

Session V: Epigenomics

Armando Reyes-Palomares

Department Biochemistry and Molecular Biology

Complutense University of Madrid

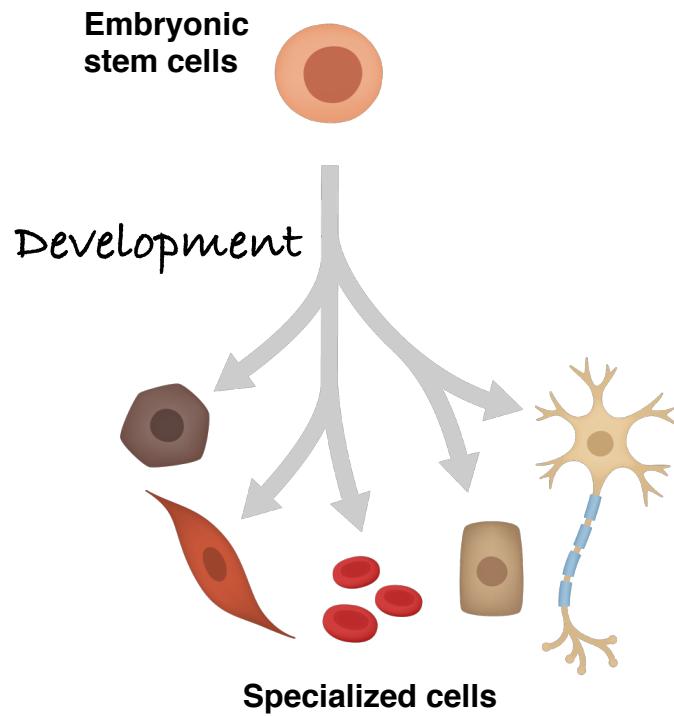
armandorp@ucm.es

Practical Sessions

1. Introduction to R
2. Bioconductor: BiomaRt and GenomicRanges
3. Whole-Genome Analysis – SARS-CoV2
4. Transcriptome Analysis (e.g. RNAseq)
5. Epigenomic Analysis

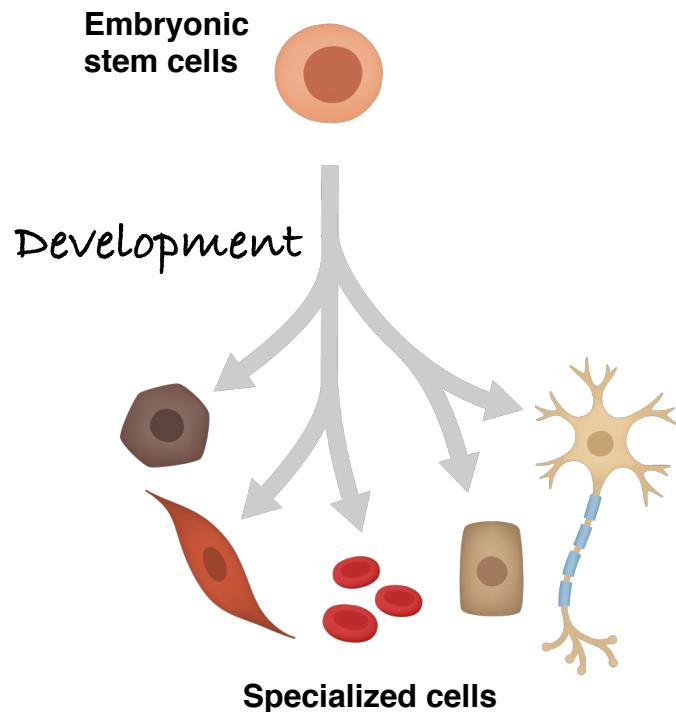
Cells are biological units of life

Responsible of functional diversity in complex organisms

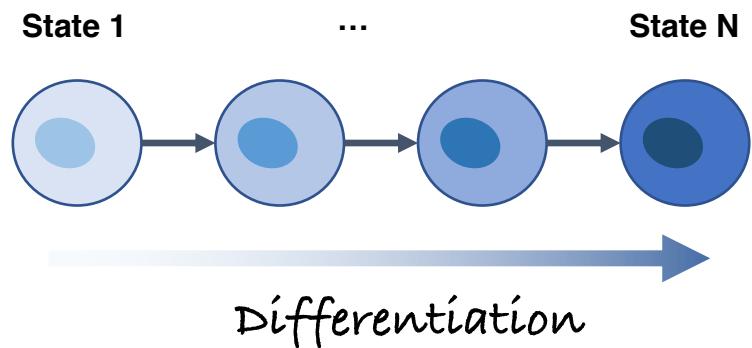


Cells are biological units of life

Responsible of functional diversity in complex organisms



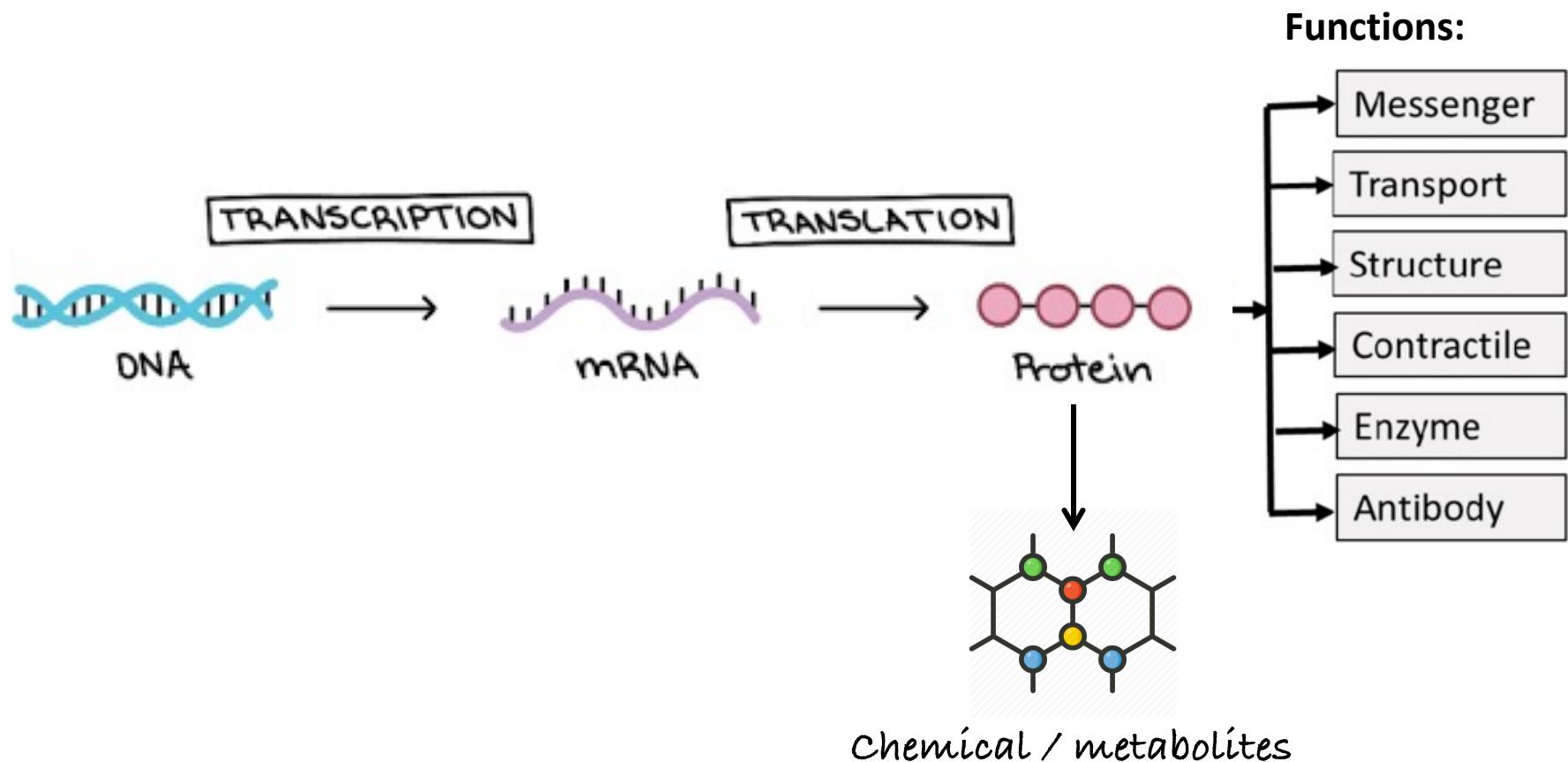
Modulate response to internal/external stimulation



