



Single-cell Epigenomic Analysis of Mammalian Brains and Neurological Disorders

Yang (Eric) Li

CSHL seqtec course

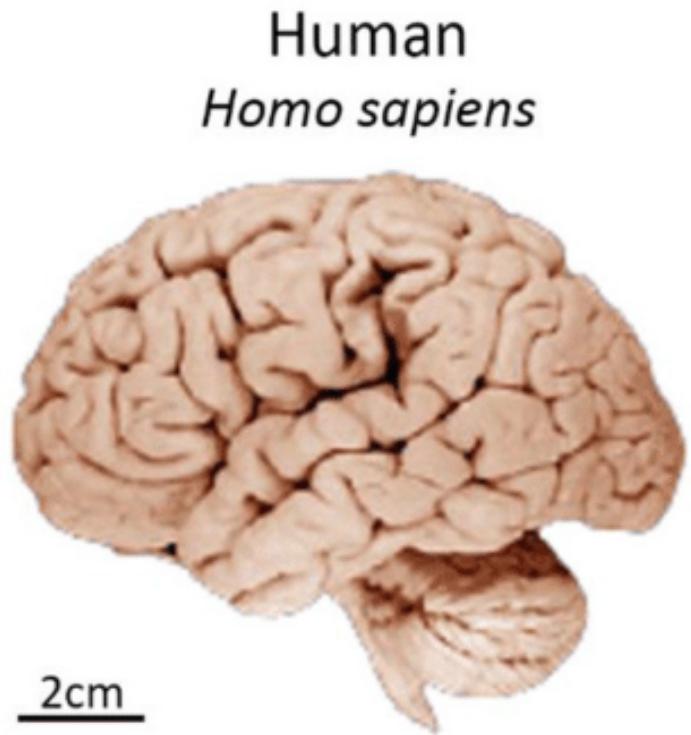
Nov 13, 2023



Yang (Eric) Li Lab
Website: yelilab.wustl.edu
Email: yeli@wustl.edu



What does the brain look like?

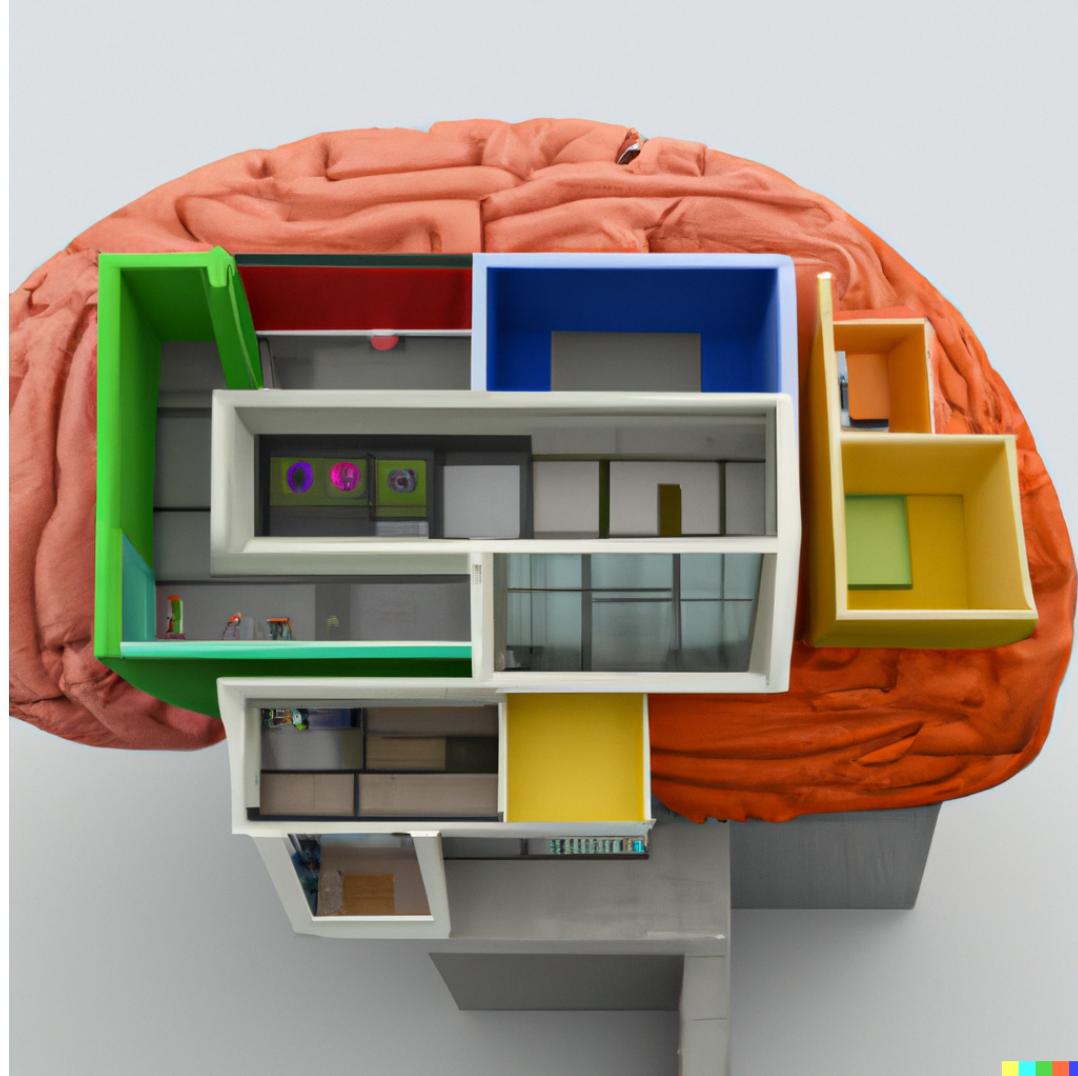


CSHL Dolan

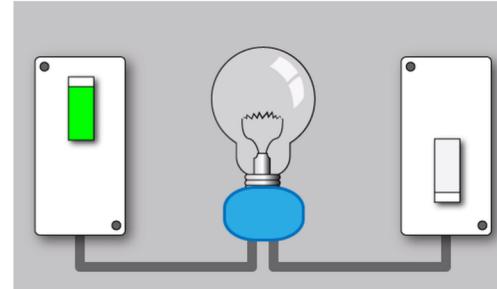


- Brains are built over time, from the bottom up.
- Brain architecture is comprised of billions of connections between neurons across different areas of the brain.

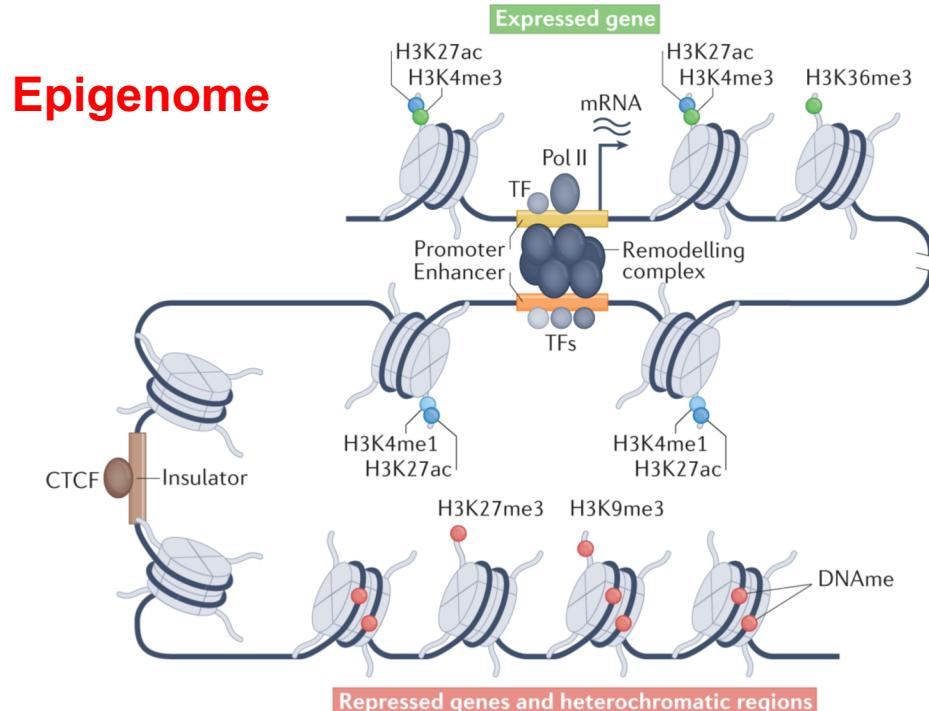
The gene regulation in glia and neurons is the key to complex functions



Created by DALL·E 2 - OpenAI

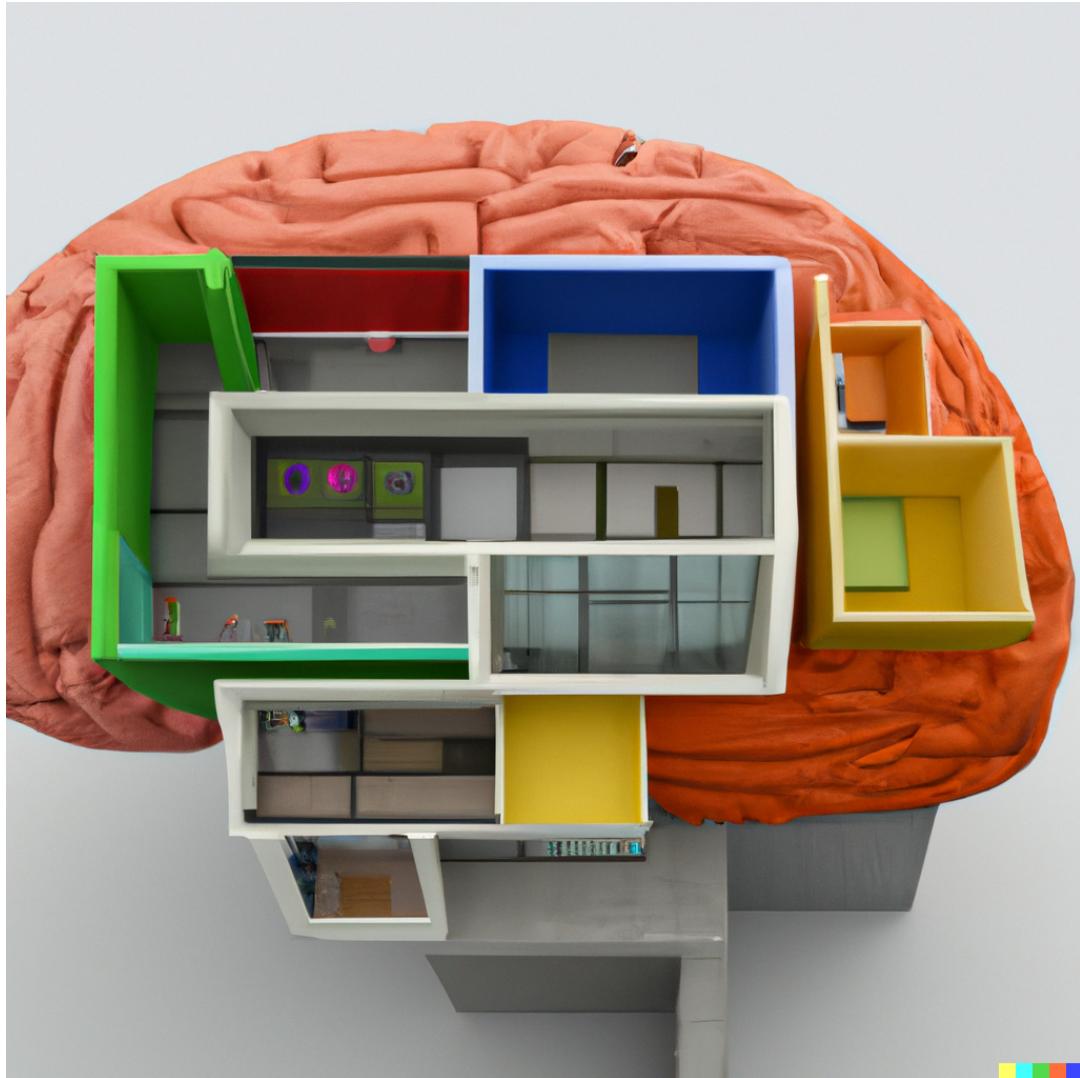


Light is “regulated” by switches.

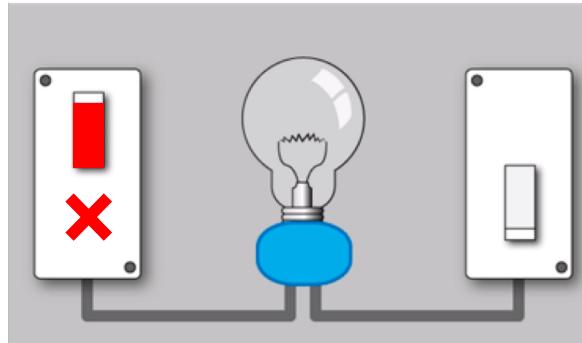


Preissl, Gaulton & Ren, Nat Rev Genetics, 2022

The gene regulation in glia and neurons is the key to complex functions



Created by DALL·E 2 - OpenAI



Schizophrenia



Bipolar Disorder



Depression



Brain tumor



Alzheimer's Disease

The burden of brain tumor and mental disorders



Brain tumor



Schizophrenia



Bipolar Disorder



Depression



Alzheimer's Disease

In 2021, **~83,500** individuals were diagnosed with brain and other CNS tumors in the US.

Average monthly cost: **~\$8,000**

Hard to treat and deliver drugs

In 2019, **1 in every 8 people**, or 970 million people around the world were living with a mental disorder.

Top 1 Burden of Diseases in the US.

20 years lived with disability.

Mental disorder, WHO, 2022
US Burden of Disease Collaborators. JAMA. 2013

Genetic variants and somatic mutations can lead to brain tumor and multiple mental disorders



Brain tumor



Schizophrenia



Bipolar Disorder



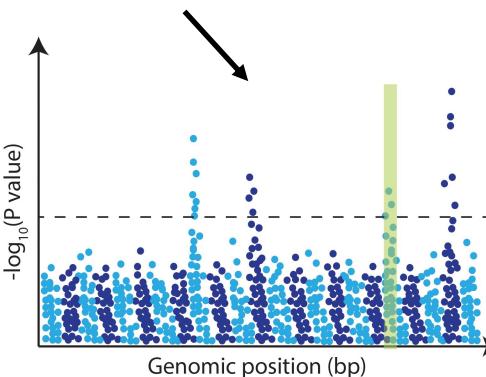
Depression



Alzheimer's Disease

Genome-wide association study
(GWAS)

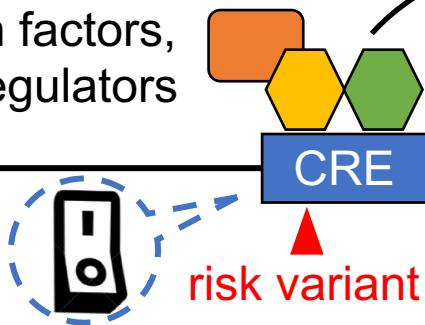
>90% of the variants are non-coding and yet to be functionally interpreted



A paradigm for understanding the noncoding disease risk variants, and the key barriers

How do risk variants affect transcription factor binding, and which ones?

Transcription factors,
Chromatin regulators



What target genes are controlled by the *cis* regulatory elements (CREs)?

Target gene



What sequences act as *cis* regulatory elements (CREs), where and when do they act?

It has been estimated that >50% of risk variants act in this way ([Nasser J, ... Engreitz, Nature, 2021](#))

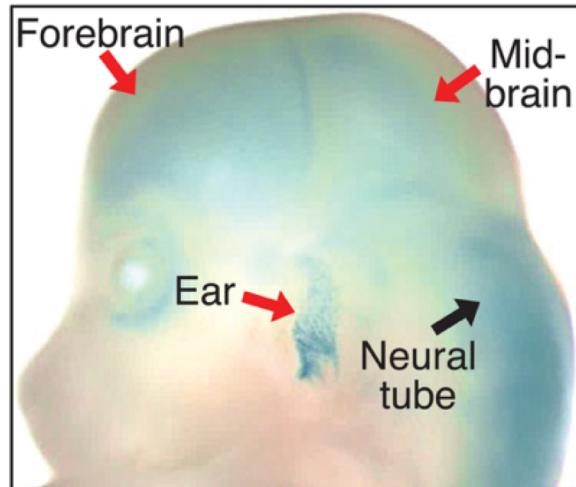


A noncoding single-nucleotide polymorphism at 8q24 drives *IDH1*-mutant glioma formation

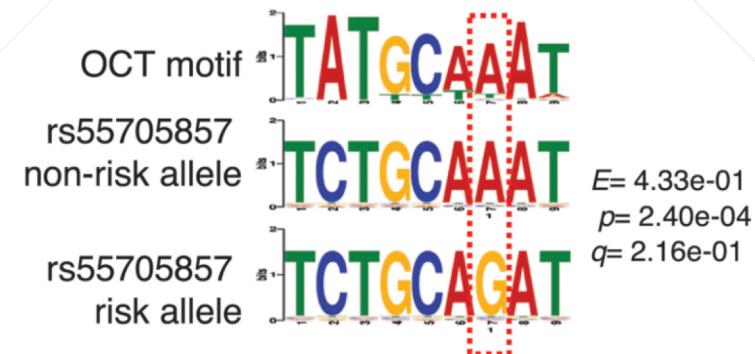
CONNOR YANCHUS , KRISTEN L. DRUCKER , THOMAS M. KOLLMAYER , RICKY TSAI , WARREN WINICK-NG , MINGGAO LIANG , AHMAD MALIK, JUDY PAWL-
ING , SILVANA B. DE LORENZO, [...], AND DANIEL SCHRAMEK +49 authors [Authors Info & Affiliations](#)

SCIENCE • 6 Oct 2022 • Vol 378, Issue 6615 • pp. 68-78 • DOI: 10.1126/science.abj2890

rs55705857 resides within a brain-specific enhancer

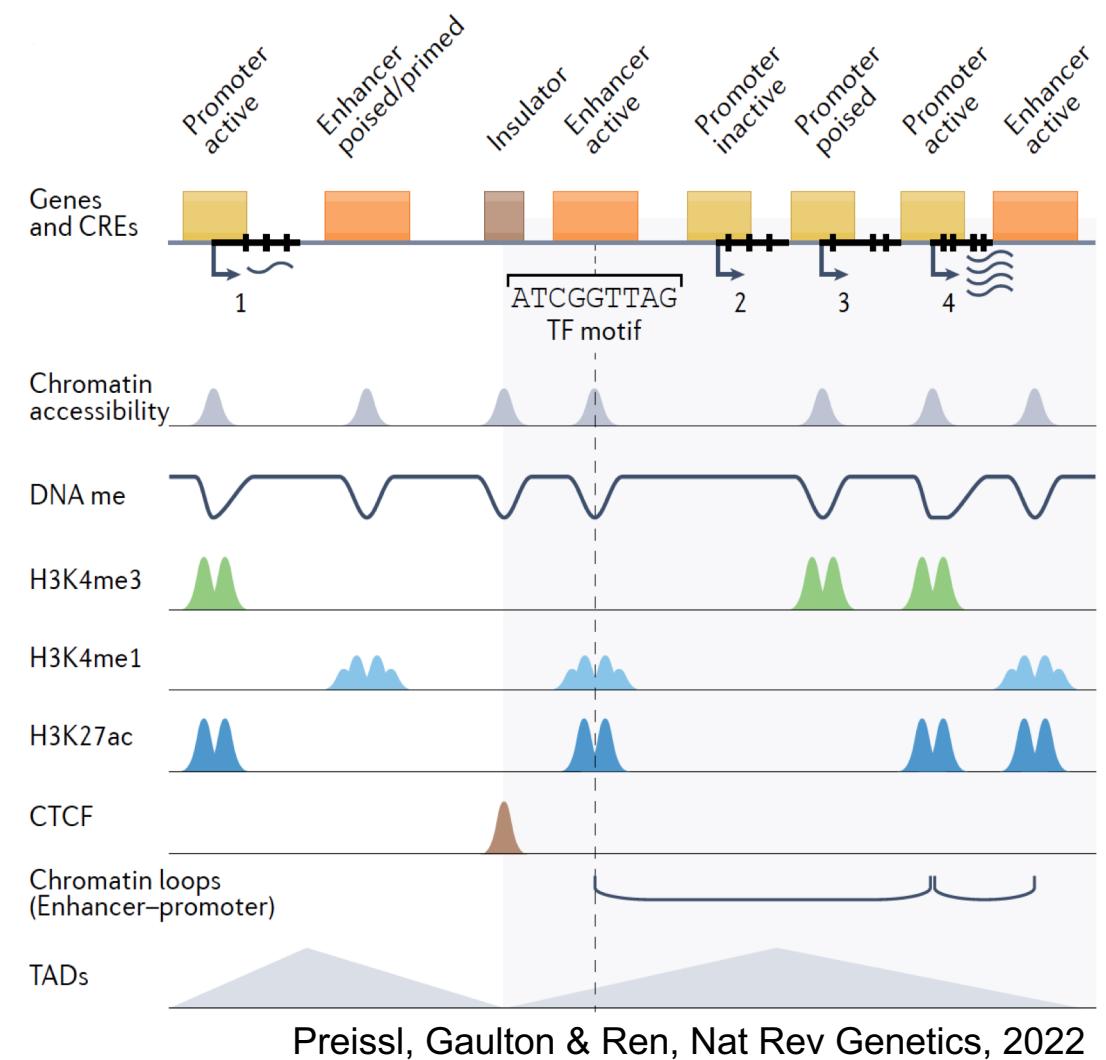
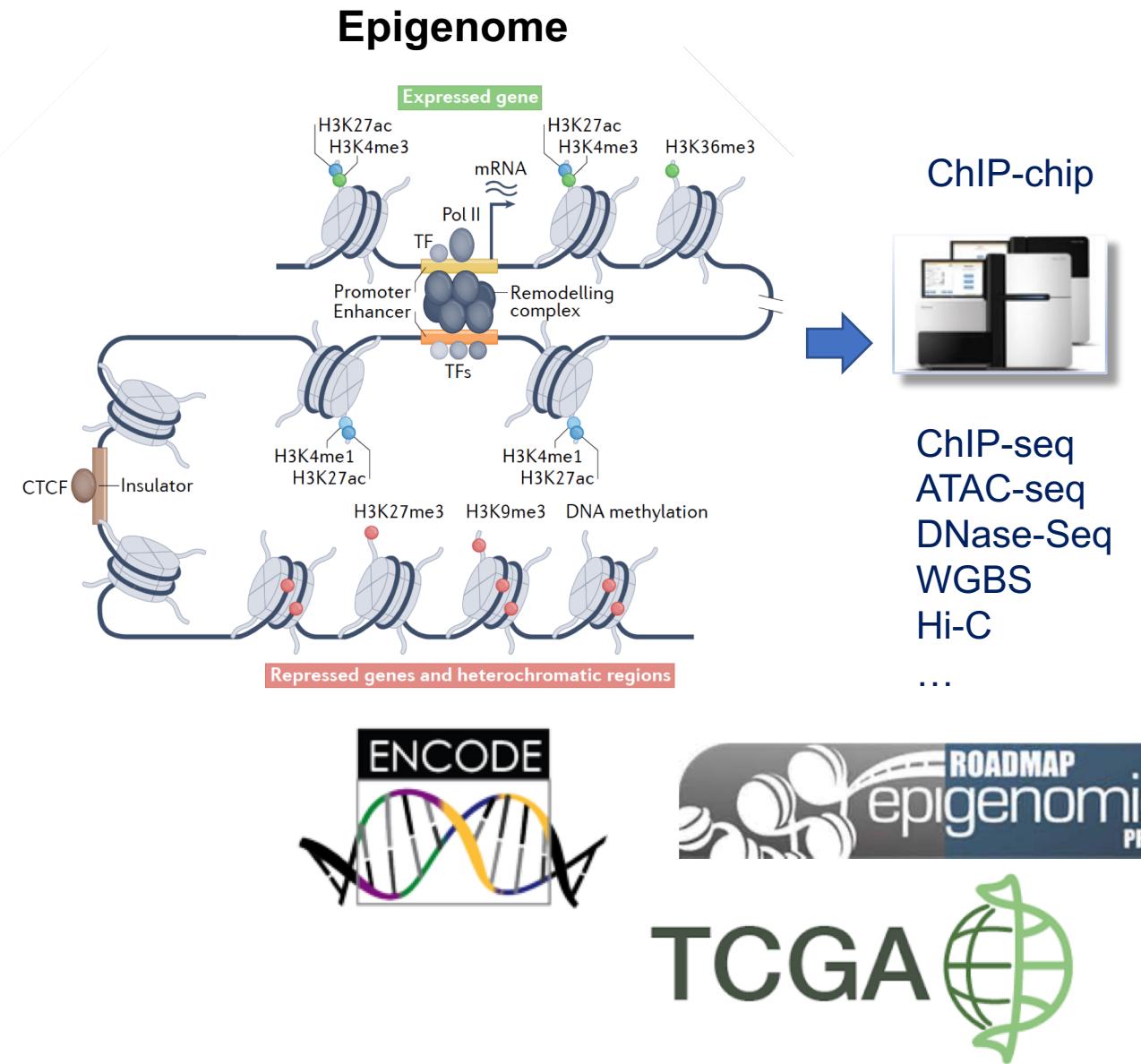


