

# **The application of epigenomic profiling strategies to study cancer and other common diseases**

Peter C. Scacheri, PhD

Gertrude Donnelly Hess Professor of Oncology

Department of Genetics & Genome Sciences

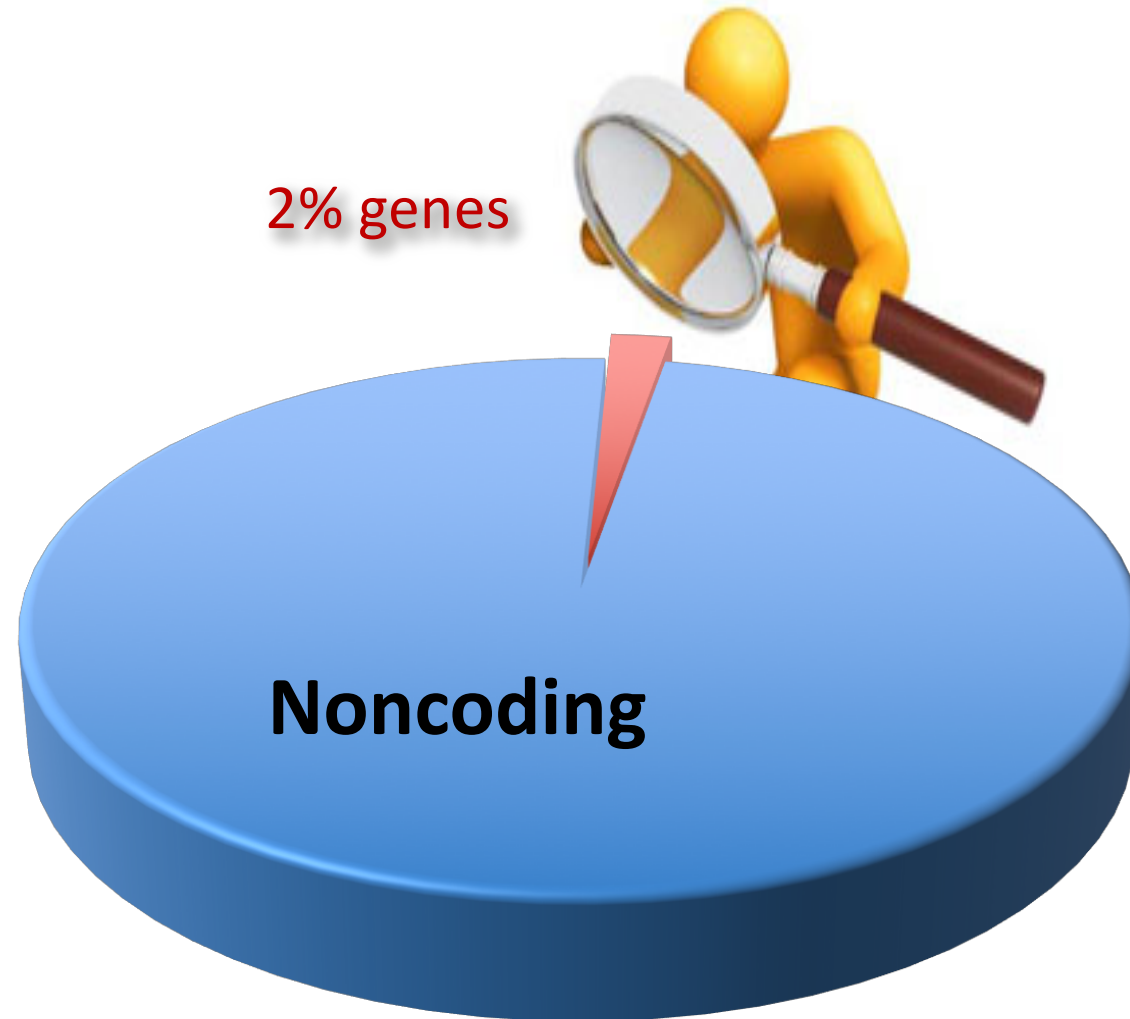
Case Western Reserve University School of Medicine

Case Comprehensive Cancer Center

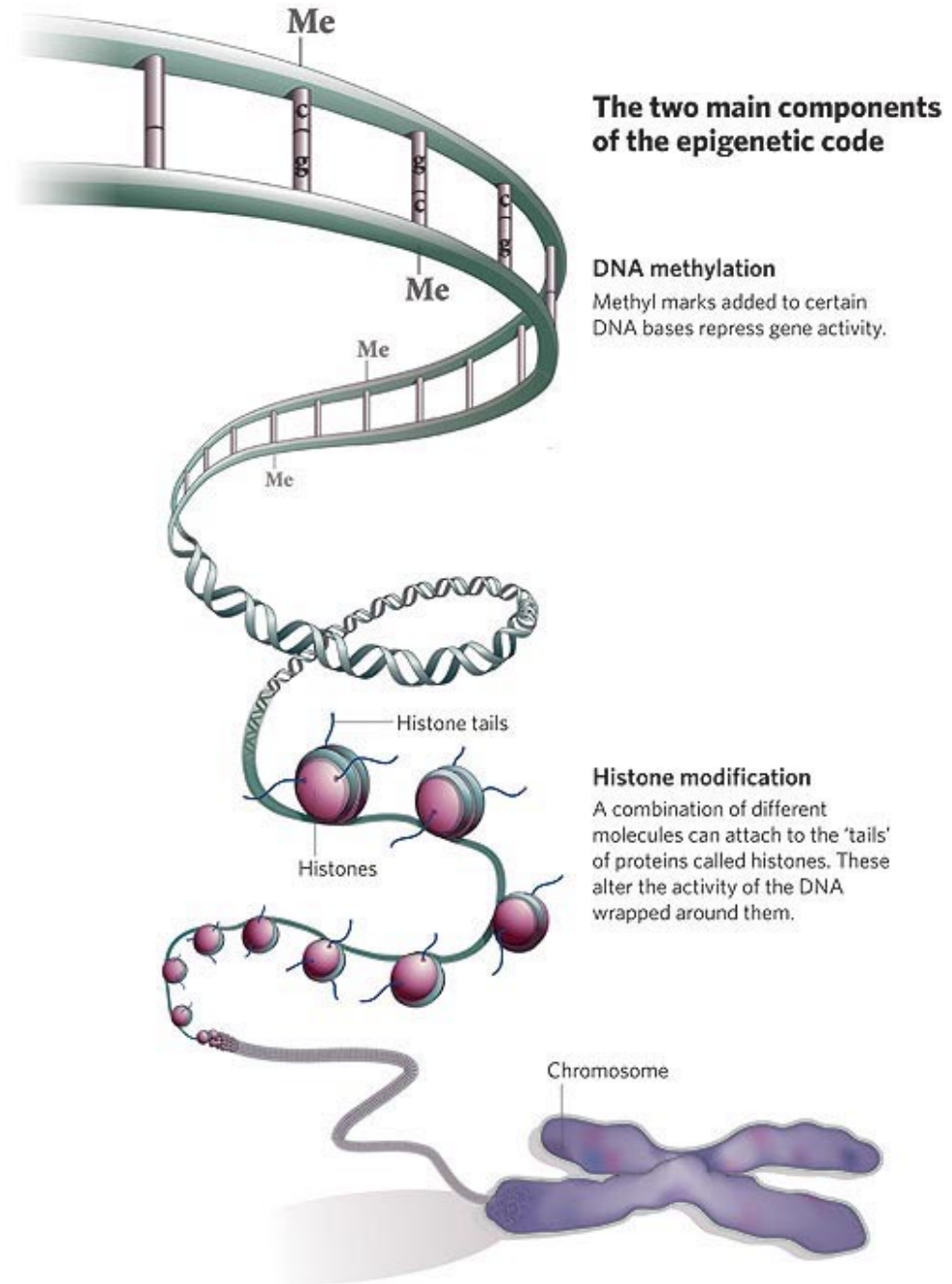
# Outline

- Introduction
  - Non-coding regions and the epigenome
  - Techniques for epigenomic profiling
  - Repositories and browsers for epigenomics datasets
  - Data analysis
- Applications of epigenomic profiling (Focus on enhancer elements and cancer)
  - Tumorigenesis
  - Clinical subtyping
  - Metastasis
  - Drug resistance
  - Identification of non-coding driver mutations
  - GWAS
  - Biomarkers
- Opportunities for Precision Oncology



Most of the human genome is non-coding DNA







# The Epigenome



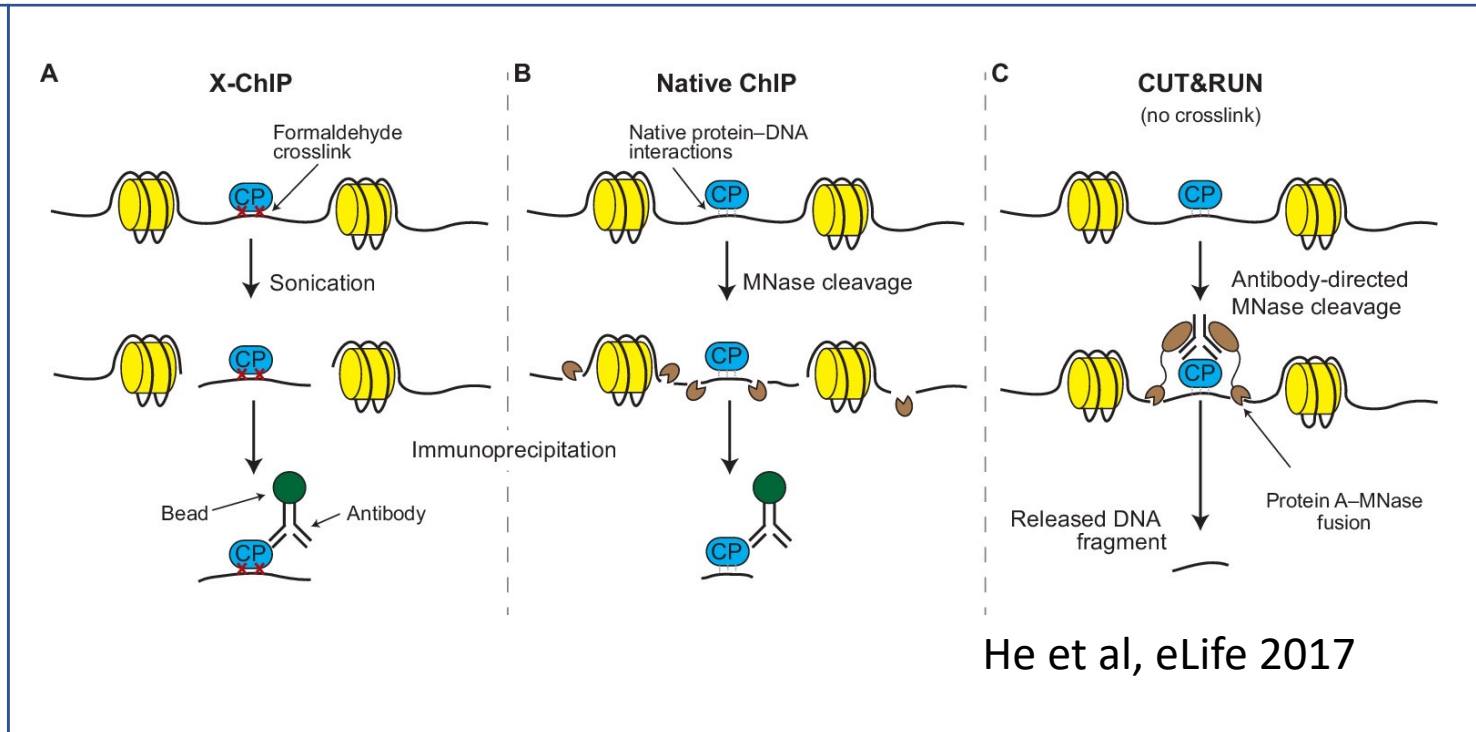
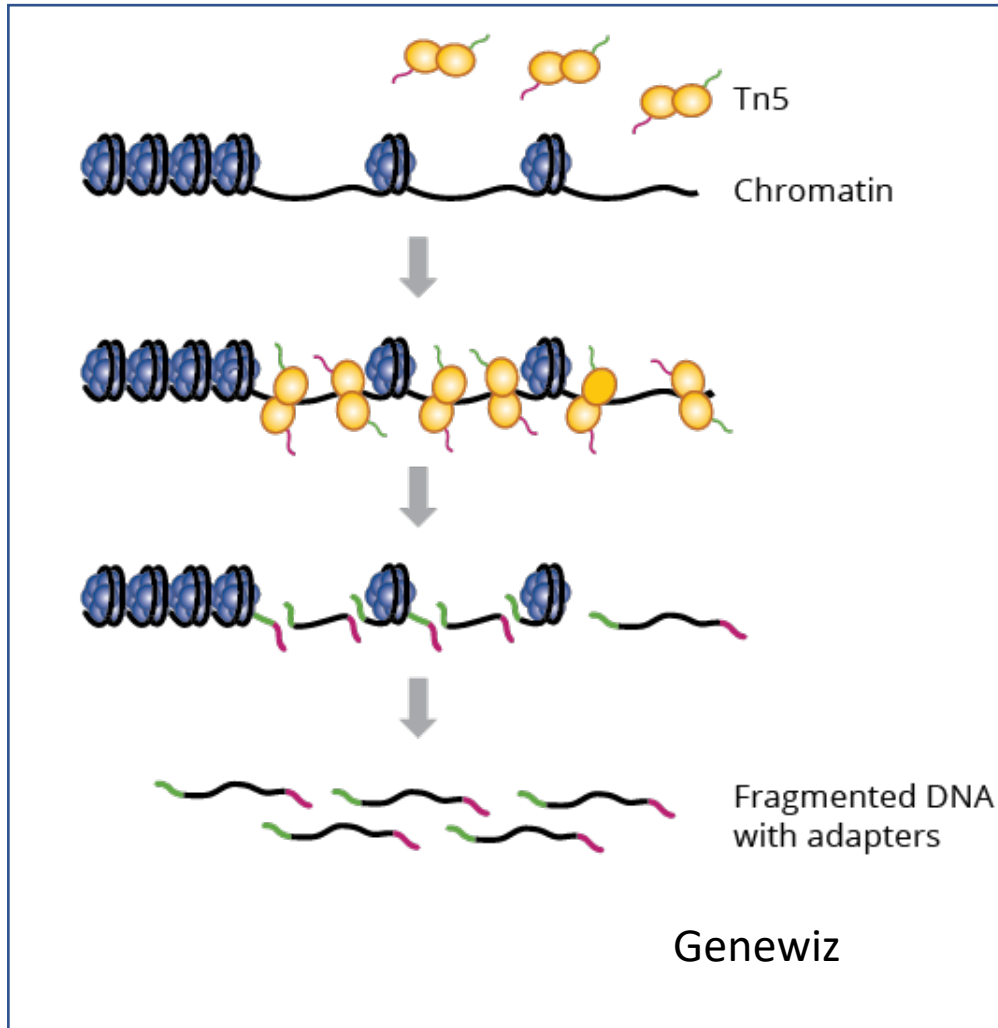
# Profiling Techniques for DNA methylation

Technique	Method	Advantages	Limitations
 Whole-Genome Bisulfite Sequencing (WGBS)	Bisulfite converted DNA is amplified and sequenced	Genome-wide, single nucleotide resolution	Costly and computationally intensive
Reduced-Representation Bisulfite Sequencing (RRBS)	Methylation-insensitive restriction enzymes digest DNA, enriching for CpG regions	Cheaper than WGBS with relatively high coverage	Enzymatic digestion covers most but not all CpG sites
Pyrosequencing	DNA is bisulfite converted, amplified, with the ratio of C/T nucleotides measured	Genome-wide or targeted, single nucleotide resolution. Allele-specific primers	Relatively expensive
Methylated DNA Immunoprecipitation (MeDip)	Methylated DNA is enriched by immunoprecipitation followed by sequencing or microarray analysis	Random fragmentation by sonication avoids restriction enzyme bias	Varying CpG density can confound methylation estimates
Methylation Sensitive Restriction Enzyme Sequencing (MSRE/MRE-Seq or Methyl-seq)	Unmethylated DNA is restriction enzyme digested while methylated DNA is amplified	No bisulfite conversion bias	DNA may be partially digested, limited coverage
Combined Bisulfite Restriction Analysis (COBRA)	Bisulfite converted DNA is amplified and restriction enzyme digested	Simple, fast, inexpensive, works on FFPE-treated DNA	DNA may be partially digested, limited coverage
Methylation Specific PCR	Bisulfite converted DNA is amplified with methylation specific primers	Simple and inexpensive	Purely qualitative
High Resolution Melt Analysis (HRM)	Bisulfite converted DNA is amplified by q-PCR	Most sensitive method for determining methylation at a specific region	Single base resolution not possible
 Illumina MethylationEPIC BeadChip Microarray (previously 450k, 27k)	Bisulfite (or oxidized + bisulfite) converted DNA is interrogated on a microarray chip	Relatively simple and inexpensive. Extremely popular	Data has limited coverage and requires pre-processing
Global DNA Methylation	Methods include LINE1, Alu, LUMA, HPLC-UV	Relatively inexpensive	Does not identify differentially methylated regions
Tet-assisted Bisulfite Sequencing (TAB-seq)	5hmC is protected then oxidized to 5caC then uracil by TET	Differentiation between 5mC and 5hmC at single base resolution	Sensitivity and specificity depends on sequencing depth
Oxidative Bisulfite Sequencing (OxBis)	DNA is oxidized then bisulfite converted to 5fC and subsequently uracil	Quantitative genome-wide coverage	Bias to regions of low 5mC. Must be performed in parallel with bisulfite techniques
APOBEC-coupled epigenetic sequencing (ACE-seq)	Non-destructive DNA deaminase enzymes discriminate between 5hmC and 5mC	Genome-wide, single nucleotide resolution. Very low DNA input required	Not yet extensively tested
Hydroxymethylated DNA Immunoprecipitation (hMeDIP)	Immunoprecipitation and sequencing of hydroxymethylated DNA	Simple and inexpensive	Only semi quantitative and bias to regions of low 5hmC

# Profiling techniques for histones and chromatin accessibility

Technique	Method	Advantages	Limitations
 Chromatin Immunoprecipitation (ChIP)	Couples highly specific antibodies for DNA-binding proteins with sequencing, microarrays or PCR	Detect DNA associated proteins and histone modifications	Requires intact cells and chromatin
Digital DNase	Enzymes digest nuclease-accessible regions, indicating open chromatin	Maps both nucleosomes and non-histone proteins	High sequencing depth required. Potential actin contamination.
NOMe-seq	Single-molecule, high-resolution nucleosome positioning assay	Maps both DNA methylation and nucleosomes at high resolution	Relies on presence of CpG residues
 Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq)	Measures chromatin accessibility based on Tn5 transposase activity. Maps nucleosomes and non-histone proteins	Simple, fast, low input of cells with single nucleotide resolution	Distance between binding sites may bias results
 Chromosome Conformation Capture (3C, 4C, 5C, Hi-C) & Hi-ChIP	Assess spatial organization of chromatin in a cell	Various modified versions	Often lack genome-wide, single nucleotide resolution
 CUT&RUN (Cleavage Under Targets and Release Using Nuclease )	antibody-targeted controlled cleavage by micrococcal nuclease releases specific protein-DNA complexes into the supernatant for paired-end DNA sequencing.	In situ, simple, fast, low input, Less sequencing depth required	Need good antibodies

