



An Experimental Model for Studying Promoter-Enhancer Regulatory Associations in Vivo

Giulio Pavesi

Bioinformatics, Evolution and Computational Genomics Crew
(BEaCOn)
Department of Biosciences
University of Milan



The Starting Point... (a few years ago)

- The traditional genetic or “gene-wise” approach:
 1. “**Inactivate**” a gene (e.g. knock it out or down)
 2. Examine the **phenotype** obtained (i.e. see what happens in “mutants”, hoping they do not die too soon, i.e. before you can perform your experiments on them)
 3. Infer a role/**function** for the gene (e.g. which processes it seems to be essential for or “involved in”)



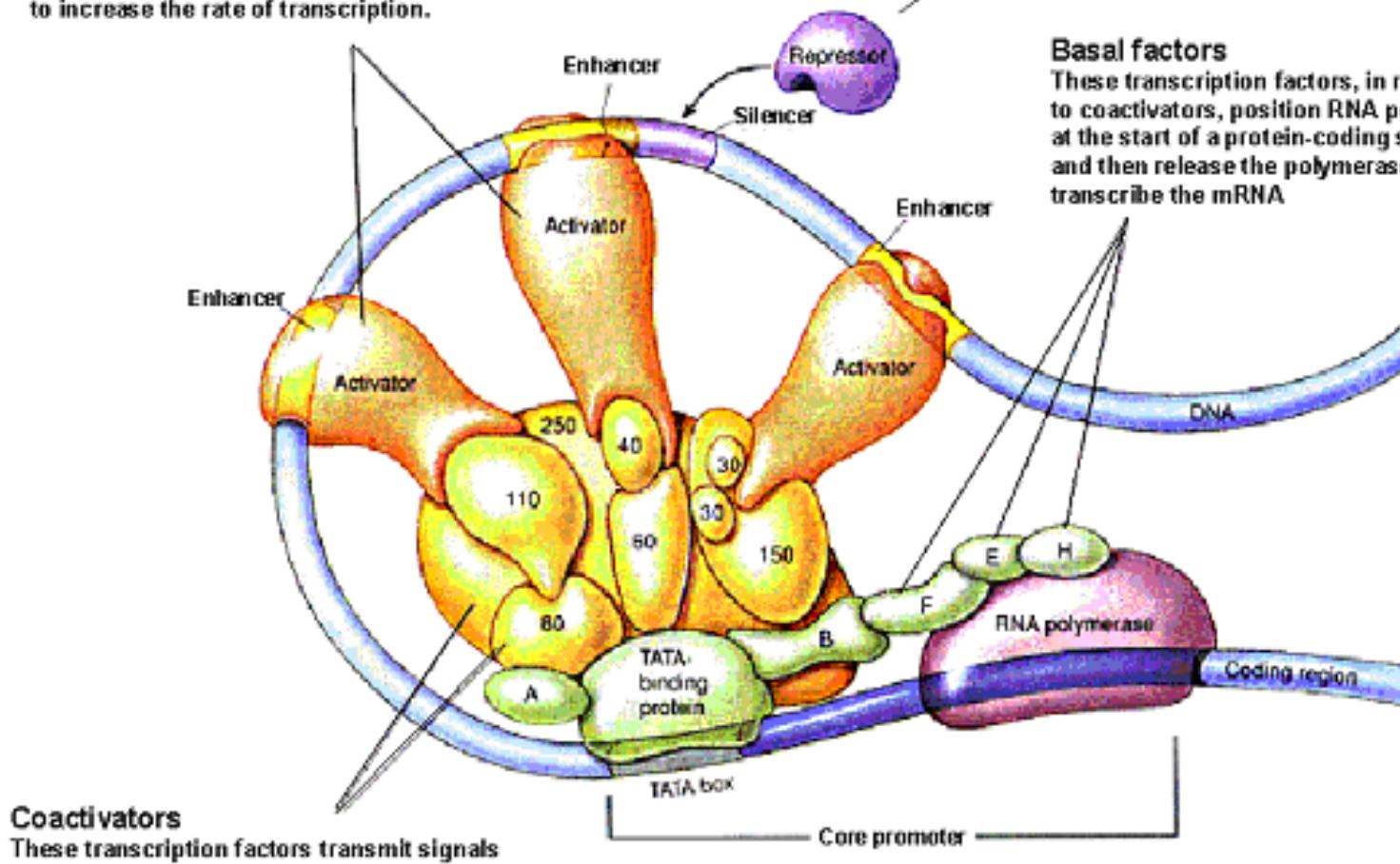
The Starting Point: Our Gene

- **SOX2:** a **transcription factor** that is essential for maintaining self-renewal, or pluripotency, of undifferentiated embryonic stem cells. Sox2 has a critical role in maintenance of **embryonic and neural stem cells (NSC)**
- SOX2 mutations cause genetically dominant nervous system disease, involving hippocampus and eye defects, epilepsy and learning disabilities
- Experimental setup: **NSC cell lines** established from forebrains at P0 (birth) wild type and Sox2-deleted (at the embryo stage) mice



Activators

The regulatory proteins bind to DNA at distant sites known as enhancers. When DNA folds so that the enhancer is brought into proximity with the transcription complex, the activator proteins interact with the complex to increase the rate of transcription.



Repressors

These regulatory proteins bind to "silencer sites" on the DNA preventing the binding of activator to nearby enhancers and so slowing transcription.

Basal factors

These transcription factors, in response to coactivators, position RNA polymerase at the start of a protein-coding sequence, and then release the polymerase to transcribe the mRNA.

Coactivators

These transcription factors transmit signals from activator proteins to the basal factors.



From **gene-wise** to **genome-wide**

- **SOX2** is a transcription factor - **activator** (but also reported as repressor) **of gene transcription**, so:
 - map its binding sites on the genome with **ChIP-Seq**
 - characterize its co-localization with histone modifications (ChIP-Seq for **H3K27ac** and **H3K4me1**)
 - do histone modifications change in mutant vs wild type cells?
 - Assess the impact on gene transcription of its inactivation (**RNA-Seq** in NSC from wild-type vs. mutant mice)



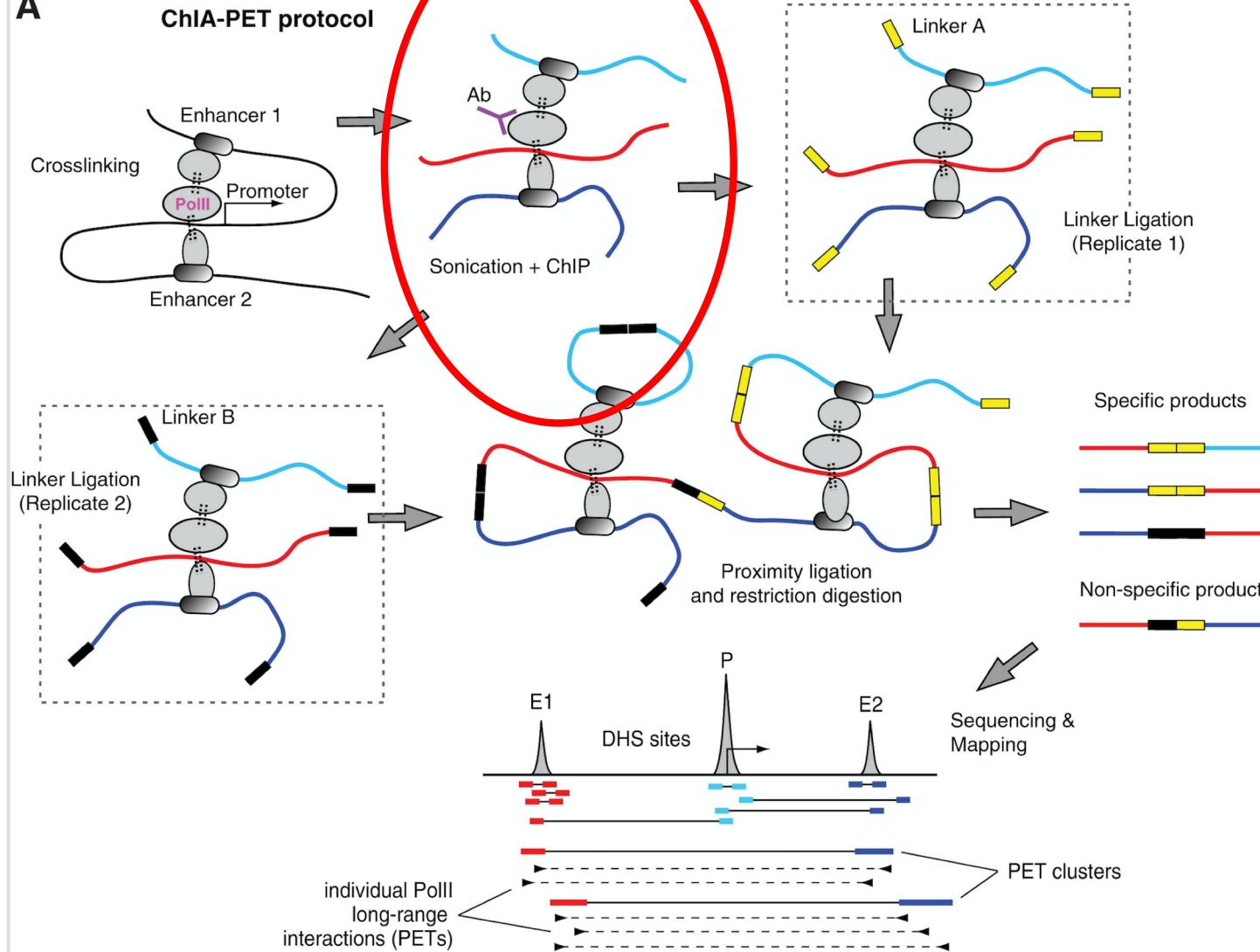
Genome-wide studies

- Determine **long-range chromatin interactions**, focusing on promoter-enhancer and promoter-promoter interactions, **in wild type and mutant NSC cells**: replicates with ChIA-PET and improved *in situ* ChIA-PET (latter requested by reviewers...)
- See whether the inactivation of SOX2 leads to changes of interactions in mutant mice
- ChIA-PET: **C**hromatin **I**nteraction **A**nalysis by **P**aired-**E**nd **T**ag sequencing