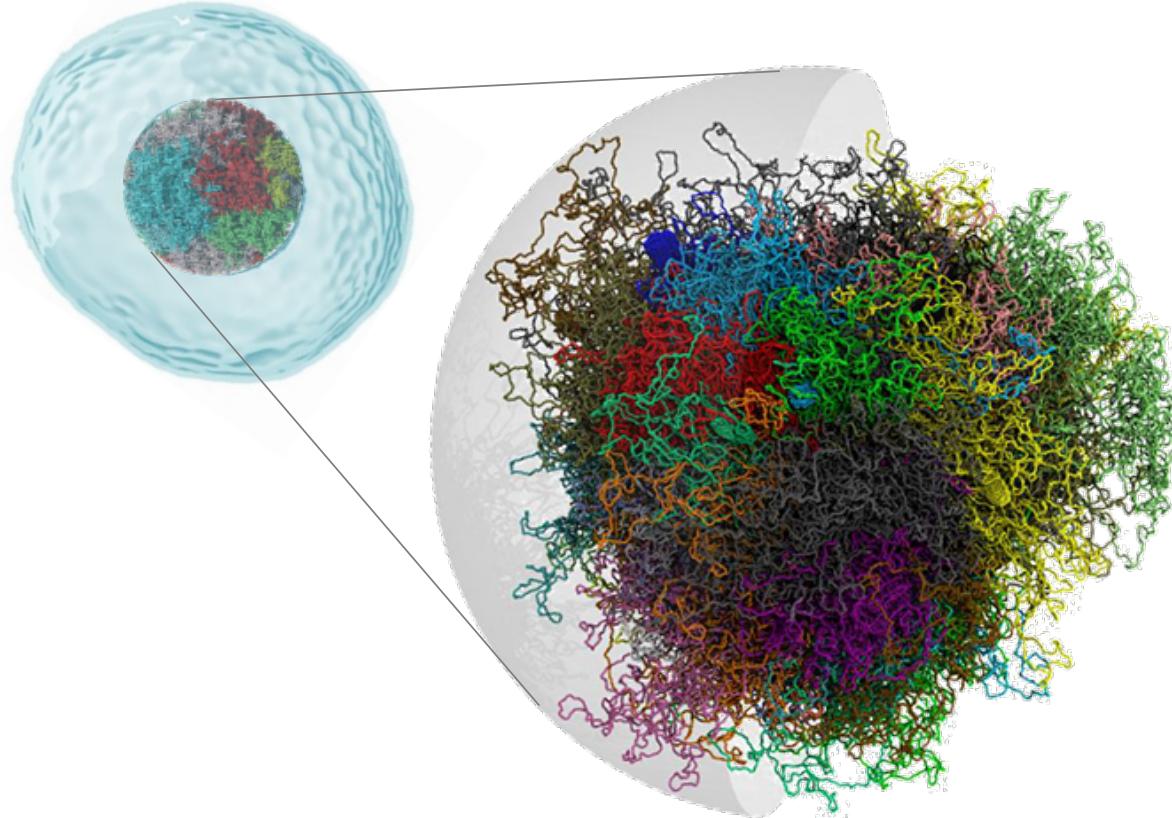
- **Stimulus**
- **Drugs**
- **Mutation**
- **Disease**



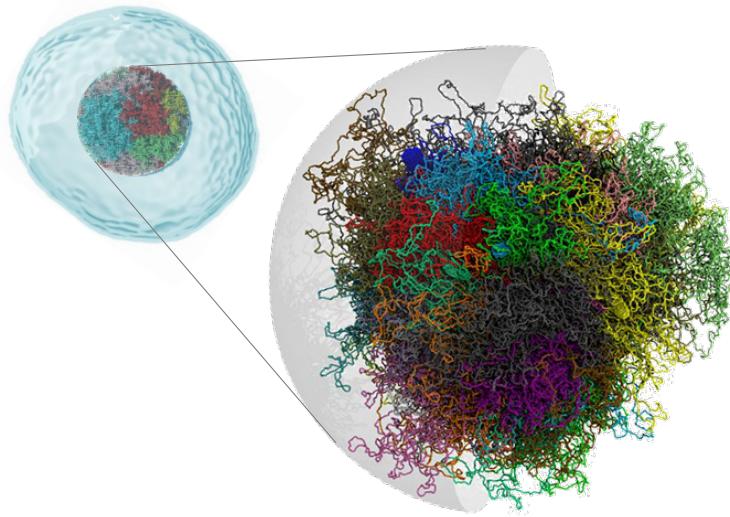
Central “Dogma” of Molecular Biology



How are cells different using the same **GENOME**?



How are cells different using the same GENOME?



...ACGTGTGCAGCTAGCTAGCTAGCTAGAAGCTCTAGCTA
GCTAAAAAAGAAACGCTAGCTATATAGAGAAAATATAAGAGA
GAGGCGCGCGATATAGCTAGTCGATCGATCGATCAGCTAT
AAAAGCGAGAGAGCTAGCTAGCTAGCTAGGAGACGCTAGCTATA
TAGAGAAAATATAAGAGAGAGAGGCCTAGCTATATAAAAAGTATA
AAAGGGCTAGCAAAAGAAGCGAGAGCTAGCTAGATATAGCT
AGTCGAAAGGCTAGCAAAAGAAGCGAGAGCTAGCTAAAGA
GGCTAGAAAAGGCTAGCTAGTCAGCTAGCAAAAGGCTAGCAAAAGAA
GCGAGAGCTAGCTAGATAGCGAGACGCTAGCTATAGAAGC
GAGAGAGCTAGGAGCTAGCTAGATTAGCGAATTAAAAGAA
GCGAGAGCTAGCTAGATTAGCGAGACGCTAGCTATATAGA
GAAAATATAAGAGAGAGGCAGCGCGATATAGCTAGTCGATC
TAGAGAAAATATAAGAGAGAAAAATATAAGAGAGAGGCCTA
GCTATATTAAAAGTATAAAAGGCTAGTAGCTAGTCGATCTAG
AGAAAATATAAGAGAGAAAAATAAAAATATAAGAGAGAGGC
CTAGCTATATTAAAAGTATAAAAGGCGATATAGCTAGTCGAT
CGAAGGCGCGCGATATAGCTAGTCGATCGATCGATCAAGG
AGCTAGCTAGATAGCGAGACGCTAGCTATATAGAGAAAATA
TAAGAGAGAGGGCGCGATATAGCTAGTCGATCGATCGAT
CAGCTATAAAACTAGAGCTATAAAAGAAGCGAGAGAGCTA
GGAGCTAGCTAGATTAGCGAATAAAAGAAGCGAG...



Thousands of Genotype-Phenotype Associations

$p < 5 \times 10^{-8}$ for 17 trait categories



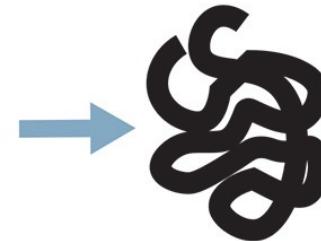
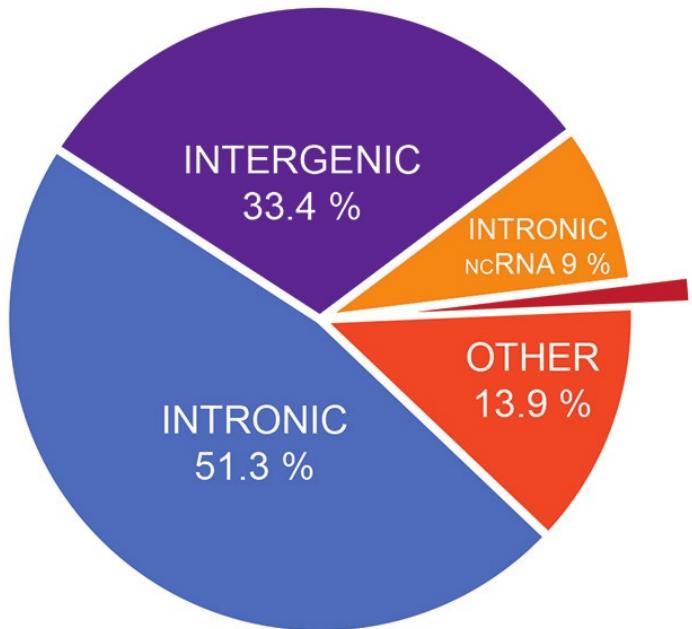
National Human Genome
Research Institute