# ATAC-seq and ChIP-seq

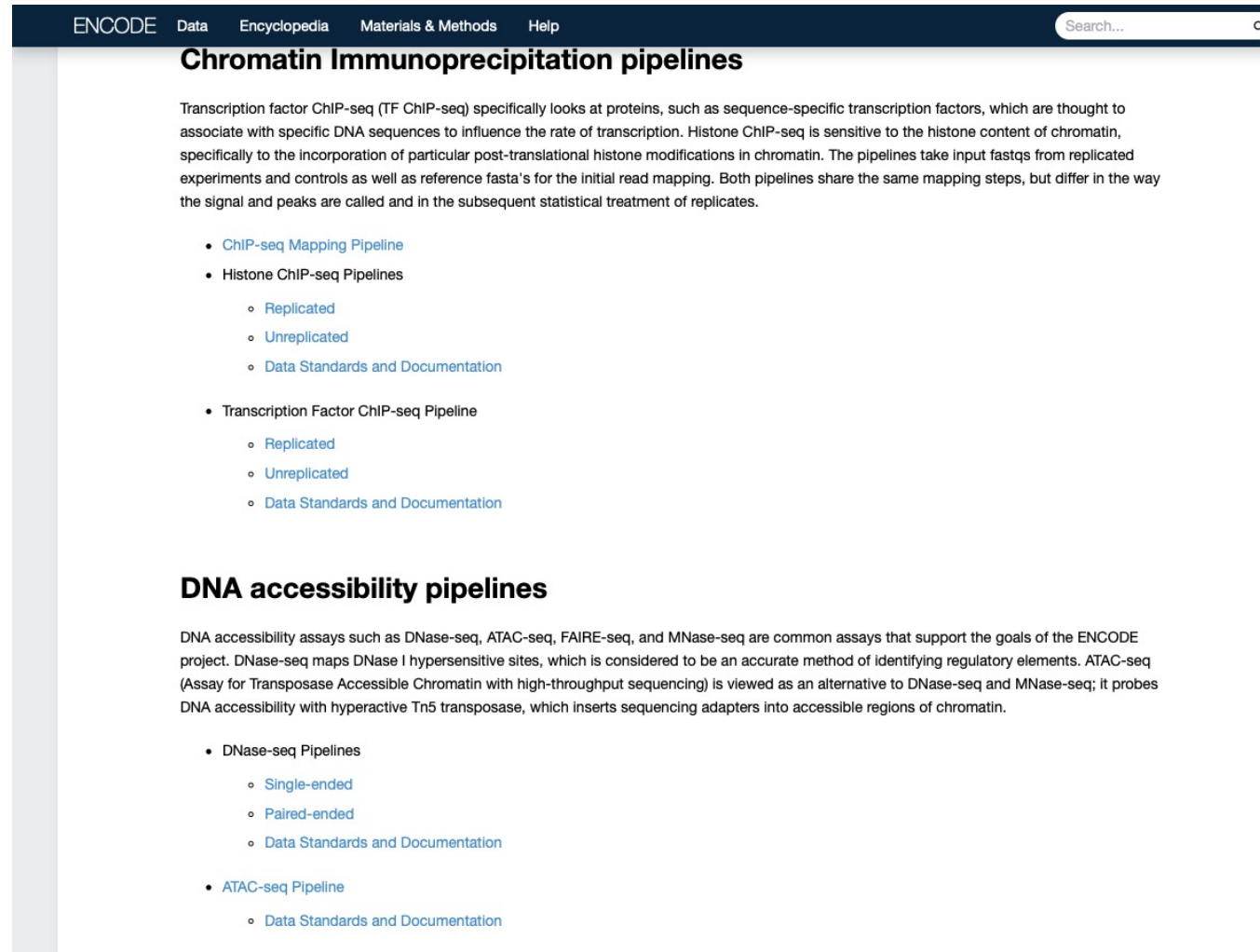


# Single cell methods for profiling chromatin

- scATAC-seq
- sc Multiome (ATAC + Gene Expression)
- single nucleus methylcytosine sequencing (snmC-seq)
- scChIP-seq
- Paired-tag (scChIP-seq + Gene expression)



# Lots of methods for processing datasets



The screenshot shows the ENCODE website interface. At the top, there is a navigation bar with links for 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located on the right side of the navigation bar. The main content area is titled 'Chromatin Immunoprecipitation pipelines'. Below the title, there is a paragraph of text explaining the process. This is followed by a bulleted list of links for 'ChIP-seq Mapping Pipeline', 'Histone ChIP-seq Pipelines' (with sub-links for 'Replicated', 'Unreplicated', and 'Data Standards and Documentation'), and 'Transcription Factor ChIP-seq Pipeline' (with sub-links for 'Replicated', 'Unreplicated', and 'Data Standards and Documentation'). Below this section, there is another section titled 'DNA accessibility pipelines'. This section also contains a paragraph of text and a bulleted list of links for 'DNase-seq Pipelines' (with sub-links for 'Single-ended', 'Paired-ended', and 'Data Standards and Documentation') and 'ATAC-seq Pipeline' (with a sub-link for 'Data Standards and Documentation').

ENCODE Data Encyclopedia Materials & Methods Help Search...

## Chromatin Immunoprecipitation pipelines

Transcription factor ChIP-seq (TF ChIP-seq) specifically looks at proteins, such as sequence-specific transcription factors, which are thought to associate with specific DNA sequences to influence the rate of transcription. Histone ChIP-seq is sensitive to the histone content of chromatin, specifically to the incorporation of particular post-translational histone modifications in chromatin. The pipelines take input fastqs from replicated experiments and controls as well as reference fasta's for the initial read mapping. Both pipelines share the same mapping steps, but differ in the way the signal and peaks are called and in the subsequent statistical treatment of replicates.

- [ChIP-seq Mapping Pipeline](#)
- [Histone ChIP-seq Pipelines](#)
  - [Replicated](#)
  - [Unreplicated](#)
  - [Data Standards and Documentation](#)
- [Transcription Factor ChIP-seq Pipeline](#)
  - [Replicated](#)
  - [Unreplicated](#)
  - [Data Standards and Documentation](#)

## DNA accessibility pipelines

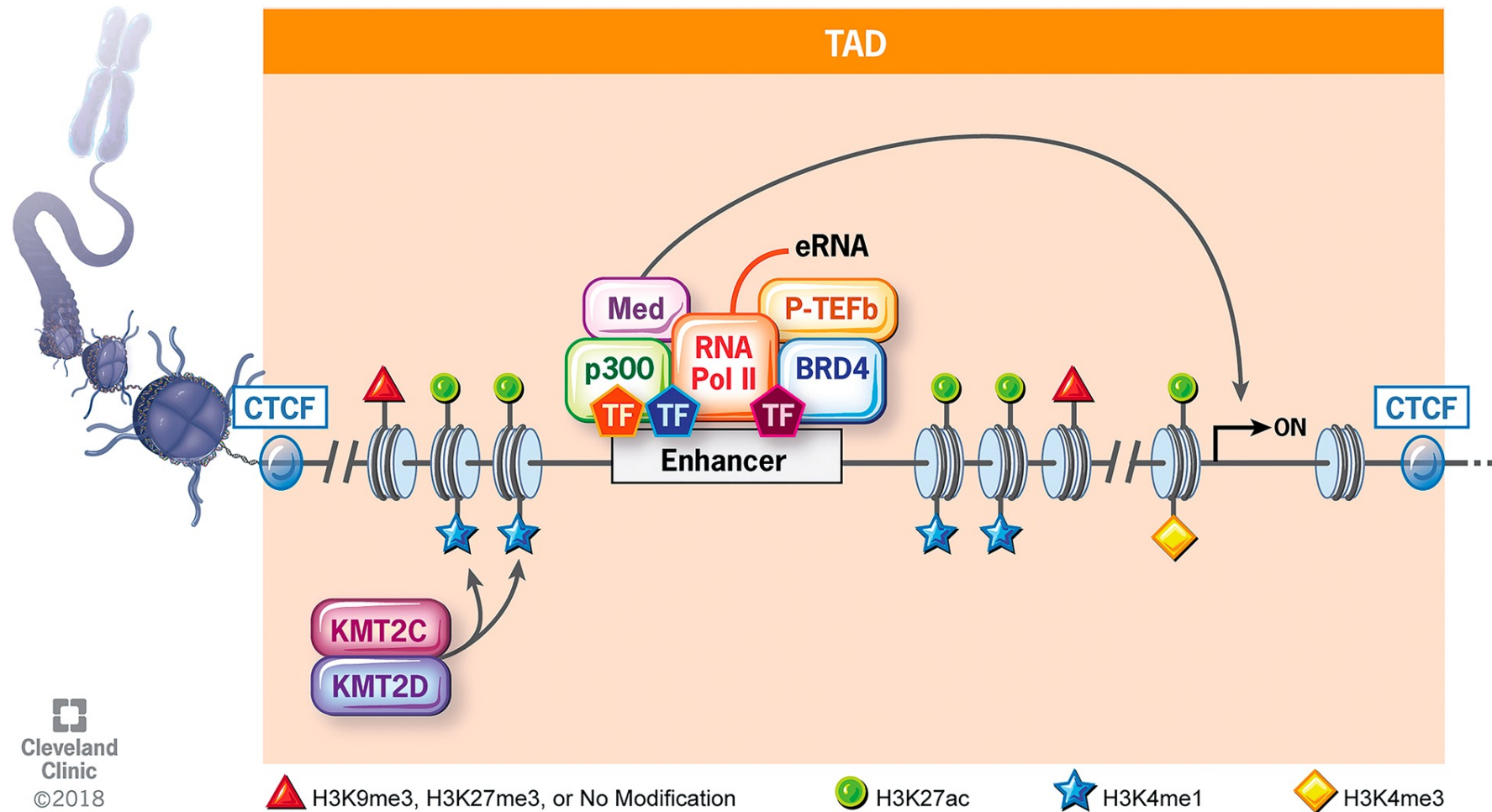
DNA accessibility assays such as DNase-seq, ATAC-seq, FAIRE-seq, and MNase-seq are common assays that support the goals of the ENCODE project. DNase-seq maps DNase I hypersensitive sites, which is considered to be an accurate method of identifying regulatory elements. ATAC-seq (Assay for Transposase Accessible Chromatin with high-throughput sequencing) is viewed as an alternative to DNase-seq and MNase-seq; it probes DNA accessibility with hyperactive Tn5 transposase, which inserts sequencing adapters into accessible regions of chromatin.

- [DNase-seq Pipelines](#)
  - [Single-ended](#)
  - [Paired-ended](#)
  - [Data Standards and Documentation](#)
- [ATAC-seq Pipeline](#)
  - [Data Standards and Documentation](#)

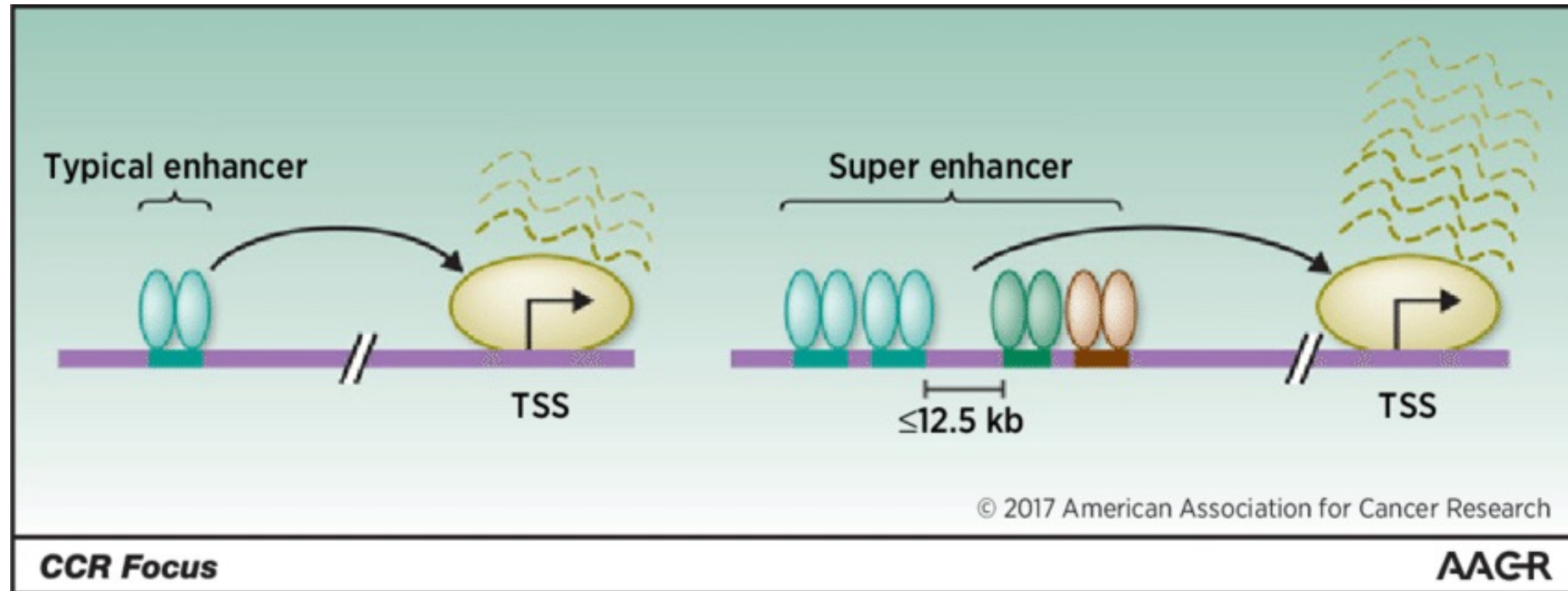
# Epigenetic data repositories and browsers

Consortia and resources	Data availability	URLs
The International Human Epigenome Consortium (IHEC)	Reference epigenomes generated by NIH Roadmap, ENCODE, CEEHRC, BLUEPRINT, DEEP, AMED/CREST, and KEP	IHEC Data Portal <a href="http://epigenomesportal.ca/ihec">http://epigenomesportal.ca/ihec</a>
NIH Roadmap Epigenomics	Maps of histone modifications, chromatin accessibility, DNA methylation and mRNA Expression in stem cells and primary <i>ex vivo</i> human tissues	VizHub <a href="http://vizhub.wustl.edu">http://vizhub.wustl.edu</a>
Canadian Epigenetics, Environment and Health Research Consortium (CEEHRC) Network	Reference epigenomes including histone modifications, DNA methylation, mRNA and miRNA of human cancer and normal cells	CEEHRC Data <a href="http://www.epigenomes.ca/site-data">http://www.epigenomes.ca/site-data</a> Software Tools <a href="http://www.epigenomes.ca/tools-and-software">http://www.epigenomes.ca/tools-and-software</a>
BLUEPRINT Epigenome	Reference epigenomes of human normal and malignant hematopoietic cells	BLUEPRINT Portal <a href="http://blueprint-data.bsc.es">http://blueprint-data.bsc.es</a>
The German epigenome programme (DEEP)	Reference epigenomes of human cells and tissues in normal and complex disease states	DEEP Data Portal <a href="http://deep.dkfz.de">http://deep.dkfz.de</a>
IHEC Team Japan (AMED-CREST)	Reference epigenomes of human gastrointestinal epithelial cells, vascular endothelial cells and cells of reproductive organs	IHEC Data Portal <a href="http://epigenomesportal.ca/ihec">http://epigenomesportal.ca/ihec</a>
Korea Epigenome Project (KEP)	Reference epigenome map for common complex diseases	IHEC Data Portal <a href="http://epigenomesportal.ca/ihec">http://epigenomesportal.ca/ihec</a>
DeepBlue	Epigenomic data server for storing and working with genomic and epigenomic data. Collection of over 30,000 experiment files from the main epigenome mapping projects available. Uploading own data allowed	DeepBlue server <a href="http://deepblue.mpi-inf.mpg.de">http://deepblue.mpi-inf.mpg.de</a>
Allelic Epigenome Project	Allelic DNA methylome, histone modifications, and transcriptome in human cells and tissues	Genboree <a href="http://genboree.org/genboreeKB/projects/allelic-epigenome">http://genboree.org/genboreeKB/projects/allelic-epigenome</a>
GTEx	Genotype and expression profiles in different tissues enabling eQTL studies	GTEx Portal <a href="http://www.gtportal.org">http://www.gtportal.org</a>
BRAINEAC	Brain eQTL Almanac provides genotype and expression profile across 10 brain regions	BRAINEAC <a href="http://braineac.org">http://braineac.org</a>
mQTLdb	Methylation and genotype data on mother-child pairs providing access to meQTL mapping across five different stages of life	mQTL Database <a href="http://www.mqtl.db.org">http://www.mqtl.db.org</a>
Fetal brain meQTLs	Epigenome-wide significant meQTLs observed in fetal brain	Fetal Brain meQTL <a href="http://epigenetics.essex.ac.uk/mQTL">http://epigenetics.essex.ac.uk/mQTL</a>
Pancan-meQTL	Database of <i>cis</i> - and <i>trans</i> - meQTLs across 23 cancer types from The Cancer Genome Atlas	Pancan-meQTL <a href="http://bioinfo.life.hust.edu.cn/Pancan-meQTL">http://bioinfo.life.hust.edu.cn/Pancan-meQTL</a>
Epigenome Browser	UCSC genome browser with tracks from ENCODE project	UCSC Epigenome Browser <a href="http://www.epigenomebrowser.org">http://www.epigenomebrowser.org</a>
WashU Epigenome Browser	Web browser with tracks from ENCODE and Roadmap Epigenomics projects	WashU Epigenome Browser <a href="http://epigenomegateway.wustl.edu">http://epigenomegateway.wustl.edu</a>
Ensembl	ENCODE data used in the regulatory build	Ensembl ENCODE <a href="https://www.ensembl.org">https://www.ensembl.org</a>
RMBase	Database listing over 100 RNA modifications	<a href="http://rna.sysu.edu.cn/rmbase/">http://rna.sysu.edu.cn/rmbase/</a>

# The Epigenome and Gene Enhancer Elements

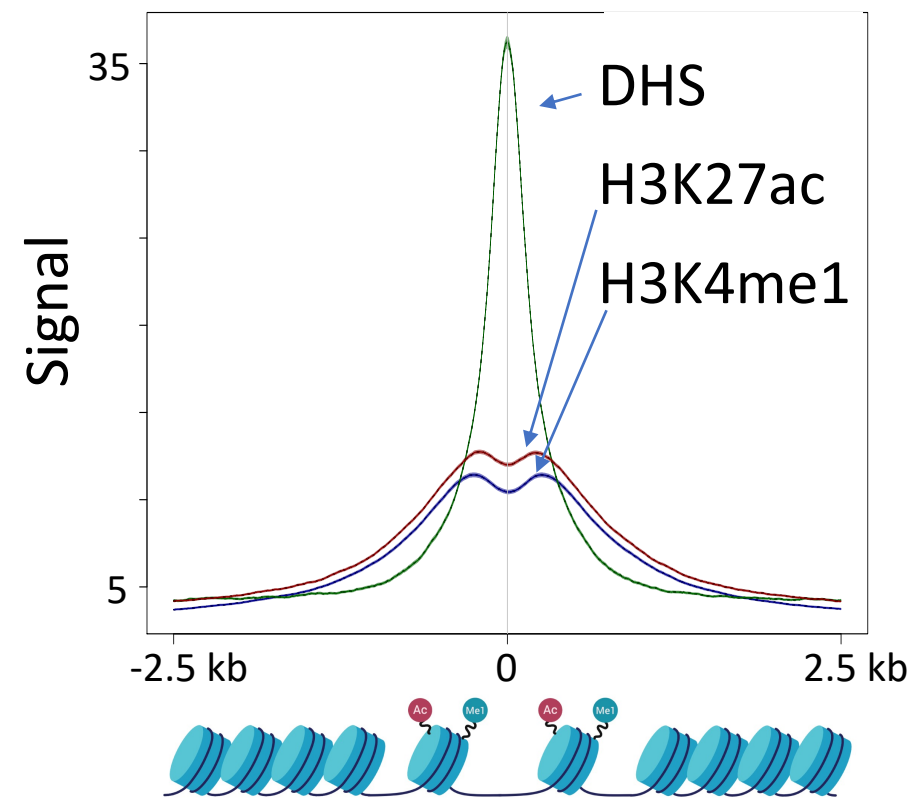
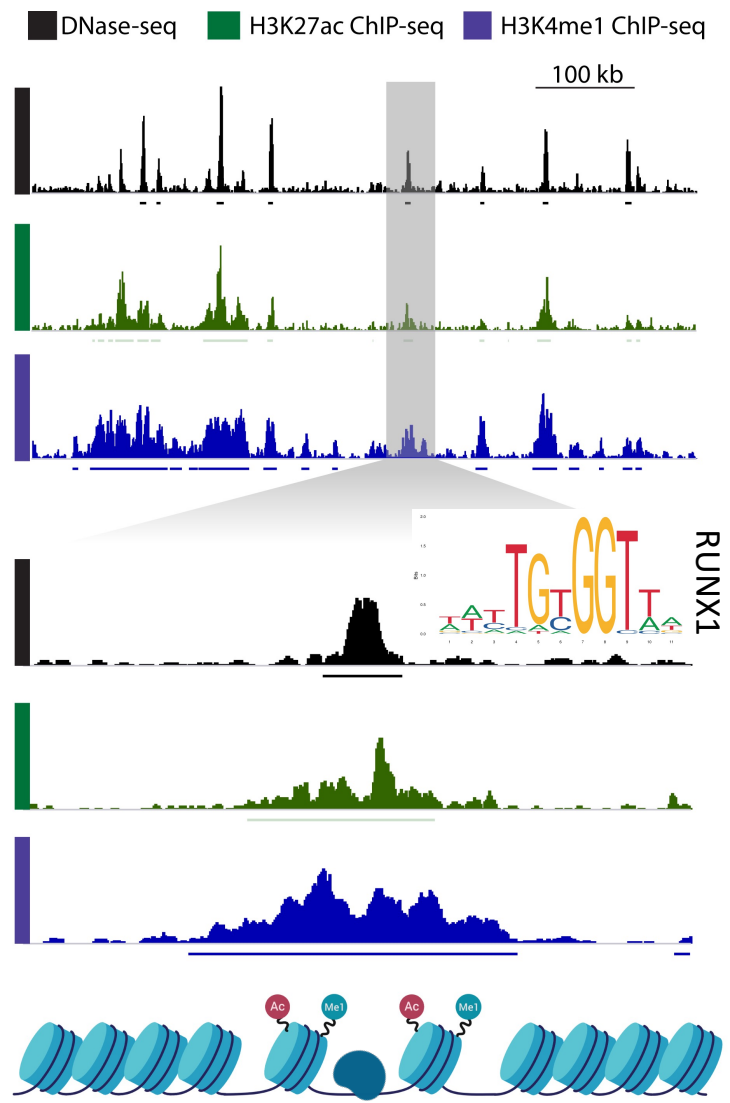


# Superenhancers



“Super enhancers are typically an order of magnitude larger than typical enhancers in size, have higher transcription factor density, and greater ability for transcriptional activation.”