ChIA-PET

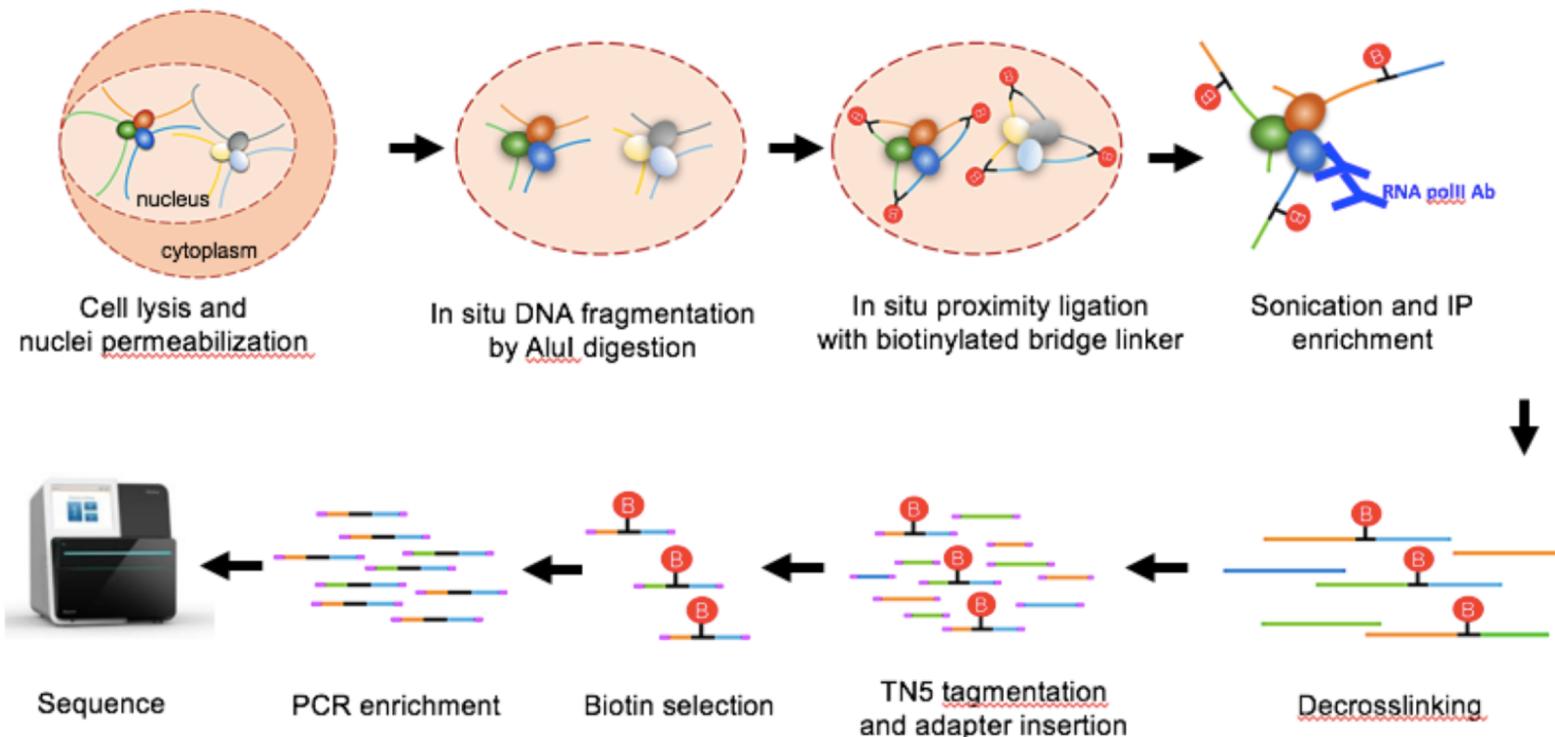
A





In situ ChIA-PET (replicate experiments)

In situ ChIA-PET workflow





ChIA-PET: the starting point (Zhang et al., Nature 2013)

LETTER

doi:10.1038/nature12716

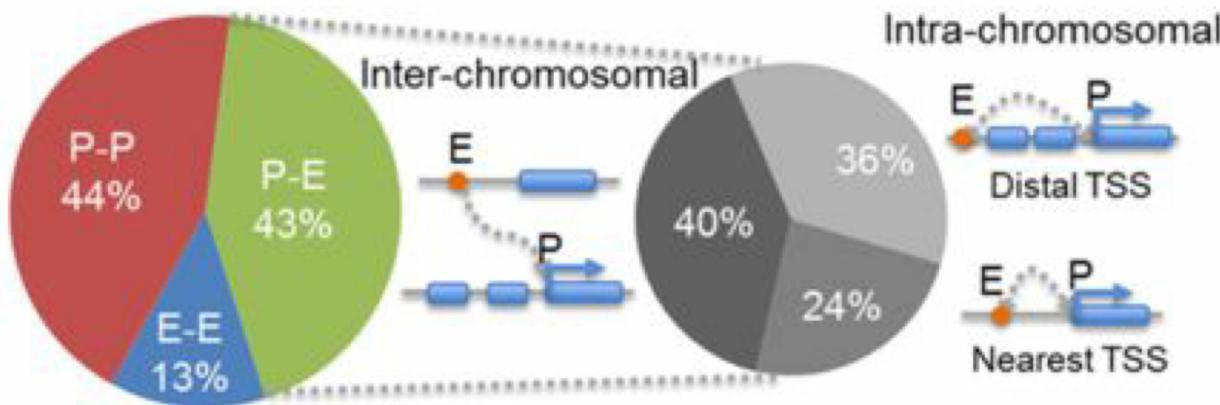
Chromatin connectivity maps reveal dynamic promoter–enhancer long-range associations

Yubo Zhang^{1*†}, Chee-Hong Wong^{1*}, Ramon Y. Birnbaum^{2*†}, Guoliang Li^{3,4}, Rebecca Favaro⁵, Chew Yee Ngan¹, Joanne Lim⁴, Eunice Tai⁴, Huay Mei Poh⁴, Eleanor Wong⁴, Fabianus Hendriyan Mulawadi⁴, Wing-Kin Sung⁴, Silvia Nicolis⁵, Nadav Ahituv², Yijun Ruan³ & Chia-Lin Wei^{1,4}

In multicellular organisms, transcription regulation is one of the central mechanisms modelling lineage differentiation and cell-fate determination¹. Transcription requires dynamic chromatin configurations between promoters and their corresponding distal regulatory elements². It is believed that their communication occurs within large discrete foci of aggregated RNA polymerases termed transcription factories in three-dimensional nuclear space³. However, the dynamic nature of chromatin connectivity has not been characterized at the genome-wide level. Here, through a chromatin interaction analysis with paired-end tagging approach^{3–5} using an antibody that primarily recognizes the pre-initiation complexes of RNA polymerase II⁶, we explore the transcriptional interactomes of three mouse cells of progressive lineage commitment, including pluripotent embryonic stem cells⁷, neural stem cells⁸ and neurosphere stem/progenitor cells⁹. Our global chromatin connectivity maps reveal approximately 40,000 long-range interactions, suggest precise enhancer–promoter associations and delineate cell-type-specific

chromatin structures. Analysis of the complex regulatory repertoire shows that there are extensive colocalizations among promoters and distal-acting enhancers. Most of the enhancers associate with promoters located beyond their nearest active genes, indicating that the linear juxtaposition is not the only guiding principle driving enhancer target selection. Although promoter–enhancer interactions exhibit high cell-type specificity, promoters involved in interactions are found to be generally common and mostly active among different cells. Chromatin connectivity networks reveal that the pivotal genes of reprogramming functions are transcribed within physical proximity to each other in embryonic stem cells, linking chromatin architecture to coordinated gene expression. Our study sets the stage for the full-scale dissection of spatial and temporal genome structures and their roles in orchestrating development.