NHGRI-EBI GWAS Catalog
www.ebi.ac.uk/gwas

Thousands of Genotype-Phenotype Associations

...but very little molecular understanding



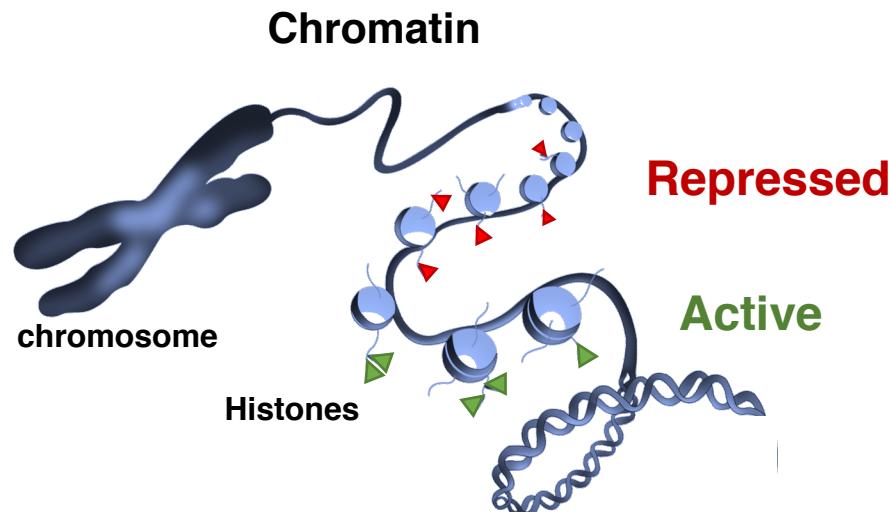
~2% protein structure

98 % non-coding regions

Most of GWAS-SNPs are on un-characterized regions of genome

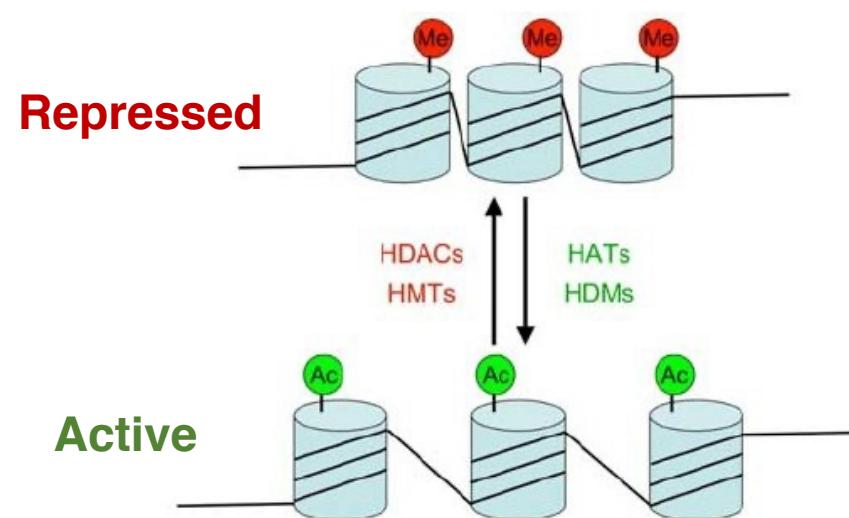
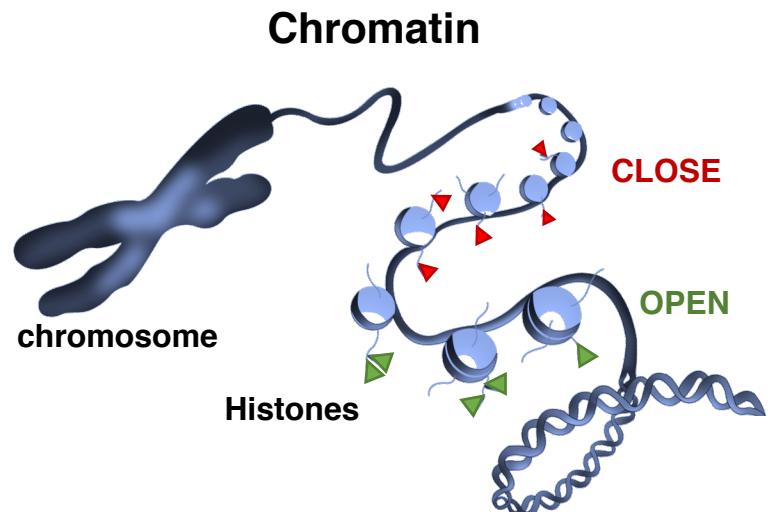
From 1D to 3D Genome (complex organisms)

...ACGTGTGCAGCTAGCTAGCTAGCTAGAAGCTCTAGCTAGCTAAAAAGAAGCGAGAGCTAGCTAGATAGCGAGACGCTAGCTATATAGAG
AAAATATAAGAGAGAGGGCGCGATATAGCTAGTCGATCGATCGATCAGCTATAAAAGCGAGAGAGCTAGCTAGCTAGGAGACGCTAGCT
ATATAGAGAAAATATAAGAGAGAGGCTAGCTATATAAAAGTATAAAAGGCTAGCAAAGAAGCGAGAGCTAGCTAGATATAGCTAGTCGA
TCTAGAGAAAATATAAGAGAGAGGGCTAGCTATATAAAAGTATTAGATCGATCGAAGGCGCGCGATATAGCTAGTCAAAGGCTAGCAAAGAAG
CGAGAGCTAGCTAGATAGCGAGACGCTAGCTATAGAAGCGAGAGAGCTAGGAGCTAGCTAGATTAGCGAATTAAAAGAAGCGAGAGCTAG
CTAGATTAGCGAGACGCTAGCTATATAGAGAAAATATAAGAGAGAGGGCGCGCGATATAGCTAGTCGATCTAGAGAAAATATAAGAGAGAAA
AATATAAGAGAGAGGGCTAGCTATATTAAAAGTATAAAAGGCTAGCGATATAGCTAGTCGATCGAAGGCGCGCGATATAGCTAGTCGATCG
ATCGATCAAGGAGCTAGCTAGATAGCGAGACGCTAGCTATATAGAGAAAATATAAGAGAGAGGGCGCGCGATATAGCTAGTCGATCGATCG
ATCAGCTATAAAACTAGTGATCGATCGATCAGCTATAAAAGAAGCGAGAGAGCTAGGAGCTAGCTAGATTAGCGAATAAAAGAAGCGAG...



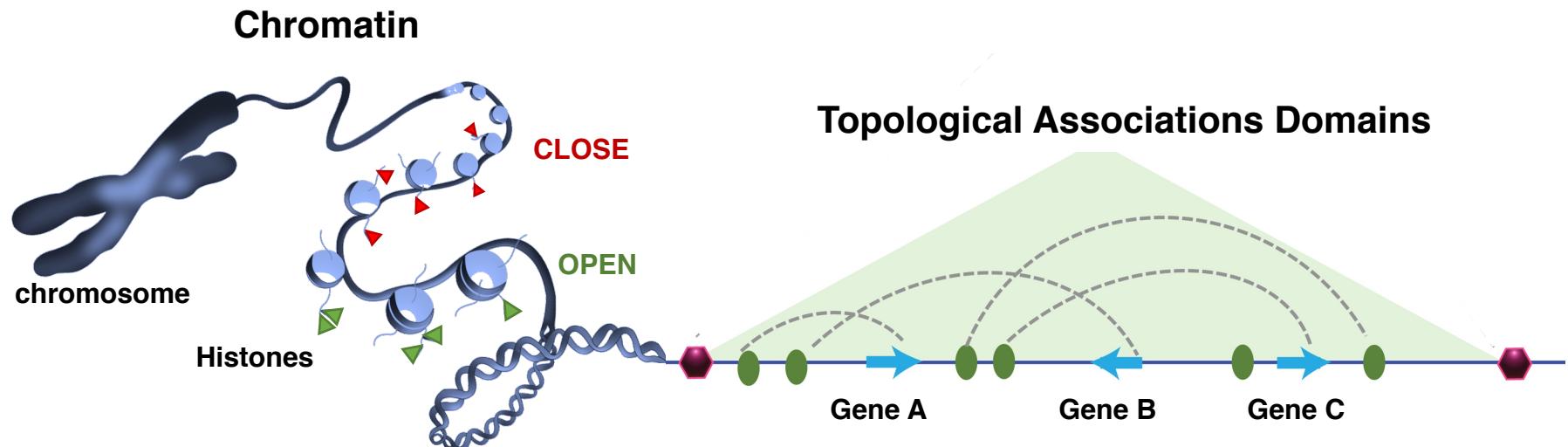
From 1D to 3D genome (complex organisms)

...ACGTGTGCAGCTAGCTAGCTAGCTAGAAGCTCTAGCTAGCTAAAAAGAAGCGAGAGCTAGCTAGATAGCGAGACGCTAGCTATATAGAG
AAAATATAAGAGAGAGGGCGCGATATAGCTAGTCGATCGATCGATCAGCTATAAAAGCGAGAGAGCTAGCTAGCTAGGAGACGCTAGCT
ATATAGAGAAAATATAAGAGAGAGGGCTAGCTATATAAAAGTATAAAAGGCTAGCAAAGAAGCGAGAGCTAGCTAGATATAGCTAGTCGA
TCTAGAGAAAATATAAGAGAGAGGGCTAGCTATATAAAAGTATTAGATCGAAGGCGCGCGATATAGCTAGTCAAAGGCTAGCAAAGAAG
CGAGAGCTAGCTAGATAGCGAGACGCTAGCTATAGAAGCGAGAGAGCTAGGAGCTAGCTAGATTAGCGAATTAAAAGAAGCGAGAGCTAG
CTAGATTAGCGAGACGCTAGCTATATAGAGAAAATATAAGAGAGAGGGCGCGCGATATAGCTAGTCGATCTAGAGAAAATATAAGAGAGAAA
AATATAAGAGAGAGGGCTAGCTATATAAAAGTATAAAAGGCTAGCGATATAGCTAGTCGATCGAAGGCGCGCGATATAGCTAGTCGATCG
ATCGATCAAGGAGCTAGCTAGATAGCGAGACGCTAGCTATATAGAGAAAATATAAGAGAGAGGGCGCGCGATATAGCTAGTCGATCGATCG
ATCAGCTATAAAACTAGTGATCGATCGATCAGCTATAAAAGAAGCGAGAGAGCTAGGAGCTAGCTAGATTAGCGAATAAAAGAAGCGAG...