# Genome-wide identification of active enhancer elements based on signature chromatin features

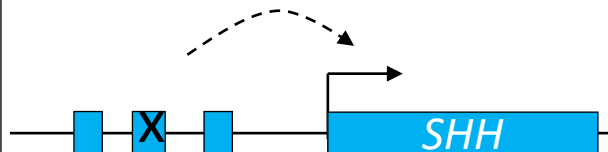


# ”Enhanceropathies”

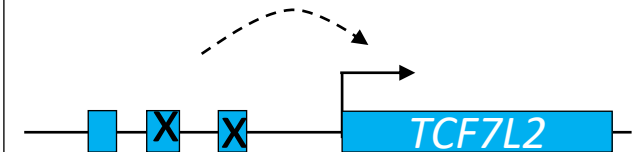
## Kabuki syndrome



## Preaxial Polydactyly



## Type II diabetes

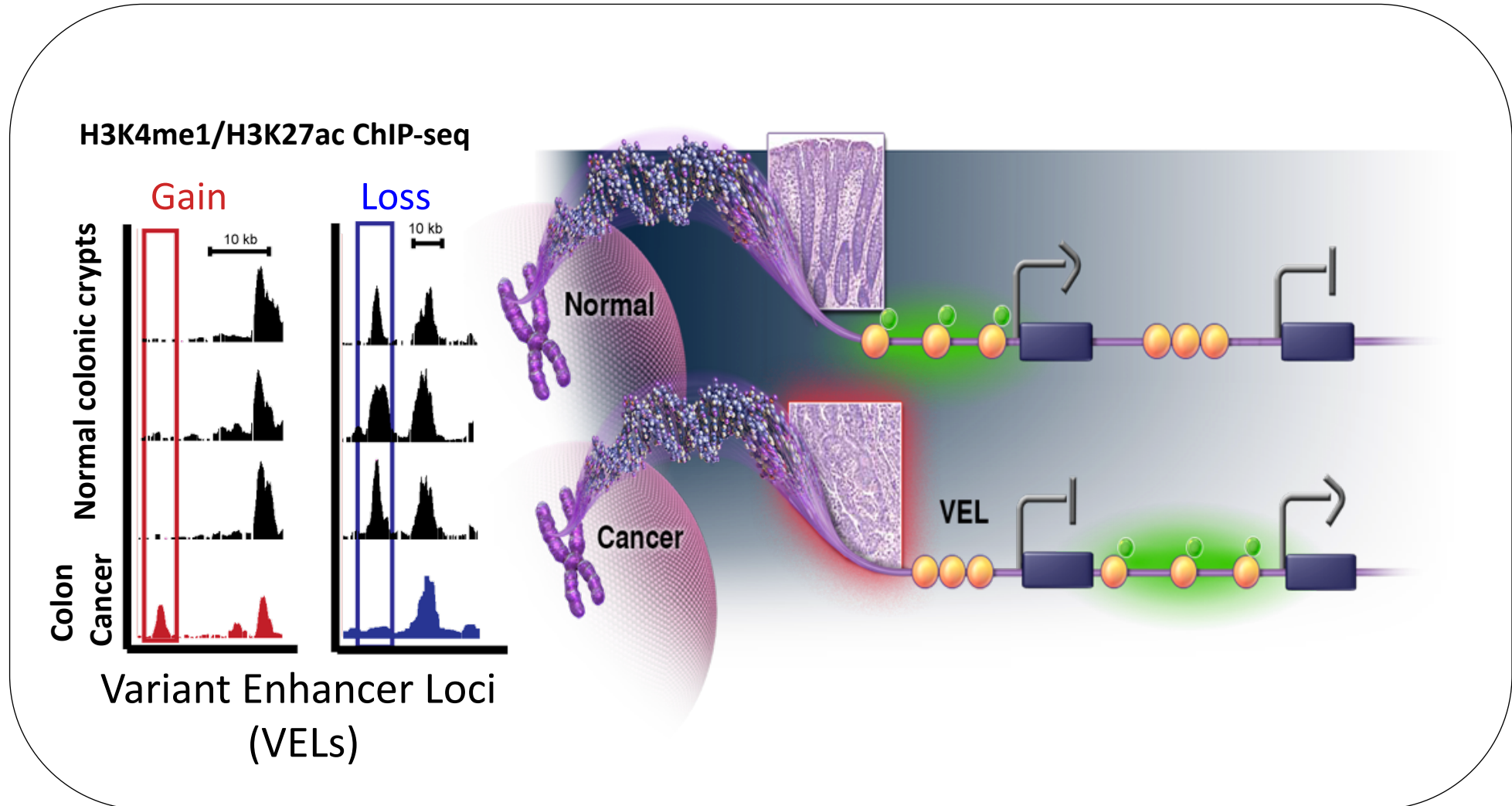


Karnuta and Scacheri, *HMG* 2018  
Scacheri and Scacheri, *Curr Opin Pediatr* 2015  
Corradin and Scacheri, *Genome Med* 2014

# Studying the epigenome can provide insights into various aspects of cancer

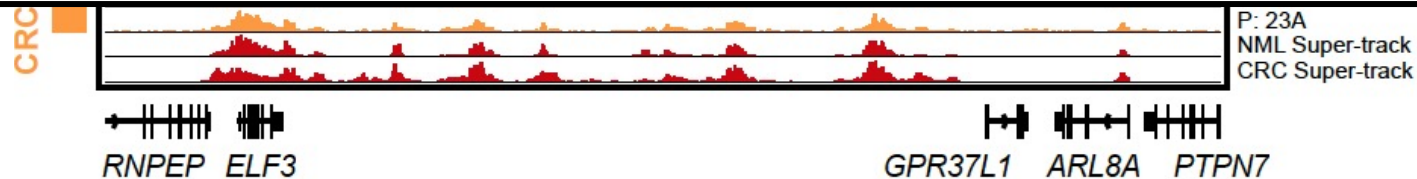
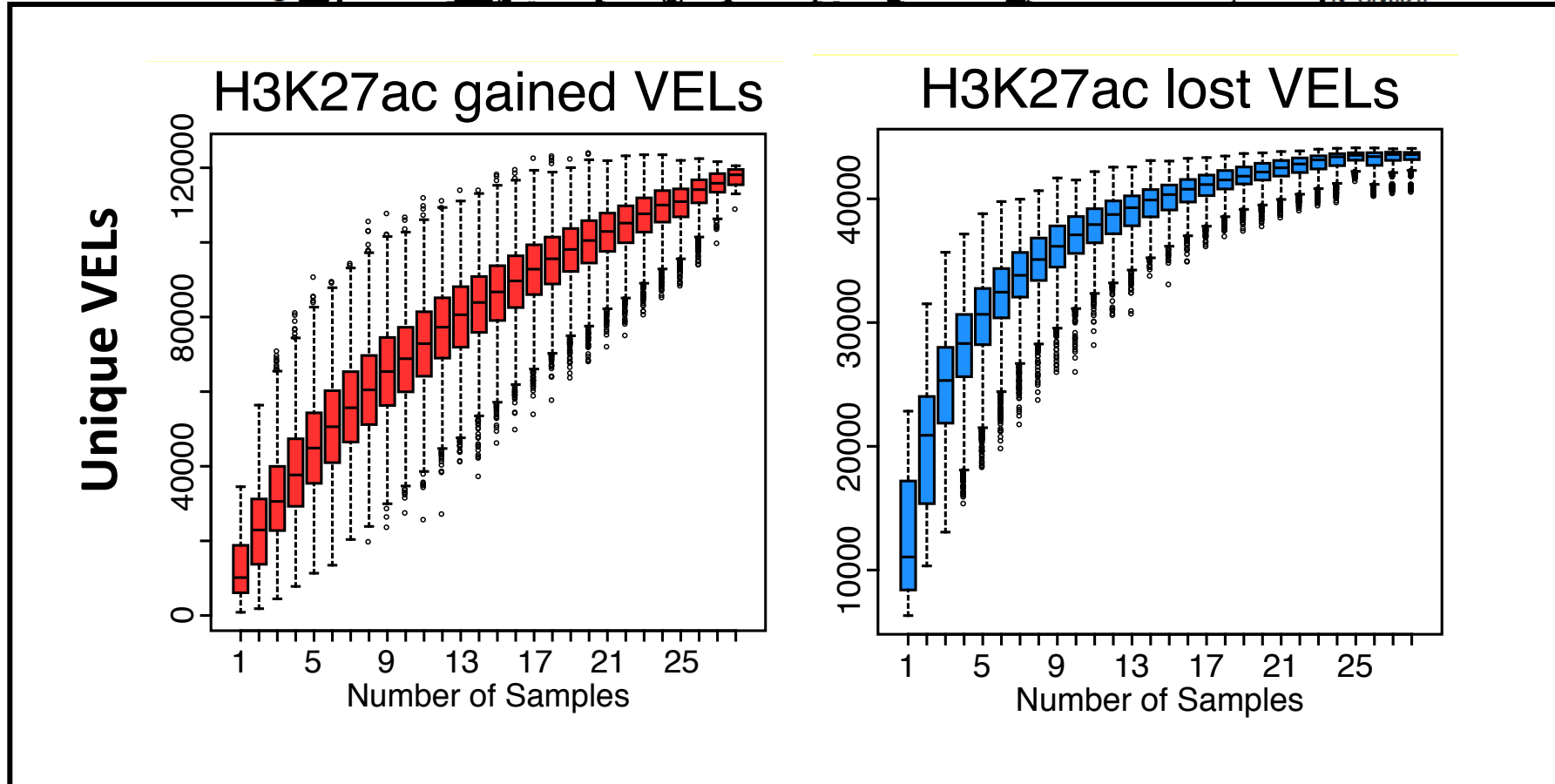
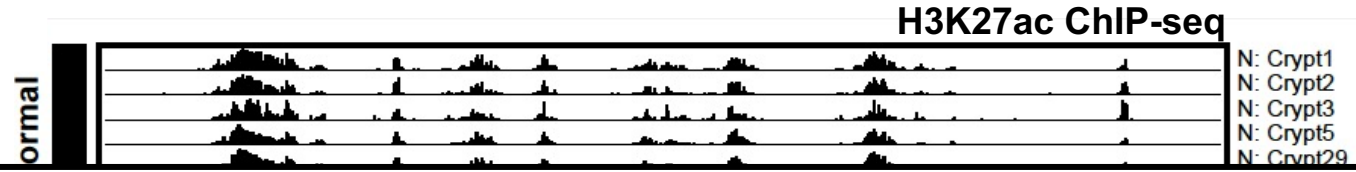
- Tumorigenesis
- Clinical subtyping
- Metastasis
- Drug resistance
- Identification of non-coding driver mutations
- GWAS
- Biomarkers

# Enhancers alterations in cancer

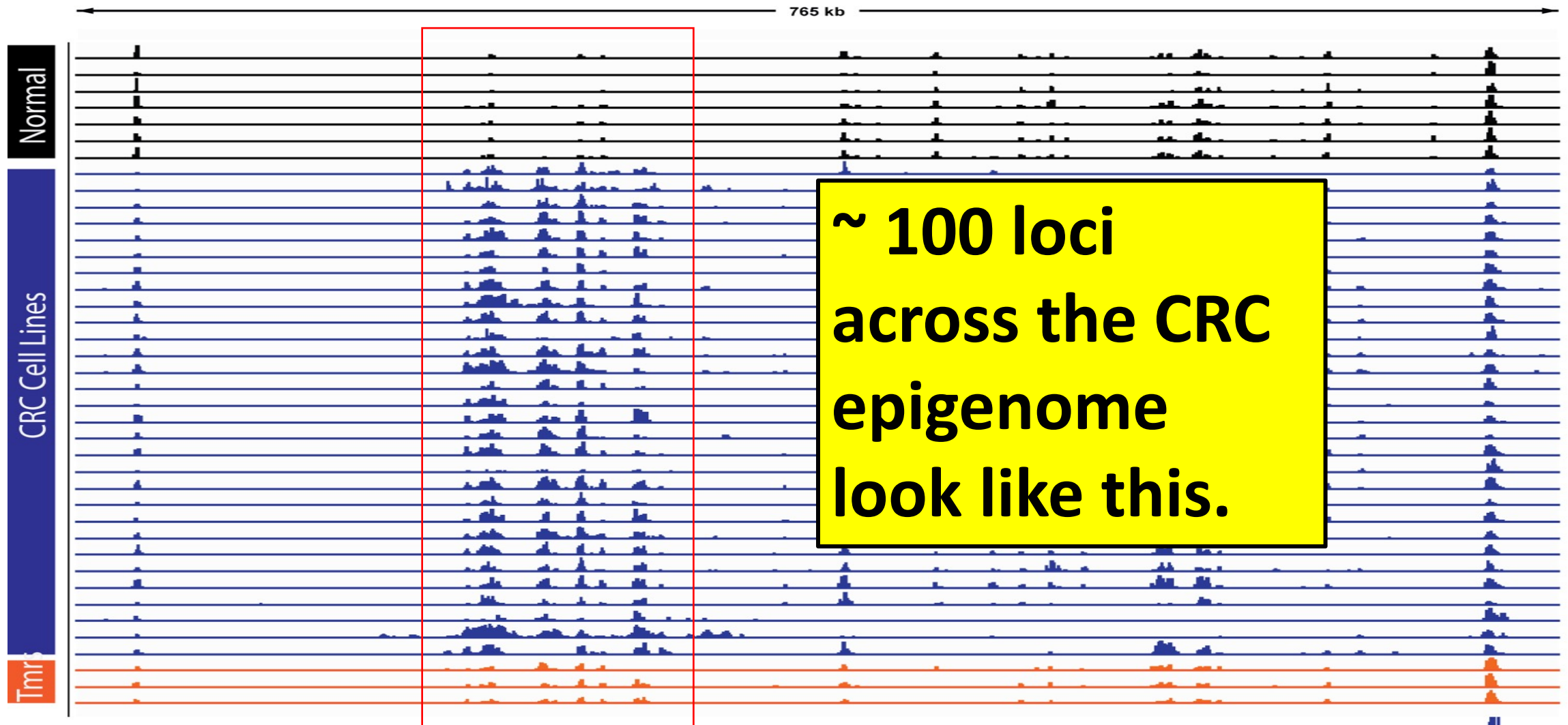




# Expansion of enhancer profiling studies



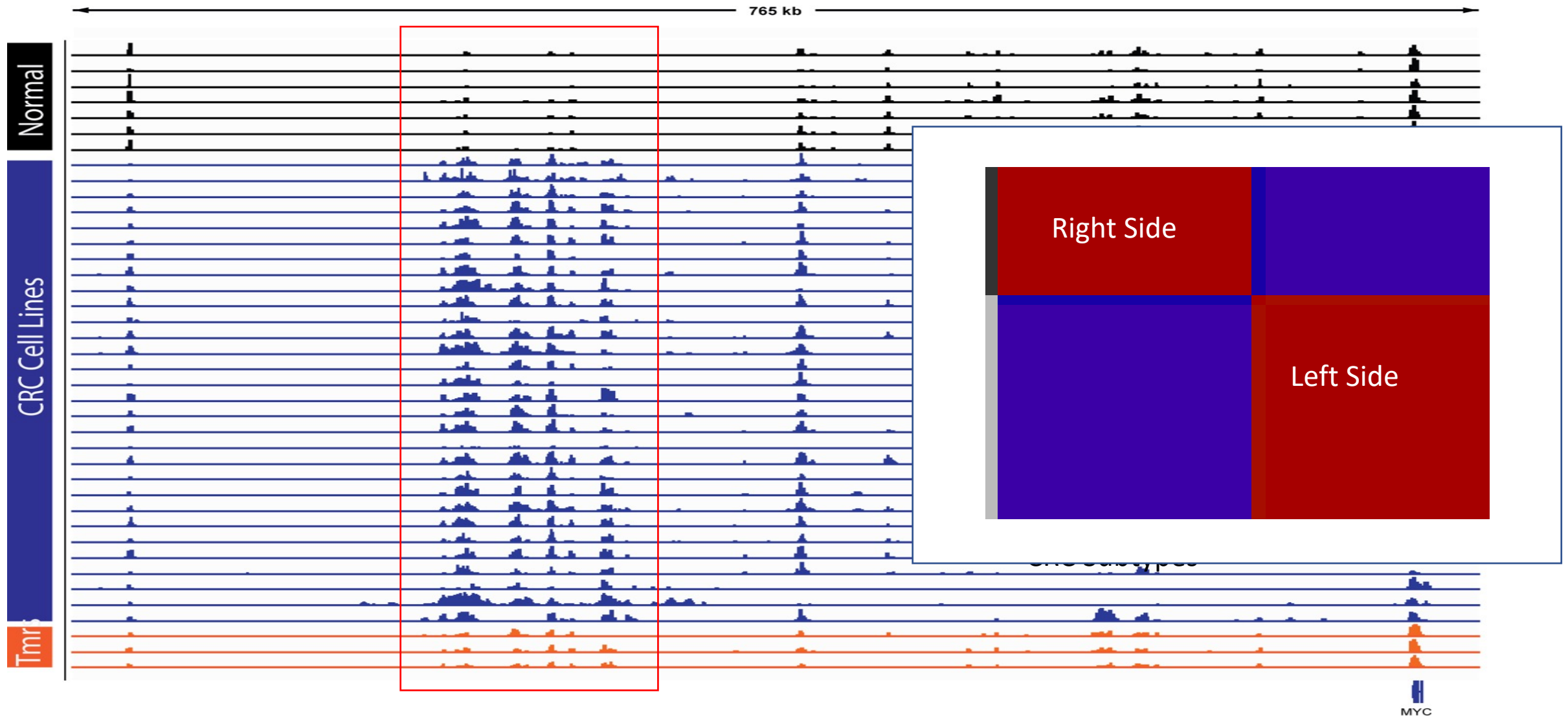
# Recurrent enhancer alterations in CRC



# Studying the epigenome can provide insights into various aspects of cancer

- Tumorigenesis
- **Clinical subtyping**
- Metastasis
- Drug resistance
- Identification of non-coding driver mutations
- GWAS
- Biomarkers

# Recurrent enhancer alterations in CRC



# Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling

Nature 2018

Stephen C. Mack<sup>1,2,3,4\*</sup>, Kristian W. Pajtler<sup>5,6,7\*</sup>, Lukas Chavez<sup>5,6,8\*</sup>, Konstantin Okonechnikov<sup>5,6</sup>, Kelsey C. Bertrand<sup>1,2,9</sup>, Xiuxing Wang<sup>3,4,10</sup>, Serap Erkek<sup>5,6,11</sup>, Alexander Federation<sup>12</sup>, Anne Song<sup>3,4</sup>, Christine Lee<sup>3,4</sup>, Xin Wang<sup>13</sup>, Laura McDonald<sup>13</sup>, James J. Morrow<sup>14</sup>, Alina Saiakhova<sup>14</sup>, Patrick Sin-Chan<sup>13</sup>, Qiulian Wu<sup>3,4,10</sup>, Kulandaimanuel Antony Michaelraj<sup>13</sup>, Tyler E. Miller<sup>3,4,15</sup>, Christopher G. Hubert<sup>3,4</sup>, Marina Rvzhova<sup>16</sup>, Livia Garzia<sup>13</sup>, Laura Donovan<sup>13</sup>, Stephen Dombrowski<sup>3,4,17</sup>.