Gene regulatory networks are organized by spatial connectivity between distal regulatory elements (DREs) and their corresponding promoters³. Many of these DREs, including cell-specific enhancers, were

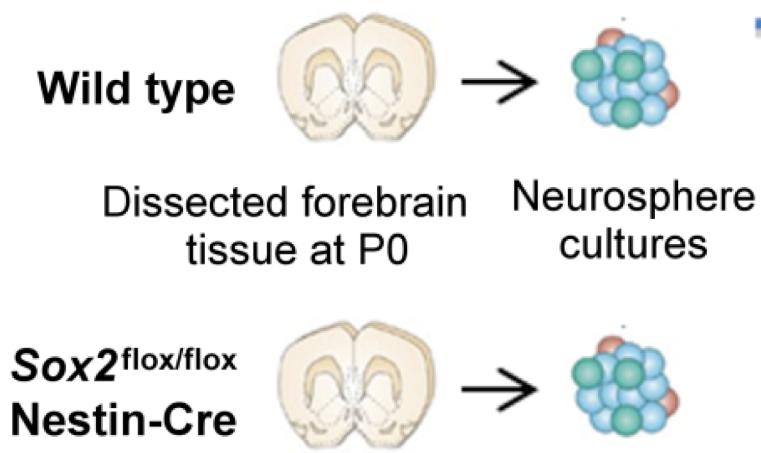
a

Types of long distance interactions detected through ChIA-Pet

Of the promoter-enhancer (green) ones:

- 24% are actually of the “nearest gene” type
- 36% of enhancers do not interact with the nearest gene
- 40% are even on different chromosomes!!!!
- (Zhang et. al. Nature. 2013 Dec 12;504(7479):306-10)

All in all

A**ChIA-PET:**

identification of long-range interactions in chromatin in wt vs *Sox2*-deleted neurospheres.

SOX2 ChIP-seq (on wt):

identification of SOX2-bound sites on wt neurospheres.

H3K27ac, H3K4me1 ChIP-seq:

on wt vs *Sox2*-deleted neurospheres.

RNA-seq:

identification of genes differentially expressed in wt vs *Sox2*-deleted neurospheres.

C



SOX2 ChIP-Seq

- Performed on wild-type NSC cells in two replicates
- 18,359 “bona fide” binding sites:
 - < 10% on promoters
 - > 90% on distal elements
(enhancers?)

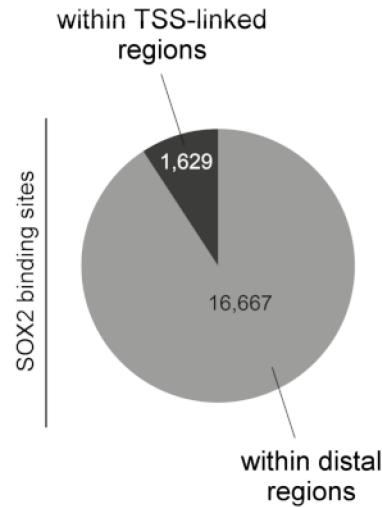
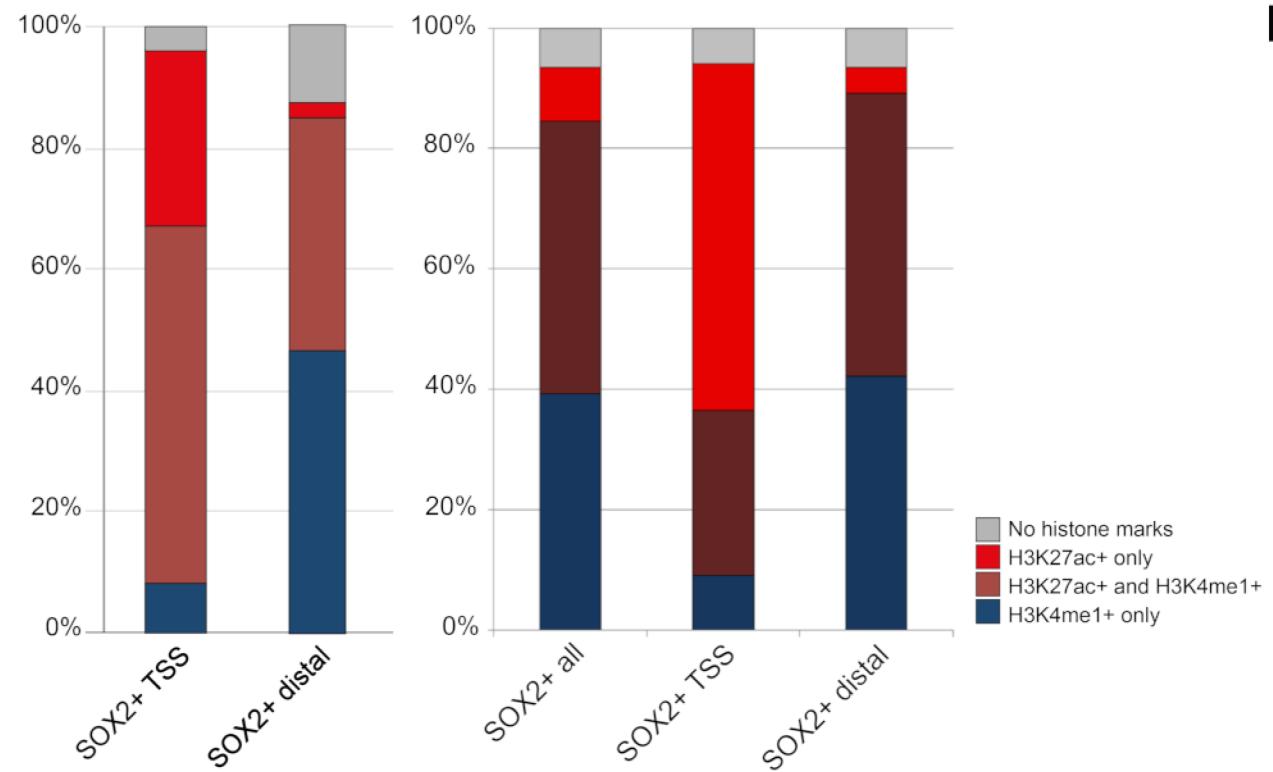


Histone Modifications ChIP-Seq

- Performed on wild-type and mutant NSC cells
- Two replicates in each for H3K4me1 and H3K27ac
- **H3K4me1+H3K27ac** mark transcriptionally active TSSs (when found together around TSS) or active enhancers (when found together on distal regions)



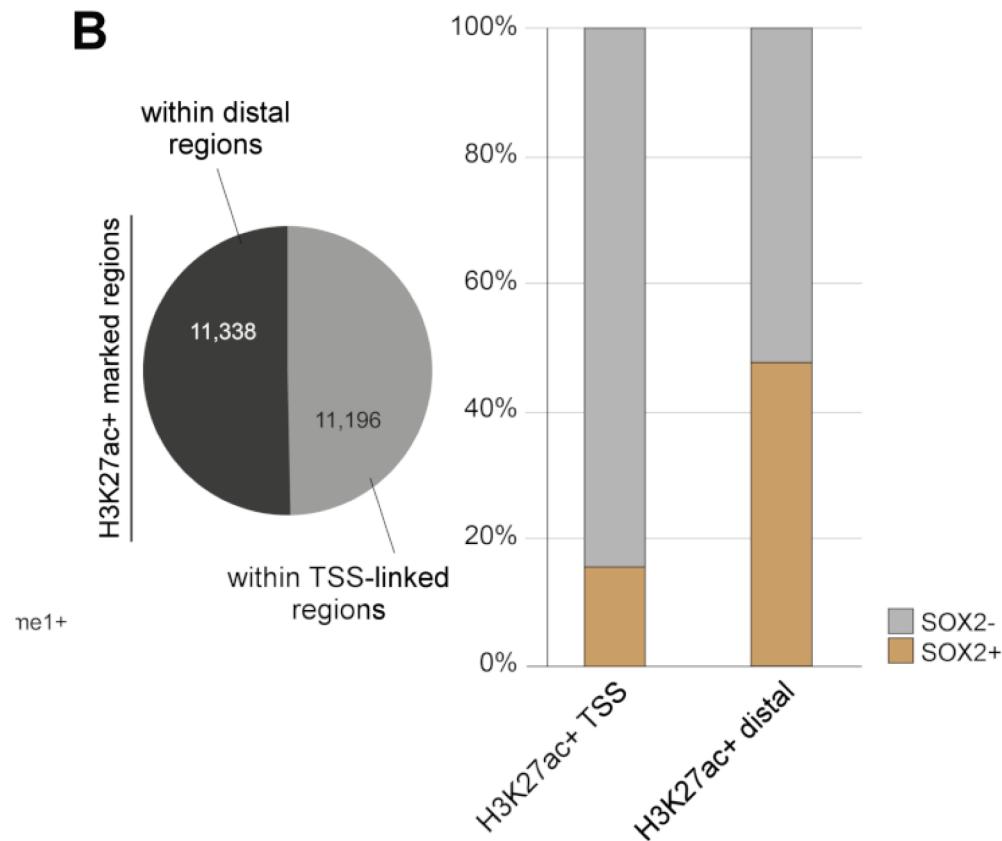
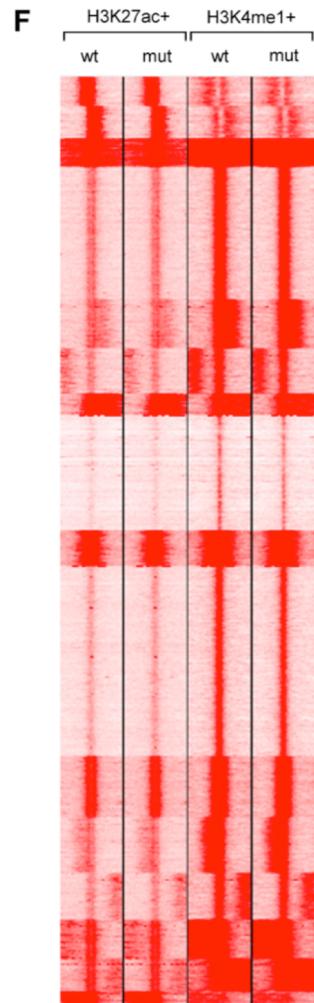
SOX2 binding and histone modifications