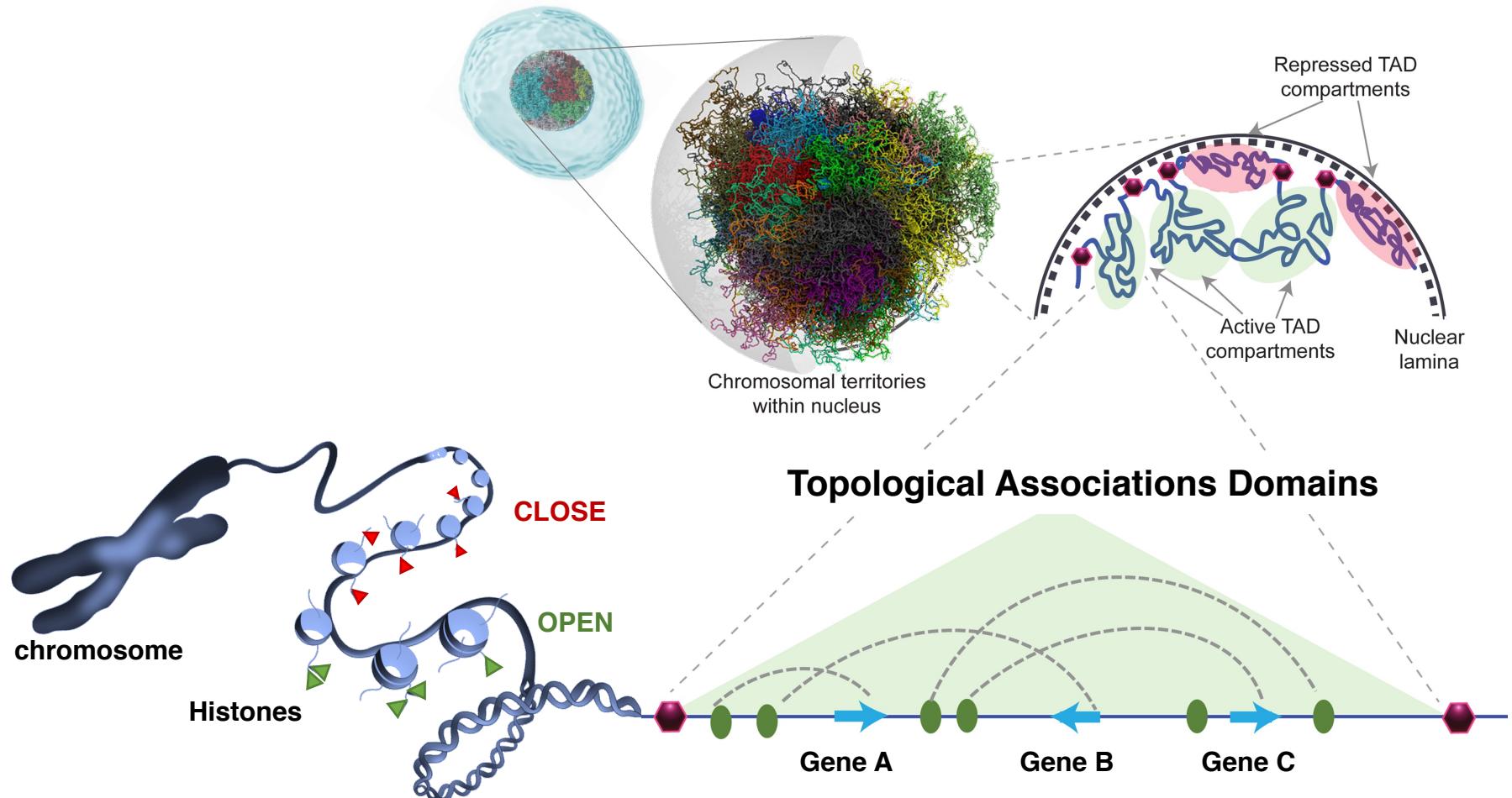
From 1D to 3D Genome (complex organisms)

...ACGTGTGCAGCTAGCTAGCTAGCTAGAAGCTCTAGCTAGCTAAAAAGAAGCGAGAGCTAGCTAGATAGCGAGACGCTAGCTATATAGAG
AAAATATAAGAGAGAGGGCGCGATATAGCTAGTCGATCGATCGATCAGCTATAAAAGCGAGAGAGCTAGCTAGCTAGGAGACGCTAGCT
ATATAGAGAAAATATAAGAGAGAGGGCTAGCTATATAAAAGTATAAAAGGCTAGCAAAAGAAGCGAGAGCTAGCTAGATATAGCTAGTCGA
TCTAGAGAAAATATAAGAGAGAGGGCTAGCTATATAAAAGTATTAGATCGAAGGCGCGATATAGCTAGTCAGCTAGCTAGCTAG
CGAGAGCTAGCTAGATAGCGAGACGCTAGCTATAGAAGCGAGAGAGCTAGGAGCTAGCTAGATTAGCGAATTAAAAGAAGCGAGAGCTAG
CTAGATTAGCGAGACGCTAGCTATATAGAGAAAATATAAGAGAGAGGGCGCGATATAGCTAGTCGATCTAGAGAAAATATAAGAGAGAAA
AATATAAGAGAGAGGGCTAGCTATATAAAAGTATAAAAGGCTAGCGATATAGCTAGTCGATCGAAGGCGCGATATAGCTAGTCGATCG
ATCGATCAAGGAGCTAGCTAGATAGCGAGACGCTAGCTATATAGAGAAAATATAAGAGAGAGGGCGCGATATAGCTAGTCGATCGATCG
ATCAGCTATAAAACTAGTGATCGATCGATCAGCTATAAAAGAAGCGAGAGAGCTAGGAGCTAGCTAGATTAGCGAATAAAAGAAGCGAG...
...



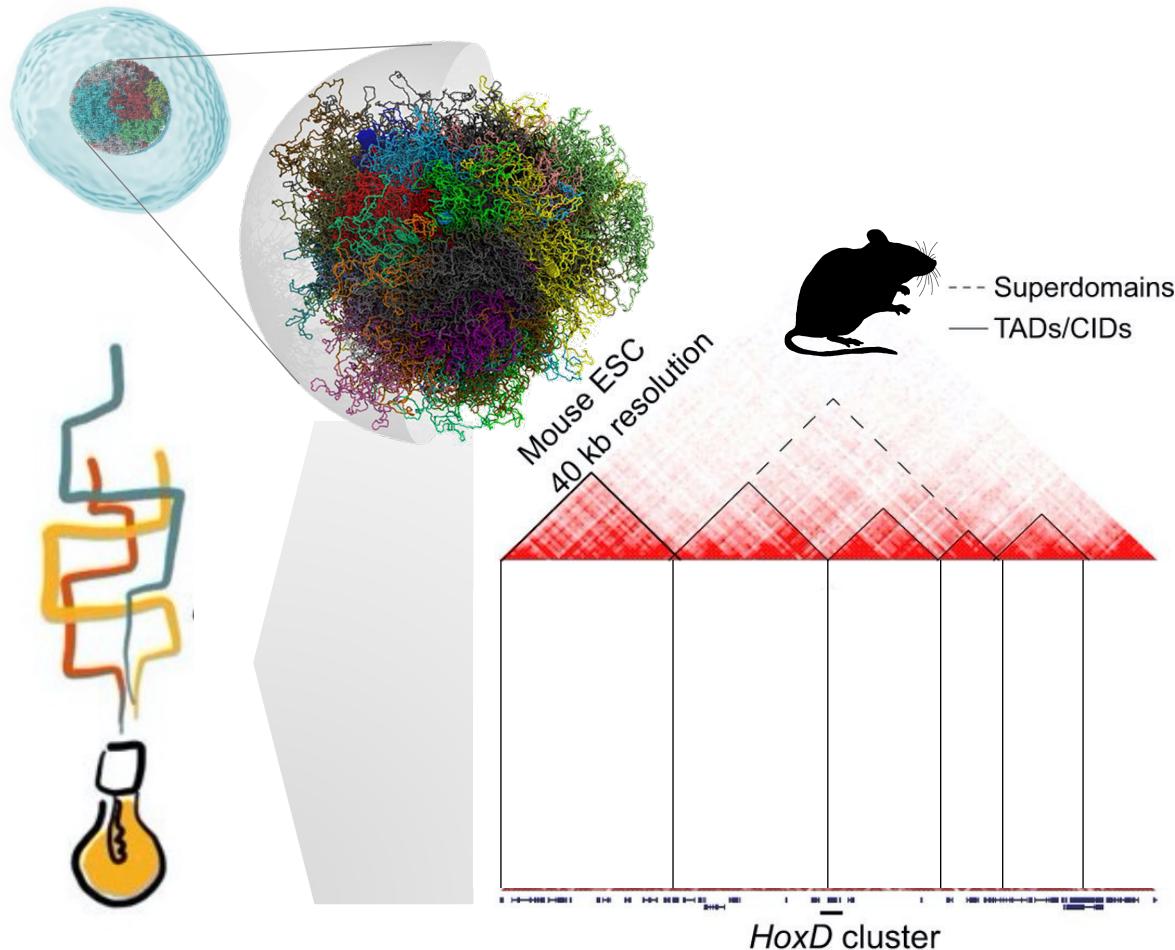
Chromatin organization conveys functionality

(quite active research on topic)