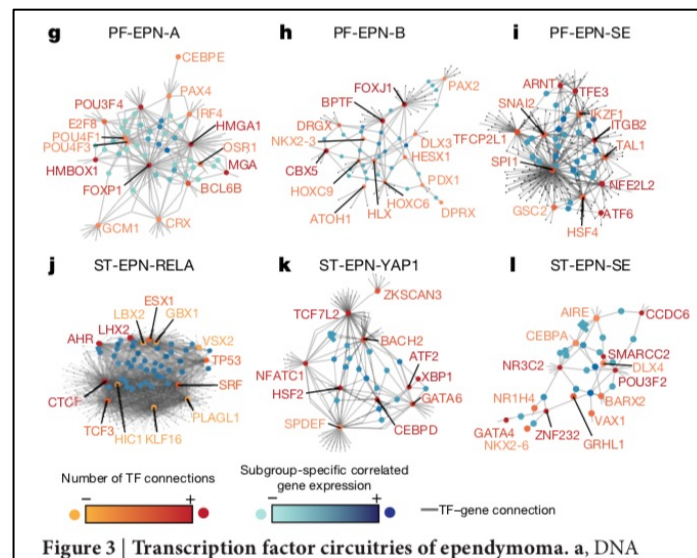
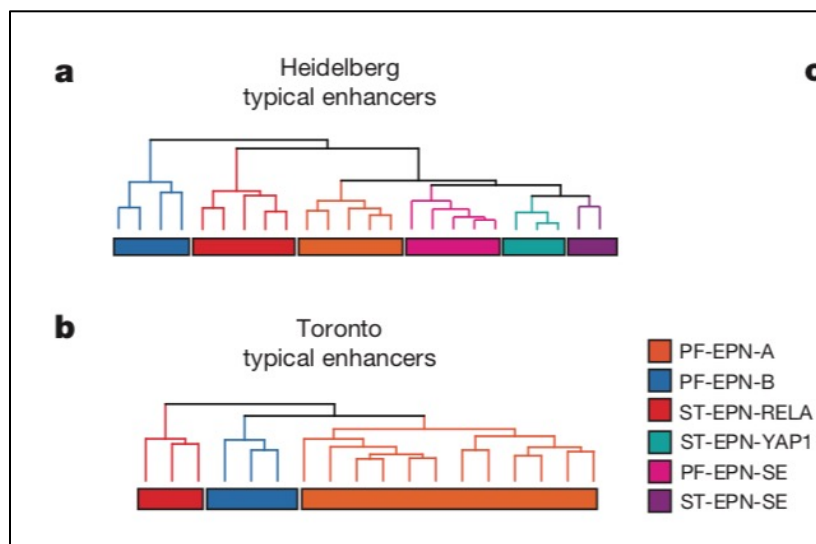
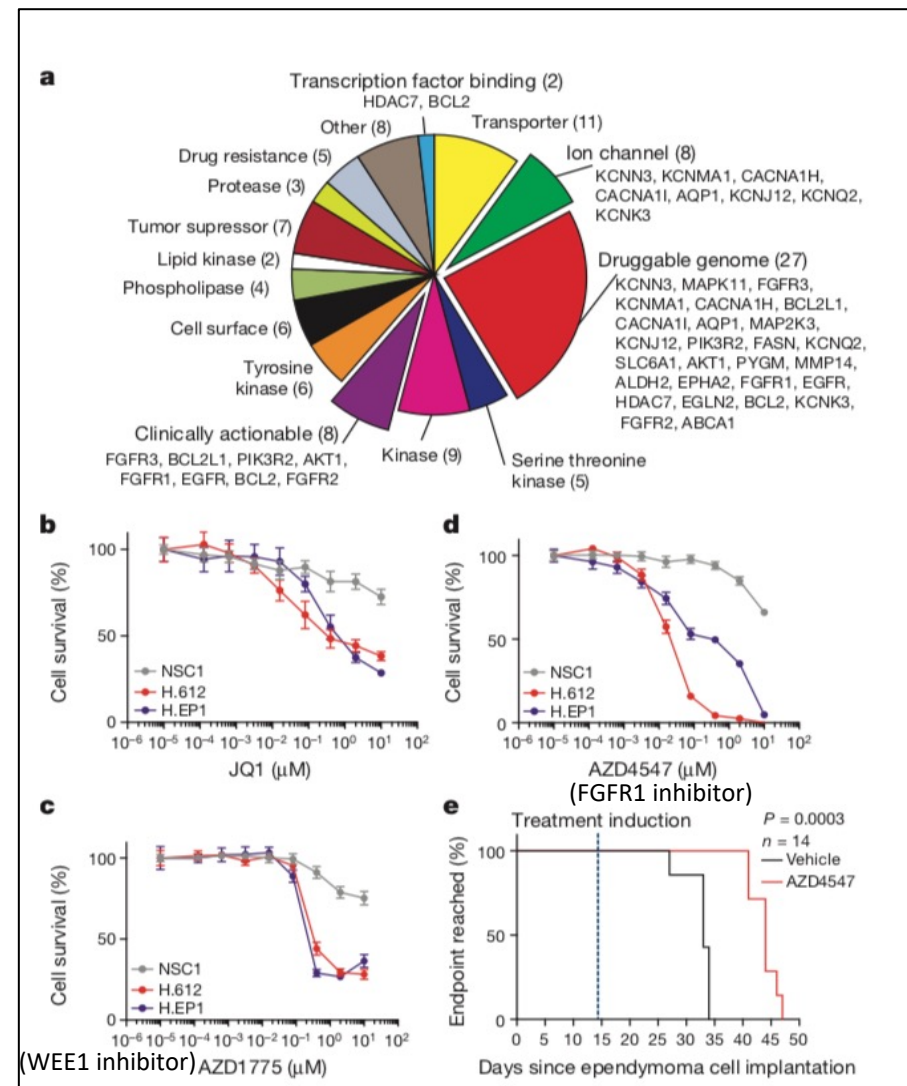
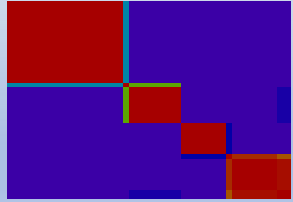


Figure 3 | Transcription factor circuitries of ependymoma. a, DNA

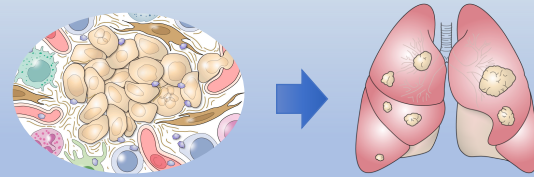


# Enhancer changes mediating acquisition of traits

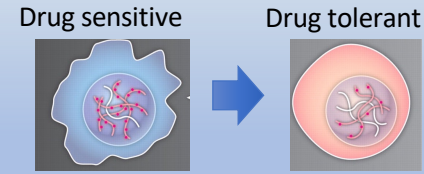
Subtypes



Metastasis



Drug resistance



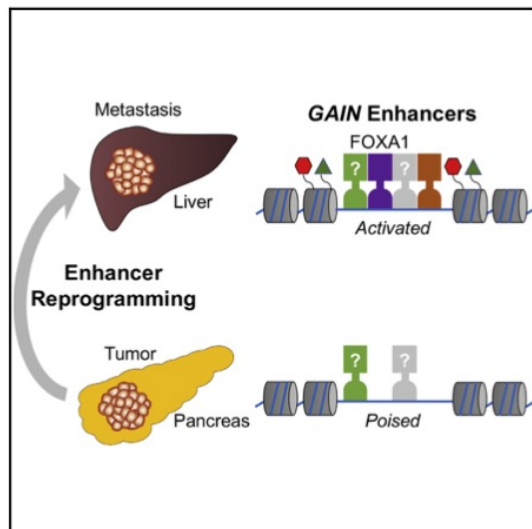
Predisposition



## Enhancer Reprogramming Promotes Pancreatic Cancer Metastasis

*Cell* 2017

Jae-Seok Roe,<sup>1,9</sup> Chang-Il Hwang,<sup>1,2,9</sup> Tim D.D. Somerville,<sup>1</sup> Joseph P. Milazzo,<sup>1</sup> Eun Jung Lee,<sup>1,2</sup> Brandon Da Silva,<sup>1,2</sup> Laura Maiorino,<sup>1</sup> Hervé Tiriac,<sup>1,2</sup> C. Megan Young,<sup>1,2</sup> Koji Miyabayashi,<sup>1,2</sup> Dea Filippini,<sup>1,2</sup> Brianna Creighton,<sup>1,2</sup> Richard A. Burkhardt,<sup>3</sup> Jonathan M. Buscaglia,<sup>4</sup> Edward J. Kim,<sup>5</sup> Jean L. Grem,<sup>6</sup> Audrey J. Lazenby,<sup>7</sup> James A. Grunkemeyer,<sup>8</sup> Michael A. Hollingsworth,<sup>8</sup> Paul M. Grandgenett,<sup>8</sup> Mikala Egeblad,<sup>1</sup> Youngkyu Park,<sup>1,2</sup> David A. Tuveson,<sup>1,2,\*</sup> and Christopher R. Vakoc<sup>1,10,\*</sup>



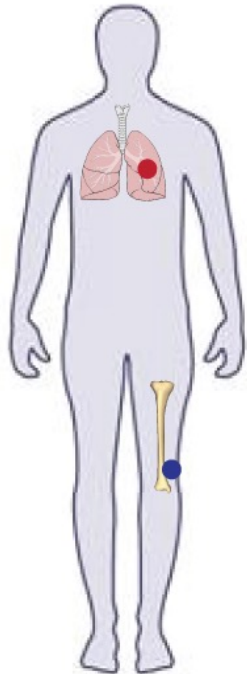
*Nature Medicine* 2018

## Positively selected enhancer elements endow osteosarcoma cells with metastatic competence

James J Morrow<sup>1,2</sup>, Ian Bayles<sup>2</sup>, Alister P W Funnell<sup>3</sup>, Tyler E Miller<sup>1</sup>, Alina Saiakhova<sup>2</sup>, Michael M Lizardo<sup>4</sup>, Cynthia F Bartels<sup>2</sup>, Maaikje Y Kapteijn<sup>5</sup>, Stevephen Hung<sup>2</sup>, Arnulfo Mendoza<sup>4</sup>, Gursimran Dhillon<sup>2</sup>, Daniel R Chee<sup>6</sup>, Jay T Myers<sup>7</sup>, Frederick Allen<sup>1</sup>, Marco Gambarotti<sup>8</sup>, Alberto Righi<sup>8</sup>, Analisa DiFeo<sup>9</sup>, Brian P Rubin<sup>10</sup>, Alex Y Huang<sup>1,7</sup>, Paul S Meltzer<sup>11</sup>, Lee J Helman<sup>4</sup>, Piero Picci<sup>8</sup>, Henri H Versteeg<sup>5</sup>, John A Stamatoyannopoulos<sup>3</sup>, Chand Khanna<sup>4,12</sup> & Peter C Scacheri<sup>2,8</sup>

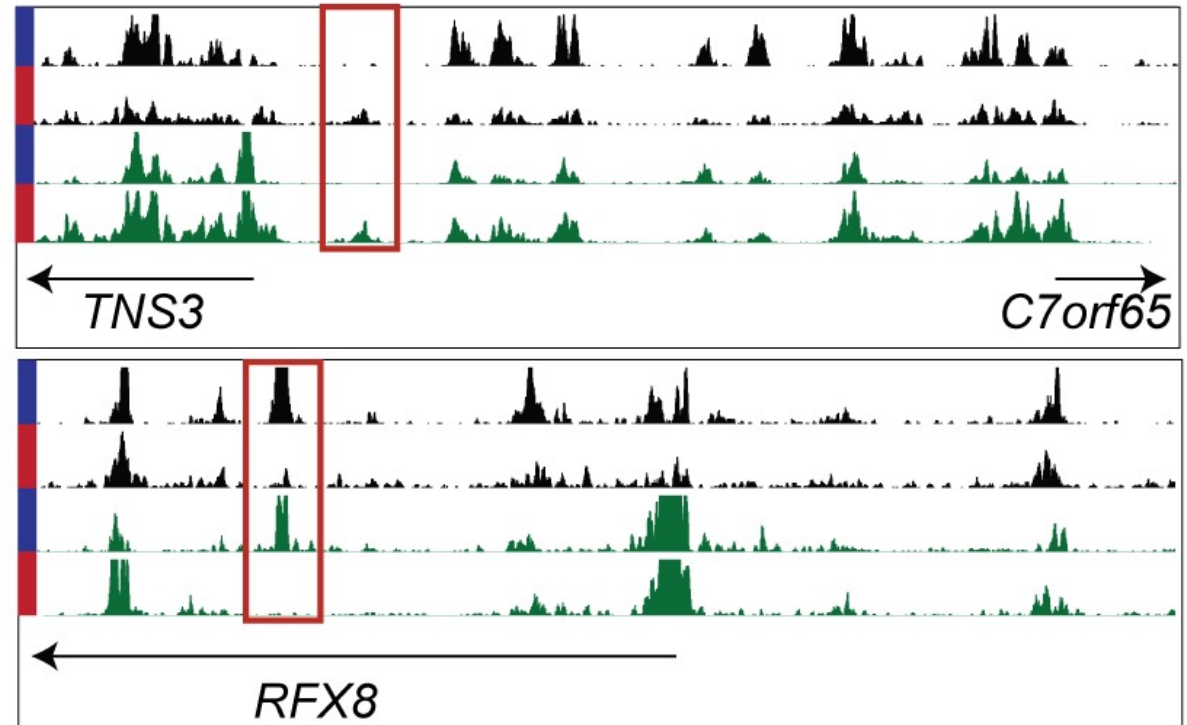
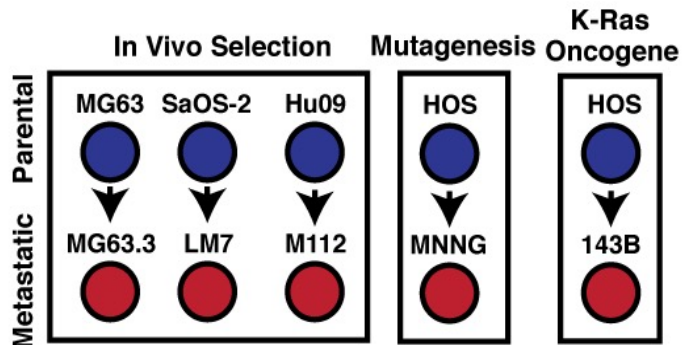
# Metastatic tumors show extensive enhancer alterations (Met-VELs)

Patient Tumors



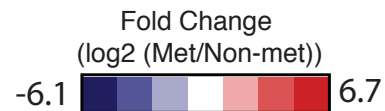
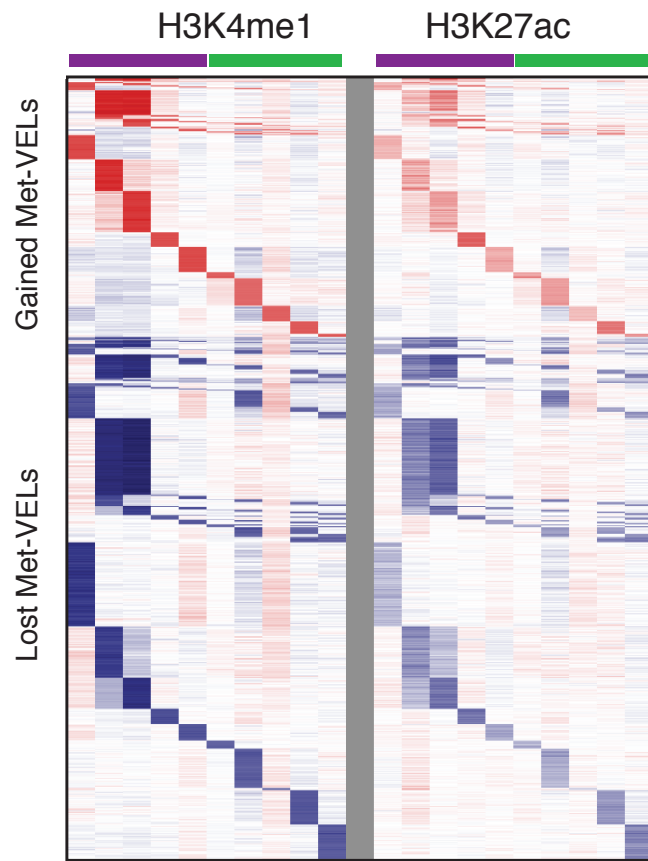
x10<sup>5</sup>

Cell Line Models



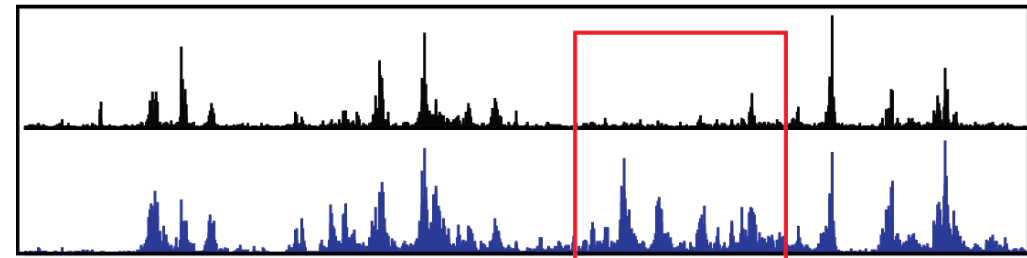
■ Primary Tumor    ■ H3K4me1  
■ Lung Metastasis    ■ H3K27ac

# Enhancer alterations (Met-VELs) in metastatic osteosarcoma



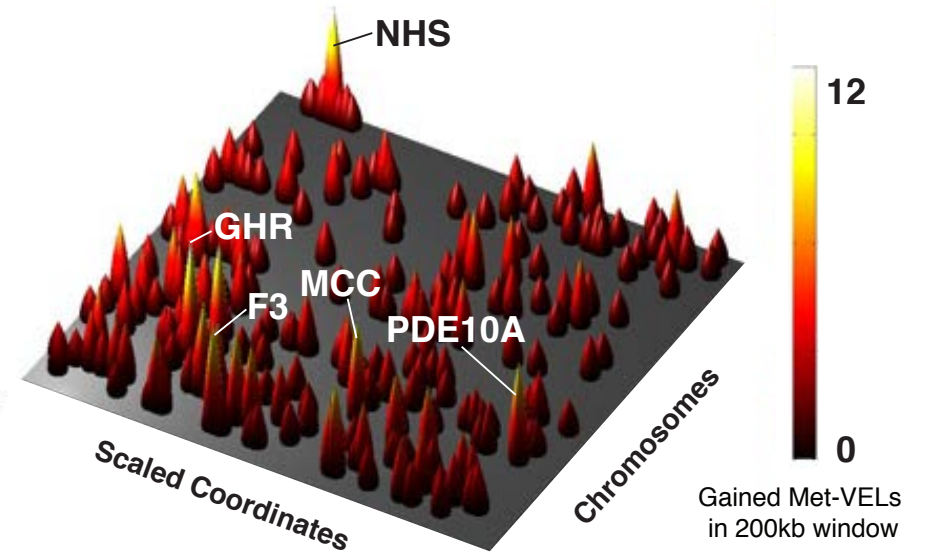
Non-Met.