A**E**

>90% of binding sites found in active chromatin signatures



SOX2 binding and histone modifications



about 50% of “enhancers”
(according to histone modifications)
are bound by SOX2

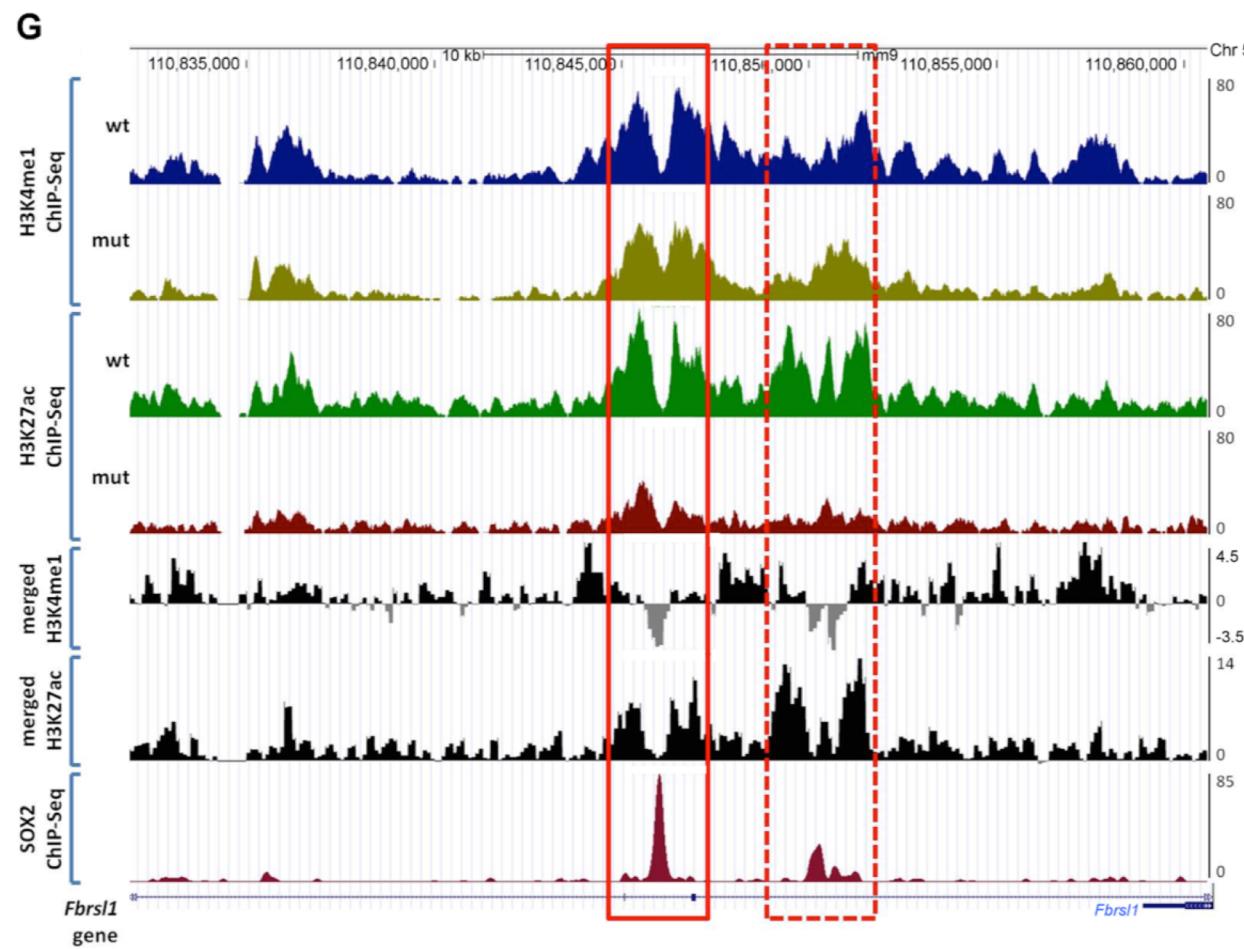


Histone modifications wt vs mut

- SOX2 binds regions with active chromatin signature (active promoters or enhancers)
- Histone modification ChIP-Seq processed both by peak calling and ChromHMM segmentation
- **No relevant change detected in the comparison of wild-type vs mutant cells**
 - only minor local changes in peak shape and height around SOX2 binding sites



Histone modifications wt vs mut



ChIA-PET wild-type vs mutant

C

	wTR1	mTR1	wTR2	mTR2	wTR3	mTR3
prom-prom interactions	3185 (45%)	1270 (43%)	7159 (40%)	1476 (51%)	3221 (44%)	1379 (43%)
prom-non prom interactions	3123 (44%)	1260 (42%)	8199 (45%)	1124 (39%)	3234 (44%)	1412 (44%)
prom-intragenic interactions						
non-promoter interactions	751 (11%)	454 (15%)	2655 (15%)	278 (10%)	886 (12%)	411 (13%)
total interactions	7059	2984	18013	2878	7341	3202

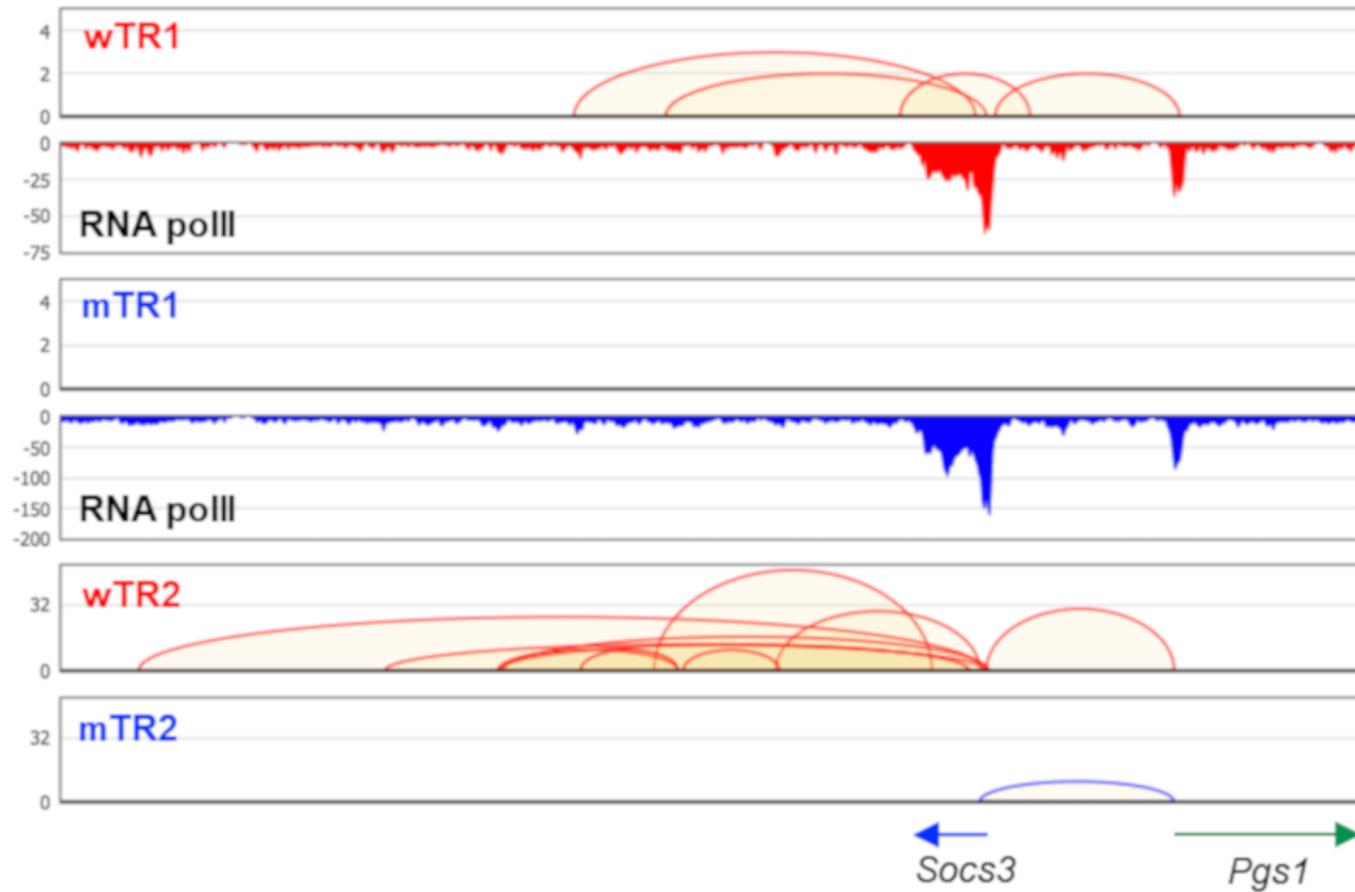
Diagram illustrating interaction types between genes X and Y:

- prom-prom interactions:** Two promoters (red dots) on the same DNA strand interact with each other.
- prom-non prom interactions:** A promoter (red dot) interacts with a non-promoter region (green box).
- prom-intragenic interactions:** A promoter (red dot) interacts with another promoter (red dot) located on the same gene (either gene X or gene Y).
- non-promoter interactions:** Two non-promoter regions (green boxes) interact with each other.