Evolution shapes Chromatin Organization

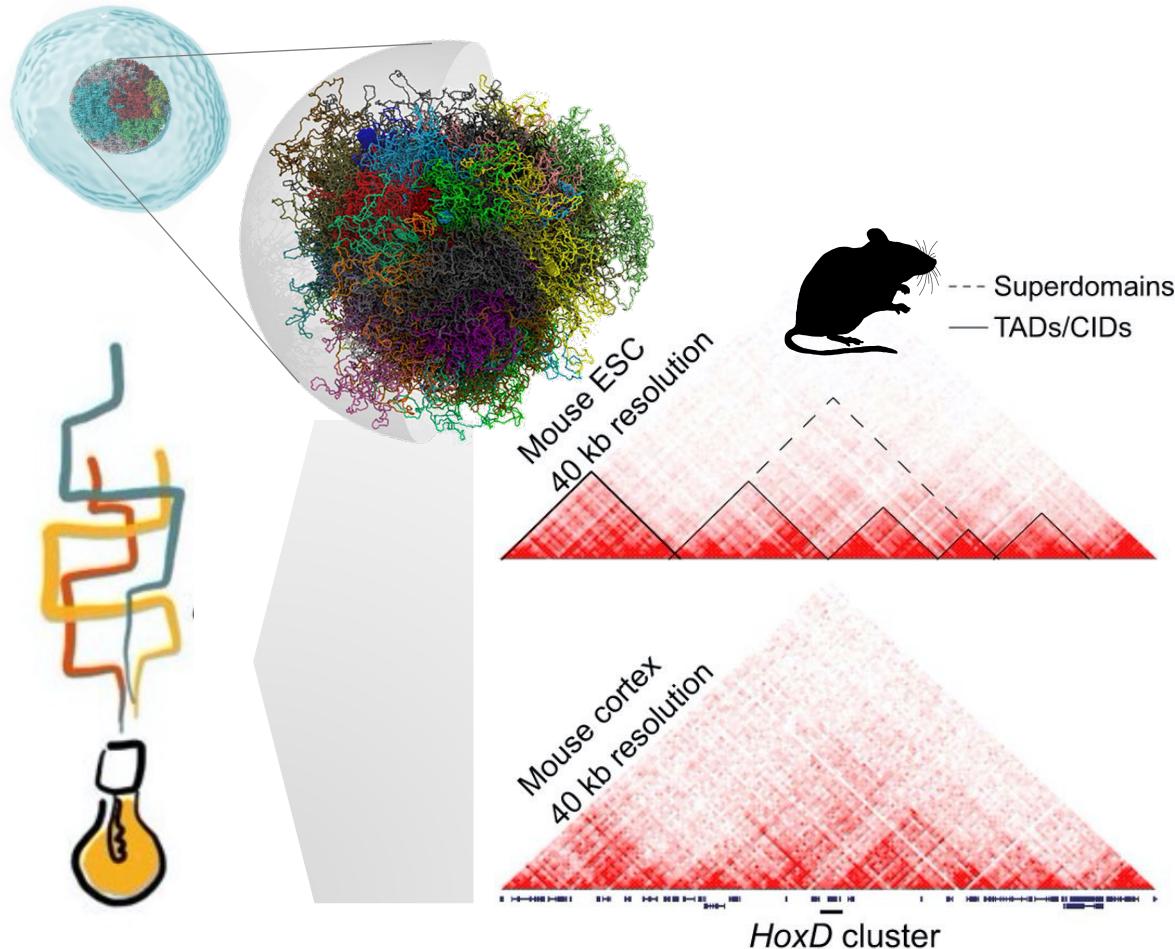
Element locations are conserved because it works well



Modified from image by Pueschel et al. 2016

Evolution shapes Chromatin Organization

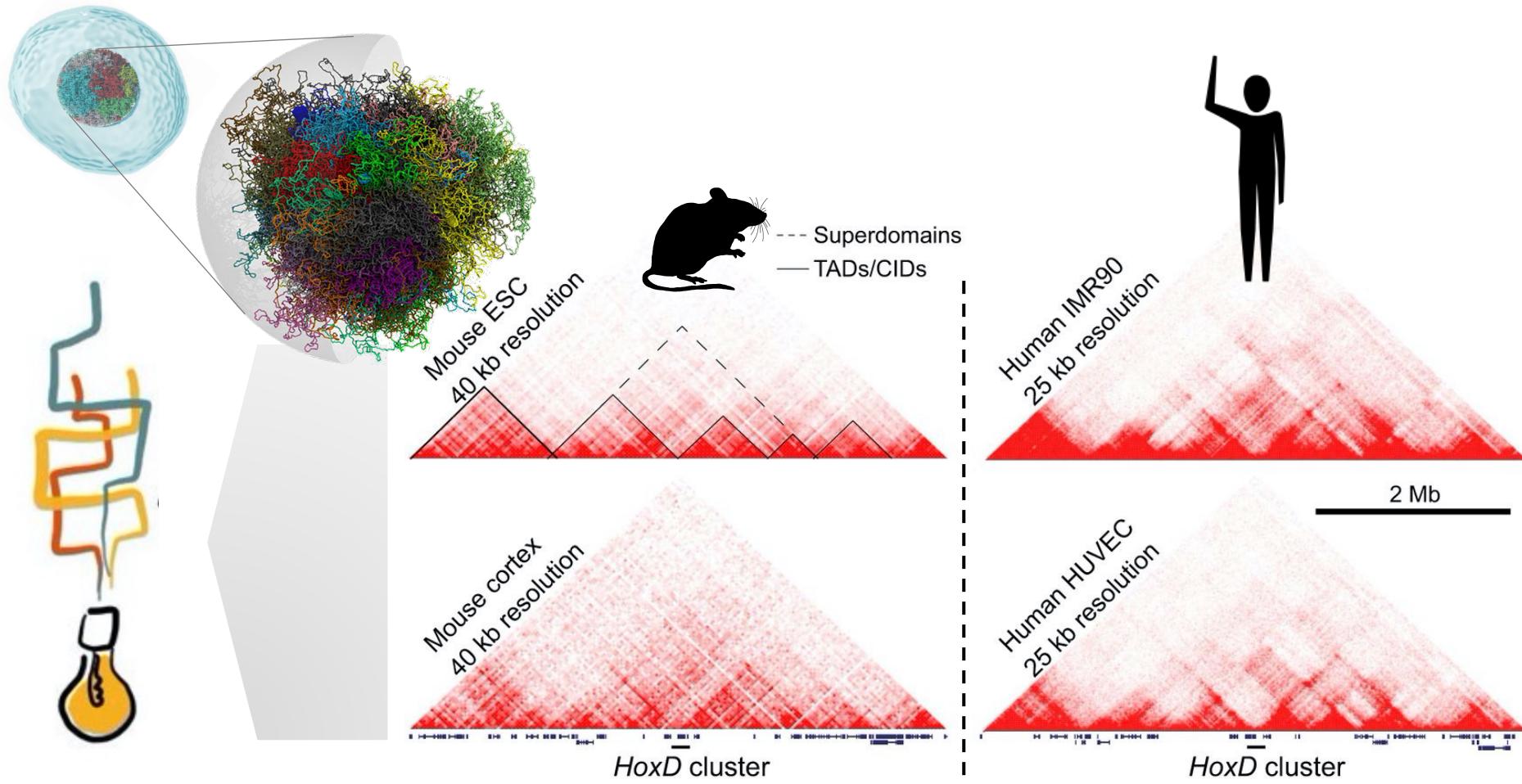
Element locations are conserved because it works well



Modified from image by Pueschel et al. 2016

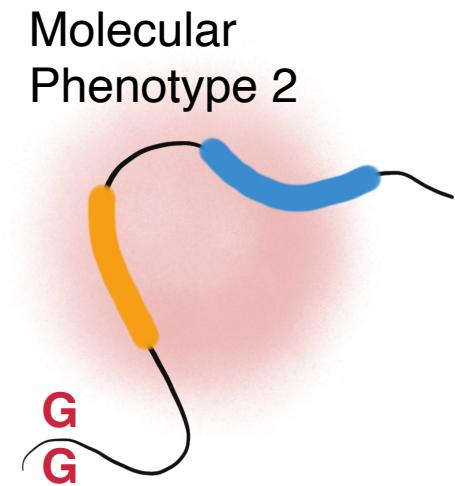
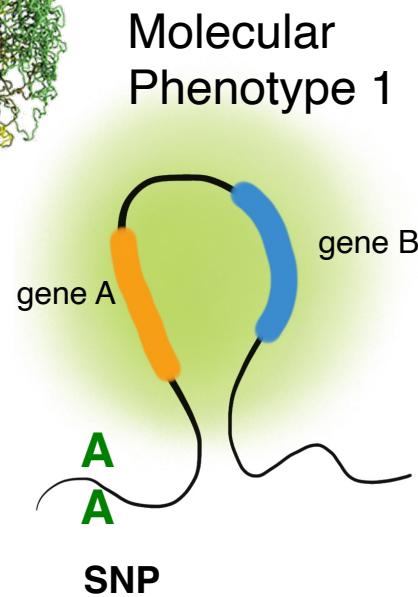
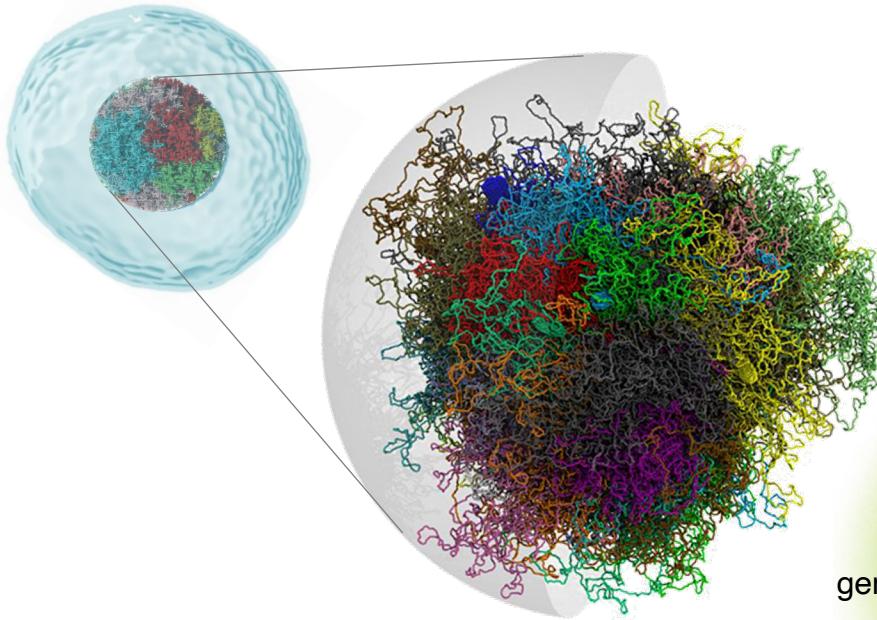
Evolution shapes Chromatin Organization