Met.

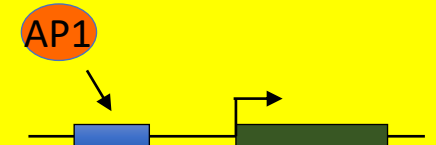


F3

Cell Line  
Patient Tumor

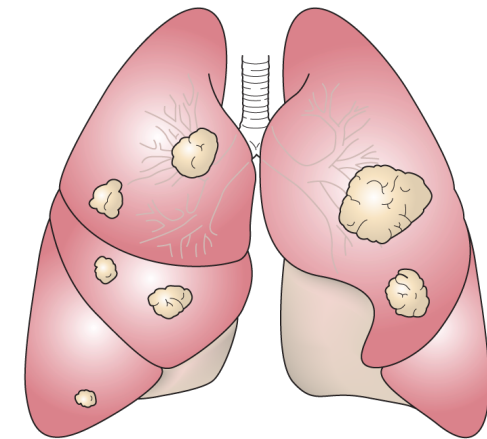
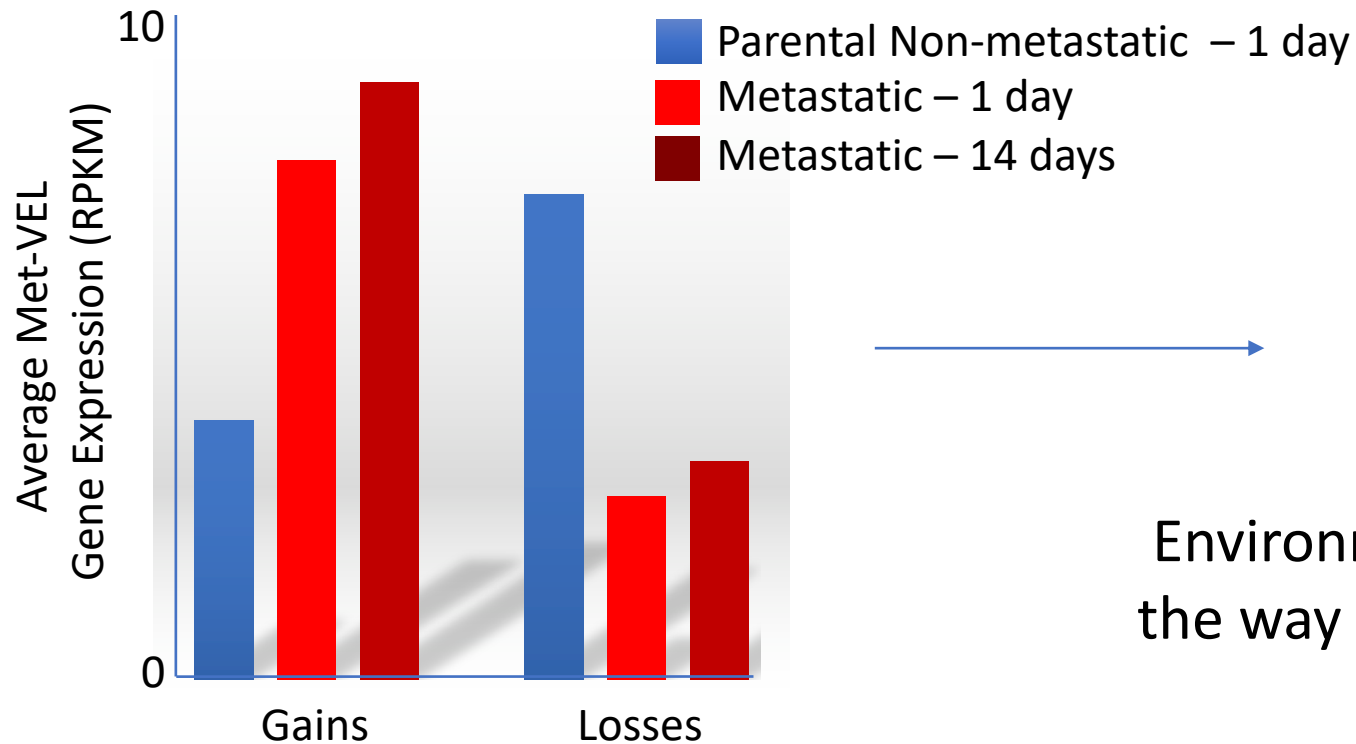
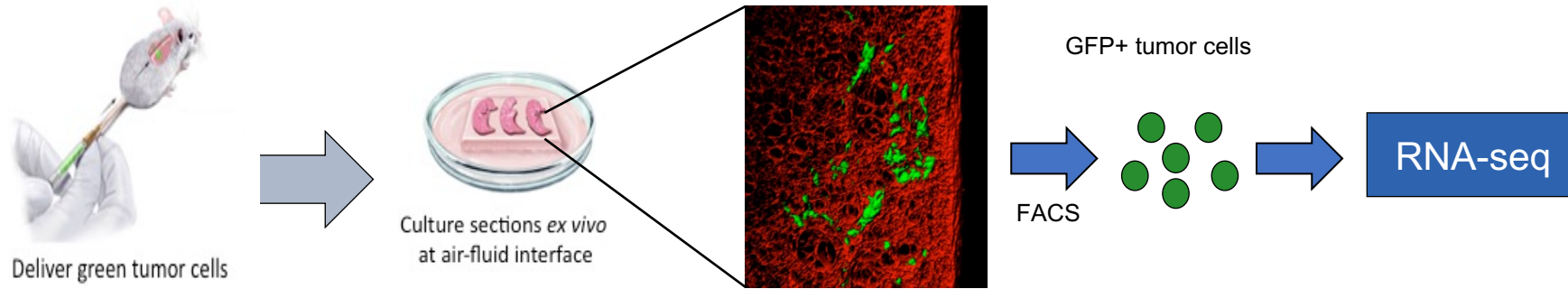


\*AP-1 transcription factors bind to Met-VELs  
(FOS, FOSL1, JUN, JUNB, JUND)



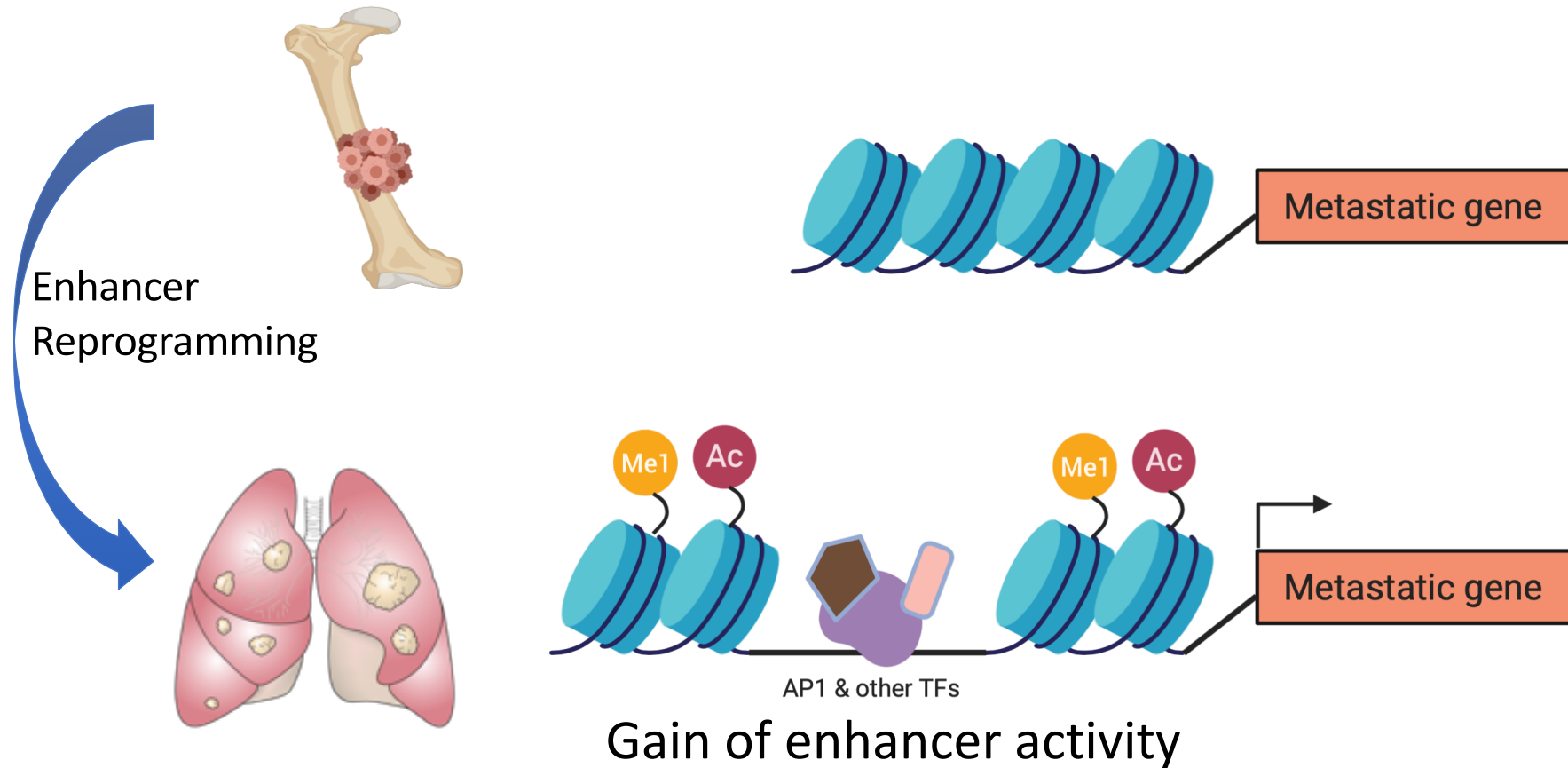


# Met-VEL genes switch on/off in the lung microenvironment

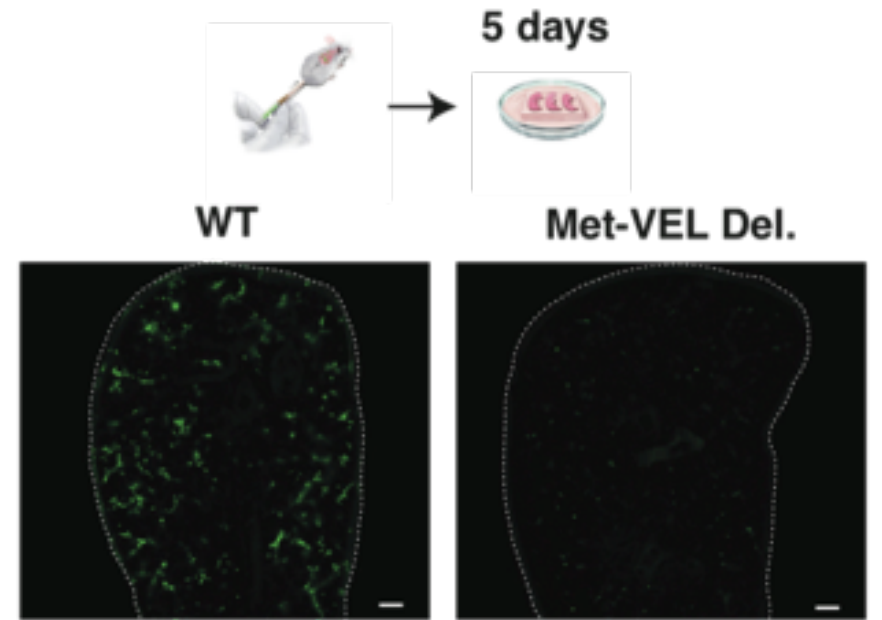
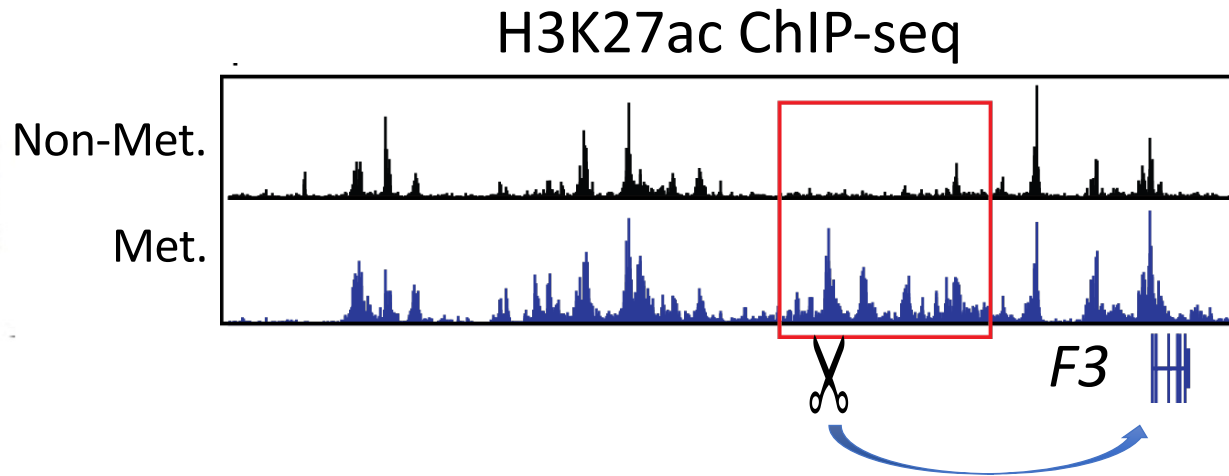
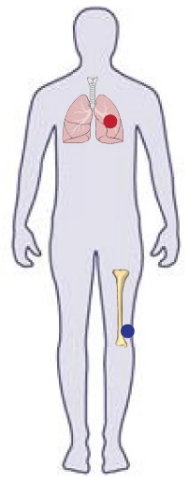


Environment of the lung is changing the way the Met-VELs regulate genes

# Model of enhancer function in metastasis

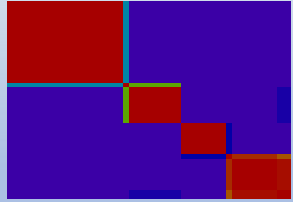


# Metastasis is dependent on the *F3* enhancer

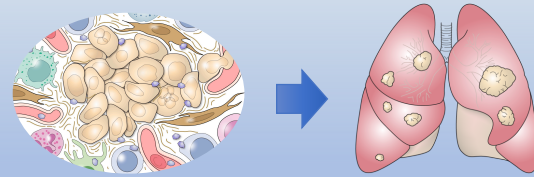


# Enhancer changes mediating acquisition of traits

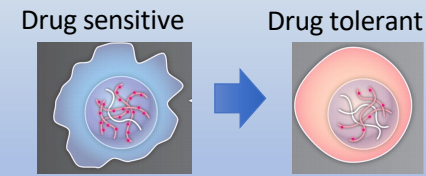
Subtypes



Metastasis



Drug resistance



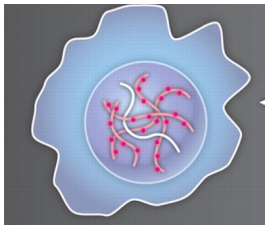
Predisposition



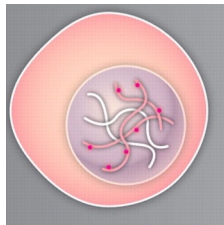
## Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer

Luca Magnani<sup>a,1</sup>, Alexander Stoeck<sup>b</sup>, Xiaoyang Zhang<sup>a</sup>, András Lánckzy<sup>c</sup>, Anne C. Mirabella<sup>a,2</sup>, Tian-Li Wang<sup>b</sup>, Balázs Györffy<sup>c,3</sup>, and Mathieu Lupien<sup>a,d,e,f,3</sup>

Drug sensitive



Drug tolerant



Genome and Epigenome

### Chemotherapy-Induced Distal Enhancers Drive Transcriptional Programs to Maintain the Chemoresistant State in Ovarian Cancer

Stephen Shang<sup>1</sup>, Jiekun Yang<sup>1</sup>, Amir A. Jazaeri<sup>2</sup>, Alexander James Duval<sup>1</sup>, Turan Tufan<sup>1</sup>, Natasha Lopes Fischer<sup>1</sup>, Mouadh Benamar<sup>1,3</sup>, Fadila Guessous<sup>3</sup>, Inyoung Lee<sup>1</sup>, Robert M. Campbell<sup>4</sup>, Philip J. Ebert<sup>4</sup>, Tarek Abbas<sup>1,3</sup>, Charles N. Landen<sup>5</sup>, Analisa Difeo<sup>6</sup>, Peter C. Scacheri<sup>6</sup>, and Mazhar Adli<sup>1</sup>

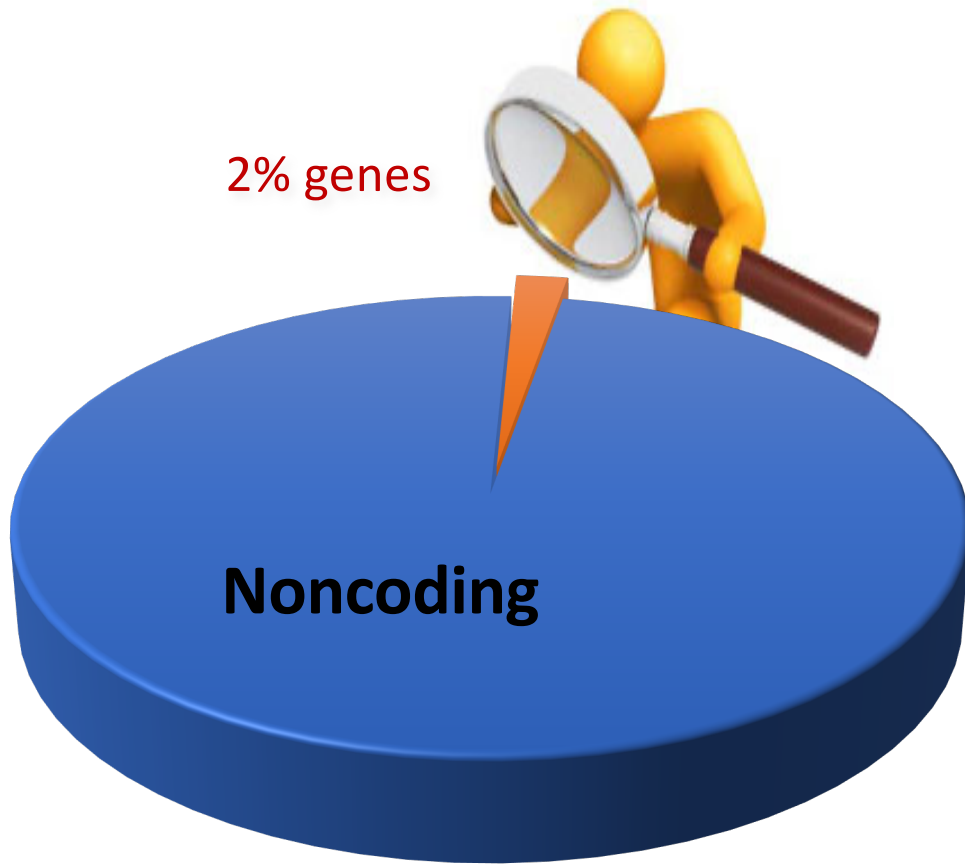
Cancer Research



# Studying the epigenome can provide insights into various aspects of cancer

- Tumorigenesis
- Clinical subtyping
- Metastasis
- Drug resistance
- **Identification of non-coding driver mutations**
- GWAS
- Biomarkers