This risk variant disrupts OCT2/4 binding



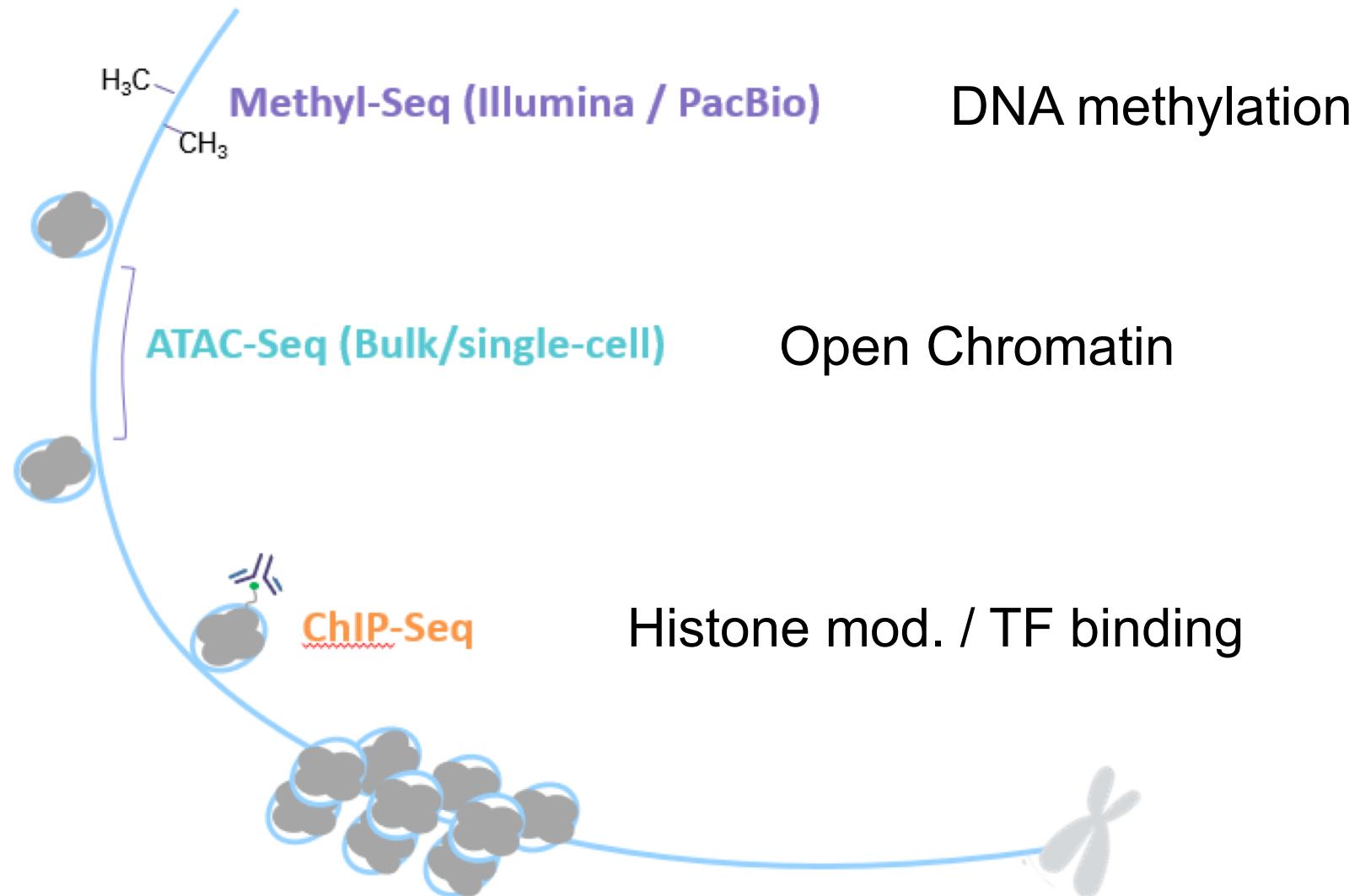
..., allowing increased interaction with the Myc promoter and increased Myc expression in ~40% of *IDH*-mutant low-grade glioma (LGG) patients

Epigenomic profiling reveals genome regulation

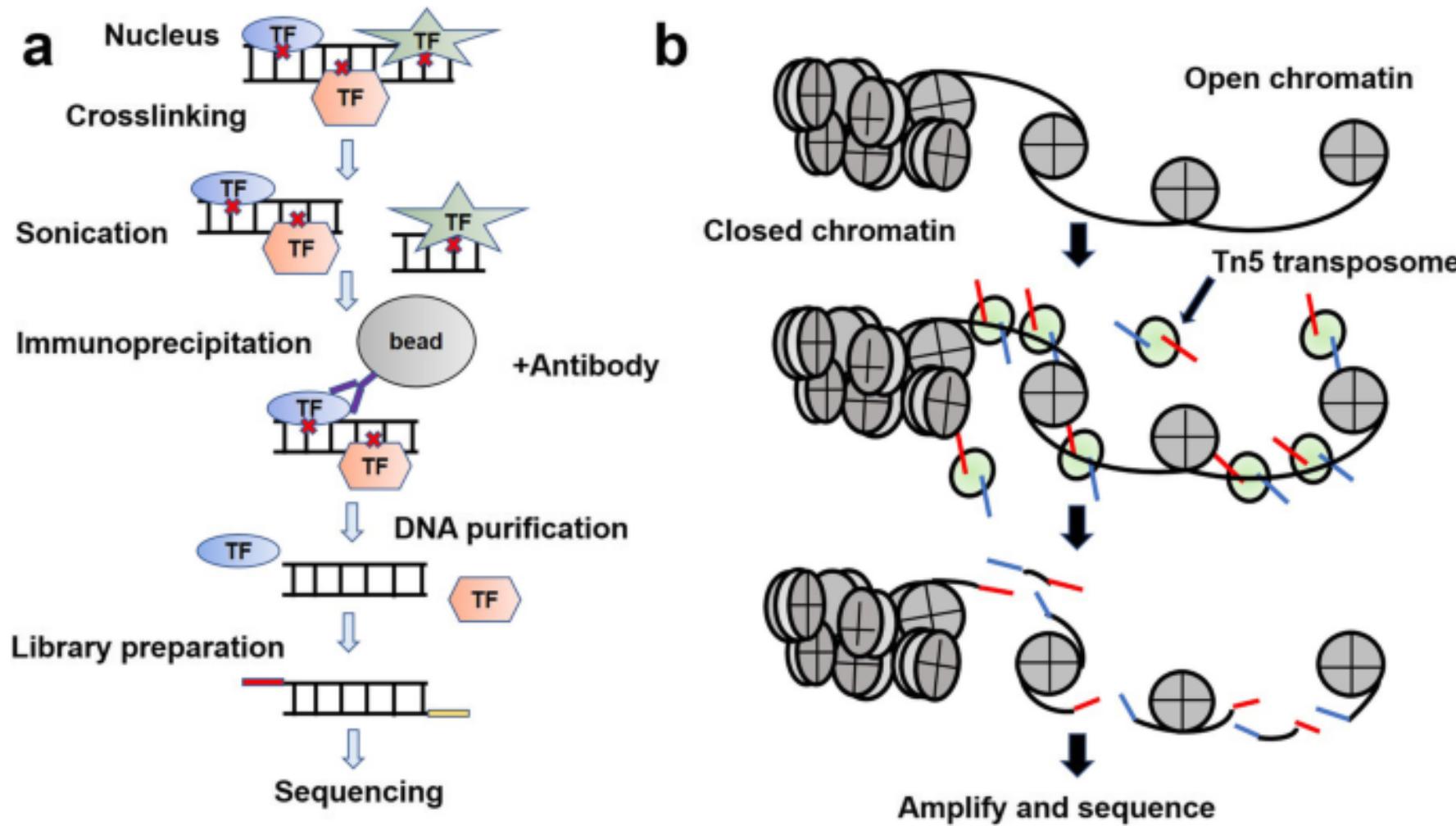


Preissl, Gaulton & Ren, Nat Rev Genetics, 2022

The difference between genomic assays



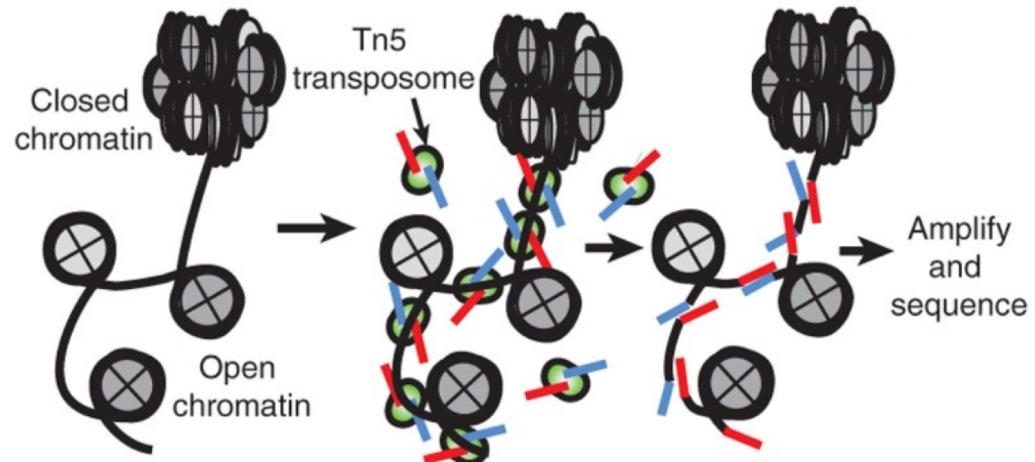
The difference between ChIP-seq and ATAC-seq



ATAC-seq identifies accessible DNA regions by probing **open chromatin** with hyperactive mutant **Tn5 Transposase** that inserts sequencing adapters into open regions of the genome.

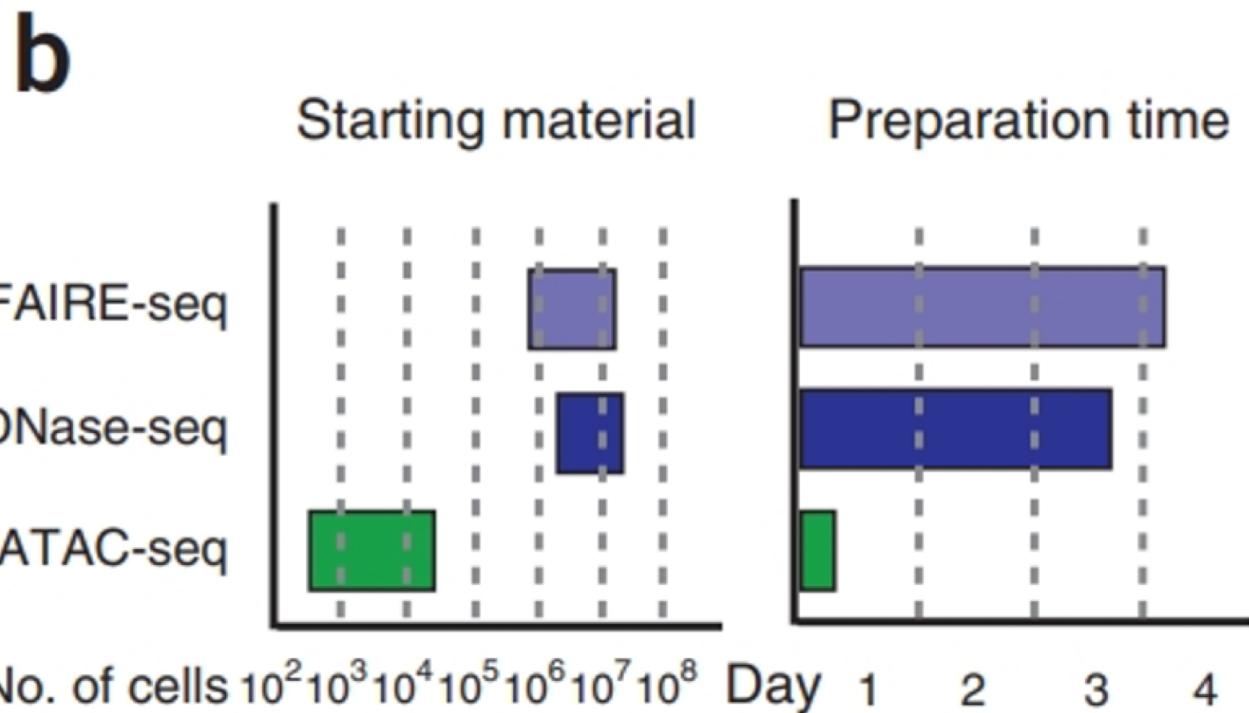
Buenrostro, et al. Nature Methods, 2013

The advantage of ATAC-seq

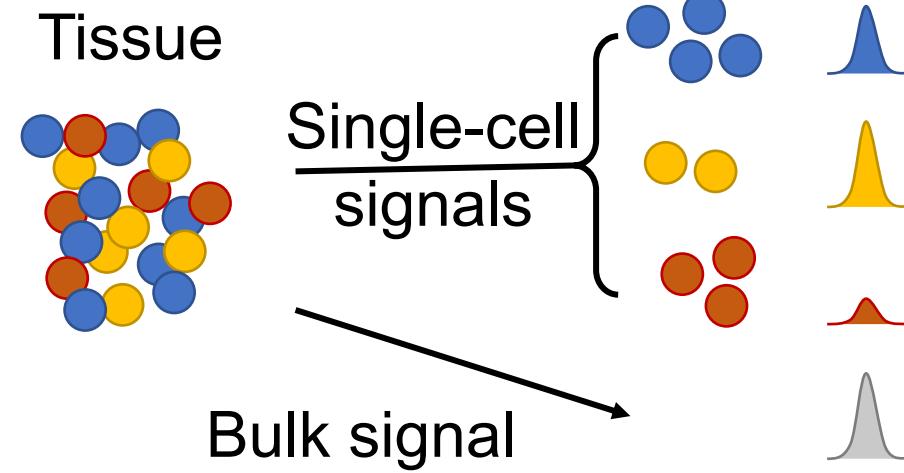
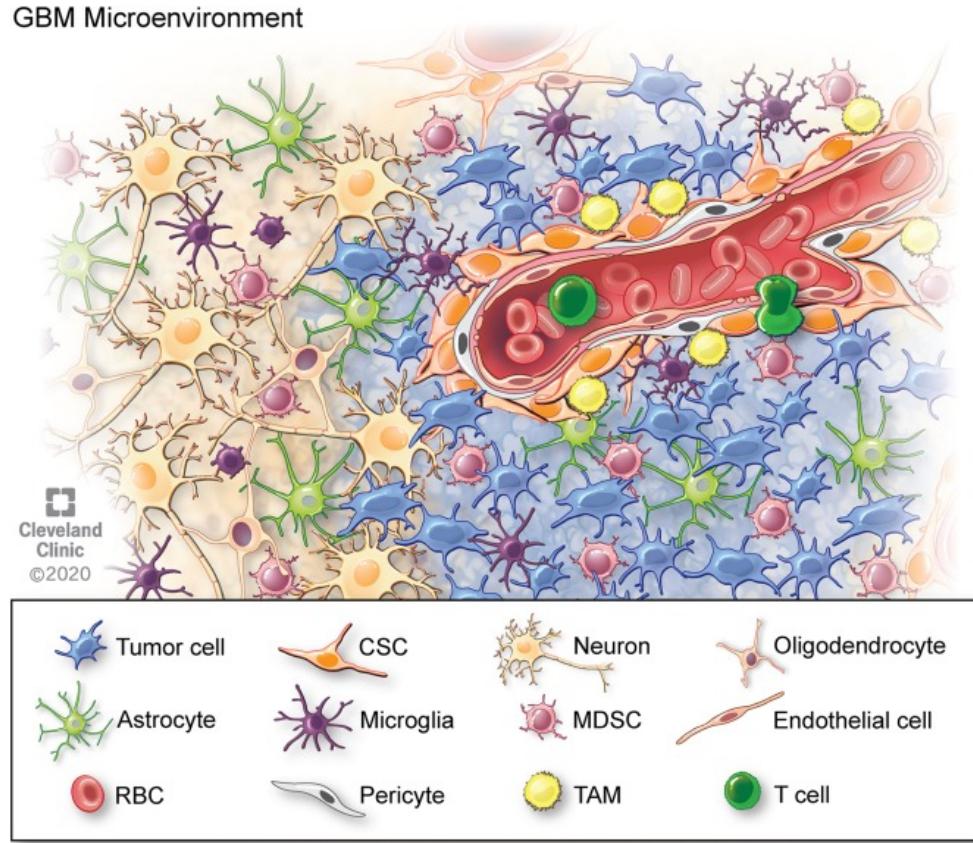


ATAC-seq identifies accessible DNA regions by probing **open chromatin** with hyperactive mutant **Tn5 Transposase** that inserts sequencing adapters into open regions of the genome.

Buenrostro, et al. Nature Methods, 2013



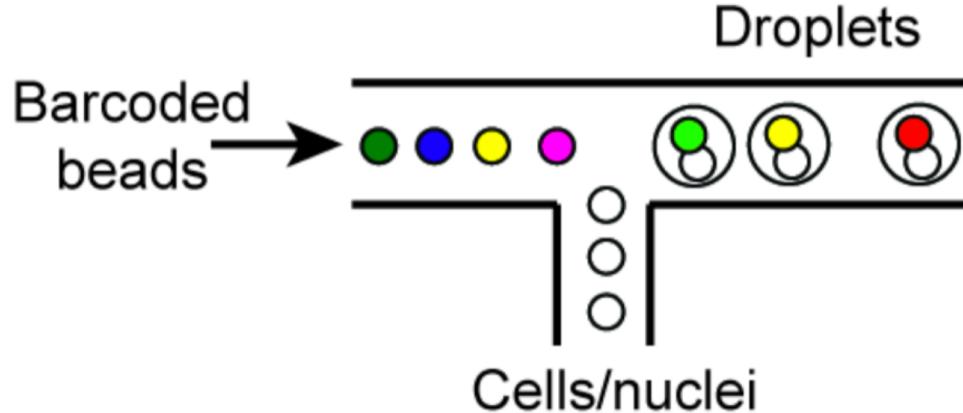
Single cell technologies solve cell heterogeneity



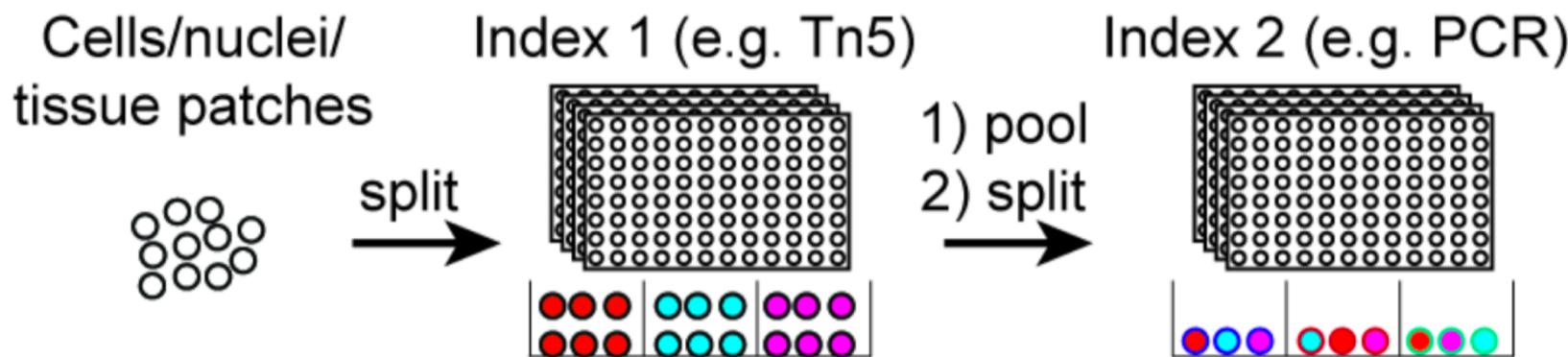
Lauko, et al. 2022

Single-nucleus ATAC-seq – different platforms

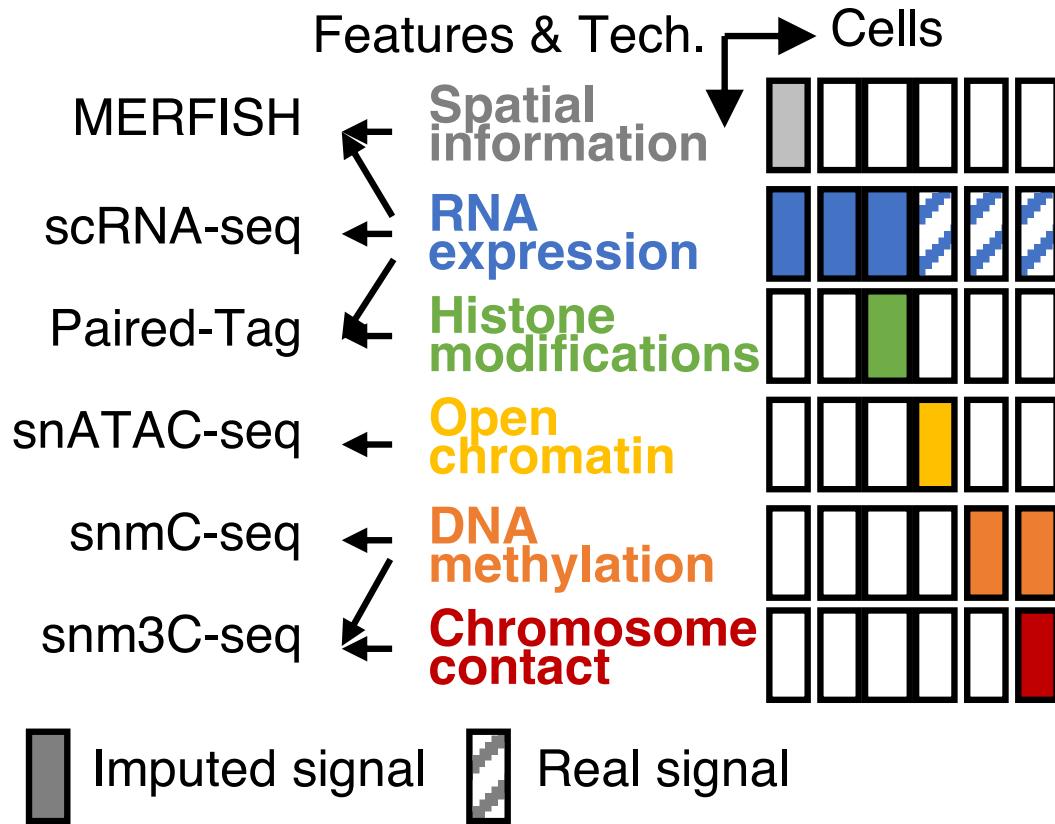
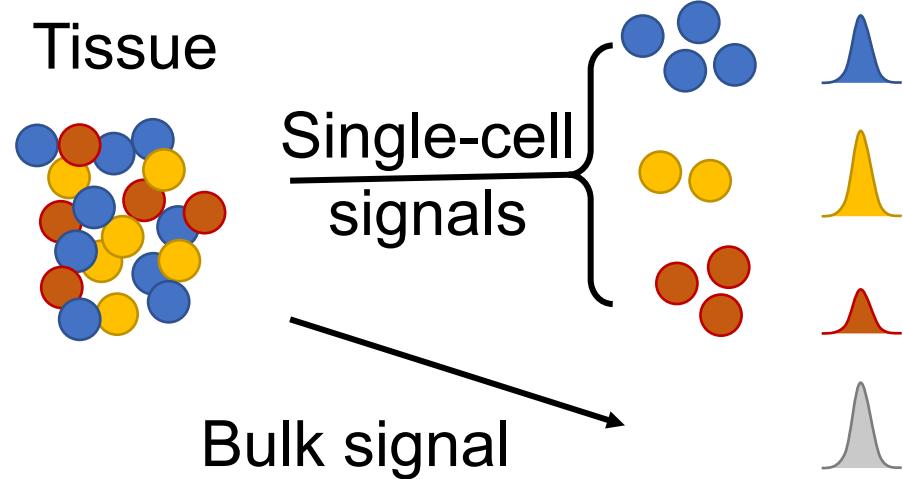
□ Droplet-based (e.g. 10x Genomics)



□ Split-pool-barcoding/single cell combinatorial indexing (sci-)



Single cell technologies solve cell heterogeneity

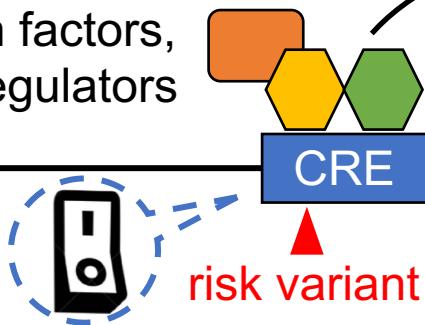


Technologies have been developed to map DNA methylation (**snmC-seq**), chromatin accessibility (**snATAC-seq**), histone modifications (**Paired-Tag**), 3D genome conformation (**snm3C-seq**), either alone or in combination with RNA-seq, in single cells.