ChIA-PET wild-type vs mutant

E

chr11:117736788-117873172

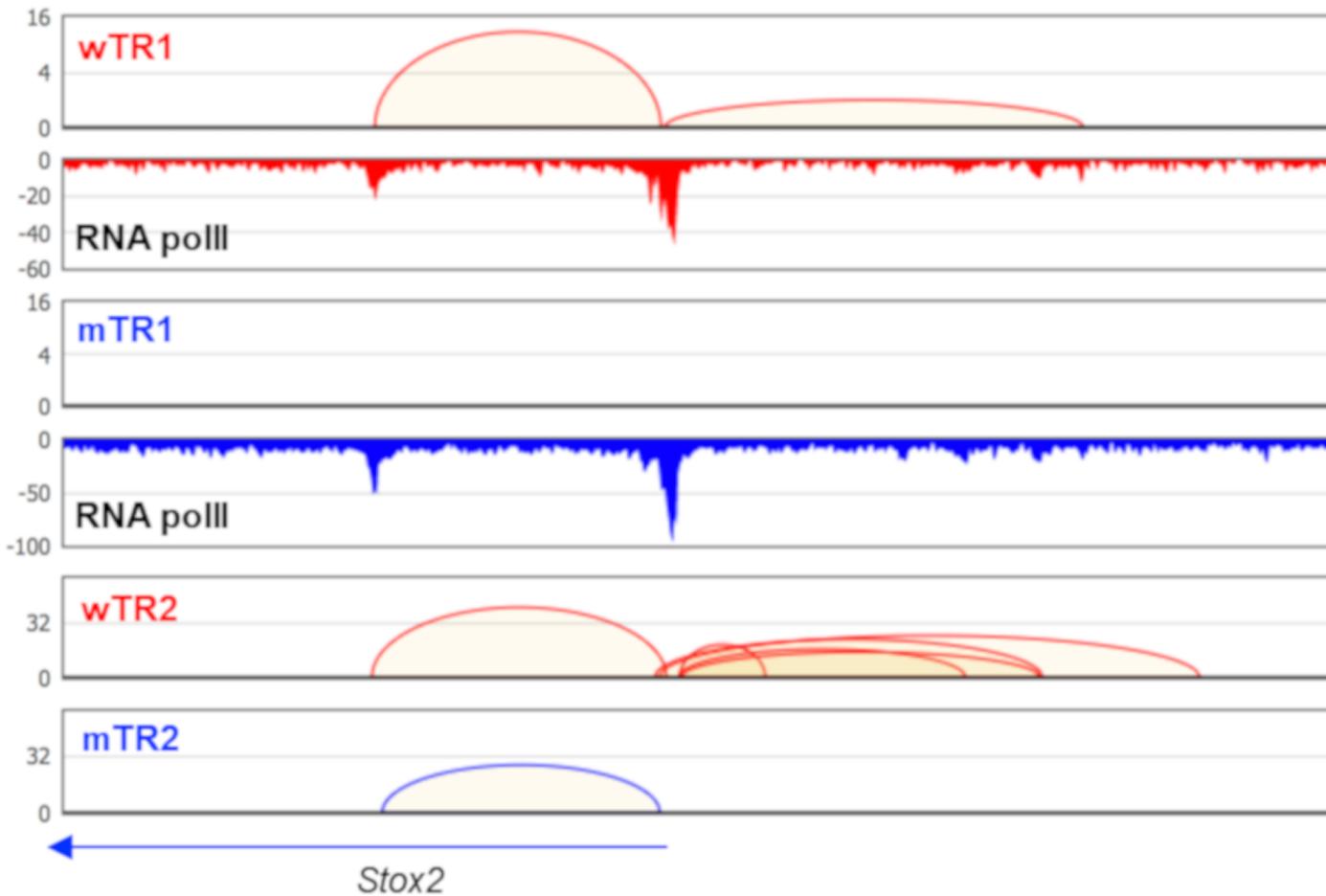




ChIA-PET wild-type vs mutant

G

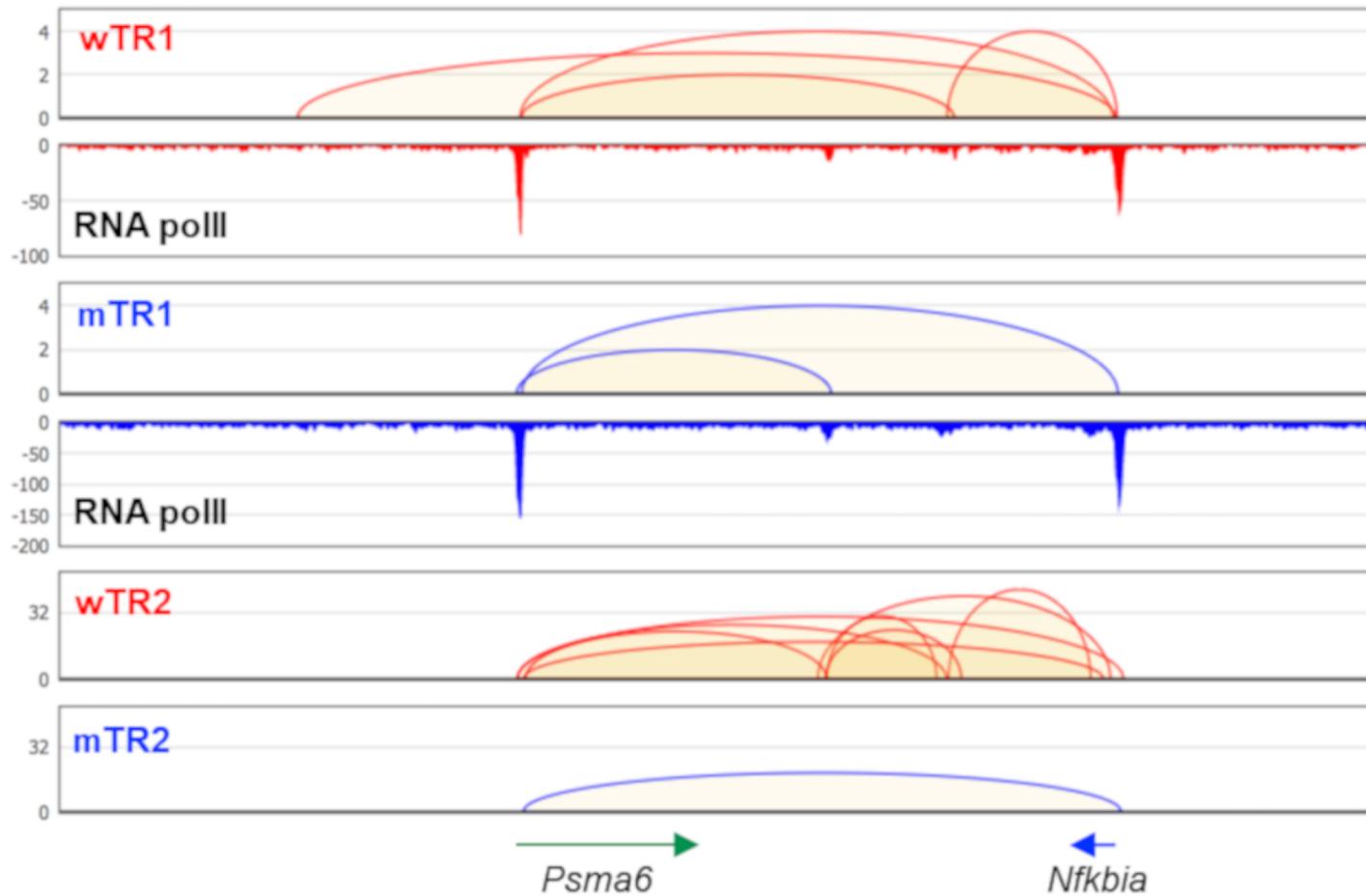
chr8:48254658-48486144



ChIA-PET wild-type vs mutant

F

chr12:56459922-56634834





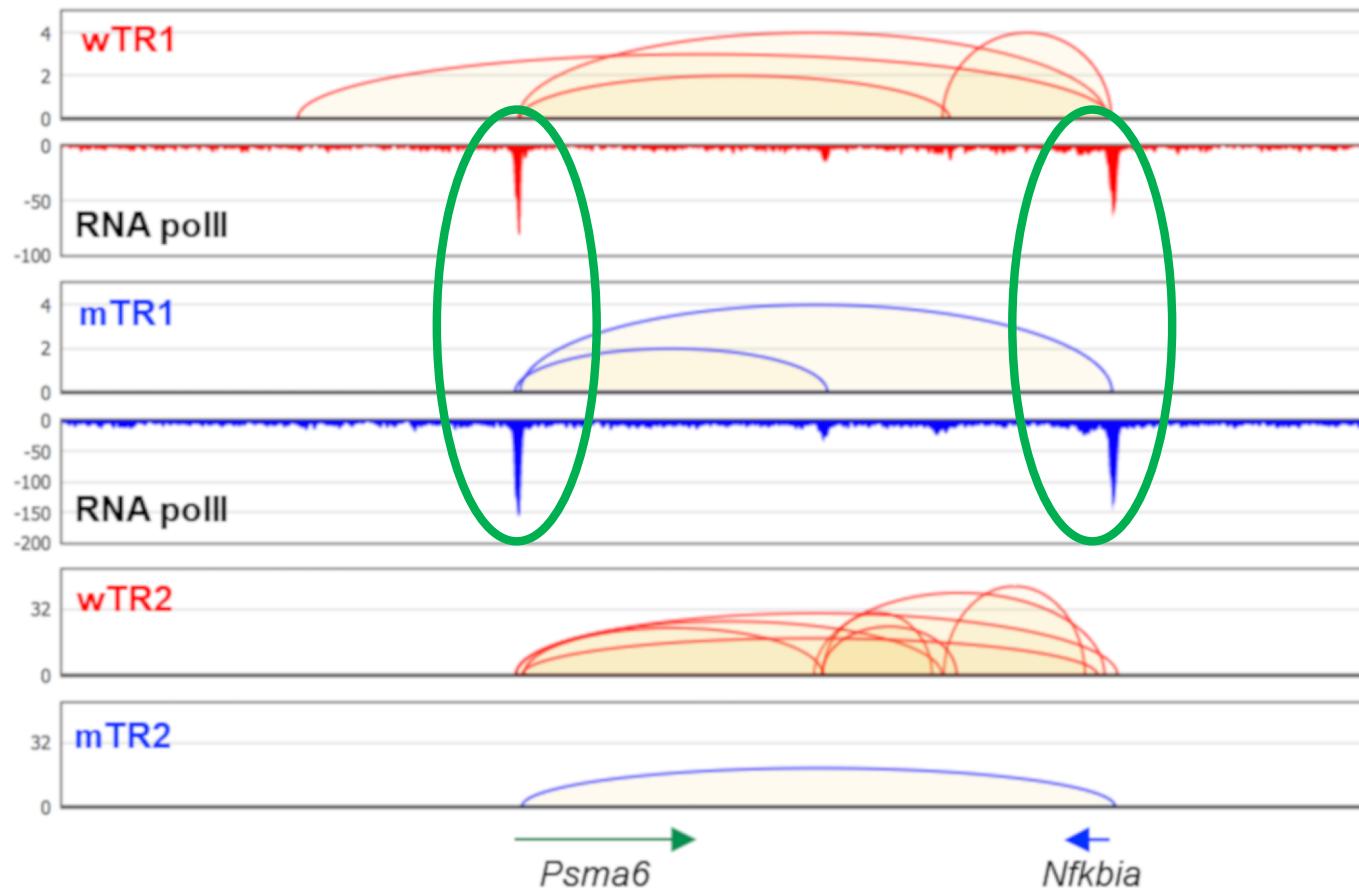
ChIA-PET wild-type vs mutant

- Take home message: **general and relevant reduction** of long-distance polII-mediated interactions in SOX2-depleted cells
- Confirmed in three “replicates”:
 - Original ChIA-PET for NSC vs mutant
 - Two more novel “in situ” ChIA-PET assays in wild type vs mutant
- **Loss of interactions is not due to loss of polII binding:** same polII-enriched regions both in wt and mutant