# Search for cancer driver mutations in the human genome



	BLCA	BRCA	COAD/READ	GBM	HNSC	KIRC	AML	LUAD	LUSC	OV	UCEC	Pan-Cancer	
MAPK signalling	0.0	0.8	<b>45.1</b>	0.7	0.3	0.2	4.0	<b>26.3</b>	1.2	0.6	<b>20.0</b>	6.7	KRAS
	7.1	2.5	1.0	<b>11.0</b>	2.7	1.7	1.0	<b>11.8</b>	<b>10.3</b>	3.8	3.5	4.4	NF1
	3.1	<b>7.2</b>	0.0	2.1	1.0	1.2	0.0	1.8	1.7	0.3	3.5	2.7	MAP3K1
	2.0	0.4	3.6	2.1	1.0	0.2	0.0	<b>6.6</b>	4.6	0.6	0.9	1.5	BRAF
	2.0	0.1	<b>8.8</b>	0.3	0.0	0.0	<b>7.5</b>	1.8	0.6	0.6	2.6	1.5	NRAS
	0.0	<b>4.1</b>	2.6	0.0	0.3	0.0	0.0	1.3	0.6	0.3	1.3	1.4	MAP2K4
PI(3)K signalling	2.0	0.3	<b>2.1</b>	0.7	0.7	0.5	0.0	1.8	1.2	0.3	0.4	0.7	MAPK8IP1
	<b>17.4</b>	<b>33.6</b>	17.6	<b>11.0</b>	<b>20.6</b>	2.9	0.0	4.4	<b>14.9</b>	0.6	<b>52.2</b>	17.8	PIK3CA
	3.1	3.8	1.0	<b>30.7</b>	1.3	4.3	0.0	2.2	8.1	0.6	<b>63.5</b>	9.7	PTEN
	1.0	2.5	2.1	<b>11.4</b>	1.7	0.5	0.0	1.3	0.6	0.3	<b>30.9</b>	4.4	PIK3R1
	2.0	1.2	0.0	0.3	2.0	0.5	0.5	<b>11.4</b>	5.8	1.0	0.4	1.9	TLR4
	2.0	0.4	0.5	2.4	2.7	0.7	0.0	<b>5.3</b>	<b>7.5</b>	1.0	1.3	1.7	PIK3CG
TGF- $\beta$ signalling	0.0	<b>2.5</b>	0.0	0.3	0.7	0.5	0.0	0.0	0.6	0.0	1.3	0.9	AKT1
	2.0	0.4	<b>9.8</b>	0.3	2.0	0.5	0.0	3.1	2.9	0.0	0.0	1.4	SMAD4
	<b>3.1</b>	0.4	2.6	0.7	3.0	0.2	0.0	0.9	1.7	1.0	1.3	1.1	TGFBR2
	0.0	0.7	<b>3.6</b>	0.0	1.3	1.0	0.0	2.2	1.2	0.3	1.7	1.0	ACVR1B
	1.0	0.5	<b>5.7</b>	0.0	1.0	0.5	0.0	0.9	1.2	0.0	1.3	0.9	SMAD2
	1.0	0.5	<b>2.6</b>	0.0	0.7	0.2	0.0	0.9	1.2	0.0	0.4	0.6	ACVR2A
Wnt/ $\beta$ -catenin signalling	4.1	0.5	<b>81.9</b>	0.3	4.0	1.4	0.0	<b>9.2</b>	4.0	2.2	<b>5.7</b>	7.3	APC
	2.0	0.1	4.7	0.3	0.7	0.2	0.0	3.5	1.7	0.6	<b>28.3</b>	2.9	CTNNB1
	3.1	0.1	<b>3.6</b>	0.3	1.7	0.2	0.0	0.9	0.6	0.3	2.6	0.9	AXIN2
	2.0	1.1	0.0	0.0	1.0	0.7	0.0	<b>2.2</b>	1.2	0.3	1.3	0.8	TBL1XR1

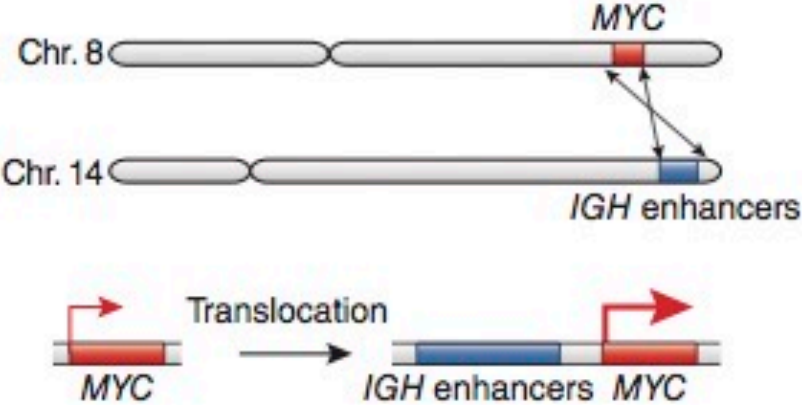
Kandoth et al, *Nature* (2013)

# Challenges with finding mutations in non-coding regions

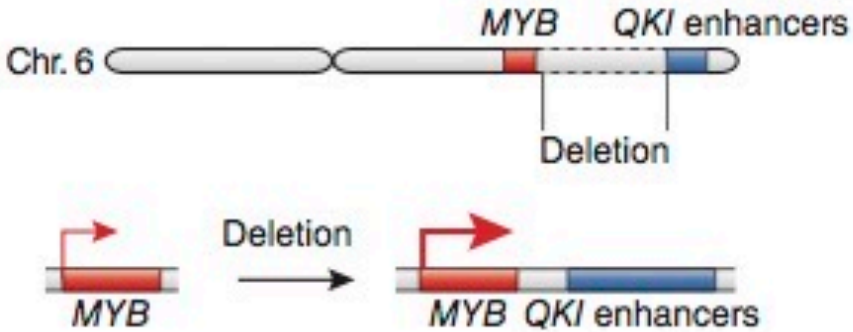
- Mutation rates vary between cancers
- Mutation rates are influenced by chromatin states
  - Active chromatin – low mutation rate
  - Inactive chromatin – high mutation rate
- Epigenome varies between tumors
- Cell type of origin is unknown or unavailable

# Cancer driver mutations in noncoding regions

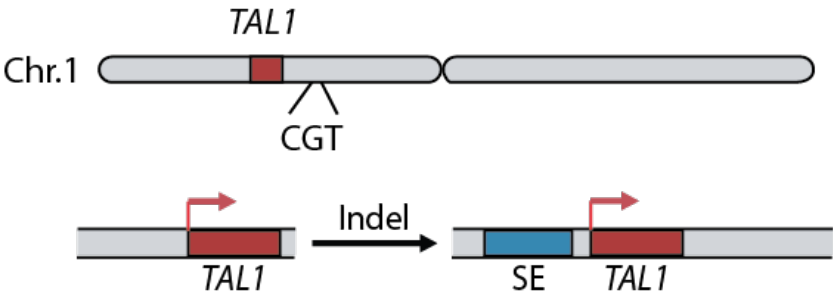
## Structural Rearrangements



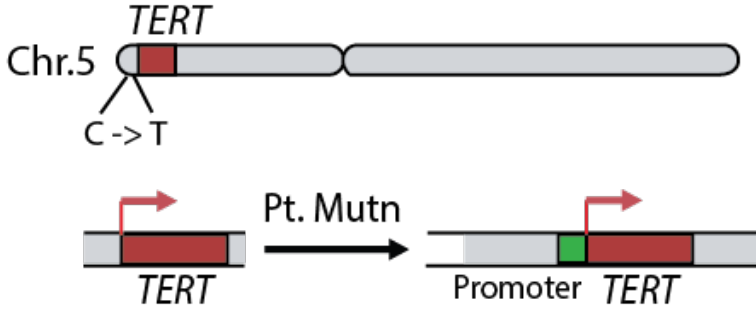
## Copy Number Alterations



## Indels



## Point Mutations





# Recurrent driver mutations in noncoding regions are rare



**PCAWG**  
PanCancer Analysis  
OF WHOLE GENOMES

New Results

Discovery and characterization of coding and non-coding driver mutations in more than 2,500 whole cancer genomes

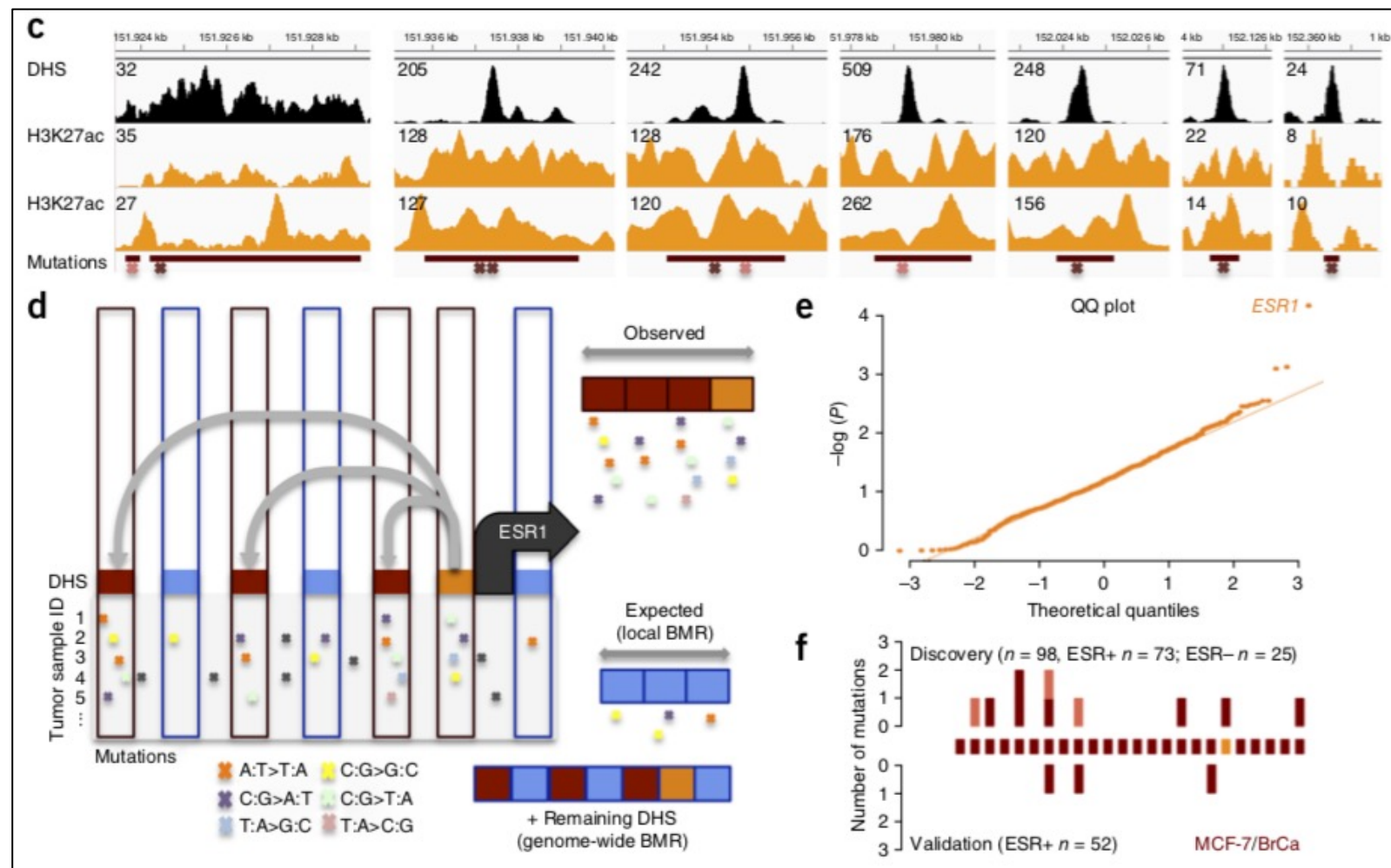
bioRxiv  
THE PREPRINT SERVER FOR BIOLOGY

“Perhaps the most striking finding is the relative paucity of point mutations driving cancer in non-coding genes and regulatory elements.”

Mark P Hamilton, Chen Hong, Andre Kanies, Youngwook Kim, Kjong-van Lehmann,  
Todd Andrew A Johnson, Abdullah Kahraman, Keunchil Park, Gordon Saksena, Lina Sieverling,  
Nicholas A Sinnott-Armstrong, Peter J Campbell, Asger Hobolth, Manolis Kellis, Michael S Lawrence,  
Ben Raphael, Mark A Rubin, Chris Sander, Lincoln Stein, Josh Stuart, Tatsuhiko Tsunoda,  
David A Wheeler, Rory Johnson, Jüri Reimand, Mark B Gerstein, Ekta Khurana,  
Nuria Lopez-Bigas, Inigo Martincorena, Jakob Skou Skou Pedersen, Gad Getz,

# Noncoding somatic and inherited single-nucleotide variants converge to promote *ESR1* expression in breast cancer

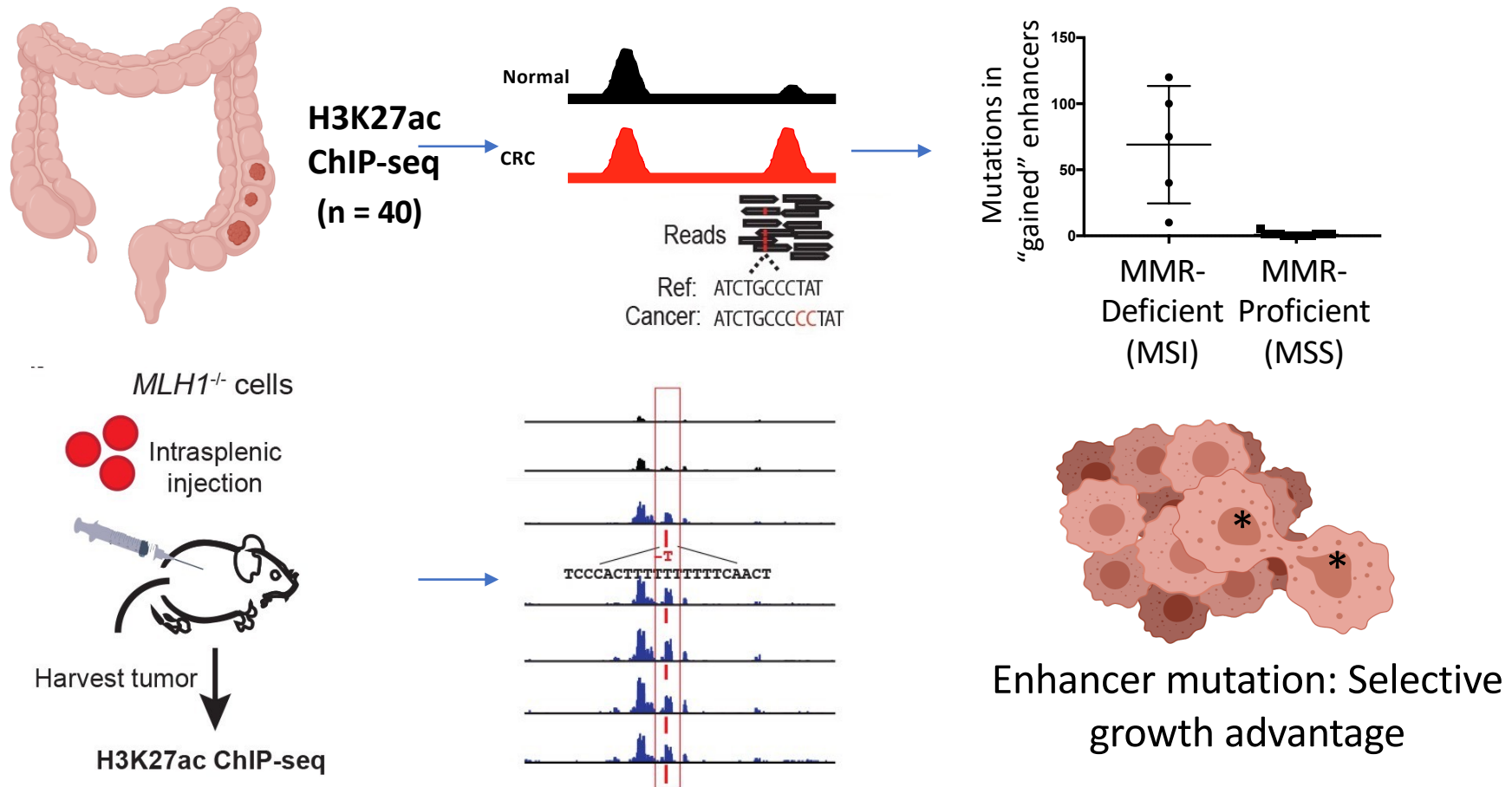
Swneke D Bailey<sup>1,2,11</sup>, Kinjal Desai<sup>3,11</sup>, Ken J Kron<sup>1,2</sup>, Parisa Mazrooei<sup>1,2</sup>, Nicholas A Sinnott-Armstrong<sup>4</sup>, Aislinn E Treloar<sup>1,2</sup>, Mark Dowar<sup>1</sup>, Kelsie L Thu<sup>5</sup>, David W Cescon<sup>1,5</sup>, Jennifer Silvester<sup>5</sup>, S Y Cindy Yang<sup>1,2</sup>, Xue Wu<sup>1,10</sup>, Rossanna C Pezo<sup>1</sup>, Benjamin Haibe-Kains<sup>1,2,6</sup>, Tak W Mak<sup>2,5</sup>, Philippe L Bedard<sup>1,7</sup>, Trevor J Pugh<sup>1,2</sup>, Richard C Sallari<sup>8</sup> & Mathieu Lupien<sup>1,2,9</sup>



# Mismatch repair-signature mutations activate gene enhancers across human colorectal cancer epigenomes

eLife 2019

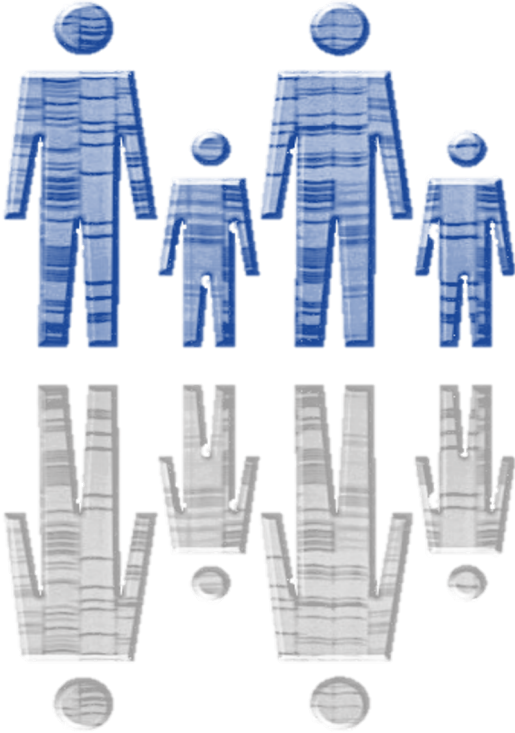
Stephen Hung<sup>1</sup>, Alina Saiakhova<sup>1</sup>, Zachary J Faber<sup>1</sup>, Cynthia F Bartels<sup>1</sup>, Devin Neu<sup>1</sup>, Ian Bayles<sup>1</sup>, Evelyn Ojo<sup>2</sup>, Ellen S Hong<sup>1</sup>, W Dean Pontius<sup>3</sup>, Andrew R Morton<sup>1</sup>, Ruifu Liu<sup>2</sup>, Matthew F Kalady<sup>3,4,5</sup>, David N Wald<sup>2,6</sup>, Sanford Markowitz<sup>1,6,7</sup>, Peter C Scacheri<sup>1,6\*</sup>



# Studying the epigenome can provide insights into various aspects of cancer

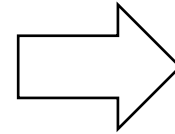
- Tumorigenesis
- Clinical subtyping
- Metastasis
- Drug resistance
- Identification of non-coding driver mutations
- **GWAS**
- Biomarkers