A paradigm for understanding the noncoding disease risk variants, and the key barriers

How do risk variants affect transcription factor binding, and which ones?

Transcription factors,
Chromatin regulators



What target genes are controlled by the *cis* regulatory elements (CREs)?

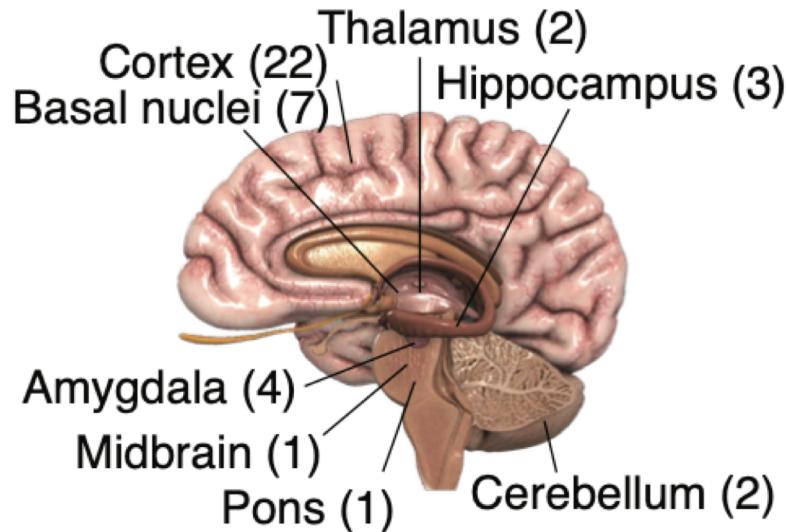
Target gene



What sequences act as *cis* regulatory elements (CREs), where and when do they act?

It has been estimated that >50% of risk variants act in this way ([Nasser J, ... Engreitz, Nature, 2021](#))

Single cell chromatin accessibility atlas of adult human brains



- 42 brain regions**
- 3 neurotypical adult donors**
- 1.1 million cells, 107 cell types**



Ed Lein, Ph.D.
(Allen Inst.)



Sebastian Preissl,
Ph.D.

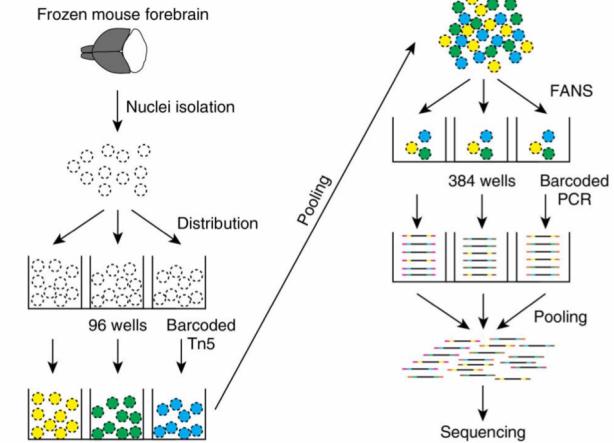


Marga Behrens,
Ph.D. (Salk)



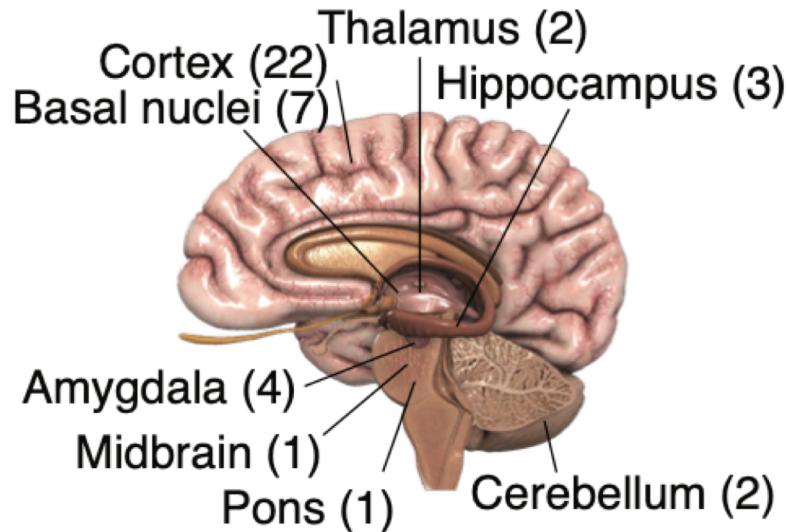
Joe Ecker, Ph.D.
(Salk/HHMI)

snATAC-seq



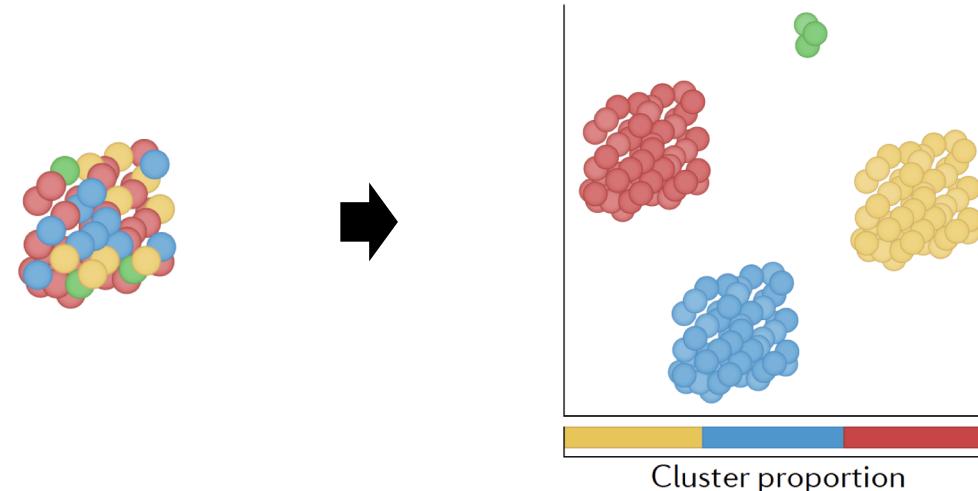
Preissl, et al. 2018

Single cell chromatin accessibility atlas of adult human brains

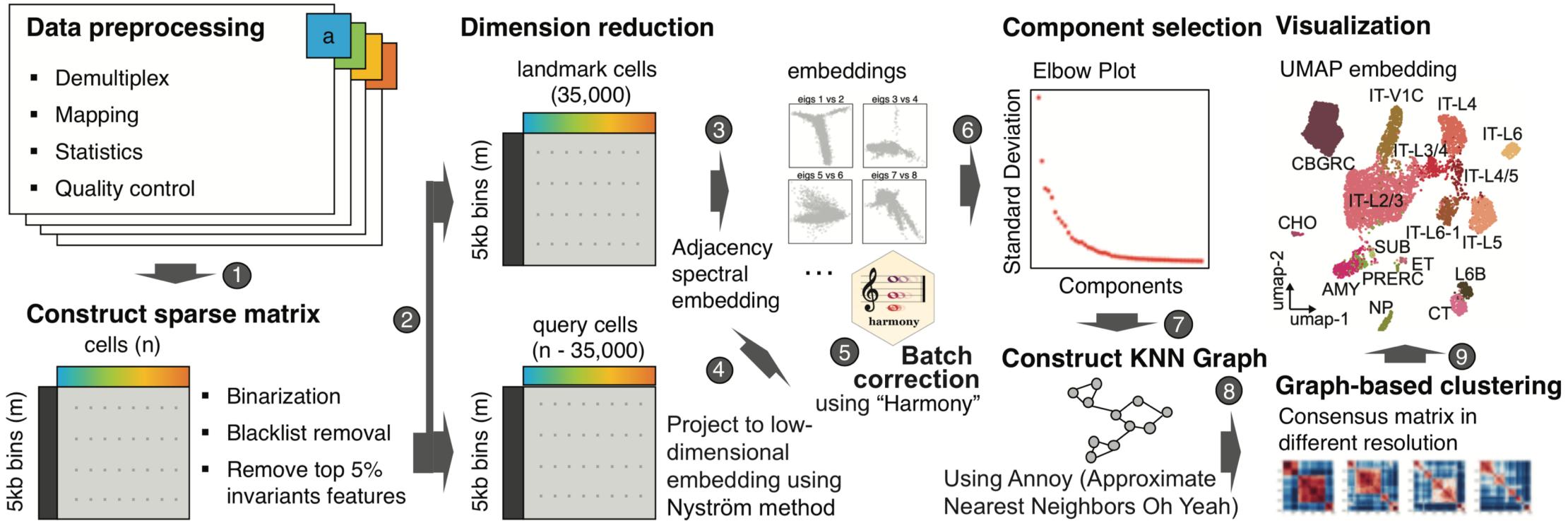


- 42 brain regions
- 3 neurotypical adult donors
- 1.1 million cells, 107 cell types

Single cell clustering

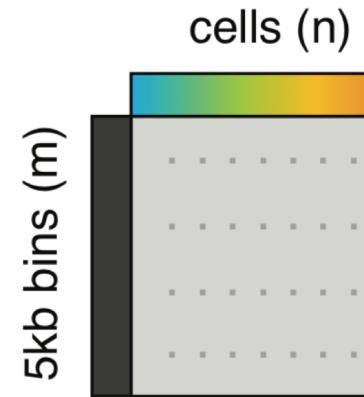


Cell clustering pipeline



Challenges in clustering extremely large sparse matrix

- ☐ Million cells
- ☐ Signals on whole genome
- ☐ Very sparse

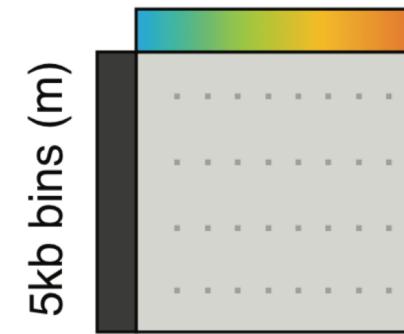


- ☐ Memory issues
- ☐ Time consuming

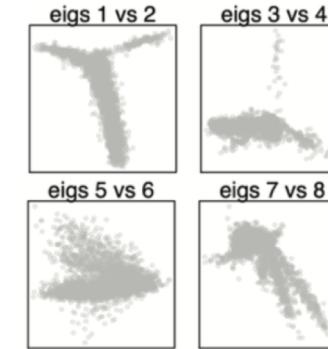


Dr. Rongxin Fang Dr. Kai Zhang

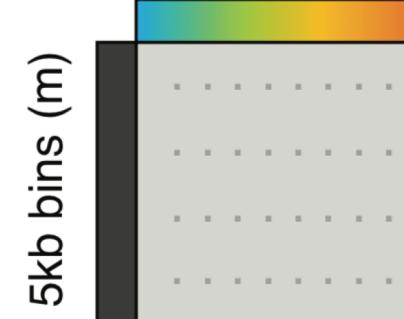
Landmark cells



Dimension reduction



Query cells



Nyström algorithm
an efficient technique
to generate low-rank
matrix approximations

Challenges in identifying robust cell clusters

Resolution parameters in community detection algorithm (Leiden)

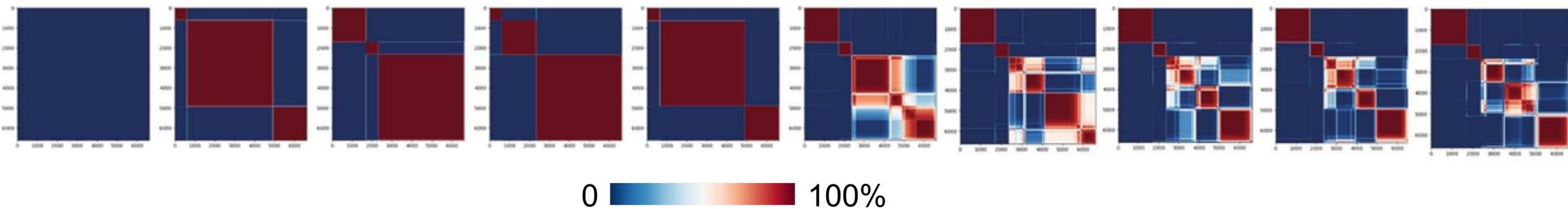
0.1

1.0

UMAP embedding by cluster

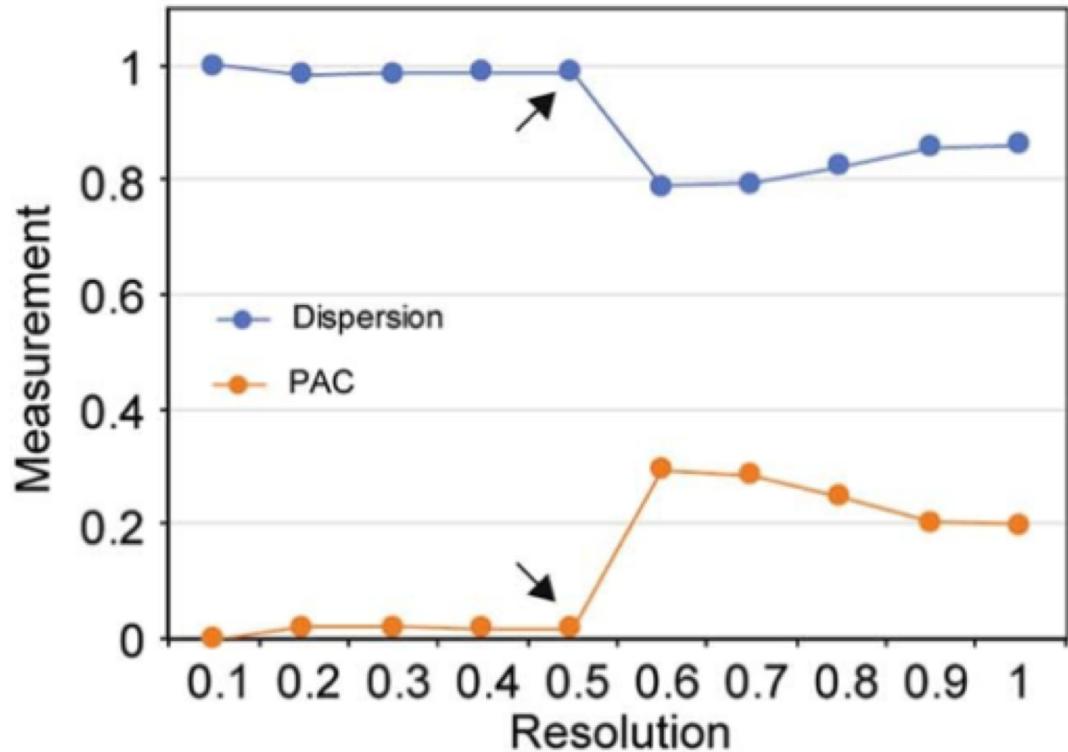


Generate consensus matrix from iterative runs



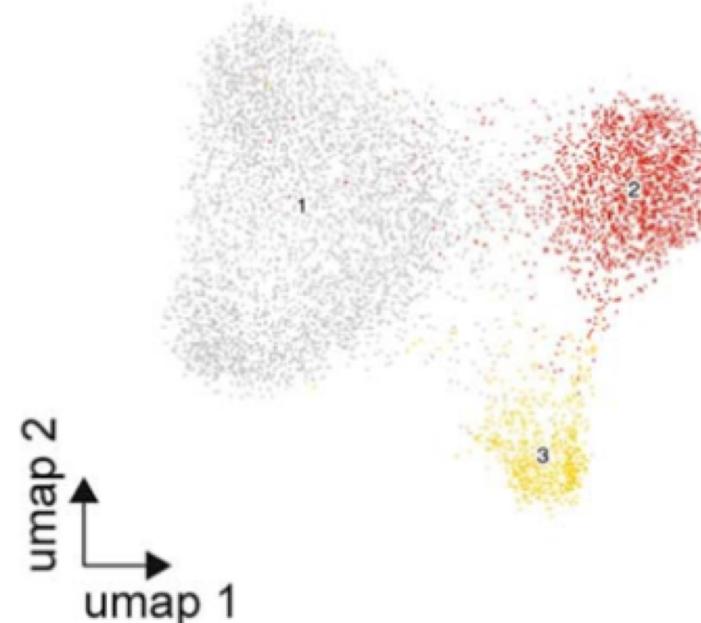
Challenges in identifying robust cell clusters

Selecting the most robust resolution

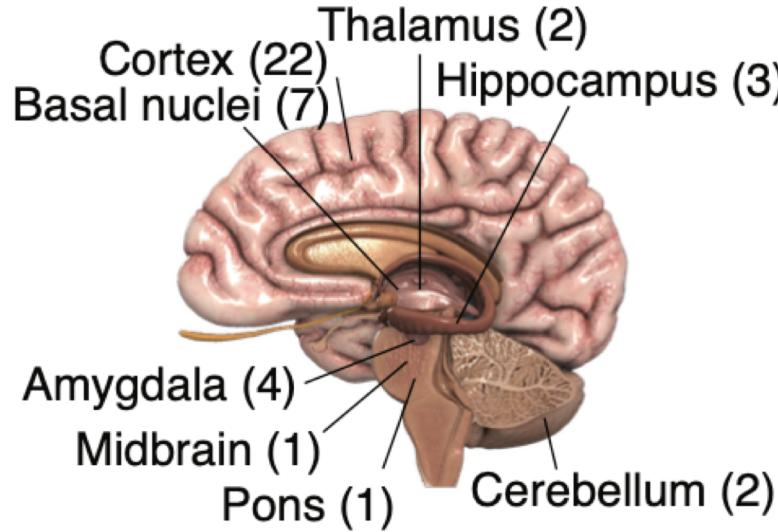


Dispersion: dispersion coefficient

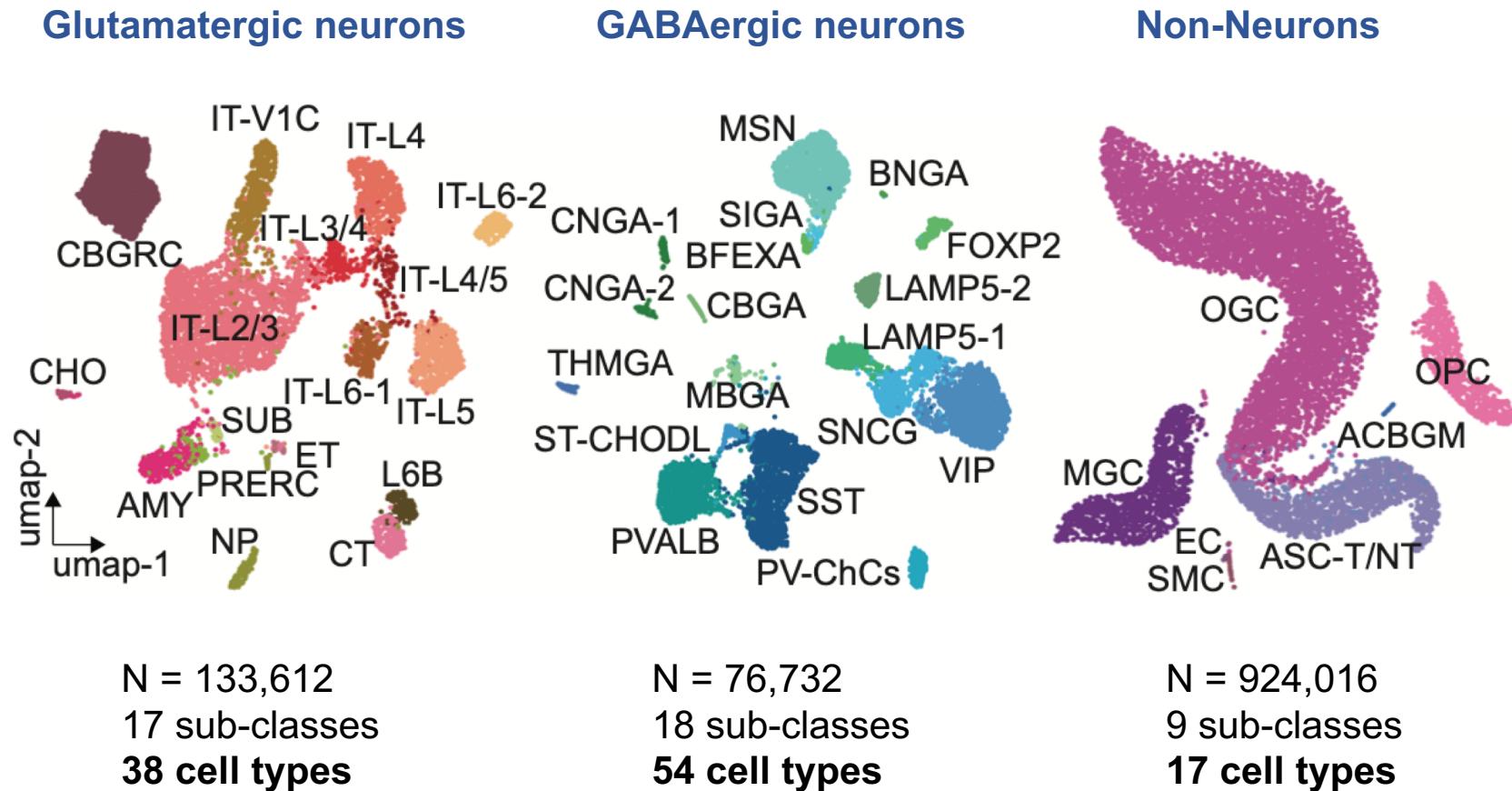
PAC: proportion of ambiguous clustering score