ChIA-PET wild-type vs mutant

F

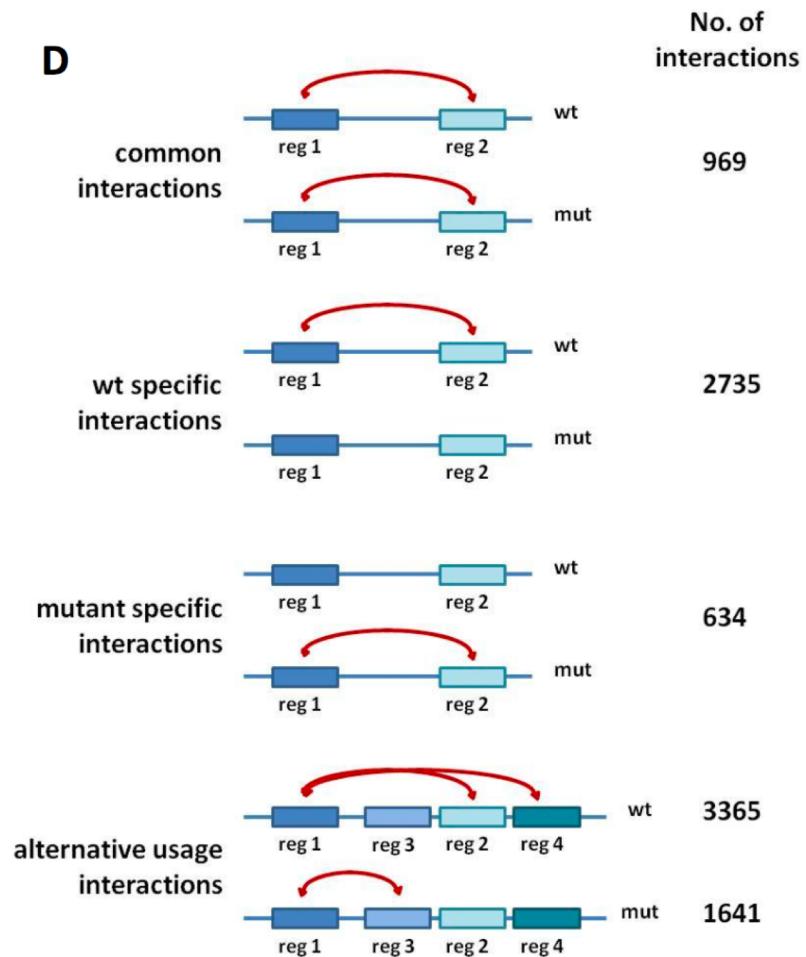
chr12:56459922-56634834





ChIA-PET wild-type vs mutant

- Different type of changes, not only loss of interactions (see right for the original ChIA-PET experiments)
- Some interactions are also gained by mutant cells





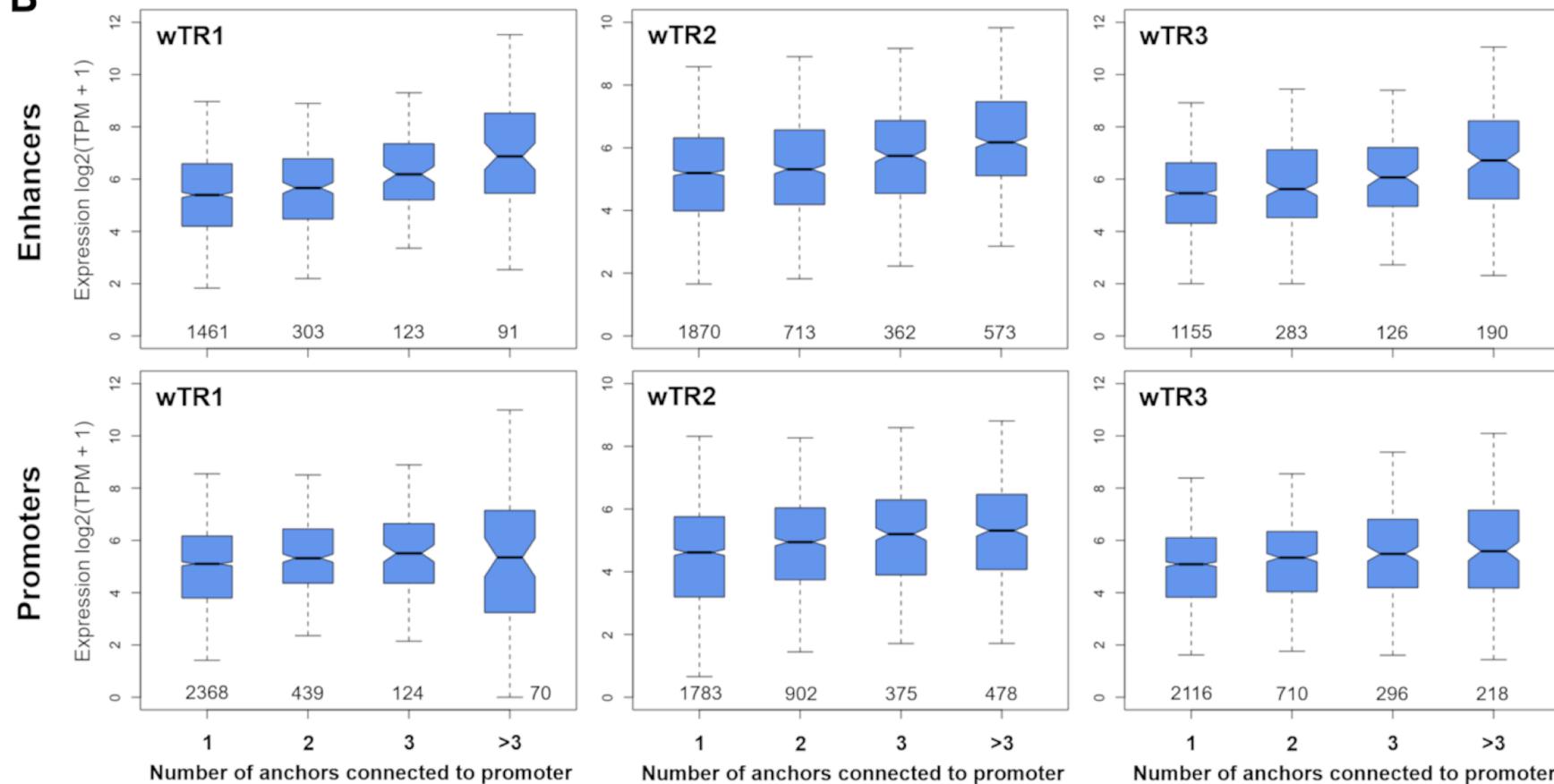
RNA-Seq wild-type vs mutant

- Transcript levels measured by (polyA) RNA-Seq in three wild type vs three mutant “replicates”
- Standard RNA-processing pipeline, quantification as “read counts” normalized in “transcripts per million” (TPM), using RefSeq gene annotation
- Identification of differentially expressed genes by using the “NOISeq” tool