# GWAS: Genome Wide Association Studies



**Cases**

**Controls**



Variants associated with  
increased risk for  
disease

# GWAS: Genome Wide Association Studies



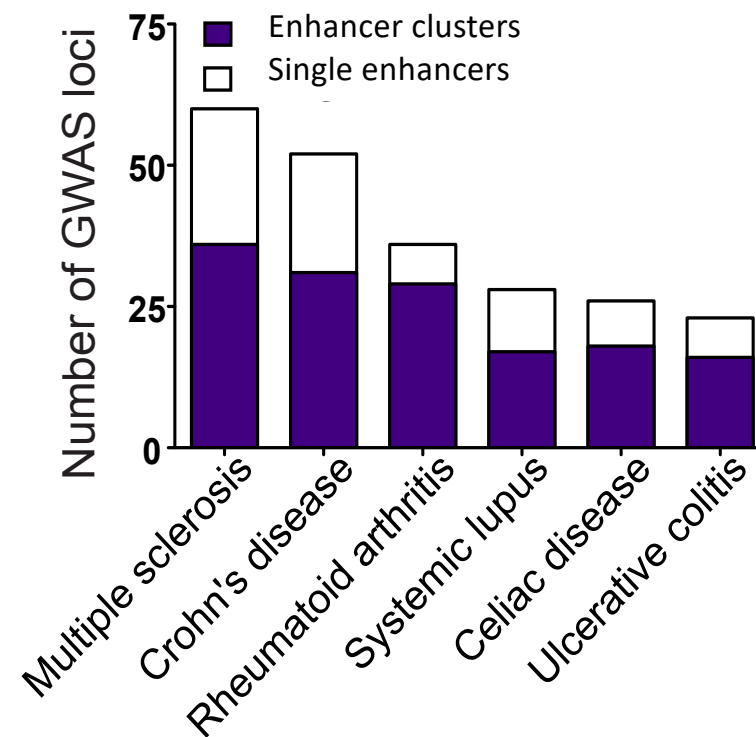
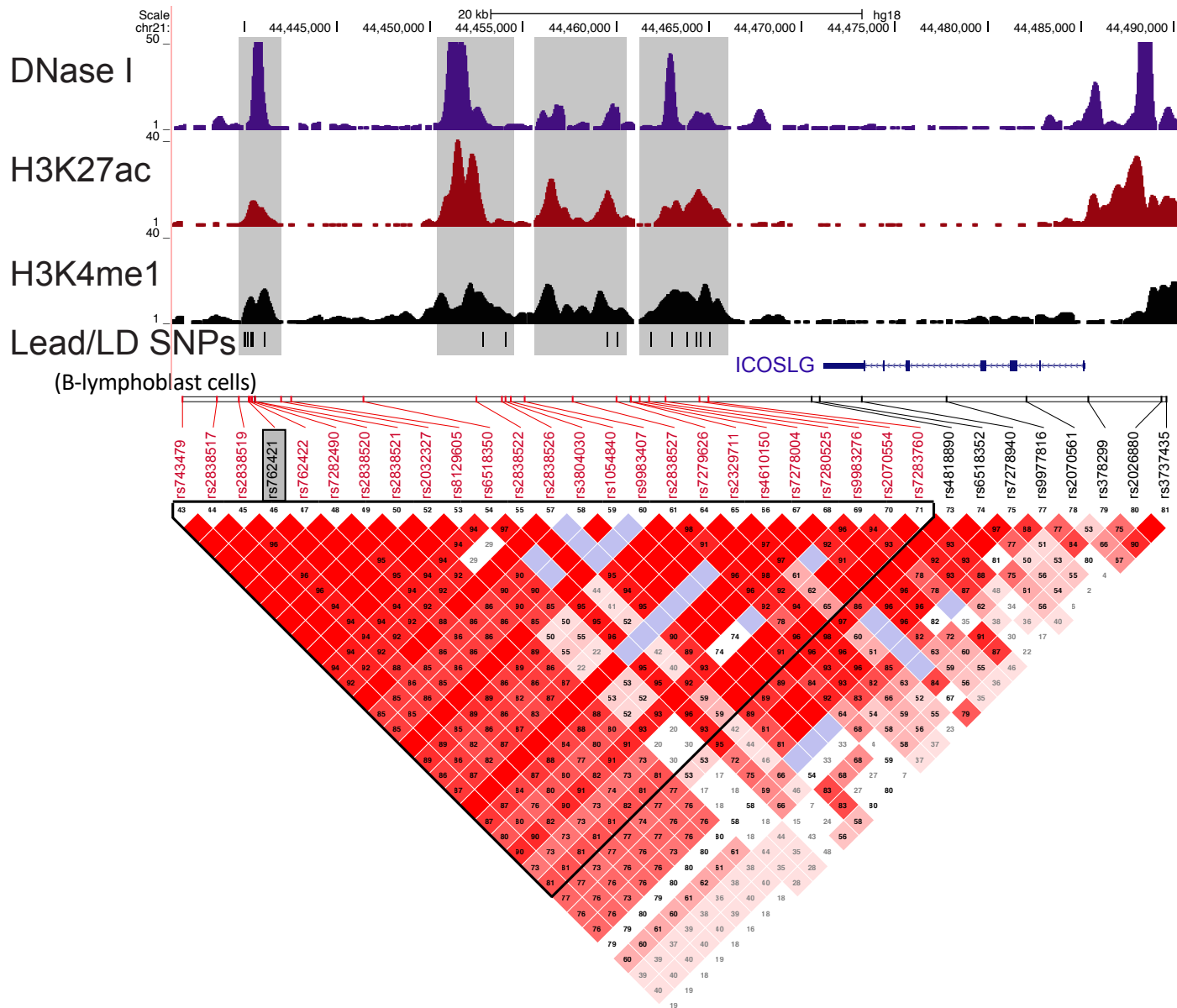
- Thousands of associations
- Most SNPs lie in non-coding regions



~80% lie in cell type-specific enhancers



# GWAS risk SNPs often lie in enhancer clusters



*Hnisz et al, Cell 2013*  
*Parker, Stitzel et al, PNAS 2013*  
*Corradin et al. Genome Res 2014*  
*Pasquali et al, Nat. Gen 2014*



# Constituents of enhancer clusters collude to regulate genes

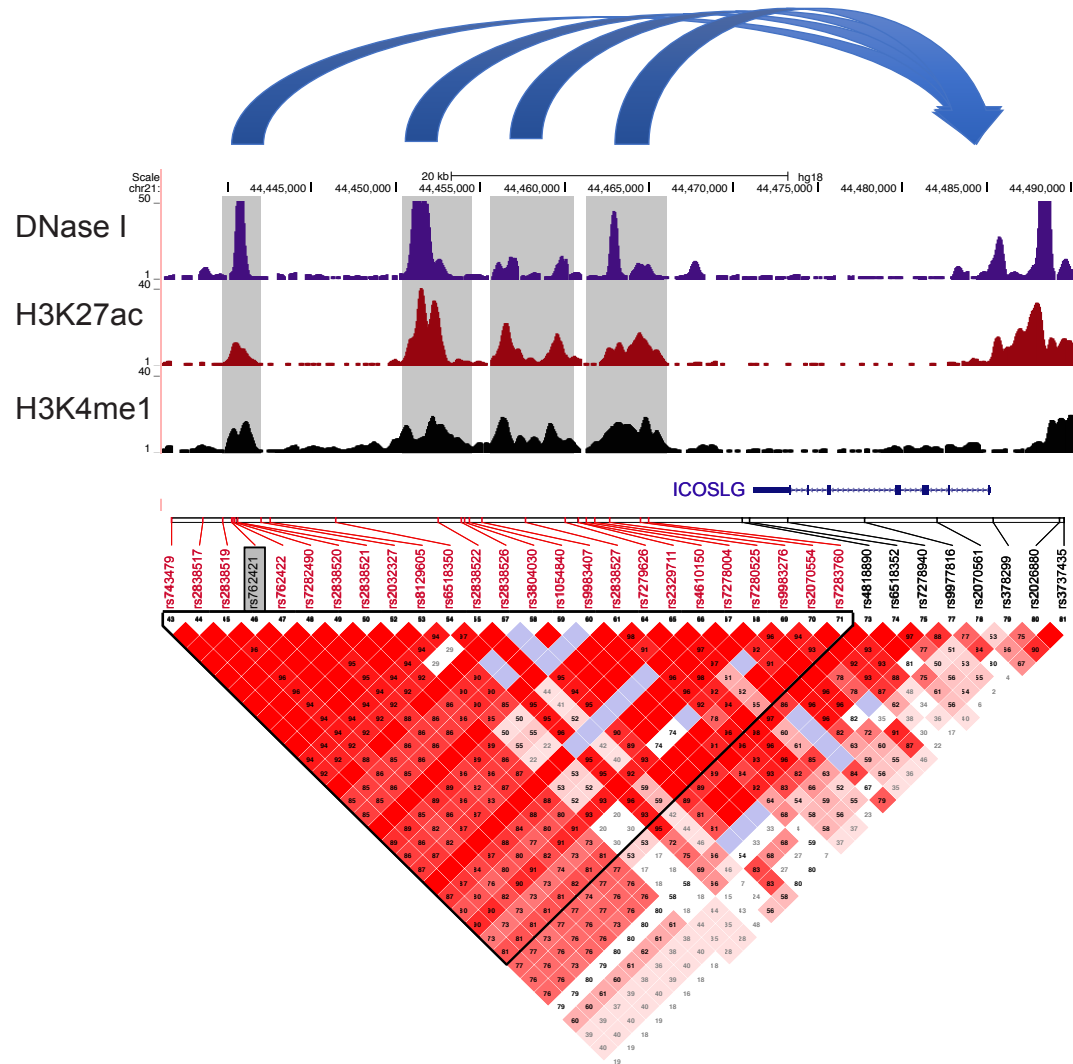
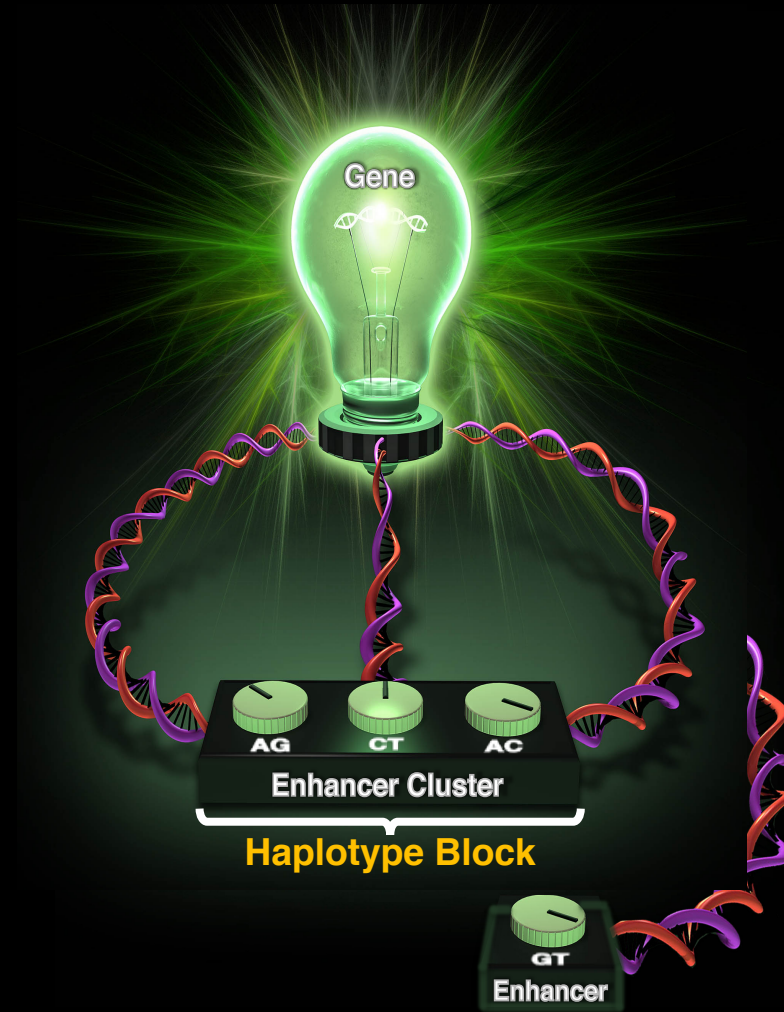


Image by Wayne Heim

May be more than 1 causal SNP at a locus!

# The Regulatory Circuitry of Gene Expression



**SNPs that lie outside the GWAS-associated region, but are part of the same regulatory circuit can influence disease risk**

# Studying the epigenome can provide insights into various aspects of cancer

- Tumorigenesis
- Clinical subtyping
- Metastasis
- Drug resistance
- Identification of non-coding driver mutations
- GWAS
- **Biomarkers**

# Sensitive tumour detection and classification using plasma cell-free DNA methylomes

Shu Yi Shen<sup>1,12</sup>, Rajat Singhania<sup>1,12</sup>, Gordon Fehrer<sup>2,12</sup>, Ankur Chakravarthy<sup>1,12</sup>, Michael H. A. Roehr<sup>1,3,4</sup>, Dianne Chadwick<sup>1</sup>, Philip C. Zuzarte<sup>5</sup>, Ayelet Borgida<sup>2</sup>, Ting Ting Wang<sup>1,4</sup>, Tiantian Li<sup>1</sup>, Olena Kis<sup>1</sup>, Zhen Zhao<sup>1</sup>, Anna Spreafico<sup>1</sup>, Tiago da Silva Medina<sup>1</sup>, Yadon Wang<sup>1</sup>, David Roulois<sup>1,6</sup>, Ilias Ettayebi<sup>1,4</sup>, Zhuo Chen<sup>1</sup>, Signy Chow<sup>1</sup>, Tracy Murphy<sup>1</sup>, Andrea Arruda<sup>1</sup>, Grainne M. O’Kane<sup>1</sup>, Jessica Liu<sup>4</sup>, Mark Mansour<sup>4</sup>, John D. McPherson<sup>7</sup>, Catherine O’Brien<sup>1</sup>, Natasha Leigh<sup>1</sup>, Philippe L. Bedard<sup>1</sup>, Neil Flesher<sup>1</sup>, Geoffrey Liu<sup>1,4,8</sup>, Mark D. Minden<sup>1</sup>, Steven Gallinger<sup>9,10</sup>, Anna Goldenberg<sup>11</sup>, Trevor J. Pugh<sup>1,4</sup>, Michael M. Hoffman<sup>1,4,11</sup>, Scott V. Bratman<sup>1,4</sup>, Rayjean J. Hung<sup>2,8\*</sup> & Daniel D. De Carvalho<sup>1,4\*</sup>

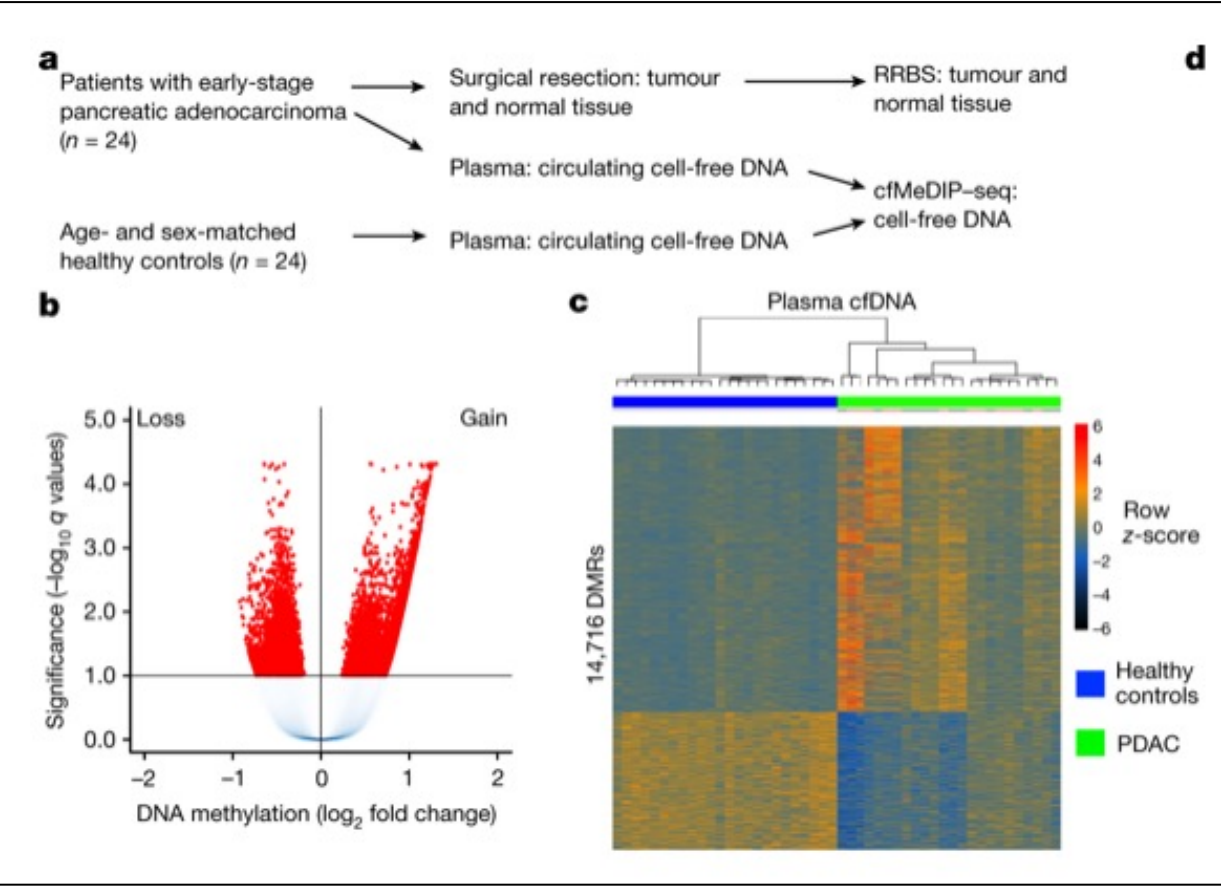
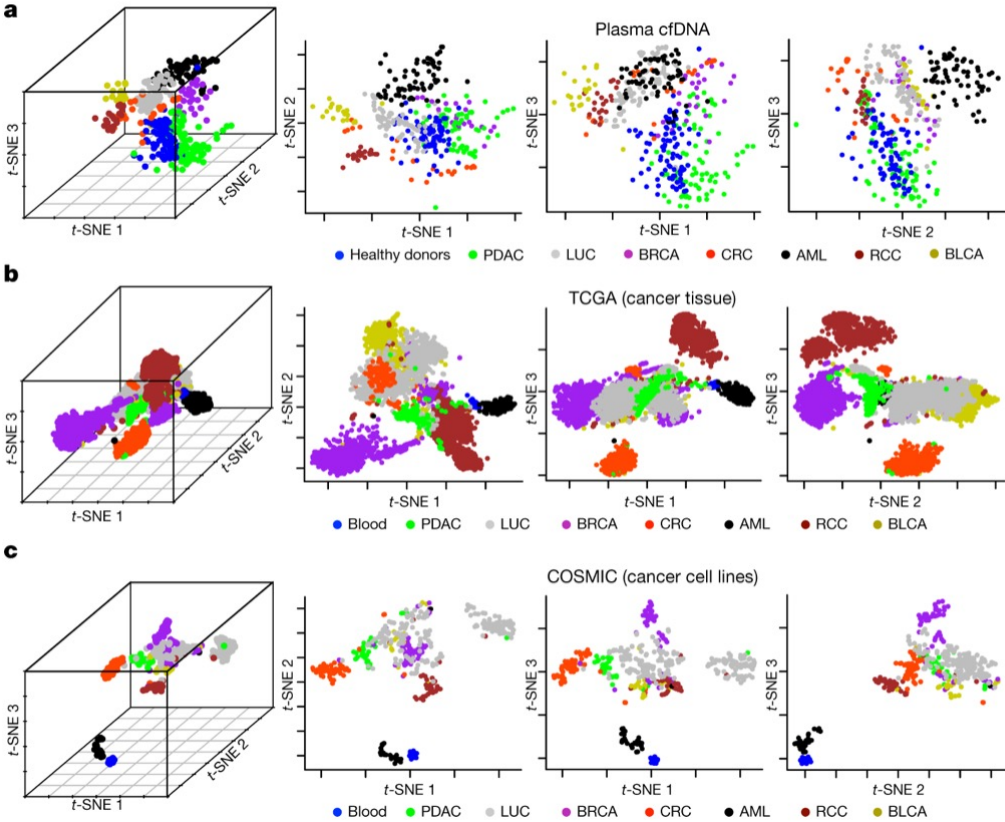


Fig. 4: Plasma-derived DMRs are informative of cancer type.