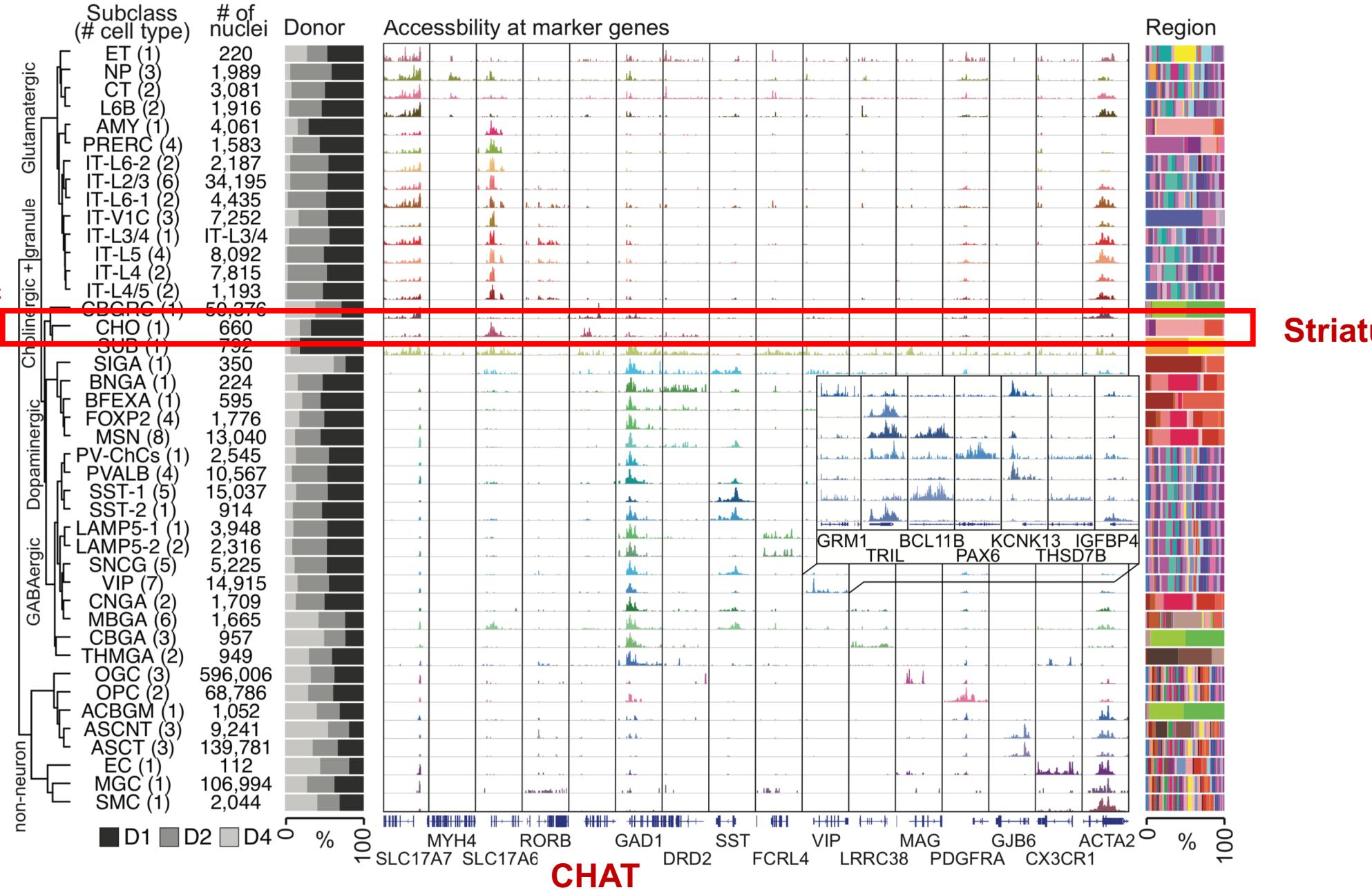
42 cell subclasses and 107 cell types identified from >1.1 million single nuclei in adult human brain



- 42 brain regions
- 3 adult donors
- 1.1 million cells, 107 cell types

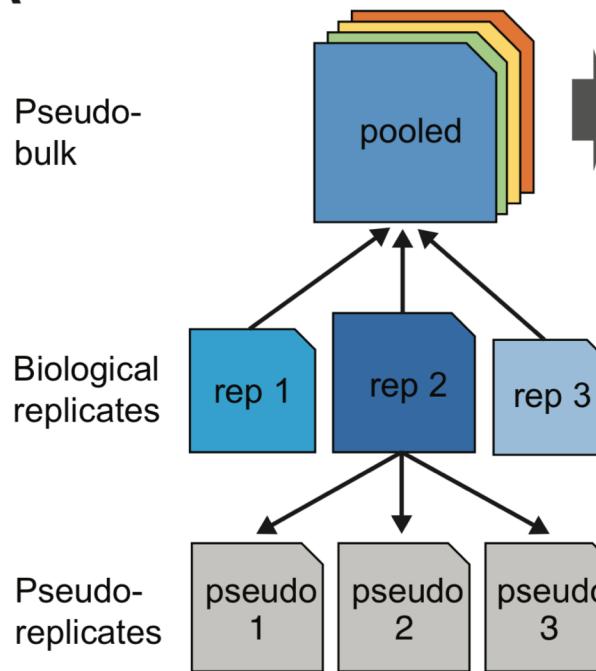


Cholinergic neurons (CHO)

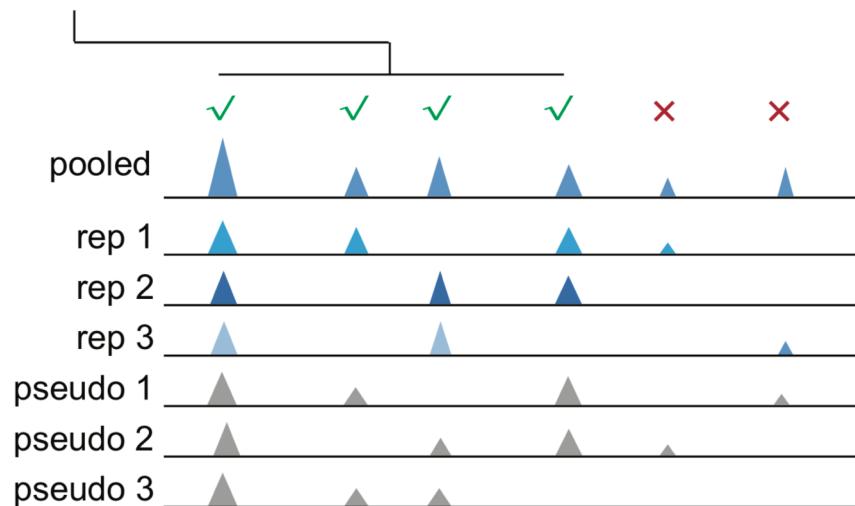


Peak calling pipeline

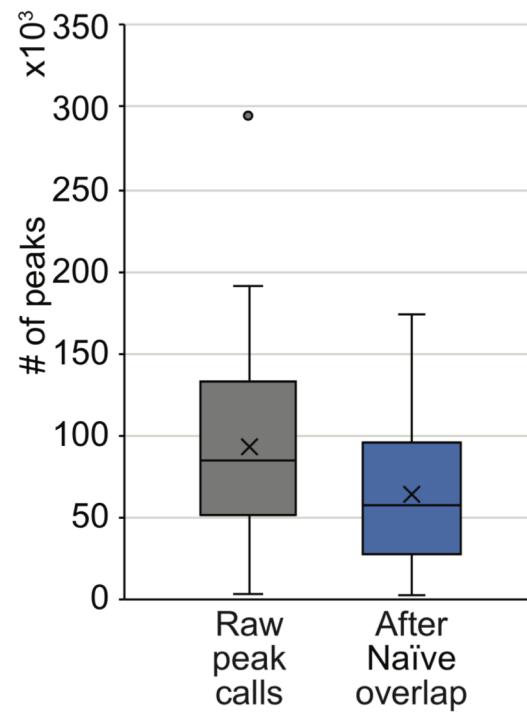
A



Naïve overlap

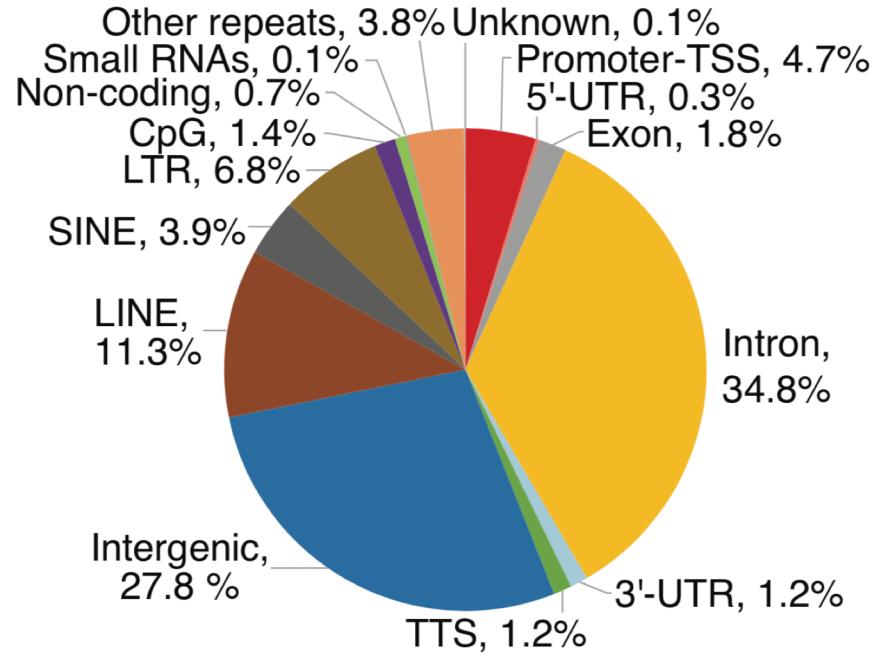


B

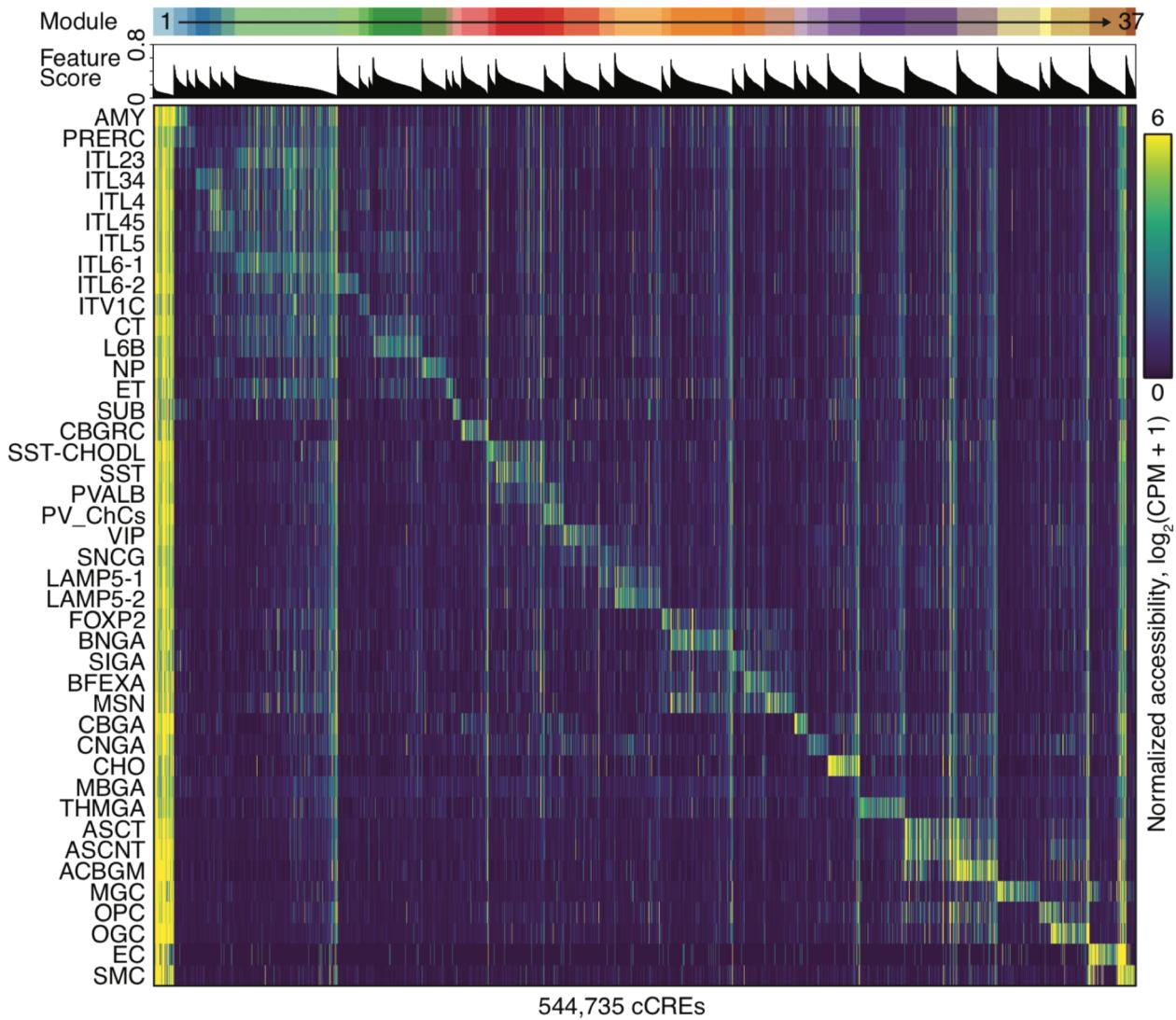


Brain cCREs display highly variable chromatin accessibility

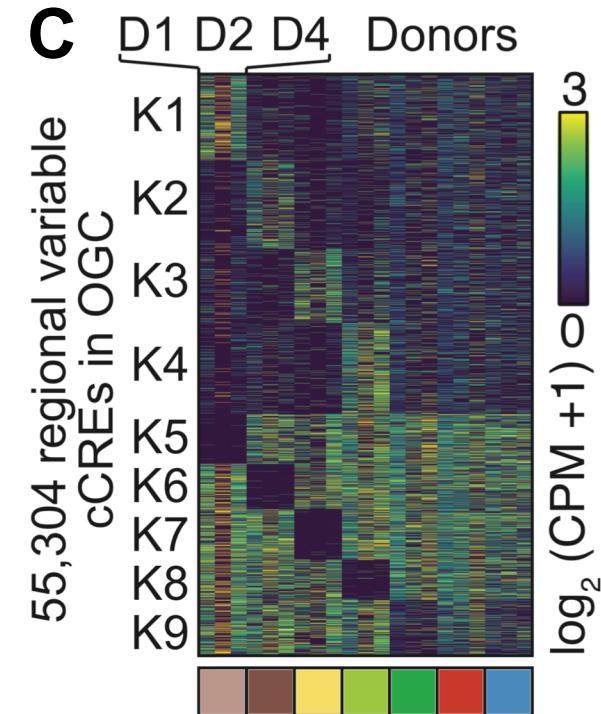
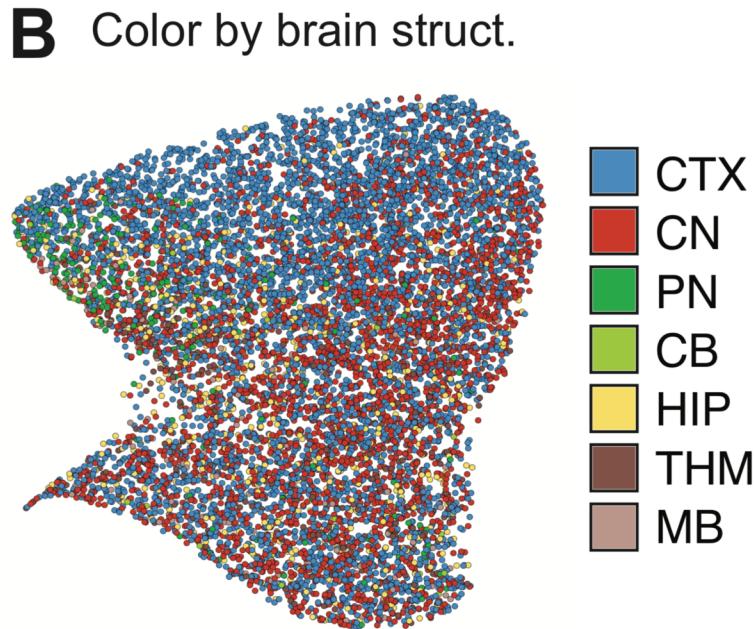
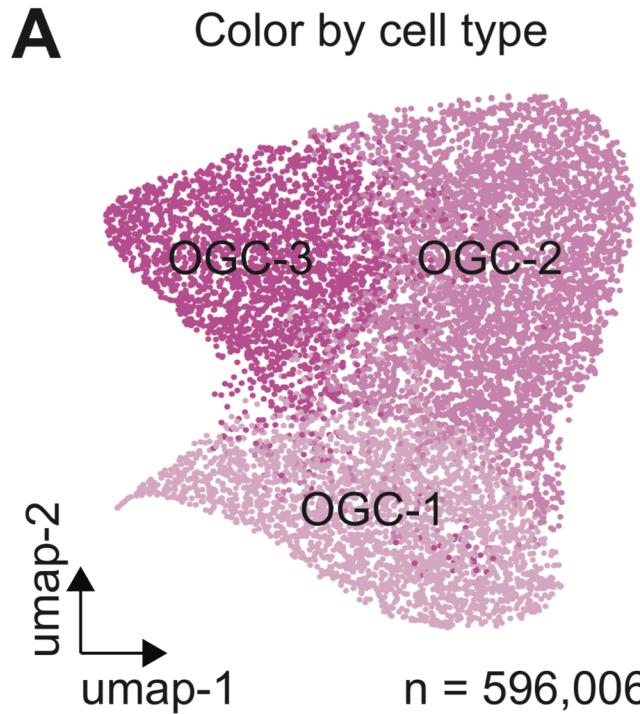
544,735 candidate cis-regulatory elements (cCREs)



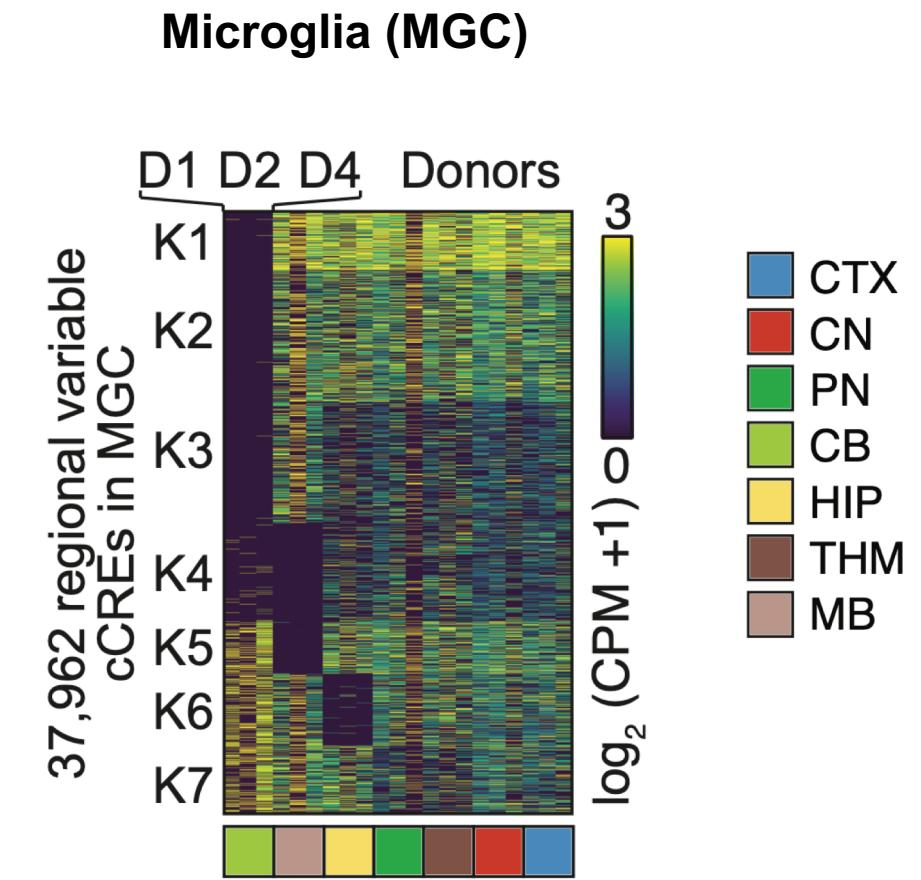
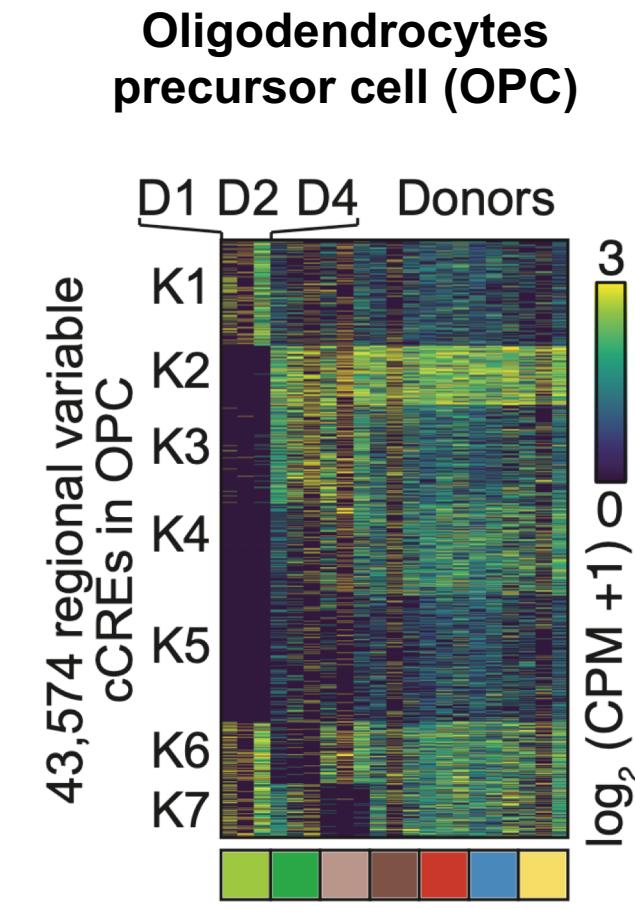
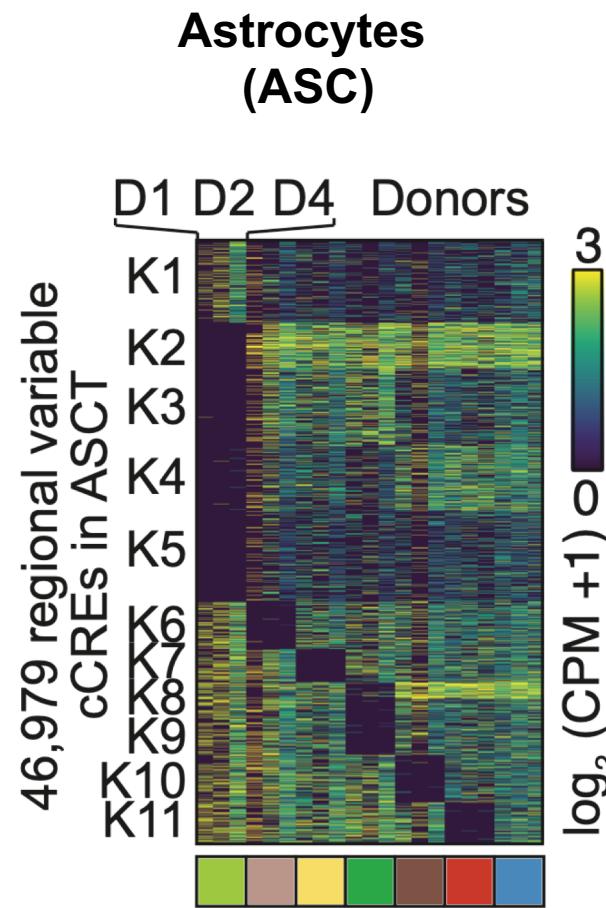
~ 46.9% are new cCREs comparing to ENCODE



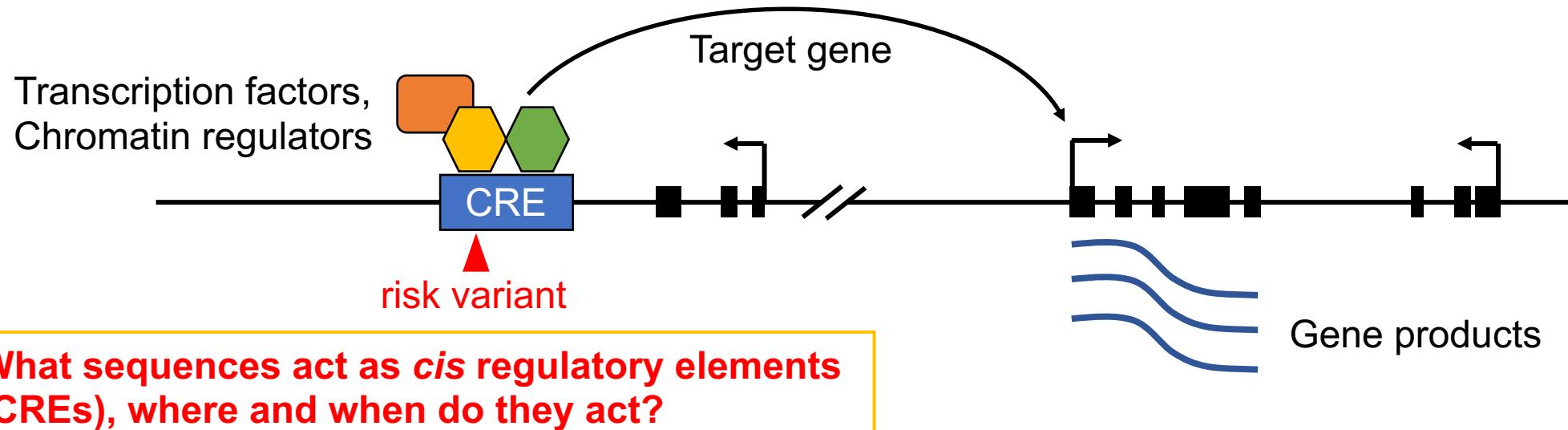
Regional specificity of Oligodendrocytes