Element locations are conserved because it works well



Modified from image by Pueschel et al. 2016

Genetic variation shapes regulation

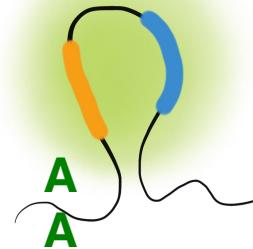


Cis-acting effects:

- histone QTLs
- Expression QTLS

A catalogue of tissue-specific co-expression QTLs

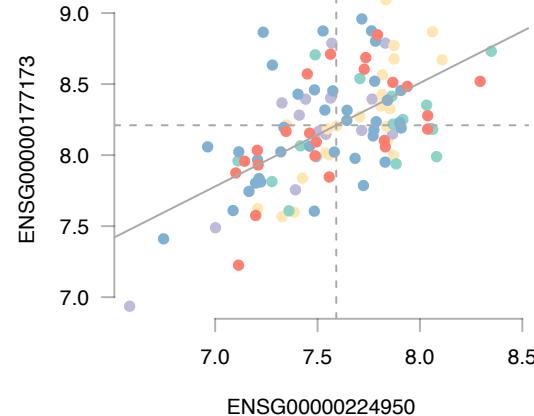
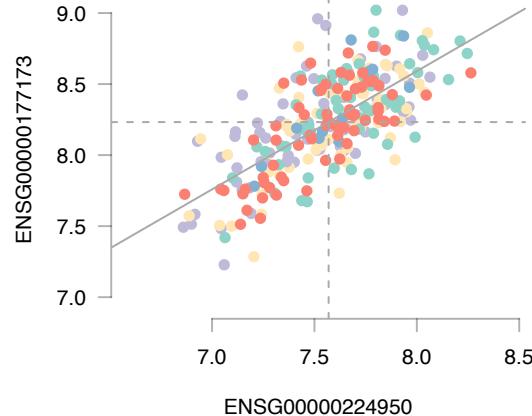
High gene co-expression



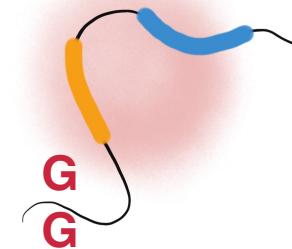
gEUVADIS and GTEx Consortium

CEU YRI GBR FIN TSI

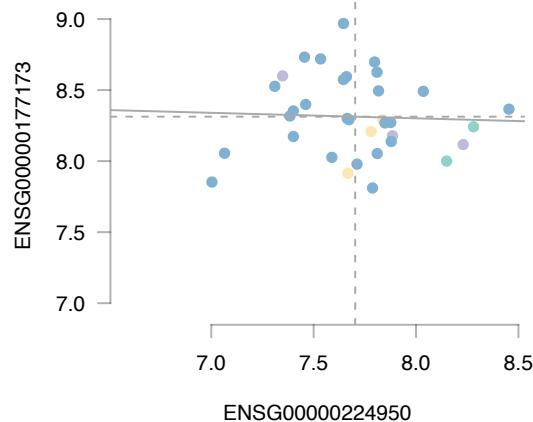
Homozygous A



Loss of gene co-expression



Homozygous G



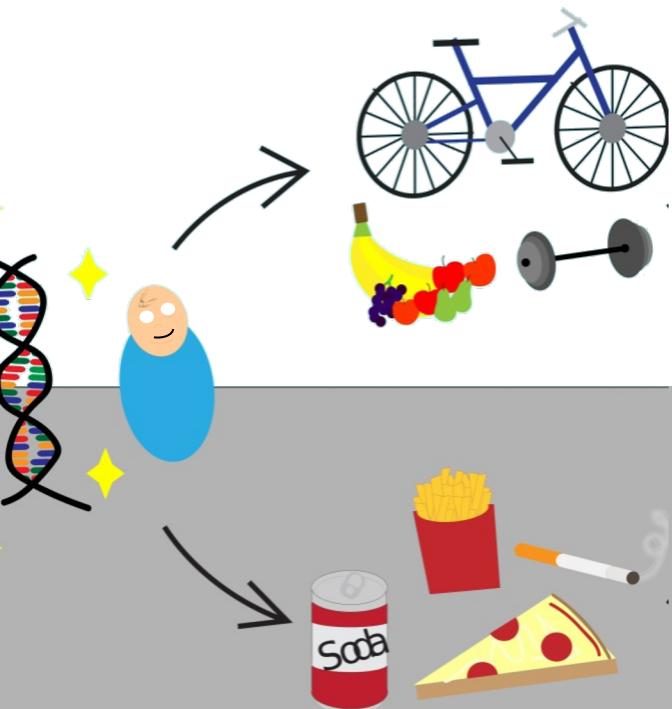
Relationship between chromatin architecture and local genome variation



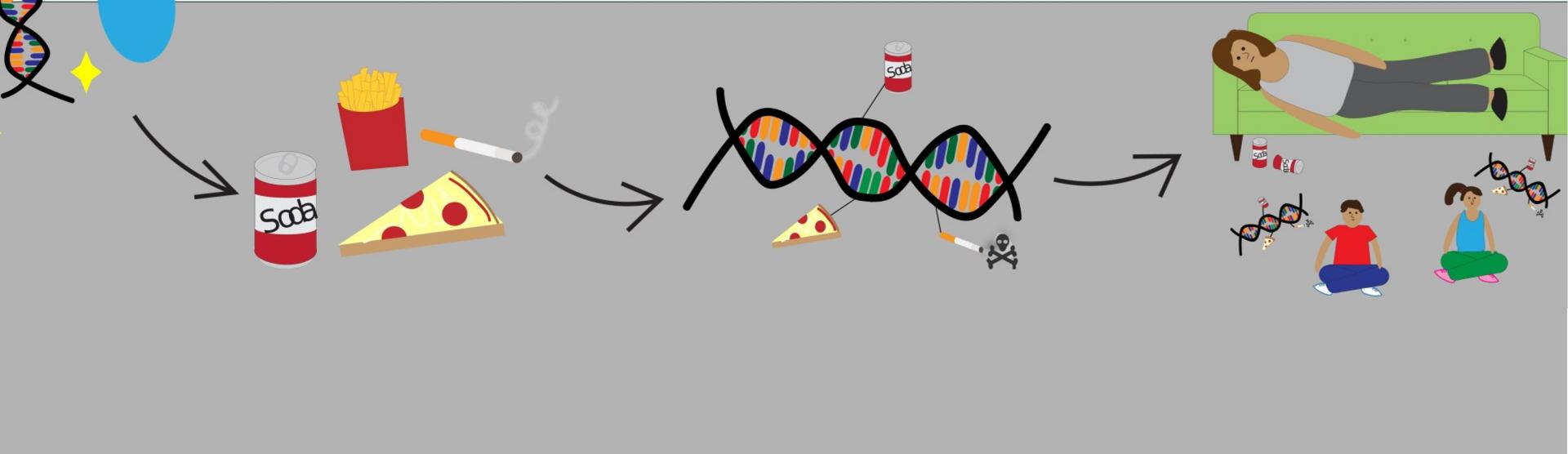
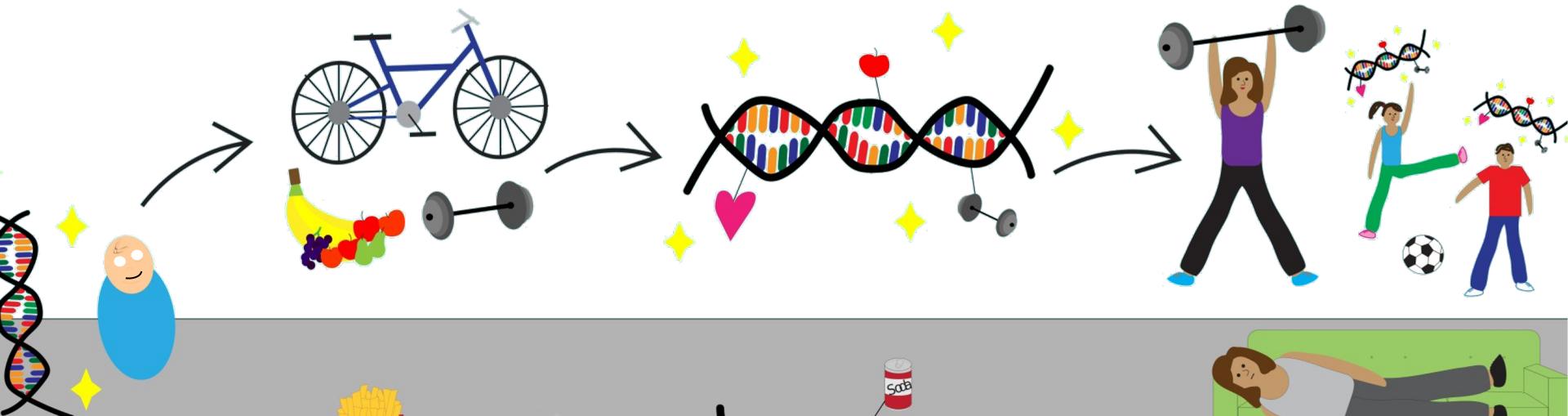
UNIVERSIDAD
COMPLUTENSE
MADRID