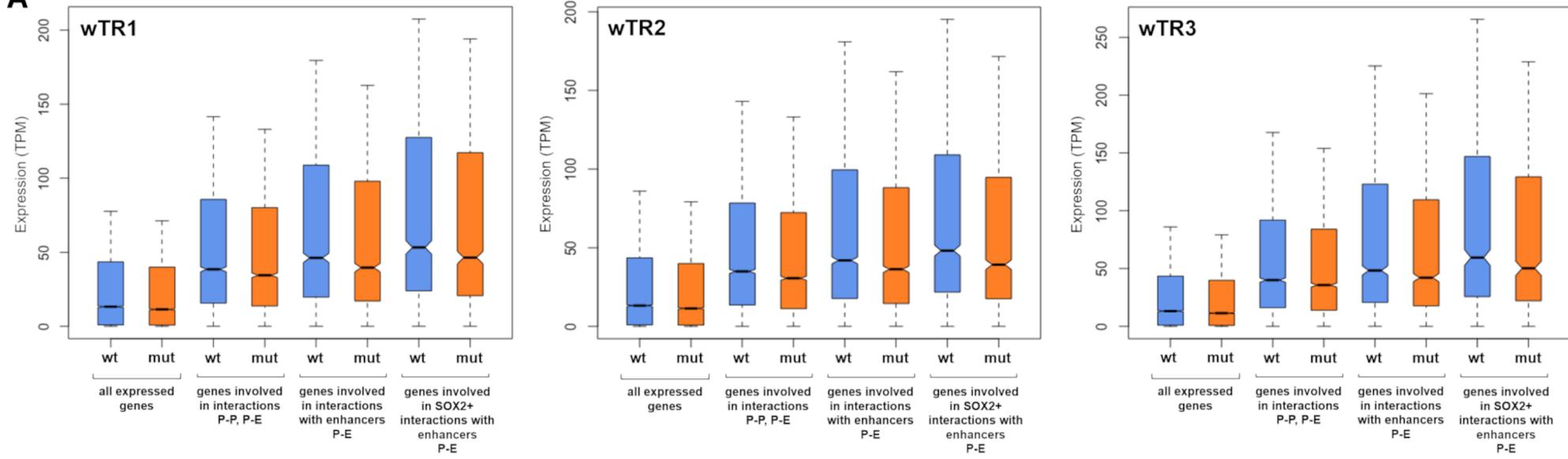
RNA-Seq wild-type vs ChIA-PET

B

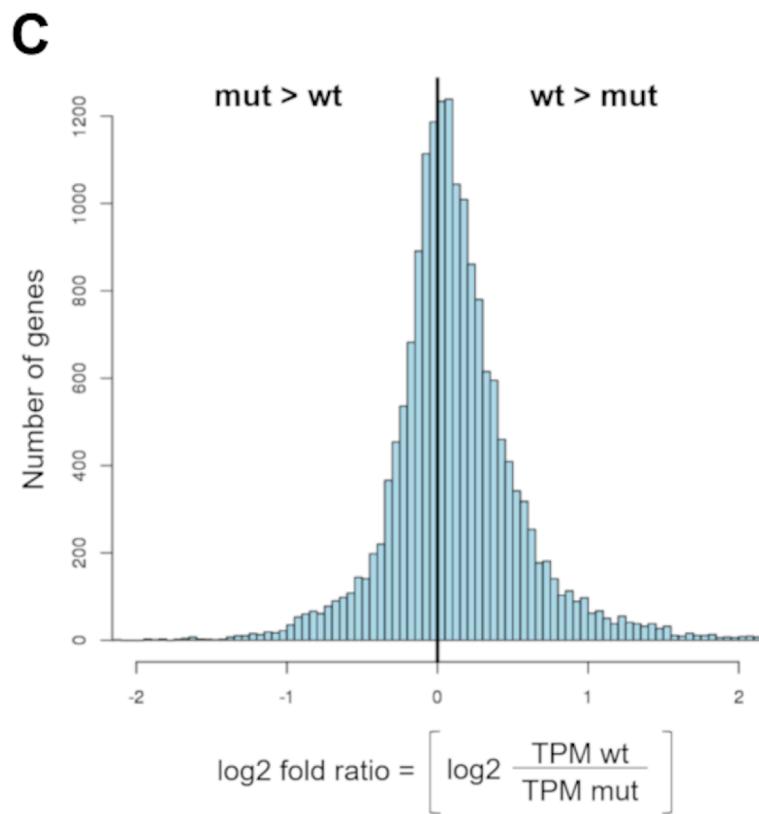




RNA-Seq wild-type vs mutant vs ChIA-PET

A

RNA-Seq wild-type vs mutant



- General trend: reduction of transcript levels in mutant cells for more (most of) the genes
- No gene switched completely “on” or “off” in wt or mut
- Hypothesis: SOX2 inactivation reduces transcript levels of its target genes



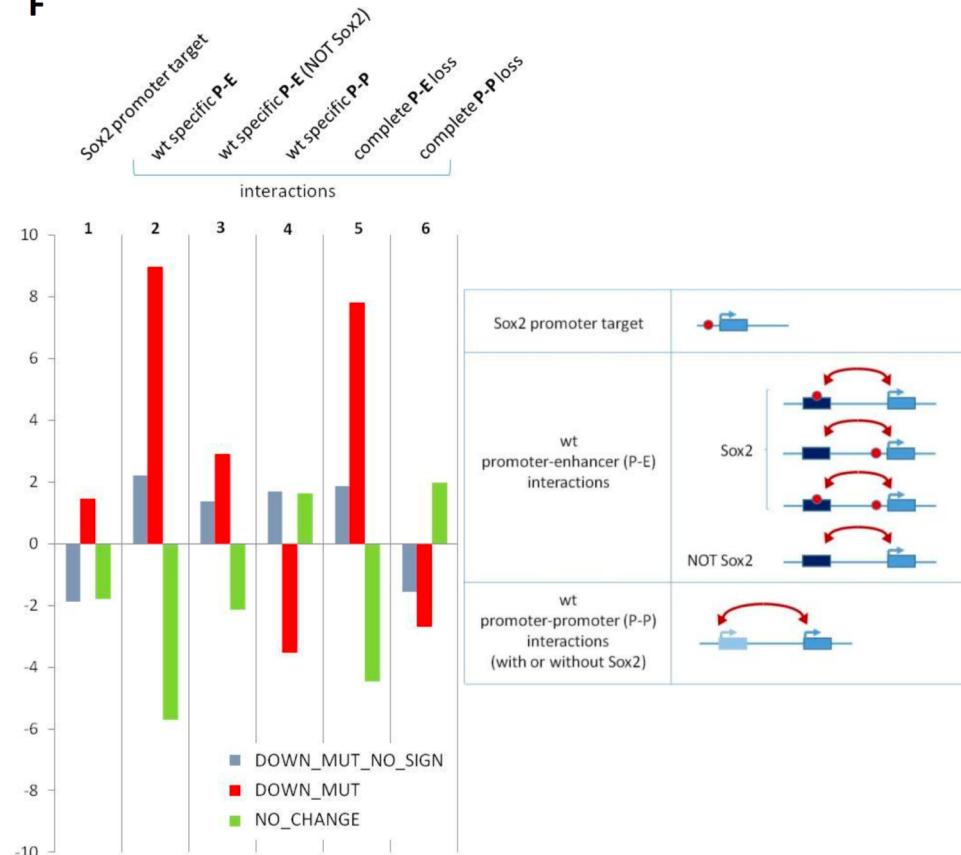
SOX2 and gene expression

- Do SOX2 targets (SOX2 on promoter or enhancer) show the most relevant change of expression?
- What is the effect of loss of interactions, with or without SOX2 on gene expression?
- Simple approach: Fisher's test on the number of genes that are simultaneously:
 - Changing their expression in a “significant” way (“differentially expressed”)
 - Changing their chromatin interactions (loss of enhancers or promoters)