33-40% of cCREs in glial cells show variations across brain regions

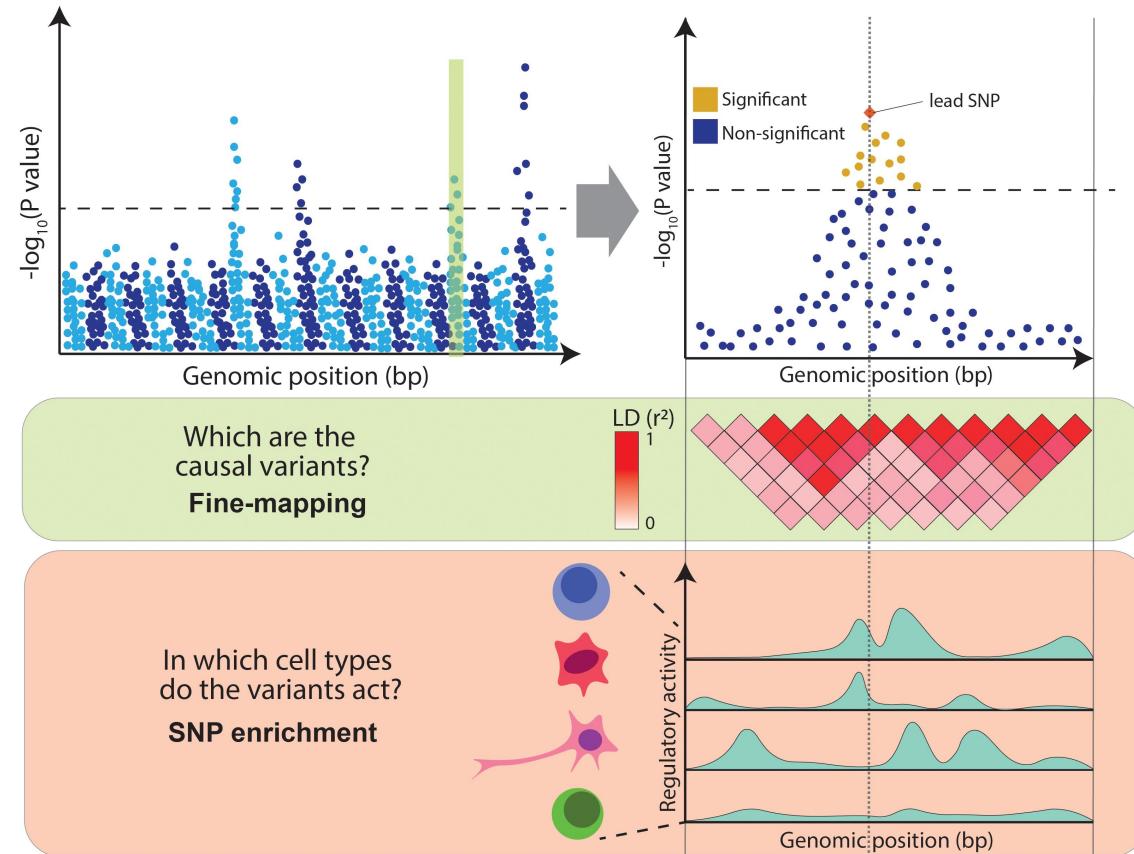


A paradigm for understanding the noncoding disease risk variants, and the key barriers



It has been estimated that >50% of risk variants act in this way (Nasser J, ... Engreitz, Nature, 2021)

Predicting Disease Relevant Cell Type(s)



Transcription factors,
Chromatin regulators

CRE
risk variant

Target gene

Gene products

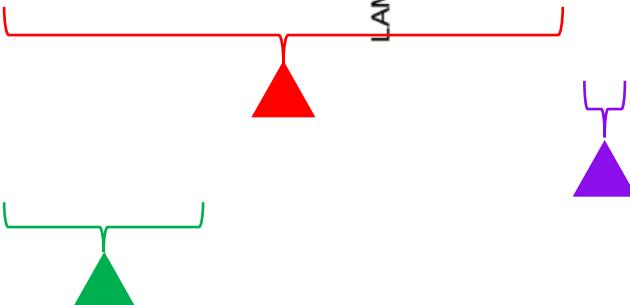
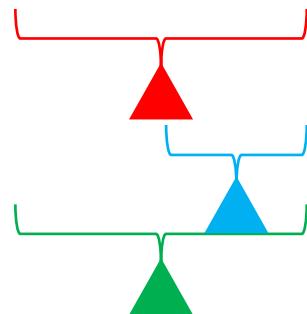
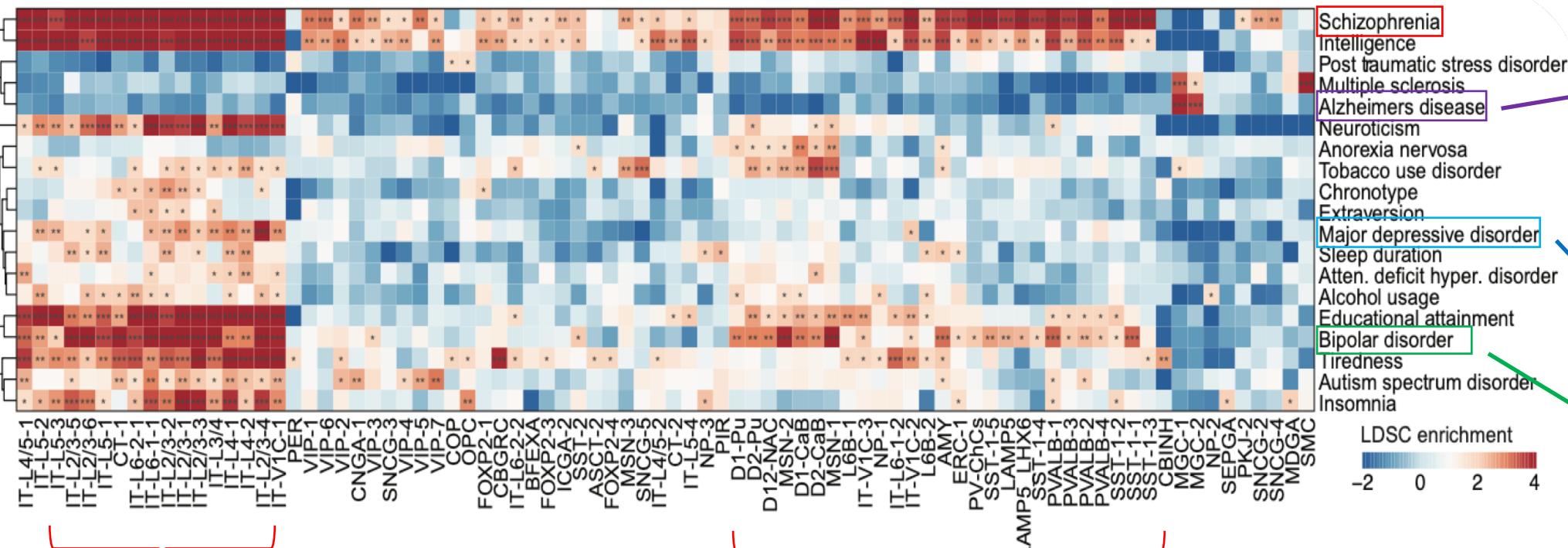
Disease risk variants are enriched in
enhancer of disease relevant cell type(s)?

Cano-Gamez and Gosia Trnka, Front. Genet., 13 May 2020

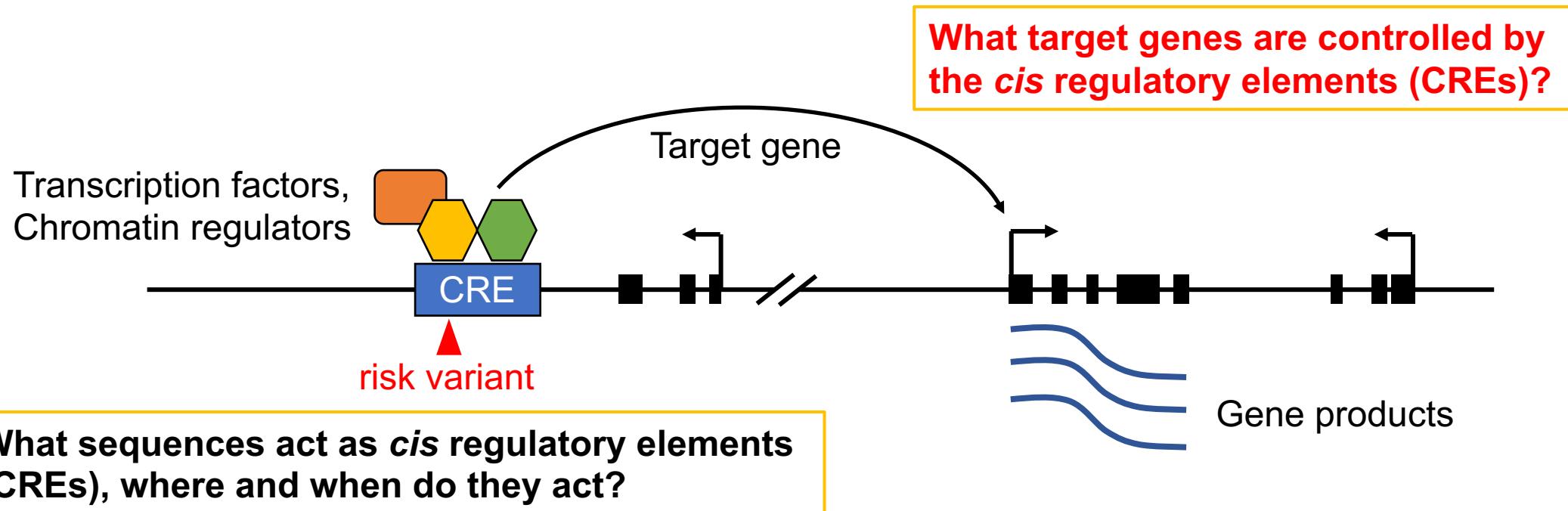
Cell-type resolved chromatin landscapes help interpret non-coding variants in mental diseases



Linkage disequilibrium score regression (LDSC) analysis



A paradigm for understanding the noncoding disease risk variants, and the key barriers

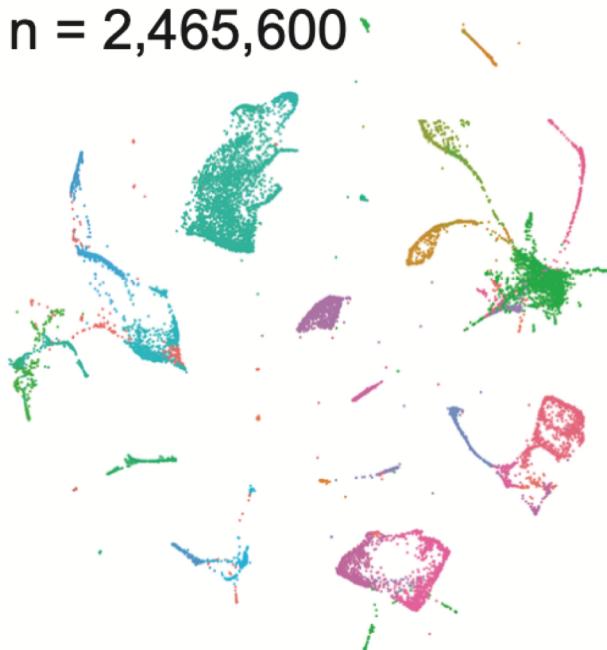


It has been estimated that >50% of risk variants act in this way (Nasser J, ... Engreitz, Nature, 2021)

Integration with snRNA-seq

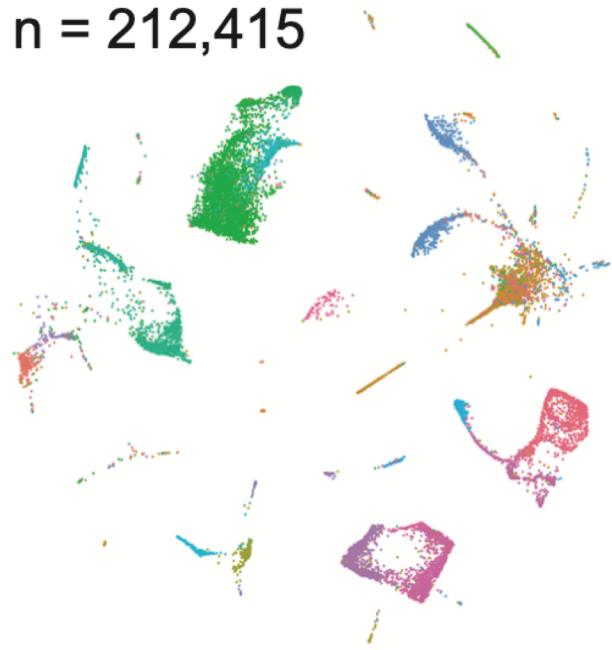
Colored by subclasses
from snRNA-seq

n = 2,465,600

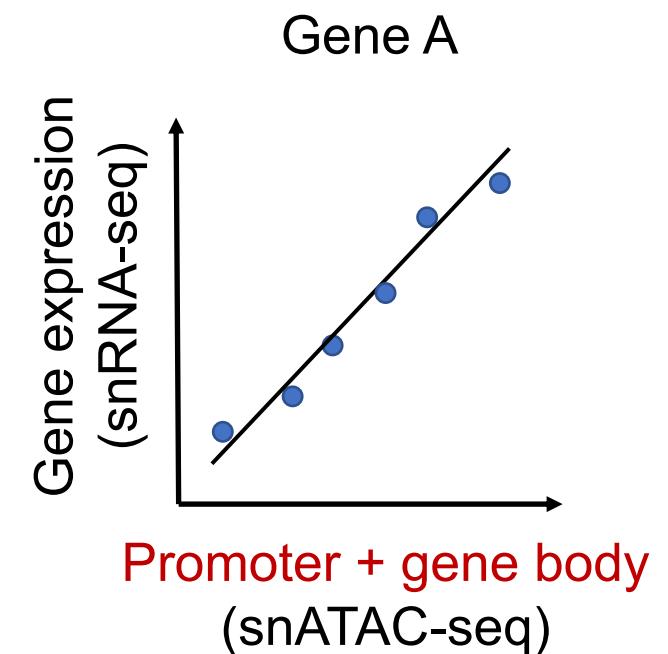


Colored by subclasses
from snATAC-seq

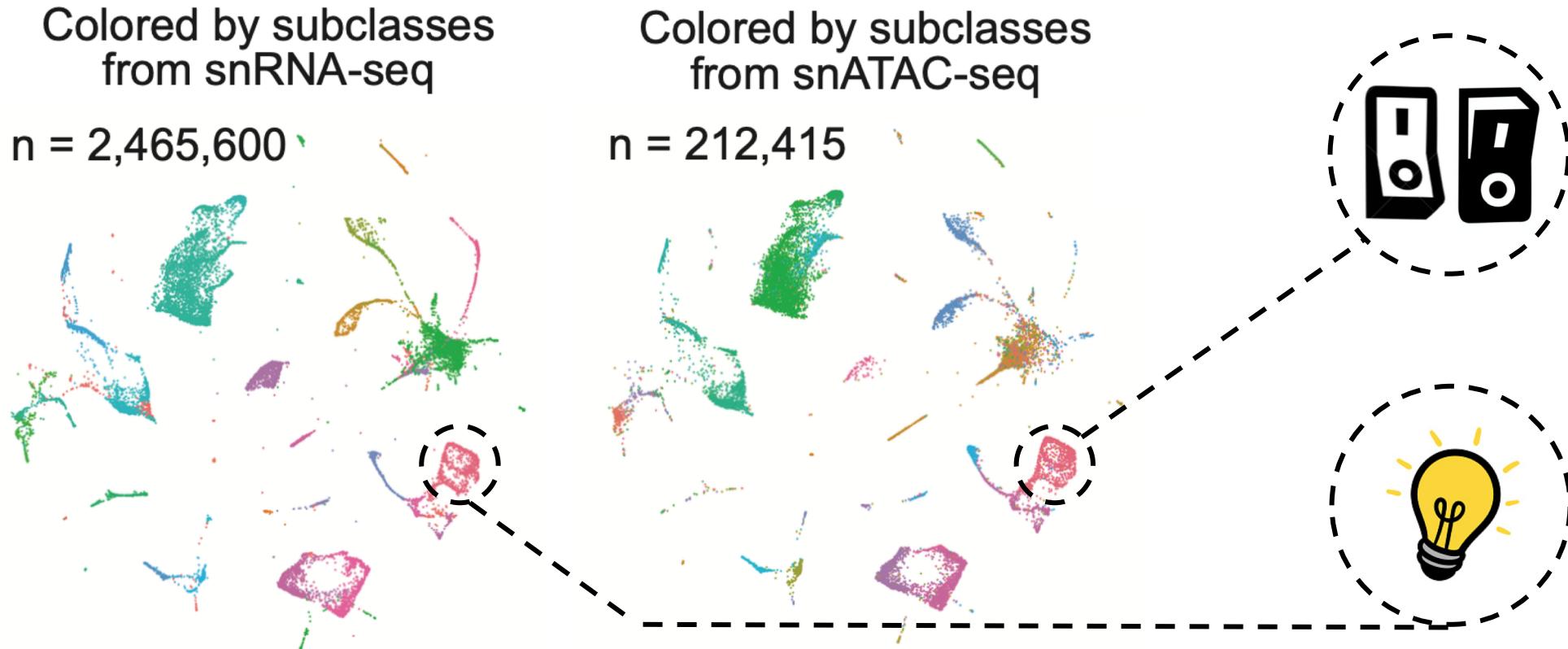
n = 212,415



scRNA-seq data: Linnarsson lab



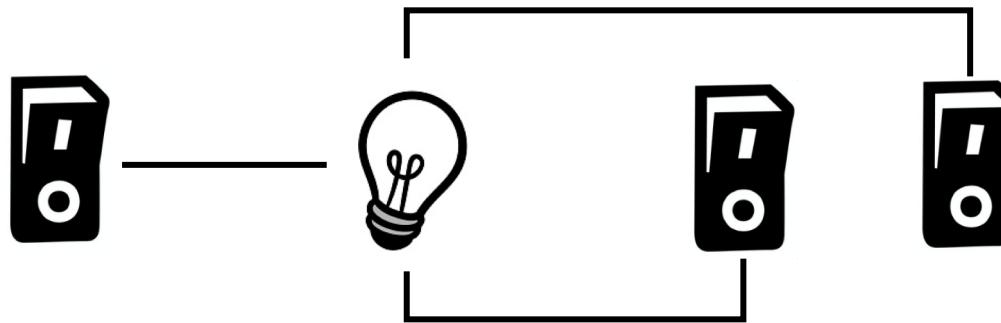
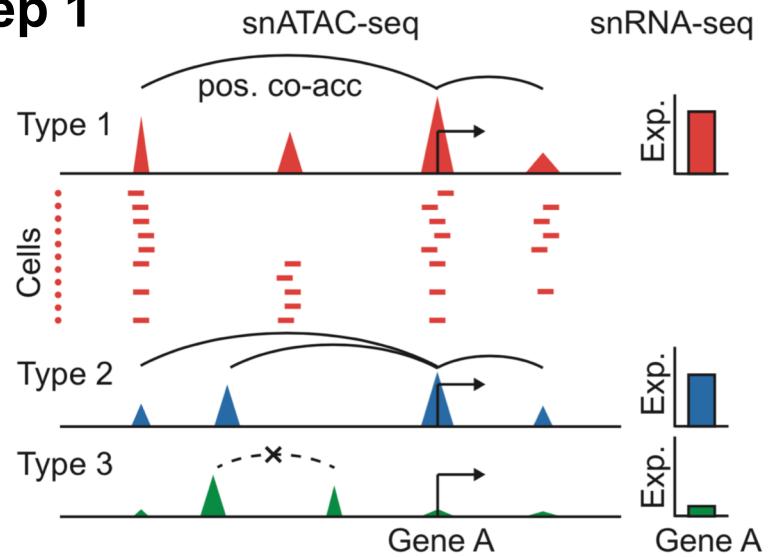
Integration with snRNA-seq



scRNA-seq data: Linnarsson lab

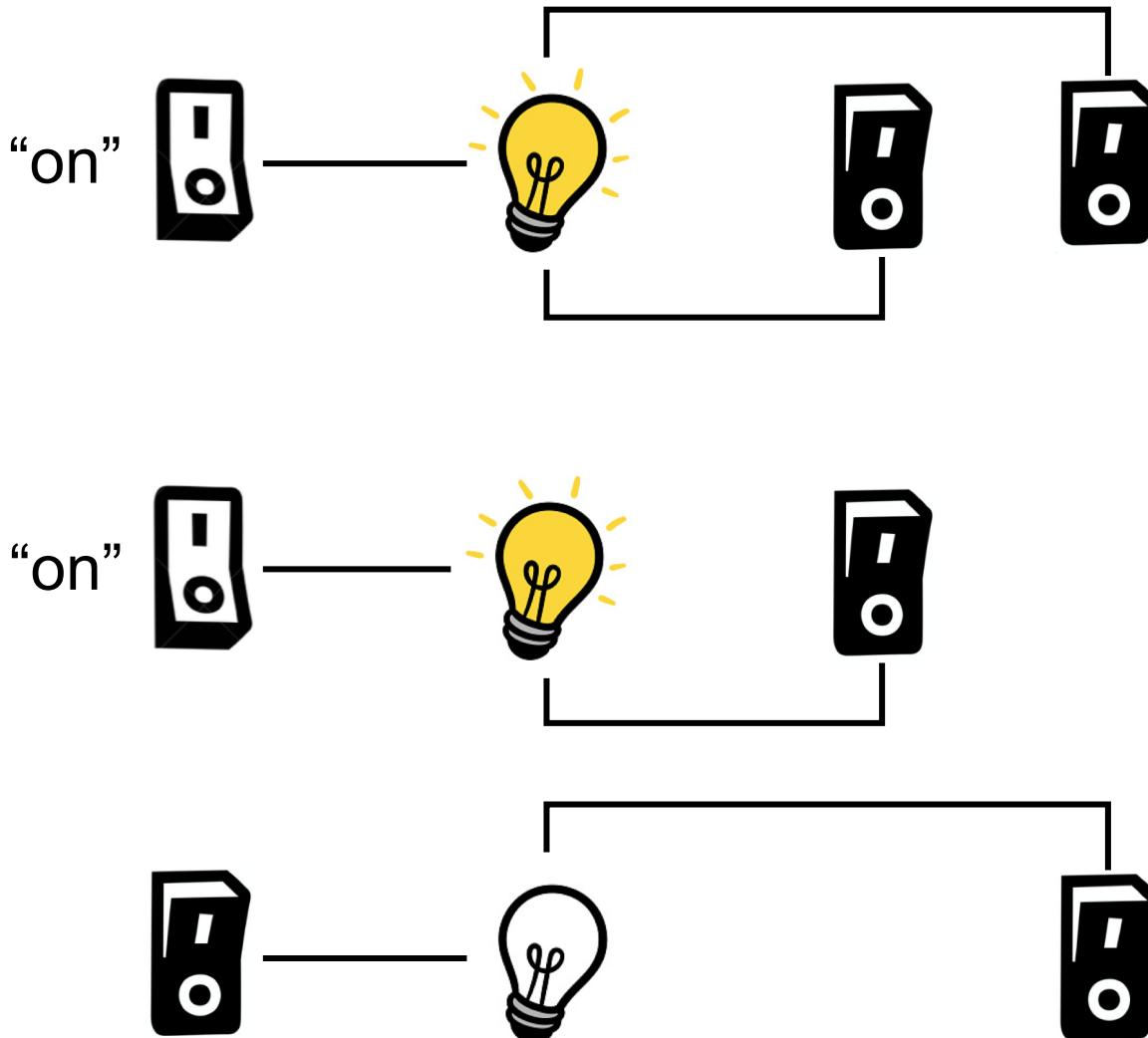
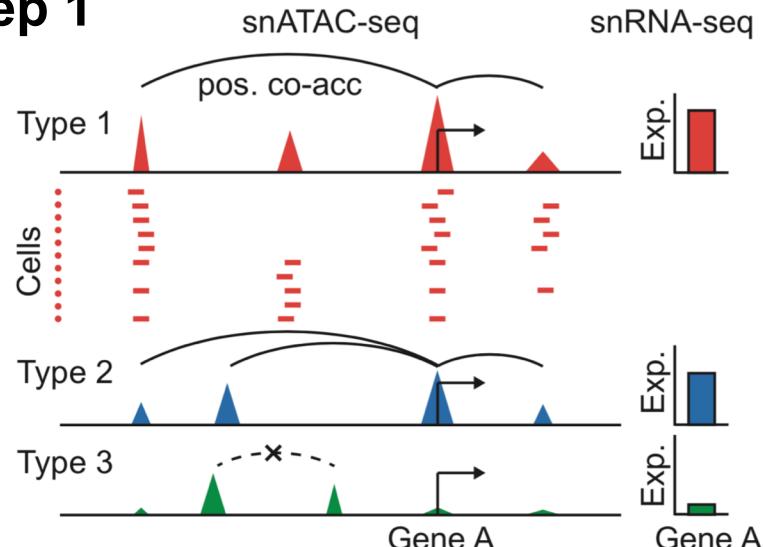
Linking distal cCREs to target genes