armandorp@ucm.es

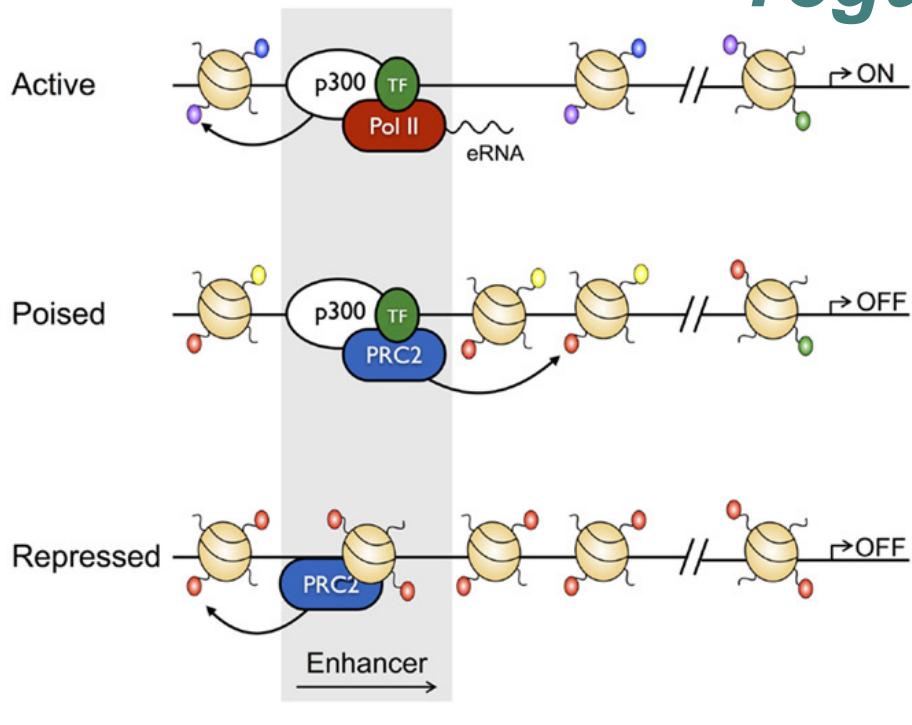
Nurture also shapes regulation



Nurture also shapes our Nature



Histone modifications are chromatin regulators



- H3K27ac
- H3K4me3
- H3K4me1

ACTIVE

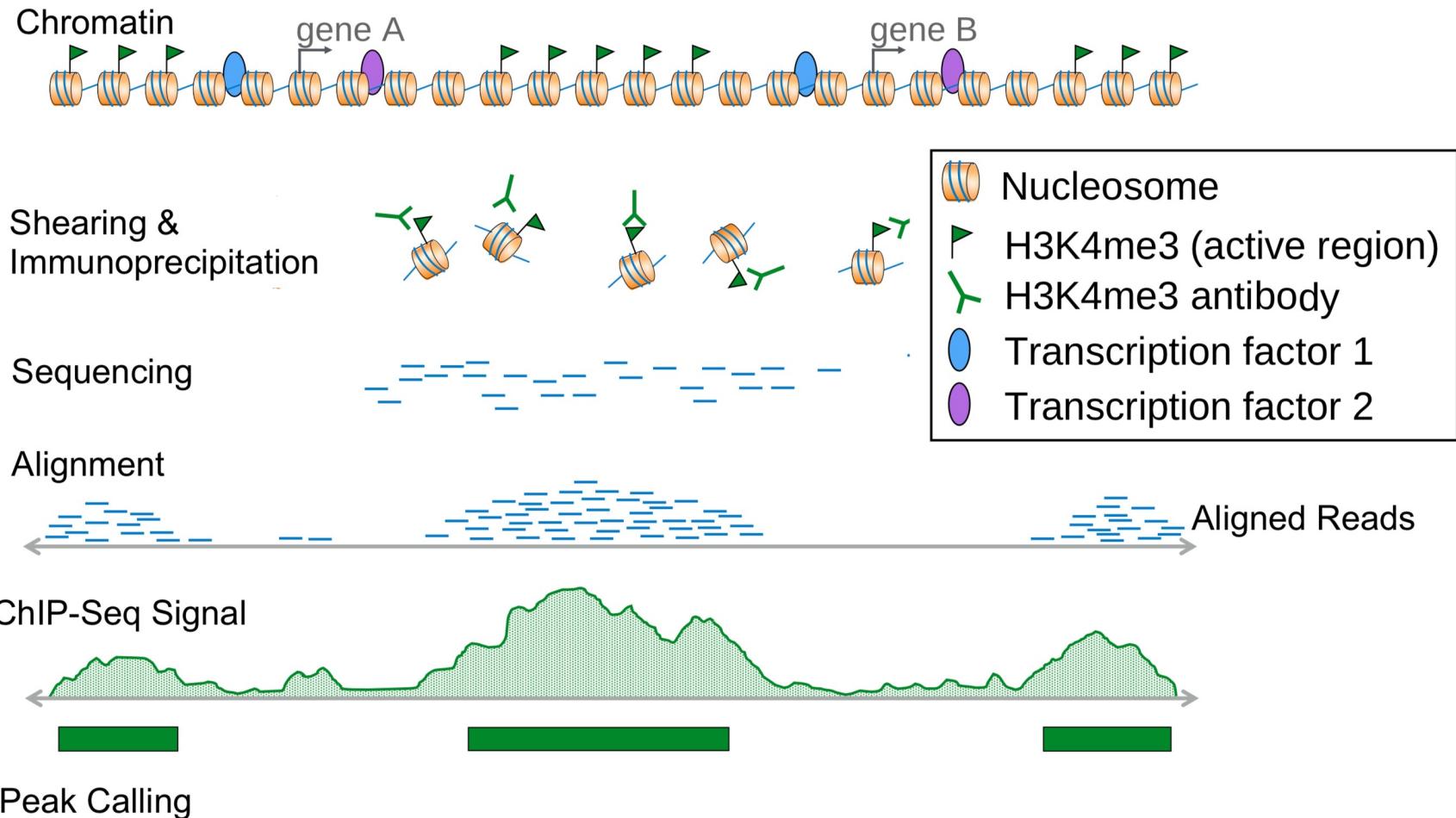
- H3K27me3
- H3K9me3

REPRESSIVE

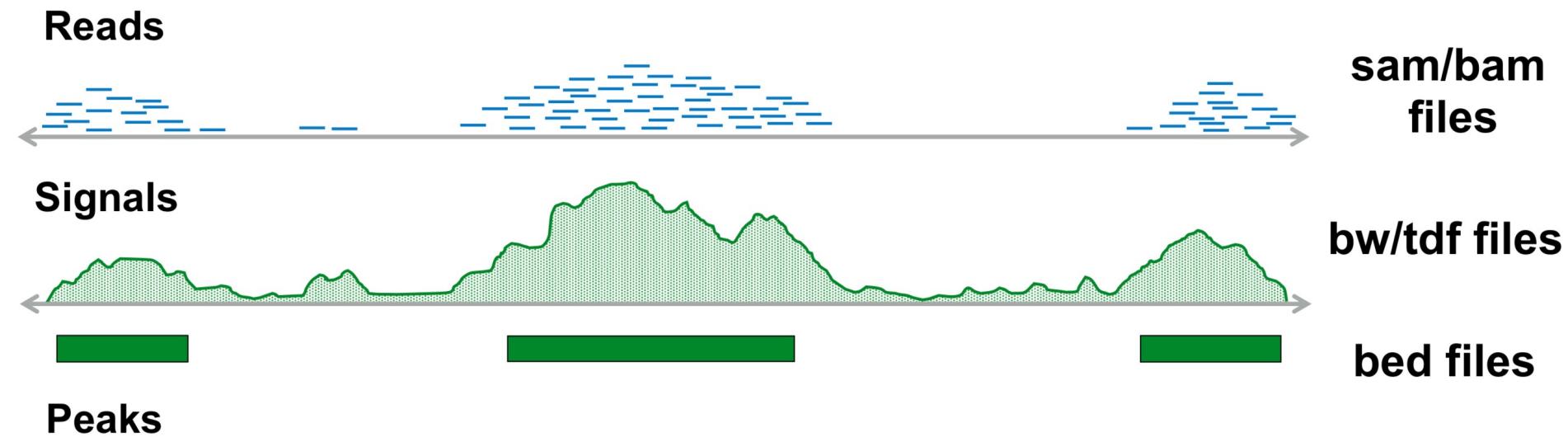
Legend:
○ H3K4me1 ○ H3K4me2 ○ H3K4me3 ○ H3K27ac ○ H3K27me3