SOX2 and gene expression

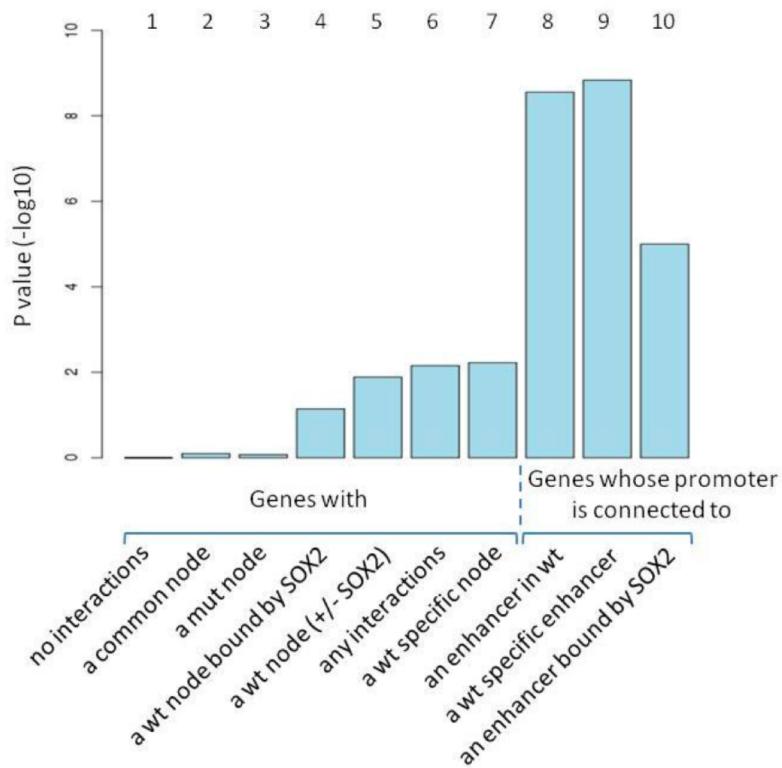
F



- Significant association with decrease in gene expression (red column) mostly with loss of enhancers, with or without SOX2 binding it
- Little or no impact of changes in promoters

SOX2 and gene expression

E

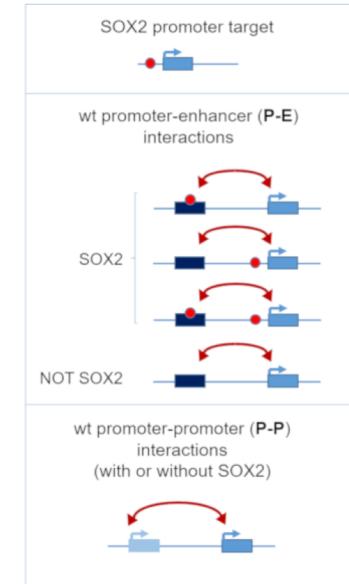
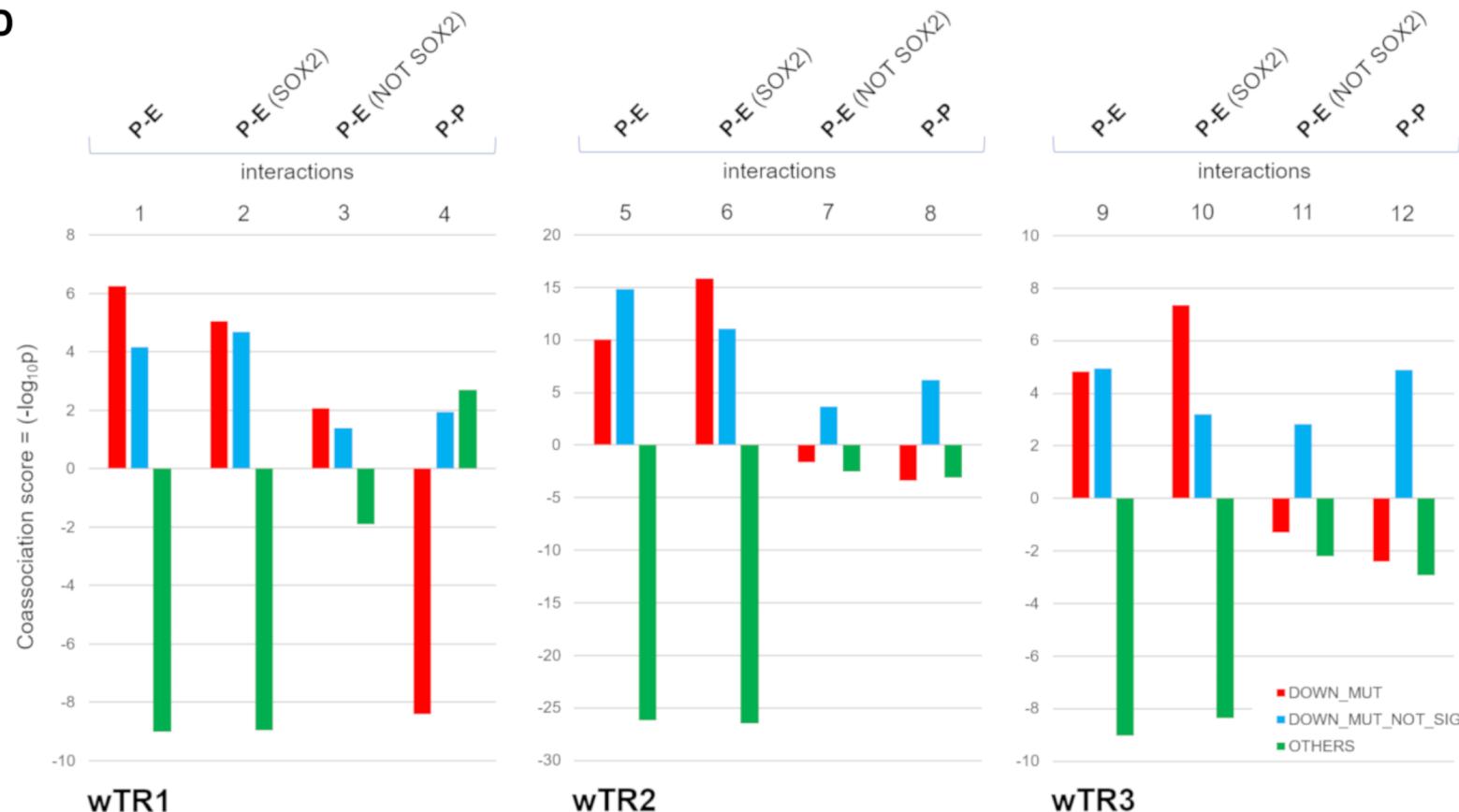


- Largest decrease in gene expression associated with the loss of an enhancer
- SOX2 binding is less relevant than the loss of interaction



Interactions and gene expression: three replicates

D





The experimental model

- Decrease in expression (but not complete silencing) due to
 - Deletion of a transcription factor
 - Loss of long distance promoter-enhancer interactions
- No change in polII recruitment on DNA
- No change in histone modification patterns



The working hypothesis

- Consider genes that:
 - Change (lose) number of connected enhancers in wt vs mutant
 - Even better, are not bound by SOX2, either in the promoters or the enhancer(s)
 - Use their difference in expression and try to model the effect of enhancers on their transcription



The emerging model

- Enhancers as “amplifiers” of a basal transcript level of a gene:

$$G = b + \alpha n b$$

where G is the overall expression of the gene, b is the basal level, α is the “*amplification factor*”, n is the number of enhancers connected to the gene



Experimental issues

- Chromatin interaction data and gene expression **data come from different individuals** (mice)
- Individual specific variation of gene expression -> individual specific patterns of enhancer-promoter associations?
- Some other “major” TFs (FOS, JUN) deregulated in mutant cells -> secondary effect?
- Low resolution in the definition of interacting regions
- Now include other TF ChIP-Seq datasets to better estimate “basal levels”
- Include TADs (thought about this 1 hour ago...)



...coming very soon...

Please cite this article in press as: Bertolini et al., Mapping the Global Chromatin Connectivity Network for Sox2 Function in Neural Stem Cell Maintenance, *Cell Stem Cell* (2019), <https://doi.org/10.1016/j.stem.2019.02.004>

Cell Stem Cell
Article

CellPress

Mapping the Global Chromatin Connectivity Network for Sox2 Function in Neural Stem Cell Maintenance

Jessica A. Bertolini,^{1,9} Rebecca Favaro,^{1,9} Yanfen Zhu,^{2,9} Miriam Pagan,¹ Chew Yee Ngan,² Chee Hong Wong,² Harianti Tjong,² Marit W. Vermunt,^{3,8} Ben Martynoga,⁴ Cristiana Barone,¹ Jessica Mariani,¹ Marcos Julián Cardozo,⁵ Noemí Tabanera,⁵ Federico Zambelli,⁶ Sara Mercurio,¹ Sergio Ottolenghi,¹ Paul Robson,^{2,7} Menno P. Creyghton,³ Paola Bovolenta,⁵ Giulio Pavesi,⁶ Francois Guillemot,⁴ Silvia K. Nicolis,^{1,10,*} and Chia-Lin Wei^{2,*}

¹Department of Biotechnology and Biosciences, University Milano-Bicocca, 20126 Milano, Italy

²The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA

³Hubrecht Institute-KNAW and University Medical Center Utrecht 3584CT, Utrecht, the Netherlands

⁴The Francis Crick Institute, Midland Road, London NW 1AT, UK

⁵Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones Científicas-Universidad Autónoma de Madrid and Ciber de Enfermedades Raras (CIBERER), ISCIII Madrid, Spain

⁶Department of Biosciences, University of Milano, 20133 Milano, Italy

⁷Stem Cell and Regenerative Biology, Genome Institute of Singapore, Singapore

⁸Present address: Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

⁹These authors contributed equally

¹⁰Lead Contact

*Correspondence: silvia.nicolis@unimib.it (S.K.N.), chia-lin.wei@jax.org (C.-L.W.)

<https://doi.org/10.1016/j.stem.2019.02.004>