Step 1



Linking distal cCREs to target genes

Step 1

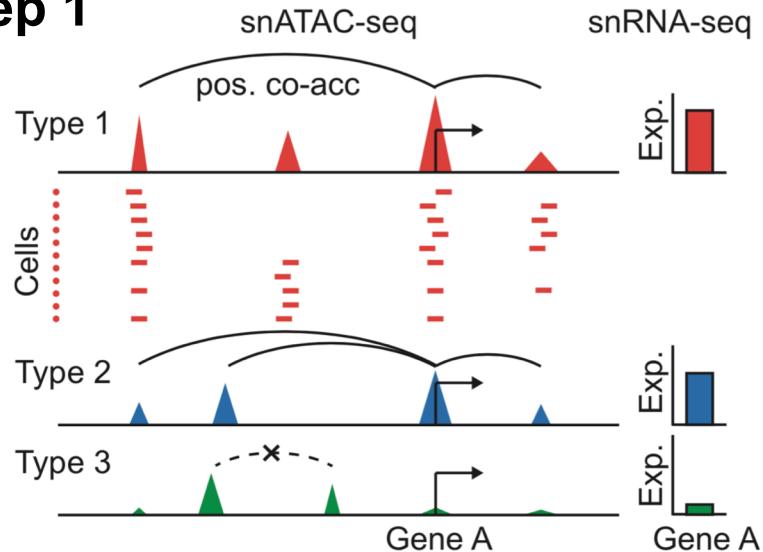


Step 2

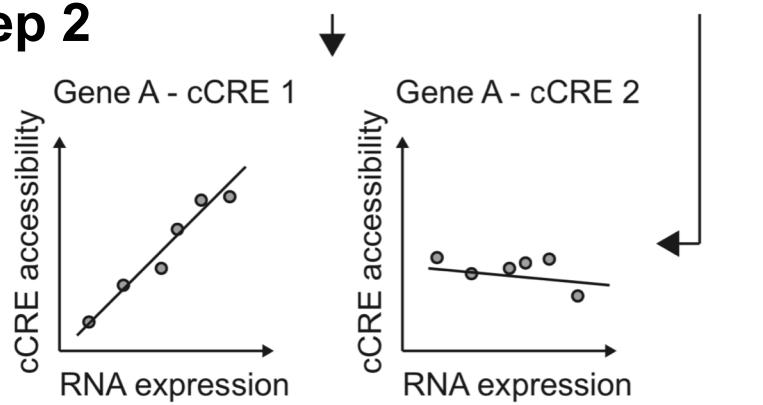
114,877 putative enhancers
13,094 potential target genes

Linking distal cCREs to target genes

Step 1

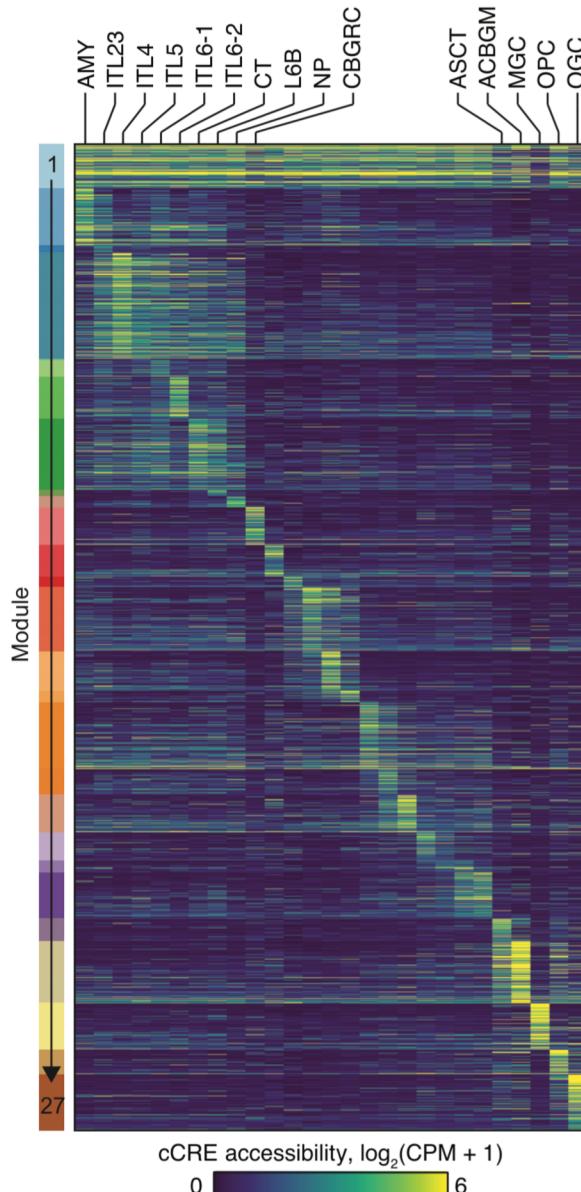


Step 2

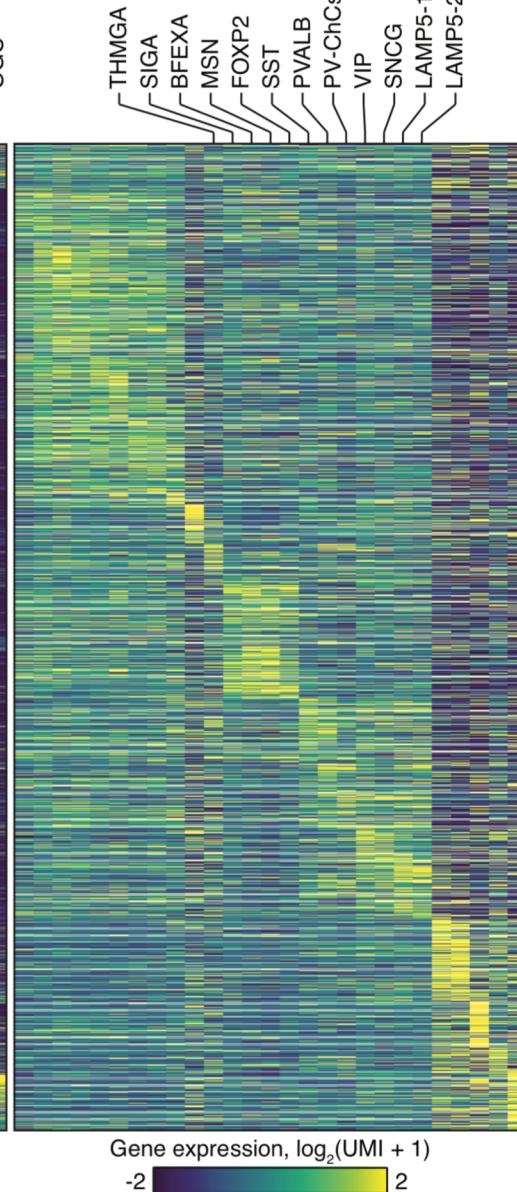


114,877 putative enhancers
13,094 potential target genes

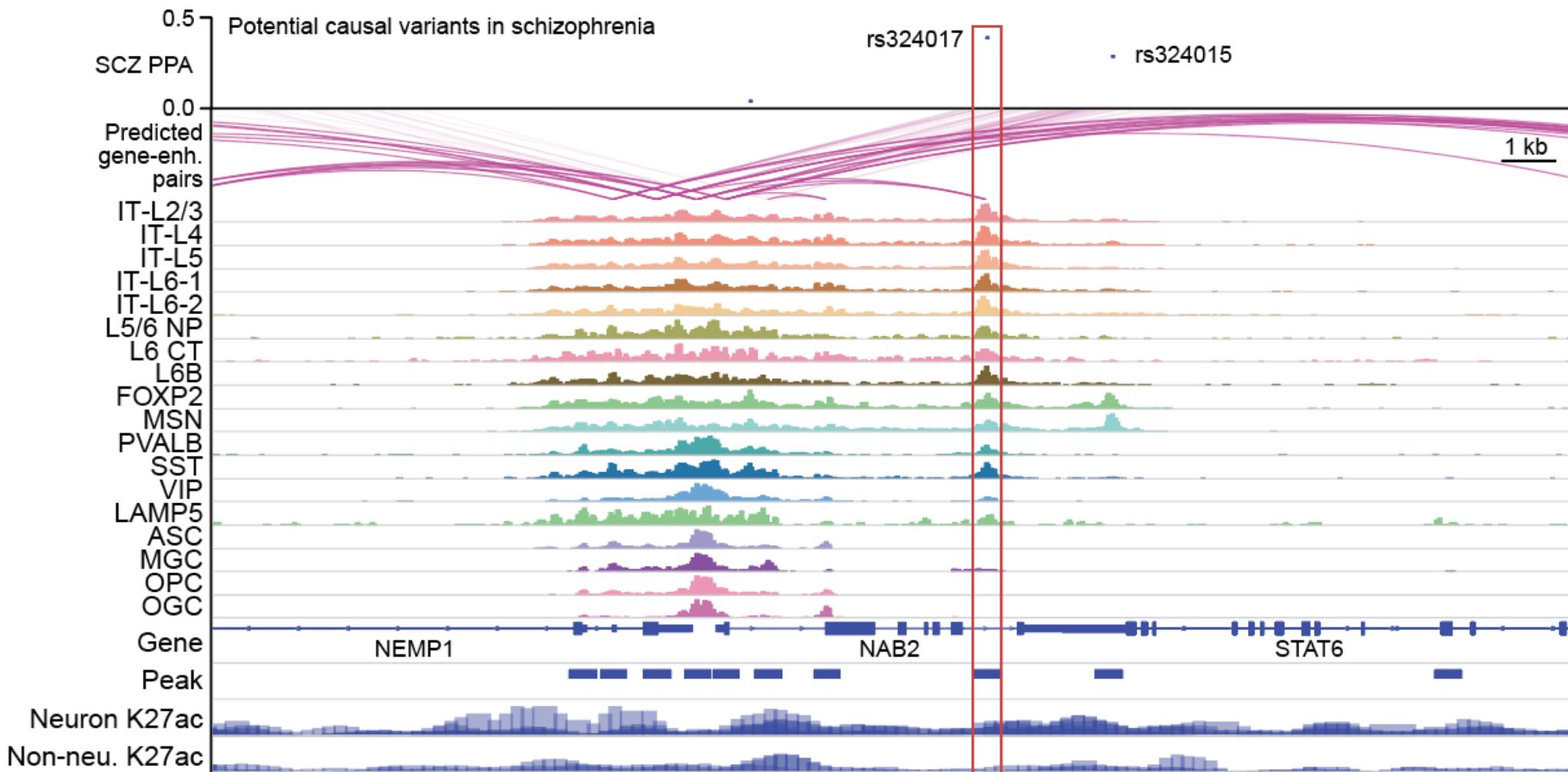
Enhancer accessibility



Expression of putative target genes



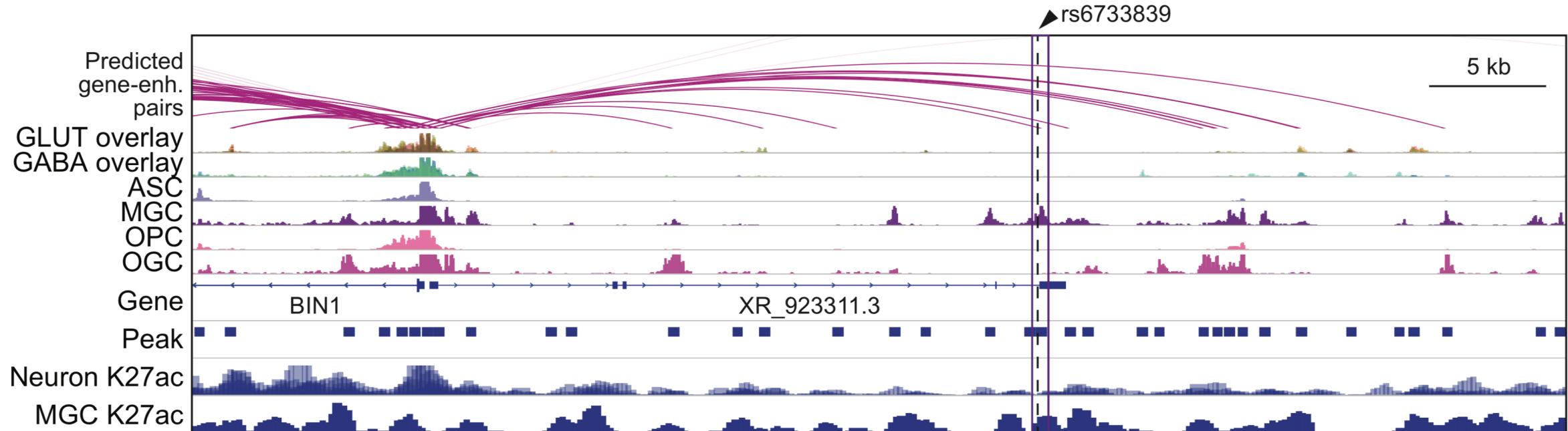
Representative risk variants in schizophrenia



NAB2: Immediate early genes anchor a biological pathway of proteins required for memory formation, long-term depression and risk for schizophrenia

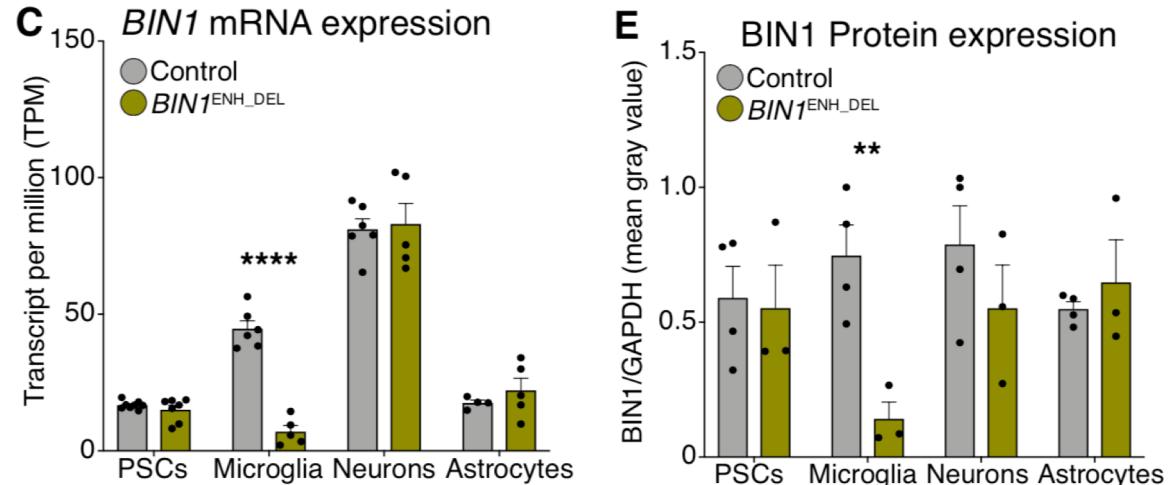
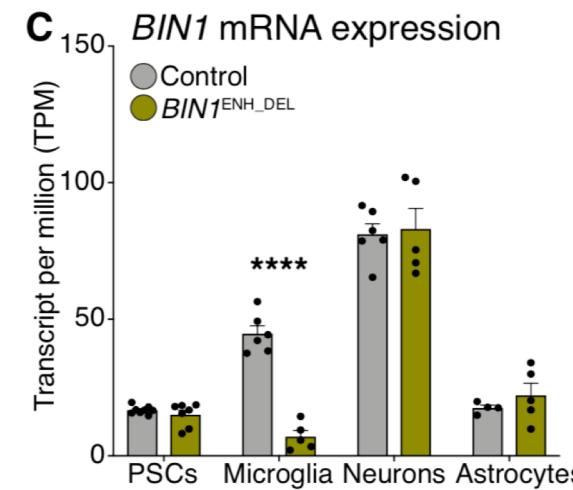
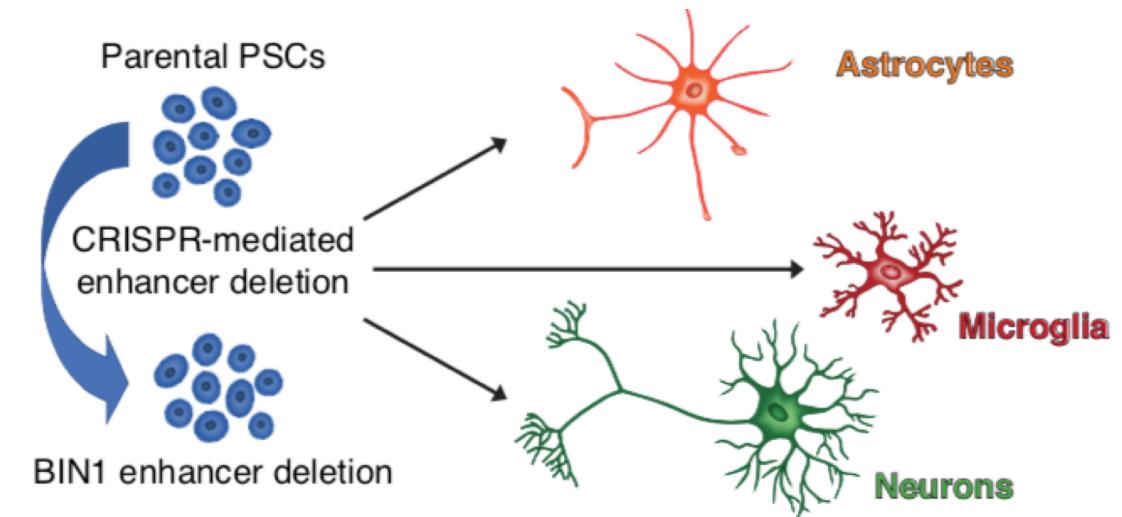
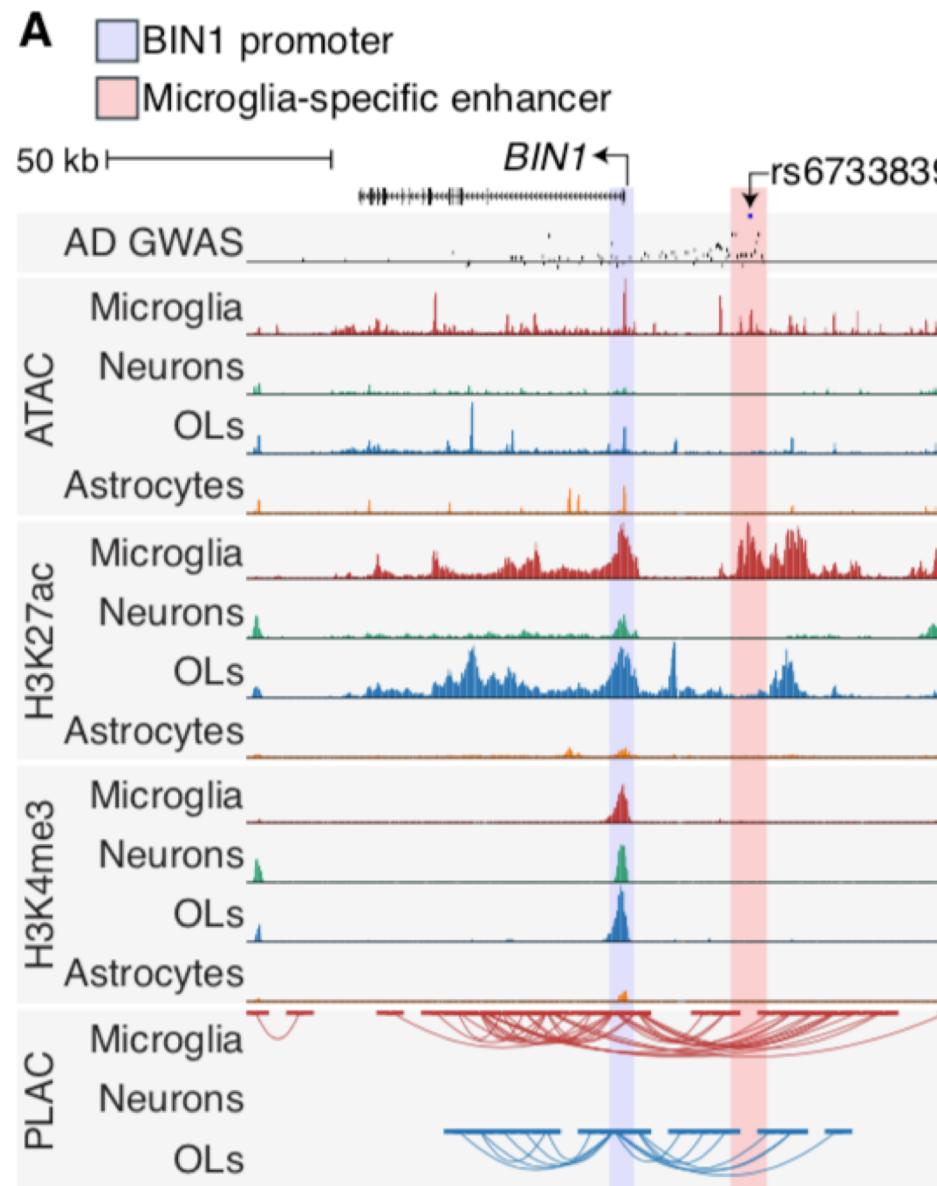
Representative risk variants in Alzheimer's disease

A microglia-specific enhancer containing AD risk variant drives microglia-specific expression of the BIN1 gene



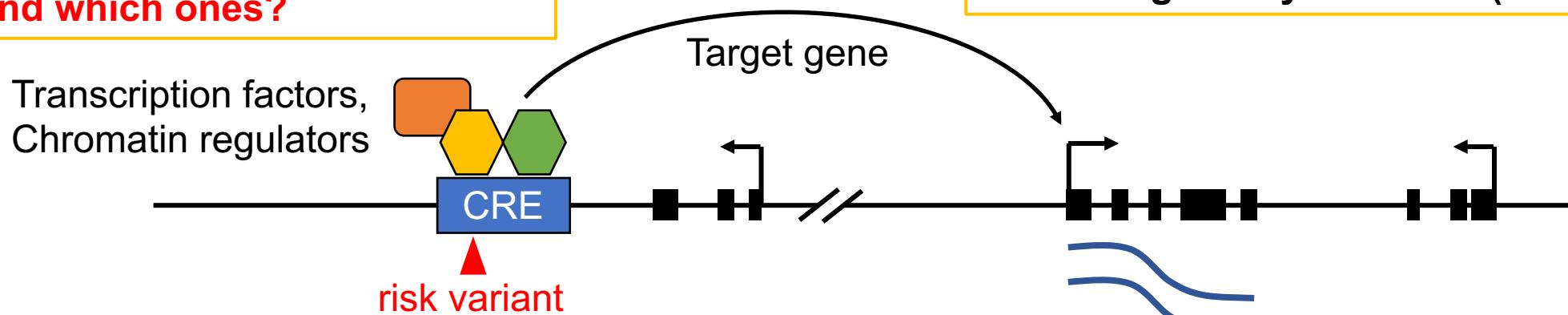
BIN1: This gene encodes several isoforms of a nucleocytoplasmic adaptor protein.
This gene is also a risk gene in Alzheimer's disease.