ChIP-seq

Chromatin Immunoprecipitation sequencing



NGS files from a ChIP-seq pipeline

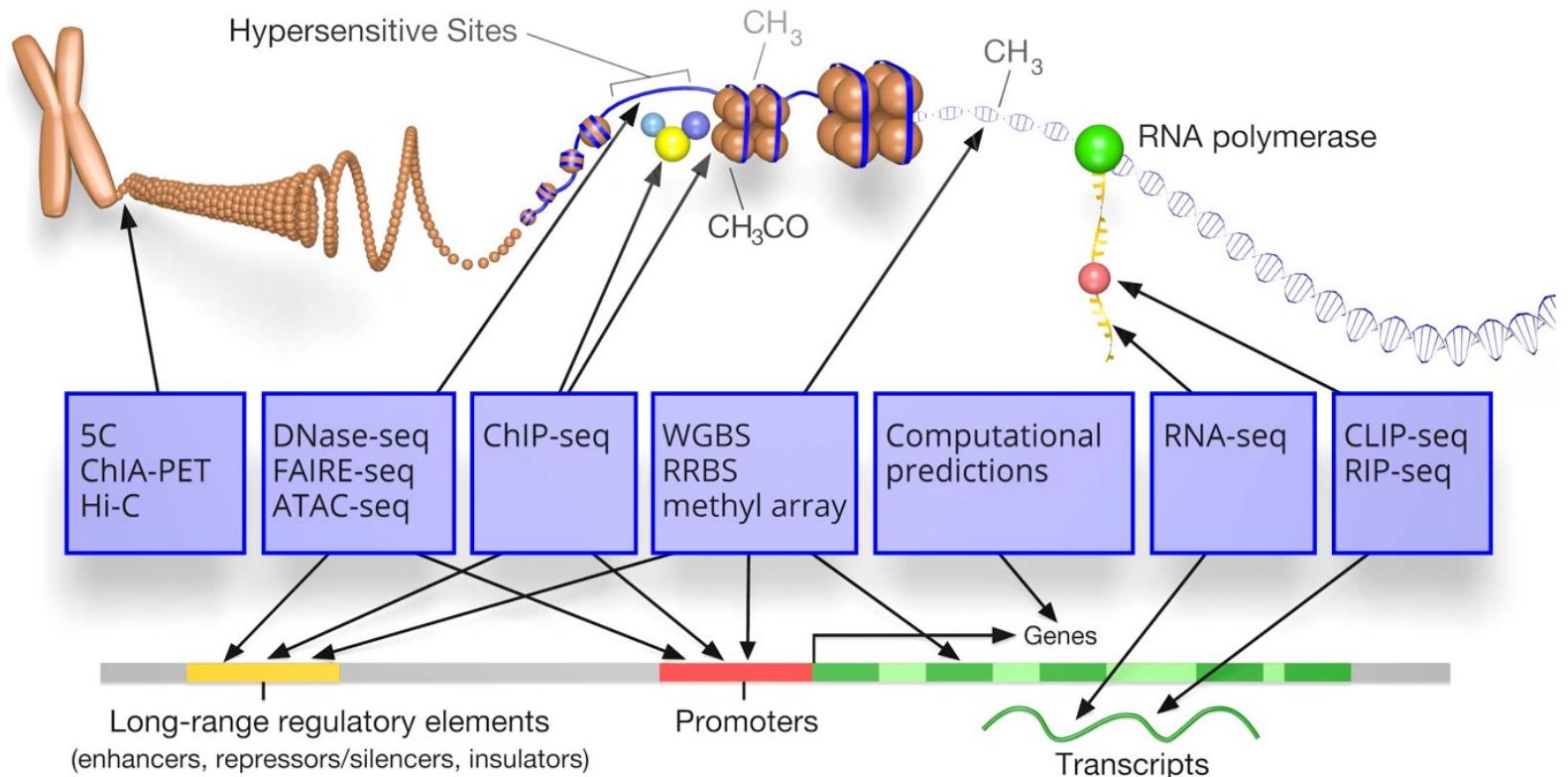


BED files (Genomic Ranges)

Chromosome	Start	End	Name	Score	Strand
chr7	127471196	127472363	Peak1	0	+
chr7	127472363	127473530	Peak2	0	+
chr7	127473530	127474697	Peak3	0	+
chr7	127474697	127475864	Peak4	0	+
chr7	127475864	127477031	Peak5	0	-
chr7	127477031	127478198	Peak6	0	-
chr7	127478198	127479365	Peak7	0	-
chr7	127479365	127480532	Peak8	0	+
chr7	127480532	127481699	Peak9	0	-

What is Cell Epigenome?

ENCODE Project Phase 3



Based on an image by Darryl Leja (NHGRI), Ian Dunham (EBI), Michael Pazin (NHGRI)

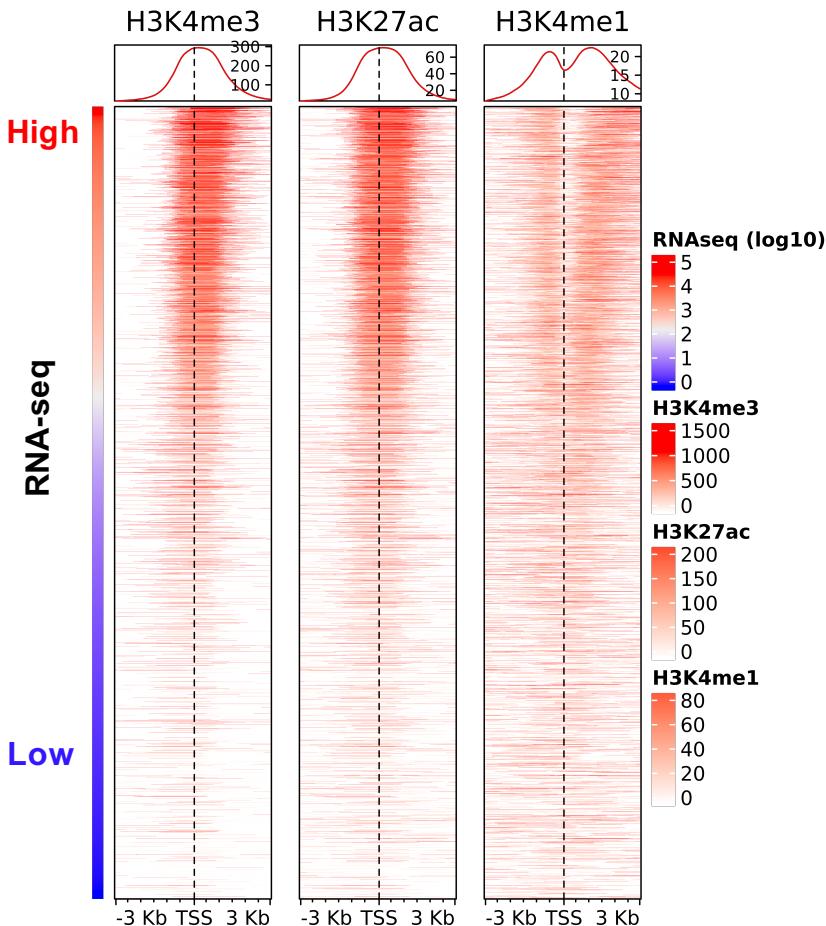


Human Epigenome in numbers

- 9 Histone modifications**
- ~1,000 Transcription factors**
- ~100 Cell types**
- ~1 RNA-seq per cell type**
- x 2 GB per assay → fastQ**

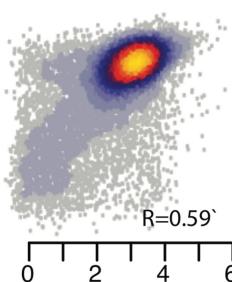
Raw Data (x2):
206 TB per person

Histone modifications correlate with RNA-seq

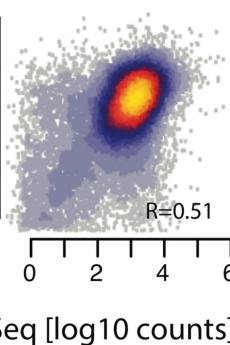


histone marks
[log₁₀ counts]

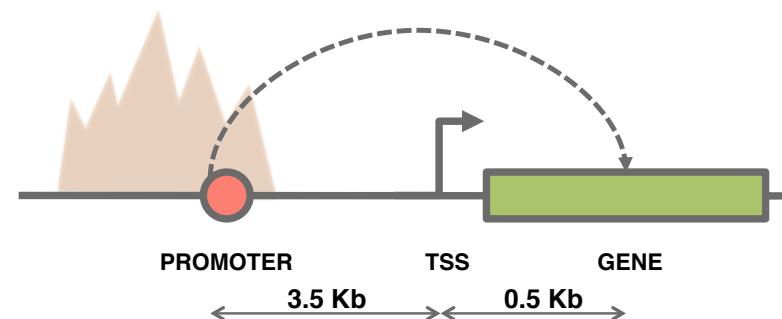
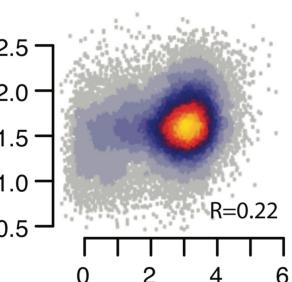
H3K4me3



H3K27ac



H3K4me1



Reyes-Palomares et al. Nature Comm.

Gene Regulatory Genomics

