and the struke is struke in the st	
thods on	scAba-seq	Dargendo Therminister 9 → → → → → → → → → → → → → → → → → → →	0
rice	scRC-Seq	Nord-Intertainer politice events Single-Contents Single-Contents Market and grite Single-Contents Market and Single-Contents Market and	DMA +
	scATAC-Seq (Cell index)	Technol And	C, C, Paul
	scATAC-Seq (Microfluidics)	Straffer of Here all and and all and all all all all all all all all all al	
Granine Mu	Drop-ChIP scChIP-seq		Biota Preparation of acylated
N N I Rotosine (Shmc)	SMDB	Engineering definition and a second and a se	RNA for biotin-streptan purification. DIBO, dibenzocyclooxt
ntine (SfC)		etics	
n too In osine (ScaC)	HELP-Seq	Buulit environment of promote DM without Methy M	
	PBAT	Not Source Adaptor Methylated DNA Boulface Post andame	
	BSPP	Buttle sequencion with patho p	
23	RRBS scRRBS	Reduced expresentation building buildin	RC
rtin (PDS)	BSAS	Rundts ongestigen Marthyland Zhak Bindhes Rody same of the provided in the provided	Unidine diphosphate glucose (UDP-Glu)
		T Mad -T	Anicose (nnu-ejn)

33

Methvl-Sea

==

.000

8-00

DNA-Protein Interactions

	DIAPPIO	Stell Interactions
ann ₂	ChIP-Seq ChIP-exo HT-ChIP Mint-ChIP	Main Main <th< td=""></th<>
t o	DNase-Seq DNaseI-Seq	Million La constance de la con
5	DNaseI-SIM	Image: Second State Logical State Logical State Logical State Image: Second State Image:
	MAINE-Seq MNase-Seq Nucleo-Seq	
	FiT-seq	The finance function of the fi
0	PAT-ChIP	Production de la construction de
	X-ChIP	Indexember of the rest of the
fick*	ORGANIC	Compare information of generative interview in
)	ATAC-Seq	Angle barrens and the Angle barrens and the Ba
tin Led tavidin Lixtyne	THS-seq	Transmers (Herman Herman Herma
	CATCH-IT	Considerational formation of the second seco
	MINCE-seq	Januare Autority
	FAIRE-seq Sono-Seq	tenencer and performance and p
	NOMe-Seq	cons linked charanter fanne Sant
	Histone methylation	
\$	ChIPmenta- tion	Terretarian Terretaria Terretarian Terretarian Terretarian Terretarian Terreta
e	Chia-PET	

We are here for you!

Illumina:

Lauren Robinson Sequencing and Informatics Specialist Irobinson@illumina.com

Robin Ball Territory Account Manager <u>rball@illumina.com</u>

Illumina Technical Support +1 800 809 4566 techsupport@illumina.com Indiana University Center for Genomics & Bioinformatics:

Matthew Hahn, PhD, Director <u>mwh@Indiana.edu</u>

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





Questions?

lrobinson@illumina.com