A microglia-specific enhancer containing AD risk variant drives microglia-specific expression of the *BIN1* gene



A paradigm for understanding the noncoding disease risk variants, and the key barriers

How do risk variants affect transcription factor binding, and which ones?



What target genes are controlled by the *cis* regulatory elements (CREs)?

What sequences act as *cis* regulatory elements (CREs), where and when do they act?

It has been estimated that >50% of risk variants act in this way (Nasser J, ... Engreitz, Nature, 2021)

Predicting chromatin accessibility from DNA sequences

Article | Open Access | Published: 04 October 2021

Effective gene expression prediction from sequence by integrating long-range interactions

Žiga Avsec  , Vikram Agarwal, Daniel Visentin, Joseph R. Ledsam, Agnieszka Grabska-Barwinska, Kyle R. Taylor, Yannis Assael, John Jumper, Pushmeet Kohli  & David R. Kelley 

Nature Methods 18, 1196–1203 (2021) | Cite this article

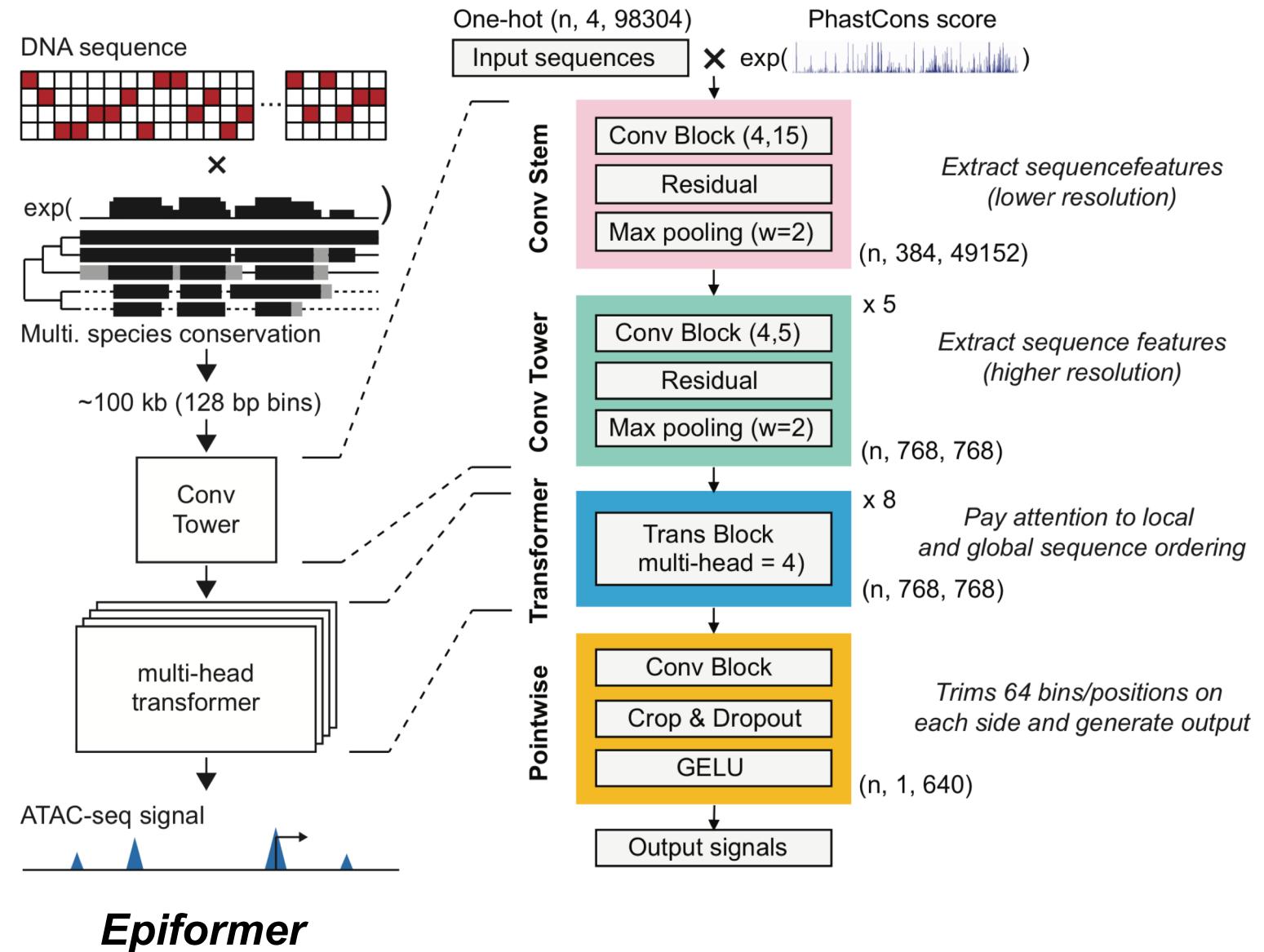
88k Accesses | 52 Citations | 396 Altmetric | Metrics



Zihan Wang
CSE, UCSD



Dr. Jingbo Shang
CSE, UCSD

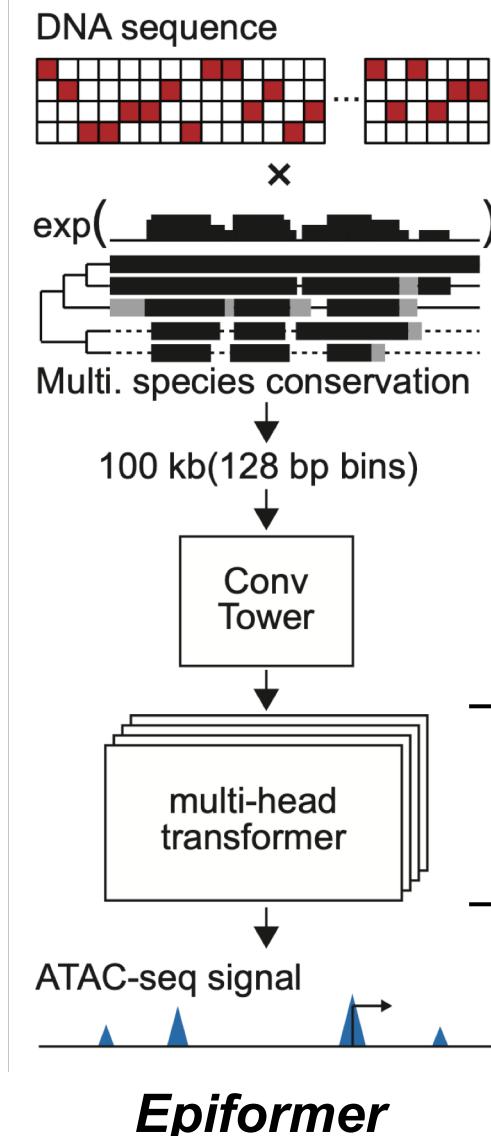


Predicting chromatin accessibility from DNA sequences

DNA sequence



Potential function
(genomic signals)



Natural language processing

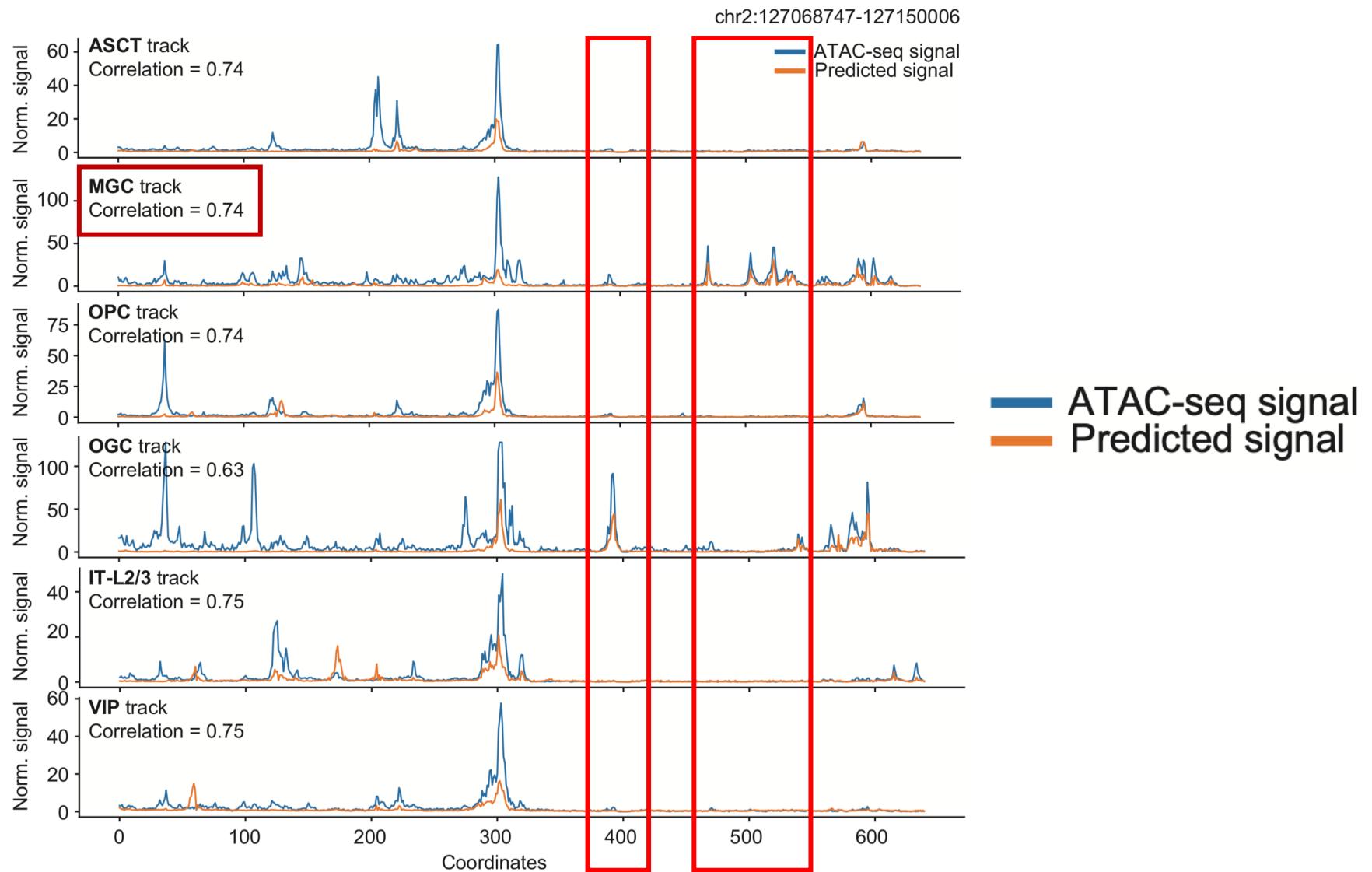
The agreement on the European Economic Area was signed in August 1992.

Translator

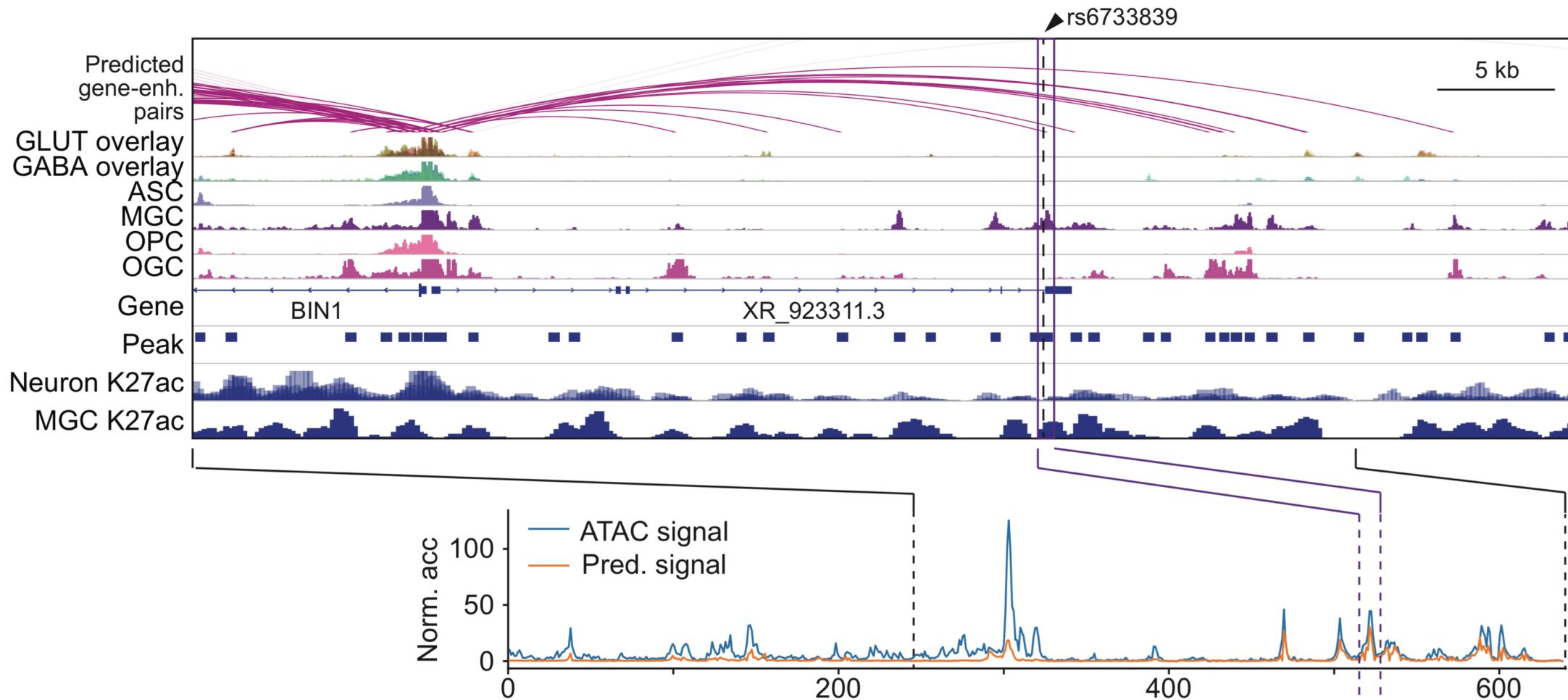
L'accord sur la zone économique européenne a été signé en août 1992.

Predicting chromatin accessibility from DNA sequences

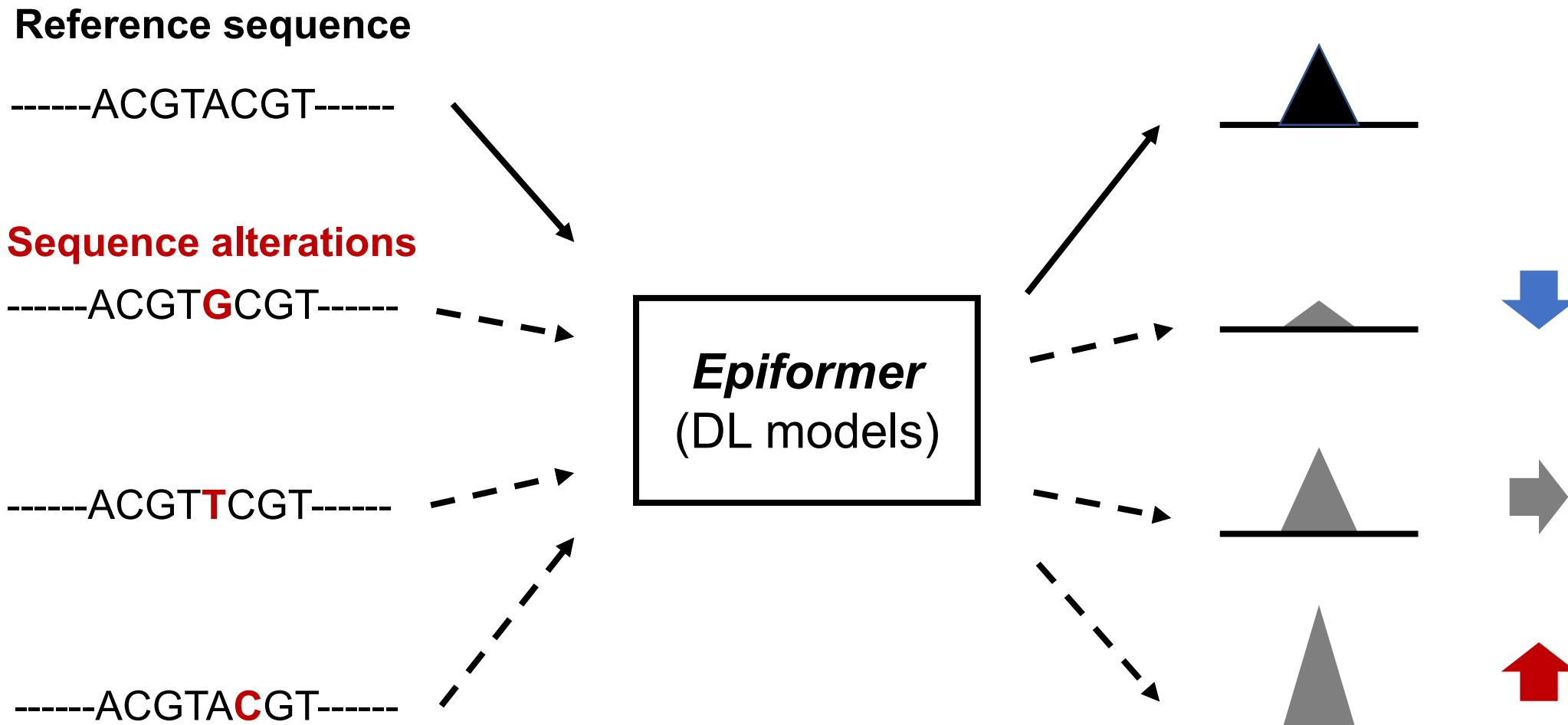
- Astrocytes (ASC)
- Microglia (MGC)**
- Oligodendrocytes precursor cell (OPC)
- Oligodendrocytes (OGC)
- Excitatory neurons at cortical layer 2/3
- VIP+ Inhibitory neurons



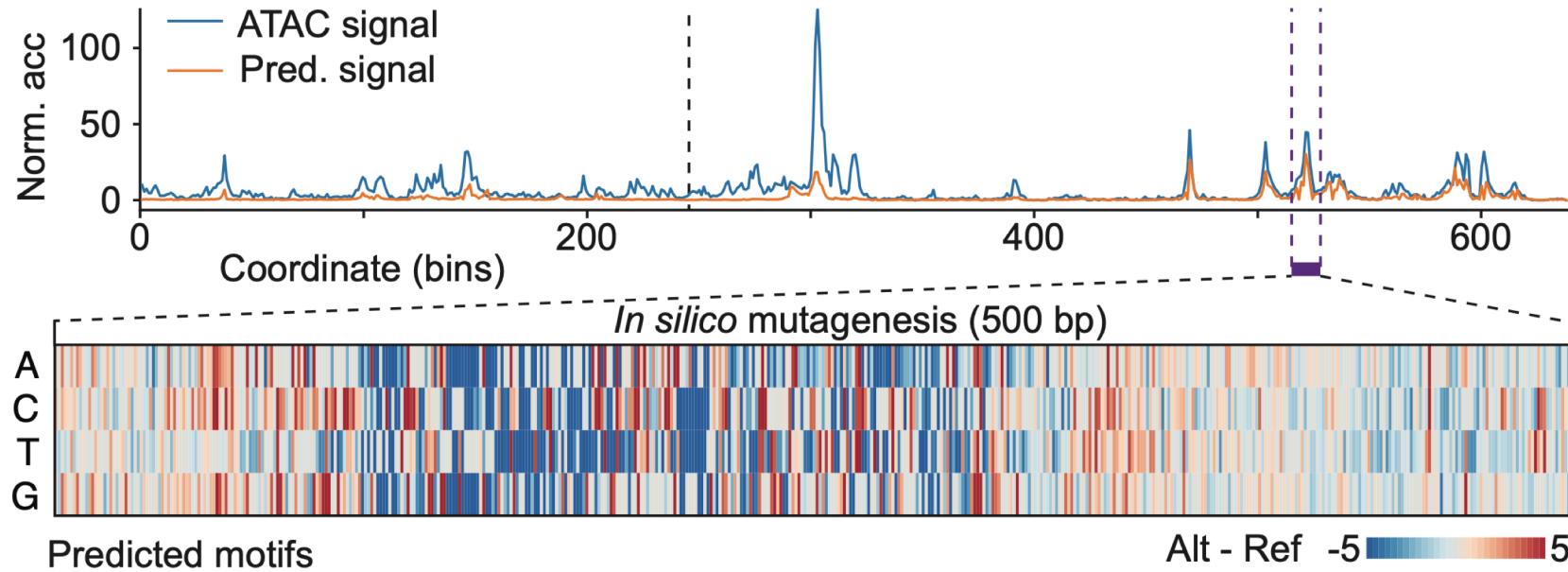
Predicting chromatin accessibility from DNA sequences