From: Sensitive tumour detection and classification using plasma cell-free DNA methylomes



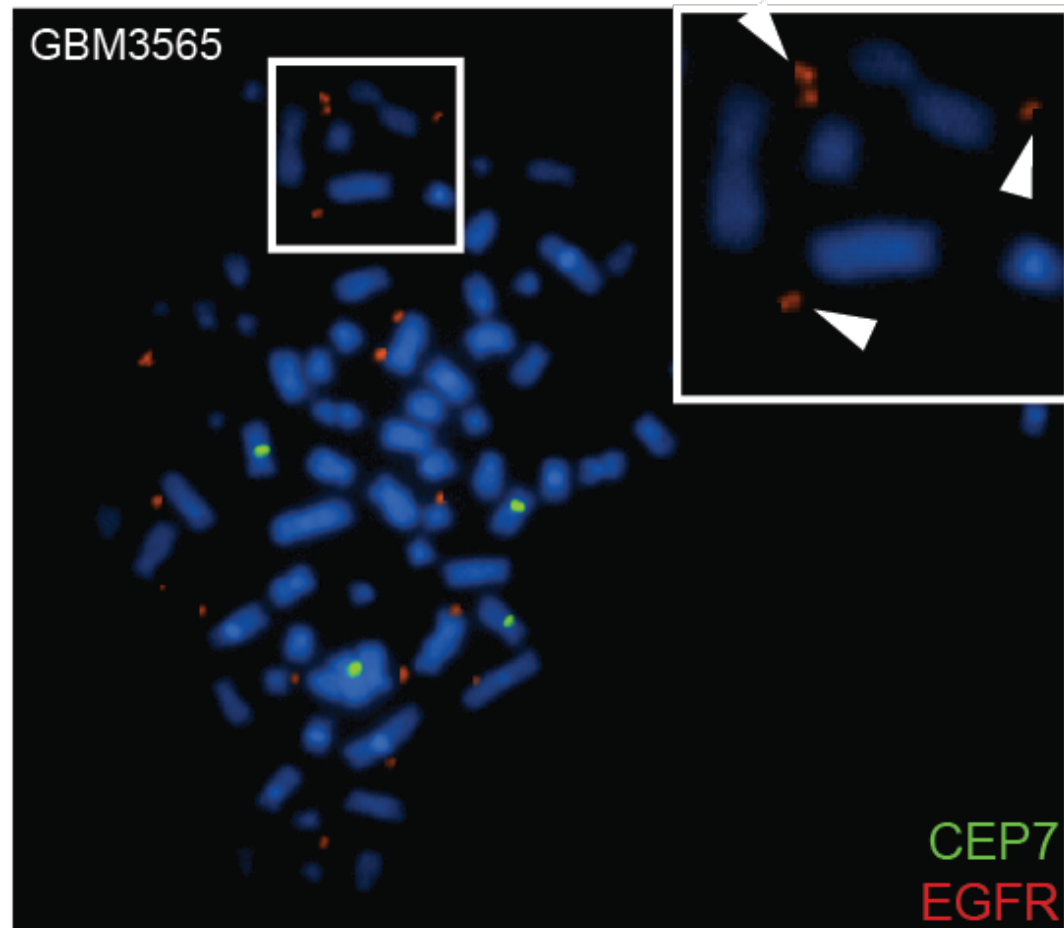
# Studying the epigenome can provide insights into various aspects of cancer

- Tumorigenesis
- Clinical subtyping
- Metastasis
- Drug resistance
- Identification of non-coding driver mutations
- GWAS
- Biomarkers
- **One final story**

# Functional enhancers on circular extrachromosomal DNA (ecDNA)

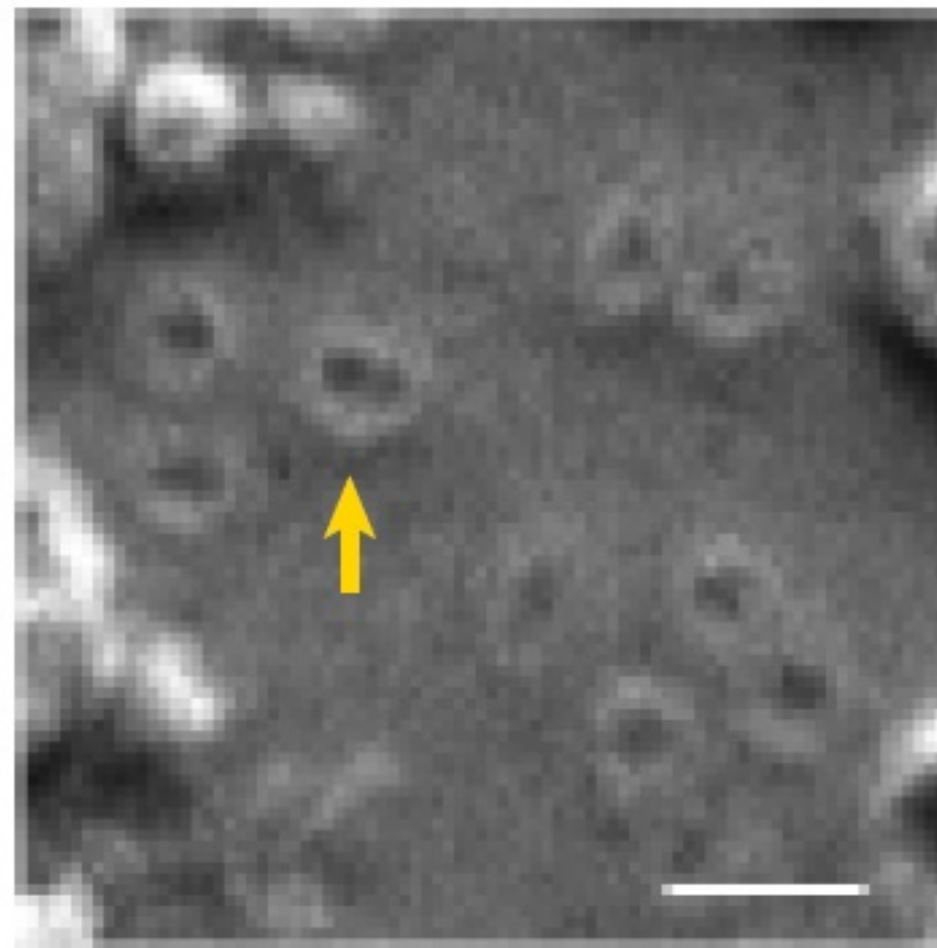


## DNA FISH



Morton et al, *Cell* 2019

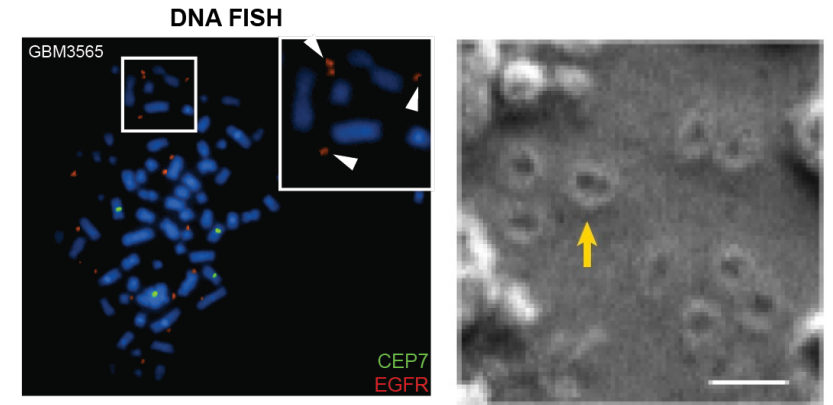
## Electron Micrograph



Wu et al, *Nature* 2019

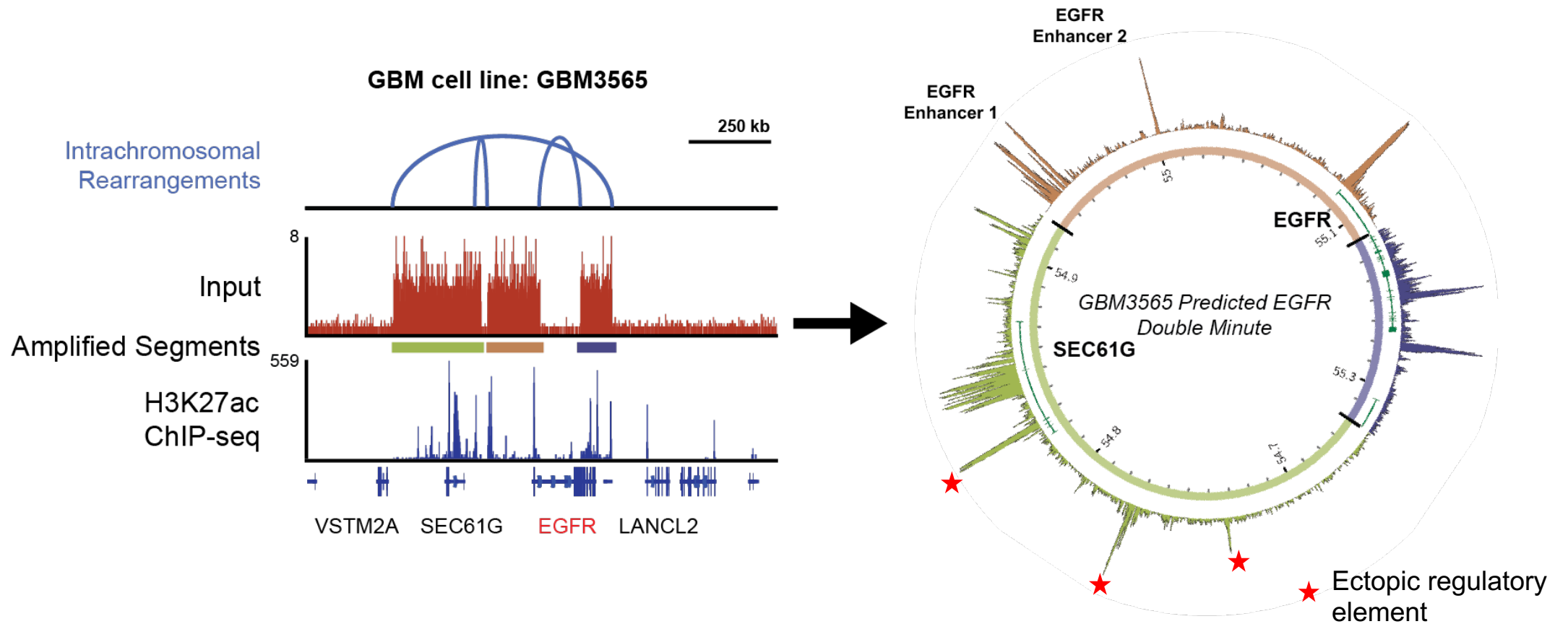
# Extrachromosomal DNA (ecDNA) AKA: “double minutes”

- Found in many forms of cancer.
  - Particularly prevalent in aggressive cancers notoriously difficult to treat
- Massive focal DNA amplifications
  - 10-100s of copies per cell
  - 0.5 – 2.5 Million base pairs in size
- Complex structures → can incorporate multiple oncogenes from different chromosomes
  - Extensive sub-clonality
- Can hop back into the genome and remodel it
- Provides a means for rapid tumor evolution and emergence of drug resistance phenotypes

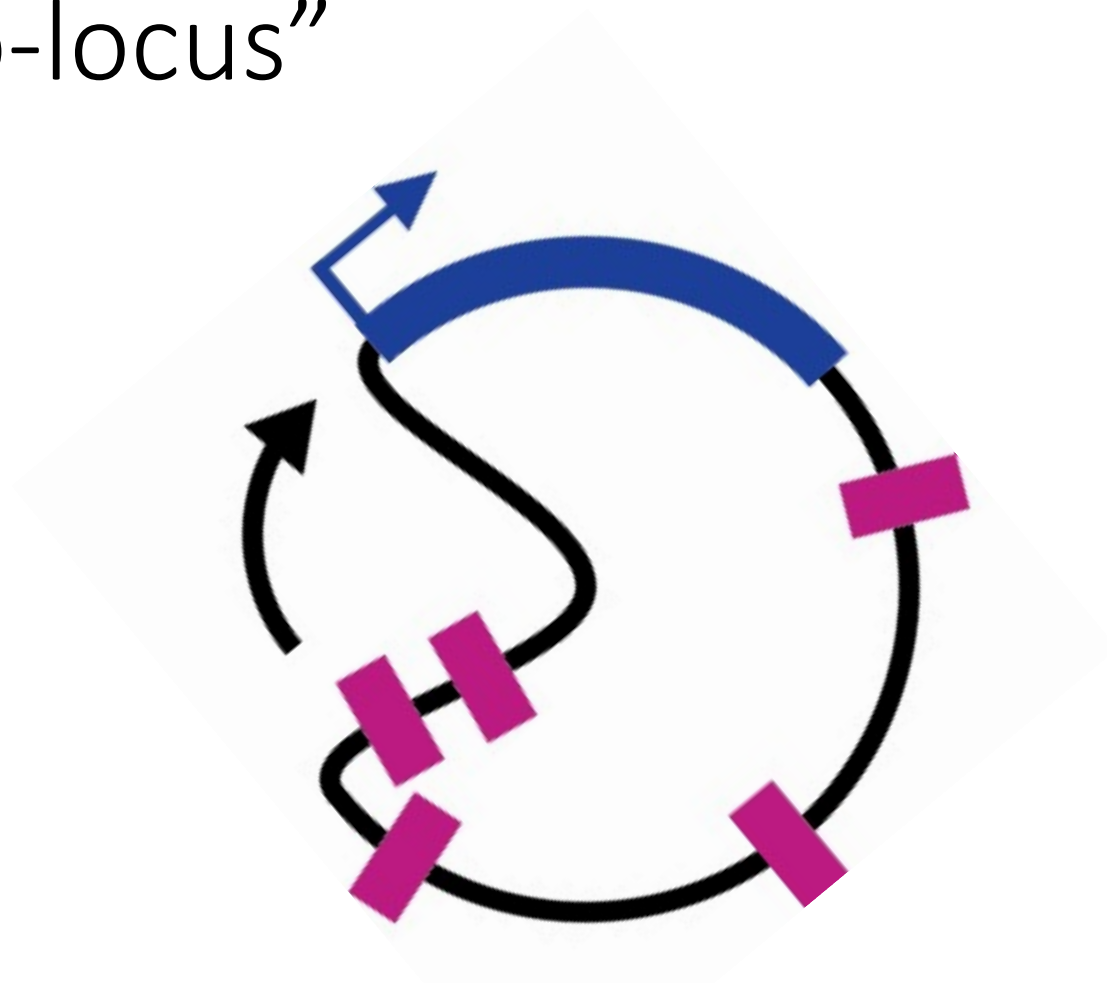




# Selection of enhancers on ecDNA



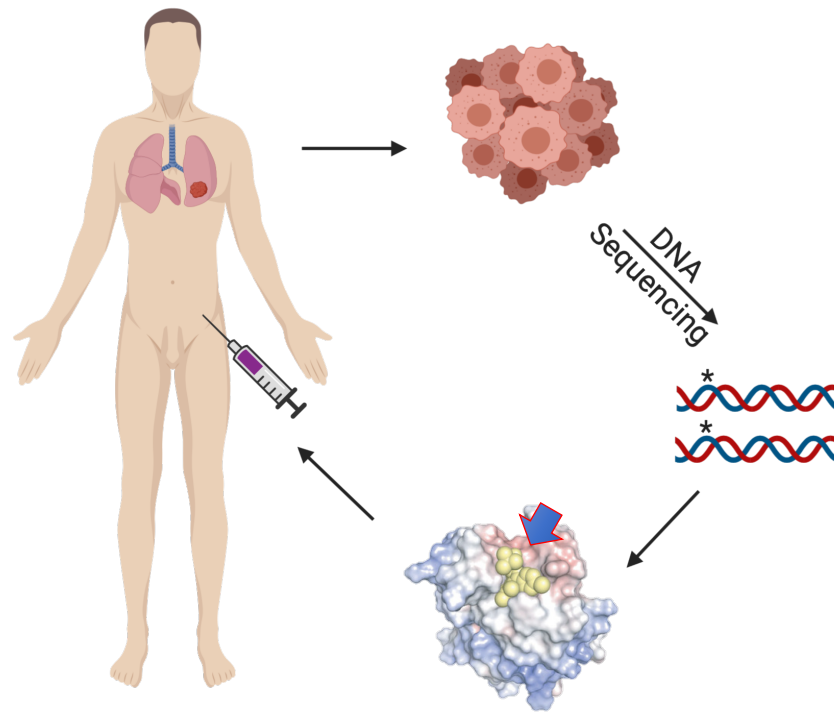
# The “Onco-locus”



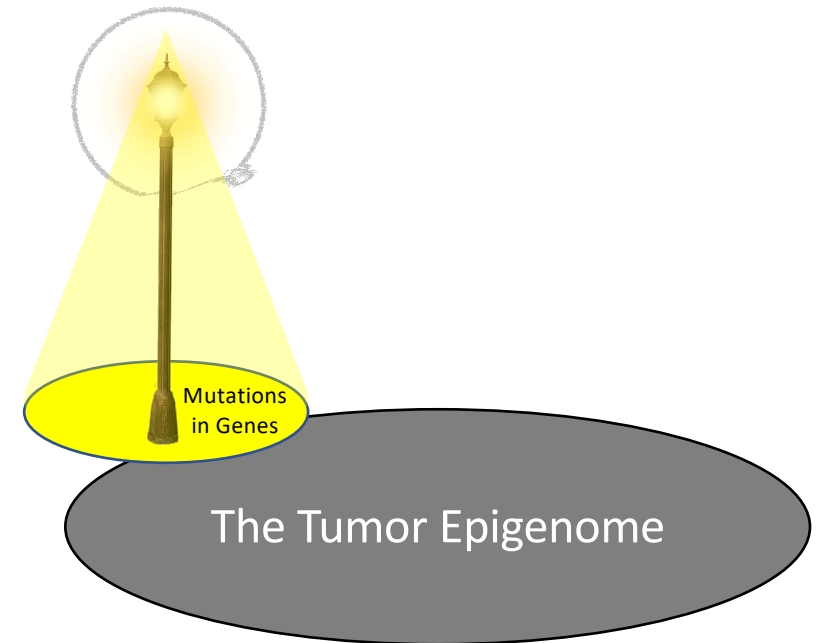
**Oncogene + enhancers +  
other selected elements contributing to fitness**

# Precision cancer medicine – A vision for an epigenomics based approach

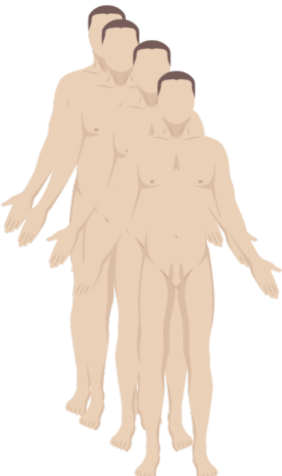
- Leverage knowledge about the genetic makeup of cancer for precisely targeted therapy.



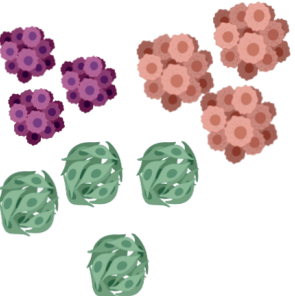
- Some success
- Most patients don't meet the clinical criteria



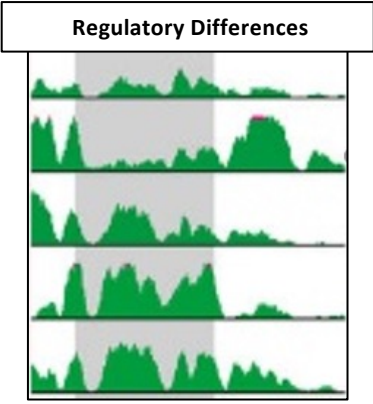
# Precision cancer medicine – An epigenomics based approach



Consented Patients



Tumor Collection



Epigenomic Profiling



Clinical Integration



Deliverables

Identification of epigenetic markers for tumor behavior and as drug targets

The End.

Questions?