In silico mutagenesis accelerates interpretation

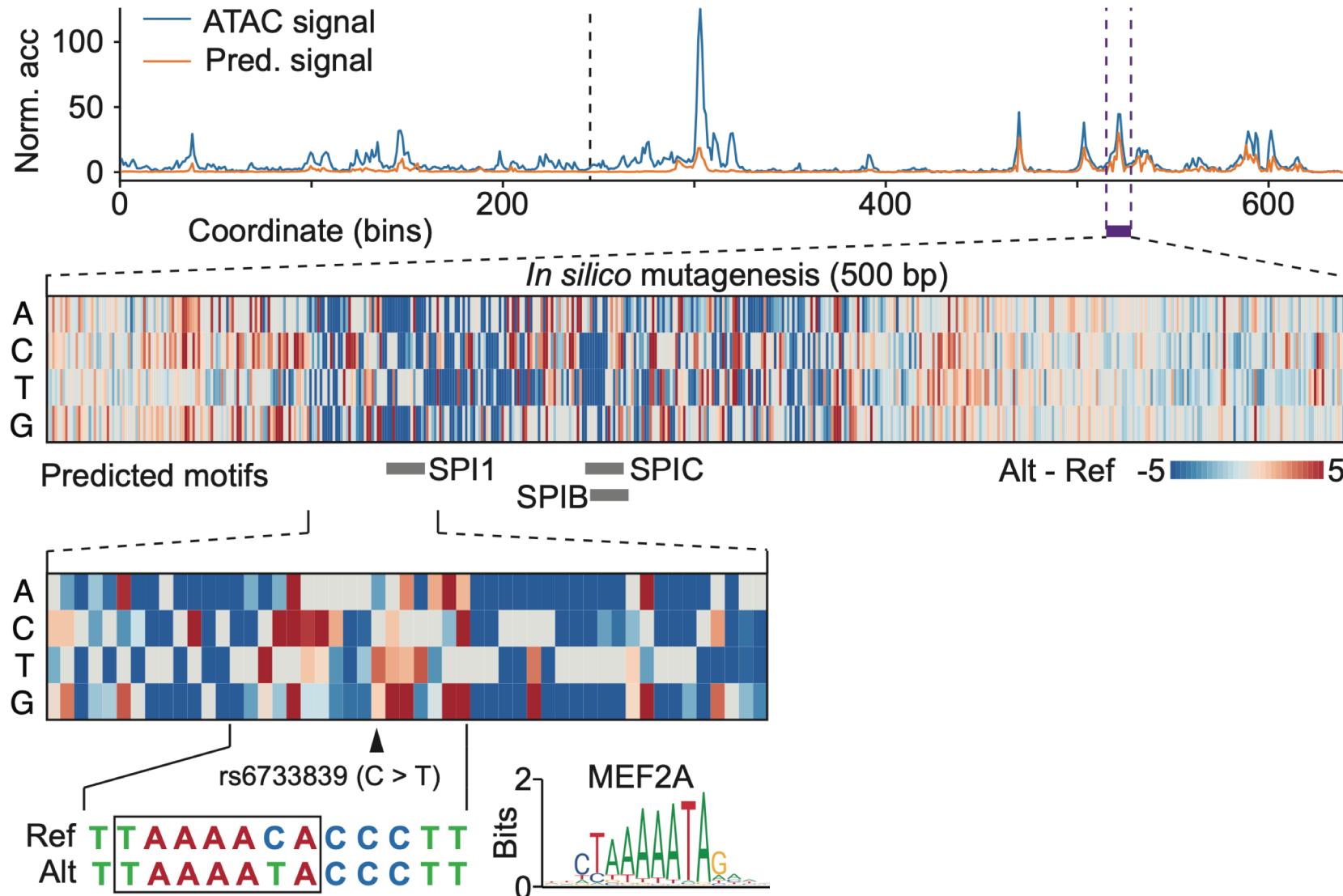


Interpret risk variants through in silico mutagenesis

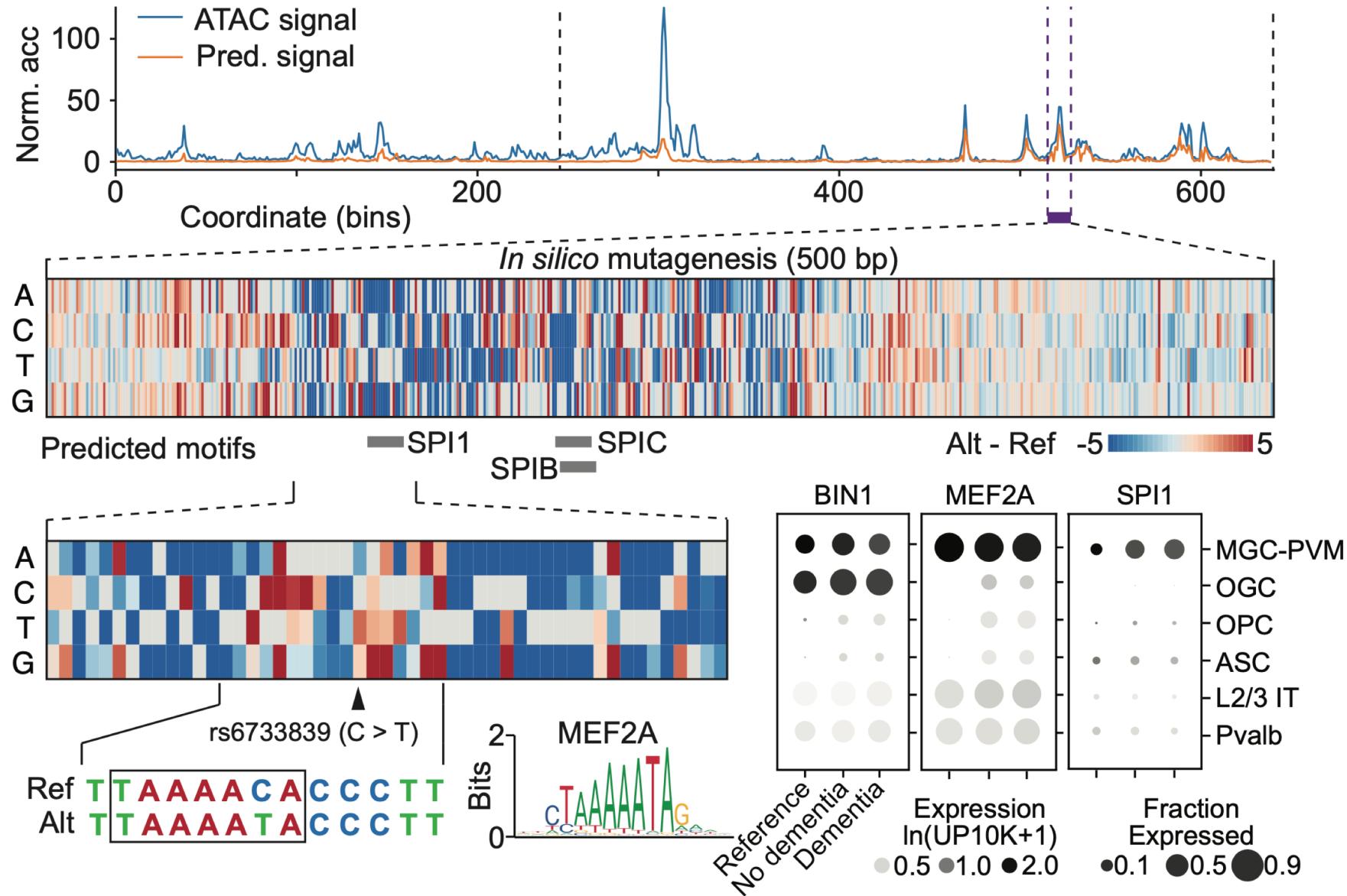


The transcription factor **PU.1 (SPI1)** regulates the normal ***development of microglial cells and their function*** in the brain (Smith et al., 2013).

Interpret risk variants through in silico mutagenesis

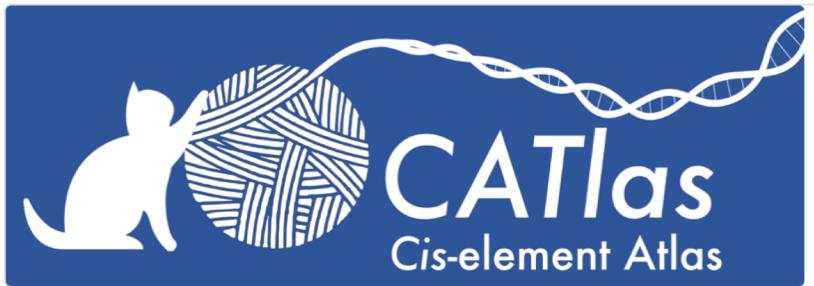


Interpret risk variants through in silico mutagenesis



Seattle Alzheimer's Disease
Brain Cell Atlas (SEA-AD)

Cis-element atlas

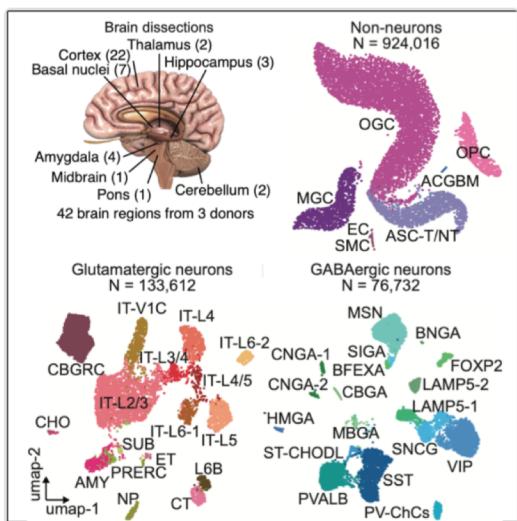


Human Brain

Sample: 42 brain regions

Method: snATAC-seq + multi-modal

Nuclei count: > 1.1 million



Cell Clusters

Explore cell clusters in different datasets

Last updated Nov,2022

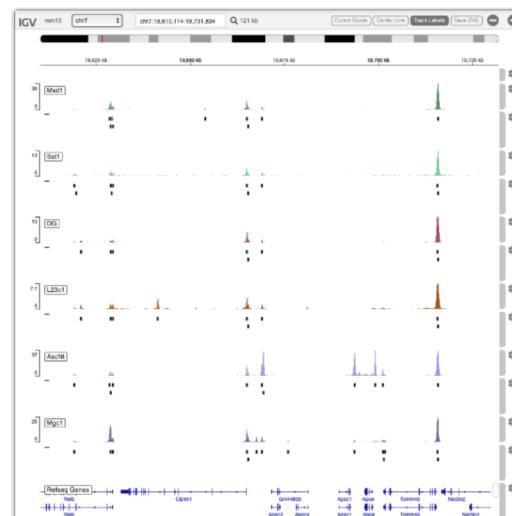
Explore

Class	L3 Cluster	Description
● GABAergic	PVGA1	MGE-derived neurogliaform cells, Pvlab
	Average TSSE:	20.89
	Average LogUMI:	3.63
	More Info	
● GABAergic	SSTGA5	MGE-derived neurogliaform cells, Sst
● GABAergic	LAMGA2	CGE-derived neurogliaform cells, Lamp5
● GABAergic	PVGA6	MGE-derived neurogliaform cells, Pvlab
● GABAergic	VIPGA3	CGE-derived neurogliaform cells, Vip
● GABAergic	LAMGA4	CGE-derived neurogliaform cells, Lamp5
● GABAergic	SSTGA1	MGE-derived neurogliaform cells, Sst
● GABAergic	VIPGA1	Cholinergic neurons, Cortex
● GABAergic	SSTGA4	MGE-derived neurogliaform cells, Sst
● GABAergic	SSTGA3	MGE-derived neurogliaform cells, Sst

Cell Types

Explore detailed information for every cell type

Explore



Explore Tracks

Load signal tracks and compare between cell types/regions

Explore

- Explore cell cluster
- Cell type annotations
- Compare signal tracks



**Data now available via our portal:
<http://catlas.org>**

> 34,599 visits

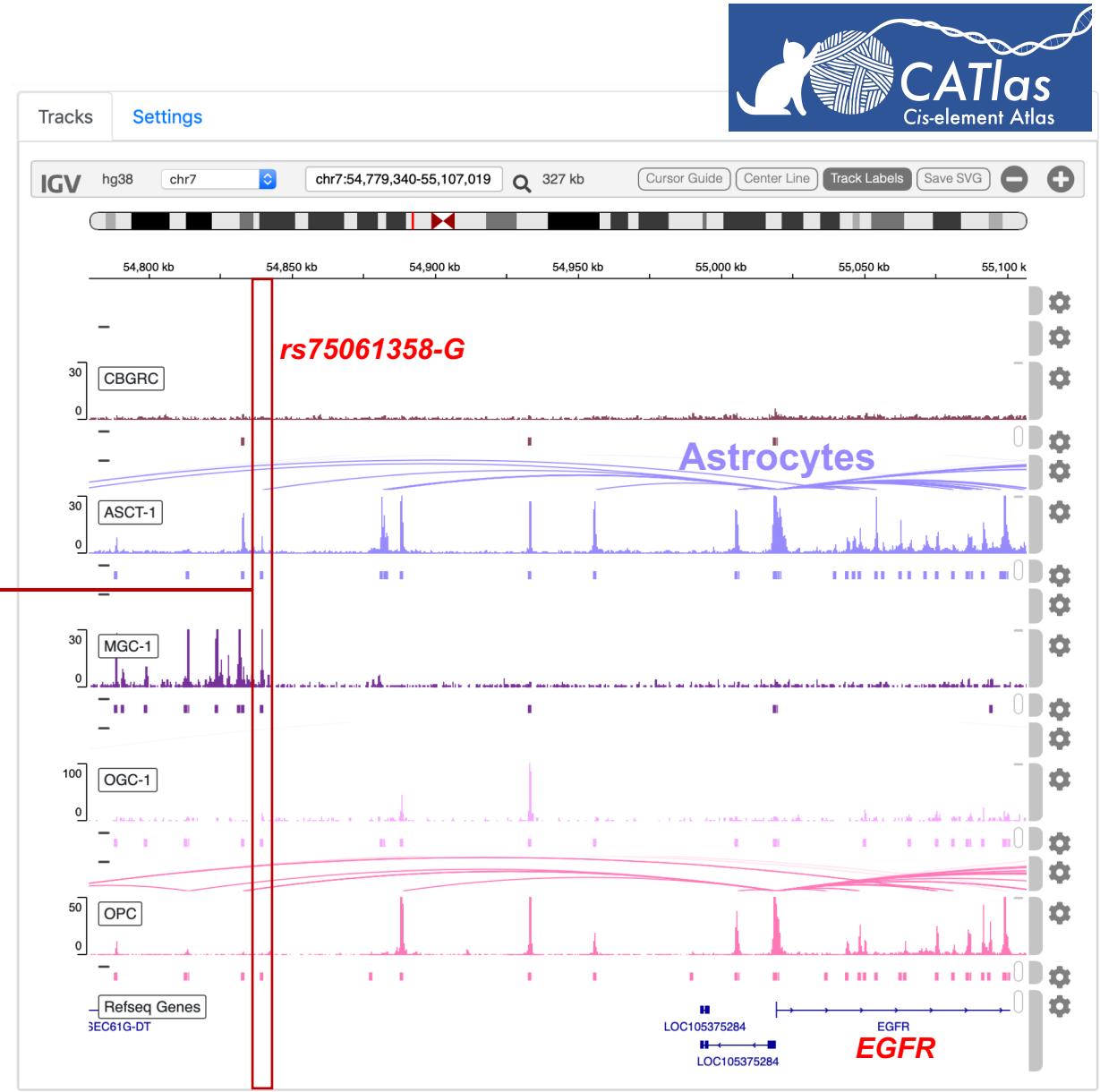
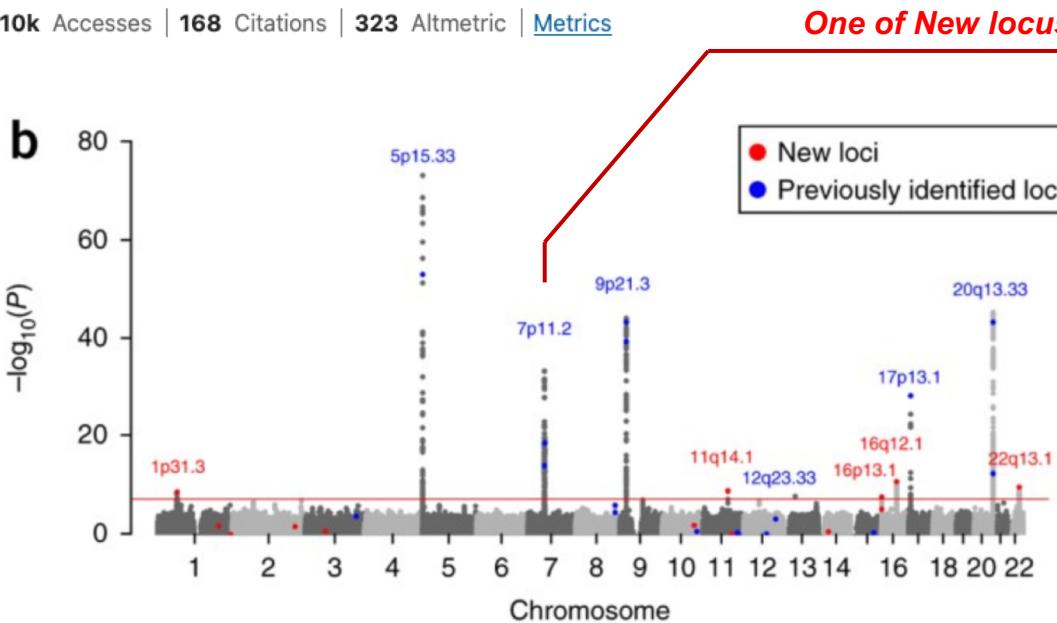
Published: 27 March 2017

Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors

Beatrice S Melin , Jill S Barnholtz-Sloan, Margaret R Wrensch, Christoffer Johansen, Dora Il'yasova, Ben Kinnersley, Quinn T Ostrom, Karim Labreche, Yanwen Chen, Georgina Armstrong, Yanhong Liu, Jeanette E Eckel-Passow, Paul A Decker, Marianne Labussière, Ahmed Idbaih, Khe Hoang-Xuan, Anna-Luisa Di Stefano, Karima Mokhtari, Jean-Yves Delattre, Peter Broderick, Pilar Galan, Konstantinos Gousias, Johannes Schramm, Minouk J Schoemaker, GliomaScan Consortium, ... Melissa L Bondy  + Show authors

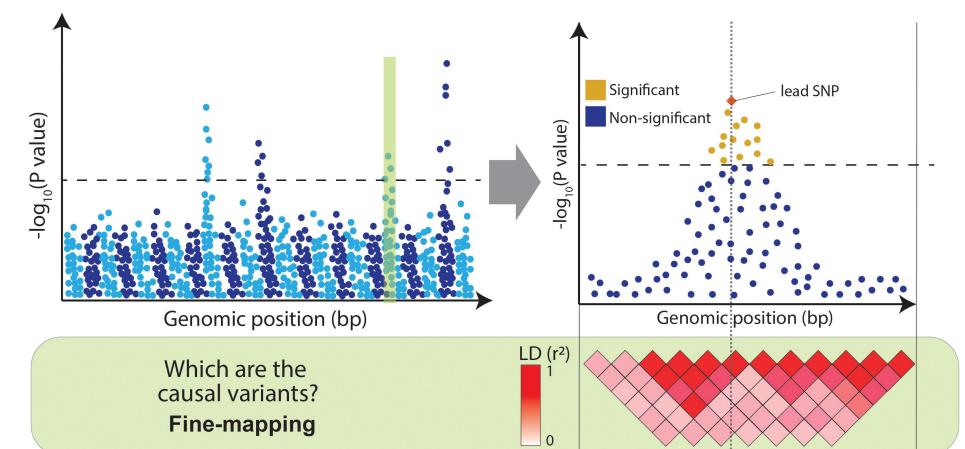
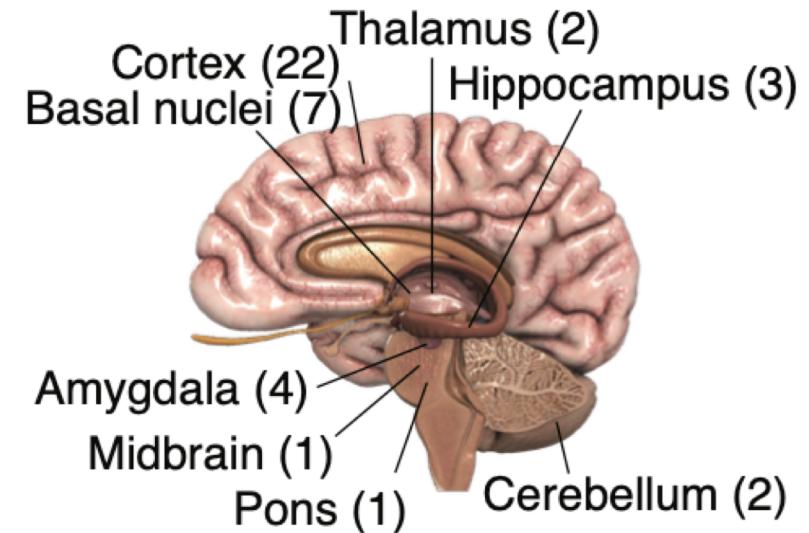
Nature Genetics 49, 789–794 (2017) | Cite this article

10k Accesses | 168 Citations | 323 Altmetric | Metrics



Summary - 1

- Identification of 42 cell subclasses and 107 cell types in > 1.1 million individual nuclei from 42 human brain regions.
- We identified 544,735 candidate cis-regulatory DNA elements (cCREs) that display highly cell-type specificity.
- We found 33-40% of cCREs in glial cells show accessibility variations in brain regions.
- We predict target genes of distal cCREs.
- We identified significant associations between 18 mental disorders and/or GWAS traits with at least one of the brain cell types.
- Deep learning model interpret potential functional of risk variants in cCREs



A naturalist goes to bat for
Mexico's biodiversity p. 144

Adapting COVID-19 booster
vaccines to variants p. 157

Spin selectivity across a
chiral bridge p. 160 & 197

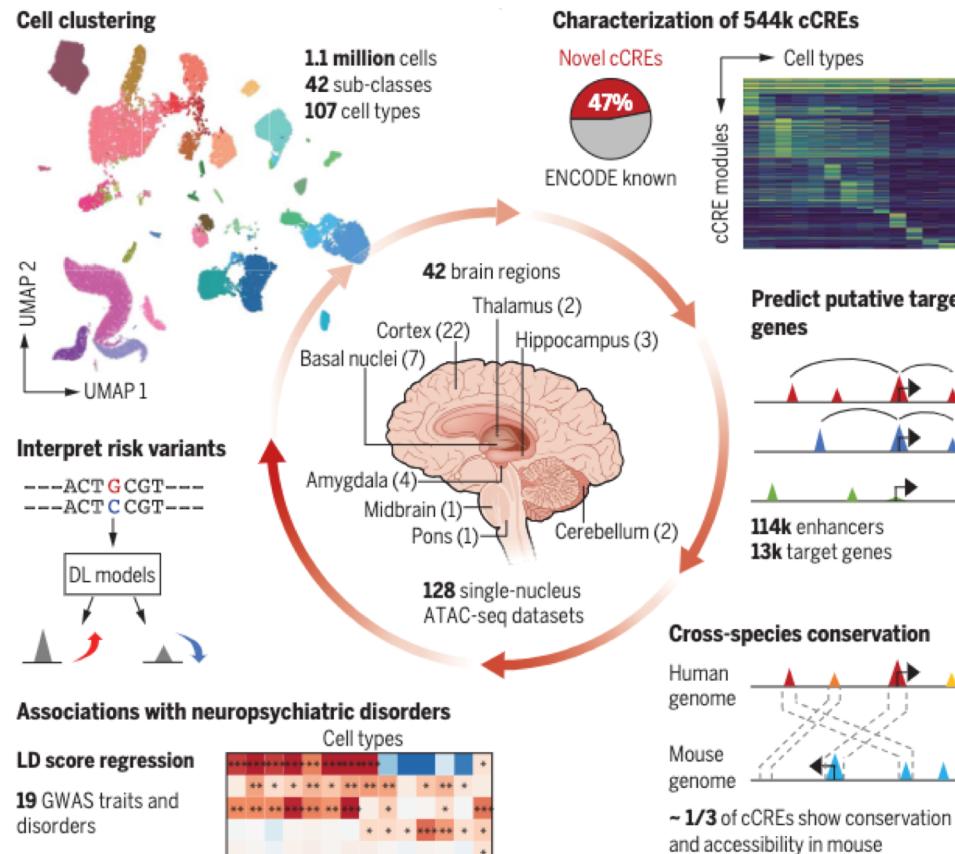


RESEARCH ARTICLE SUMMARY

BICCN

A comparative atlas of single-cell chromatin accessibility in the human brain

Yang Eric Li, Sebastian Preissl, Michael Miller, Nicholas D. Johnson, Zihan Wang, Henry Jiao, Chenxu Zhu, Zhaoning Wang, Yang Xie, Olivier Poirion, Colin Kern, Antonio Pinto-Duarte, Wei Tian, Kimberly Siletti, Nora Emerson, Julia Osteen, Jacinta Lucero, Lin Lin, Qian Yang, Quan Zhu, Nathan Zemke, Sarah Espinoza, Anna Marie Yanny, Julie Nyhus, Nick Dee, Tamara Casper, Nadiya Shapovalova, Daniel Hirschstein, Rebecca D. Hodge, Sten Linnarsson, Trygve Bakken, Boaz Levi, C. Dirk Keene, Jingbo Shang, Ed Lein, Allen Wang, M. Margarita Behrens, Joseph R. Ecker, Bing Ren*



A naturalist goes to bat for
Mexico's biodiversity p. 144

Adapting COVID-19 booster
vaccines to variants p. 157

Spin selectivity across a
chiral bridge p. 160 & 197

Science

SPECIAL ISSUE

BRAIN CELL CENSUS



13 OCTOBER 2023

Science Advances

SPECIAL ISSUE

BRAIN CELL CENSUS

21 papers

AAAS

Acknowledgment

Bing Ren Lab

Dr. Bing Ren
Dr. Chenxu Zhu
Dr. Zhaoning Wang
Yang Xie
Dr. Nathan Zemke

Ethan Armand
Bojing Blair Jia
Bin Li

Other members

Shang lab

Dr. Jingbo Shang
Zihan Wang

Center for epigenomics

Dr. Sebastian Preissl
Michael Miller
Henry Jiao
Dr. Olivier Poirion
Allen Wang
Lin Lin
Qian Yang

Dr. Quan Zhu
Sarah Espinoza
Other members

Sten Linnarsson lab

Dr. Sten Linnarsson
Dr. Kimberly Siletti

Joe Ecker lab

Dr. Joe Ecker
Dr. Wei Tian

Margarita Behrens lab

Dr. Margarita Behrens
Nick Johnson
Dr. Antonio Pinto-Duarte
Jacinta Lucero

Allen Institute

Dr. Ed Lein
Dr. Trygve Bakken
Nik Jorstad
Rebecca Hodge



Explore
data



Cis-elements Atlas
<http://catlas.org>

