

Uncoupling gene expression and chromatin organization during embryogenesis by genomic rearrangements

Aleksander Jankowski
Furlong Lab



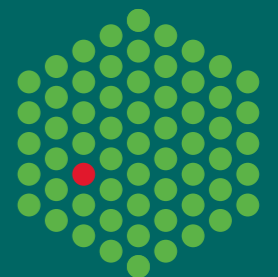
Yad Ghavi-Helm
Furlong Lab



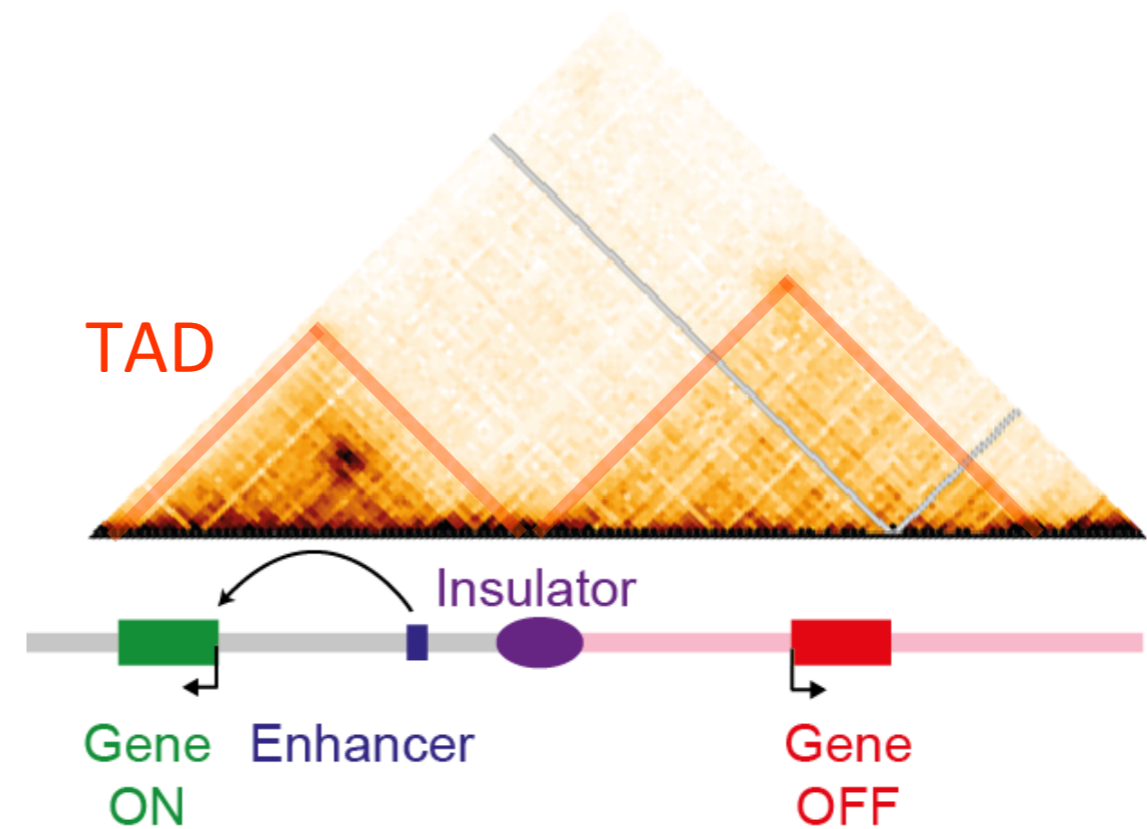
Sascha Meiers
Korbel Lab

Genome Biology Unit Seminar
March 5, 2018

EMBL

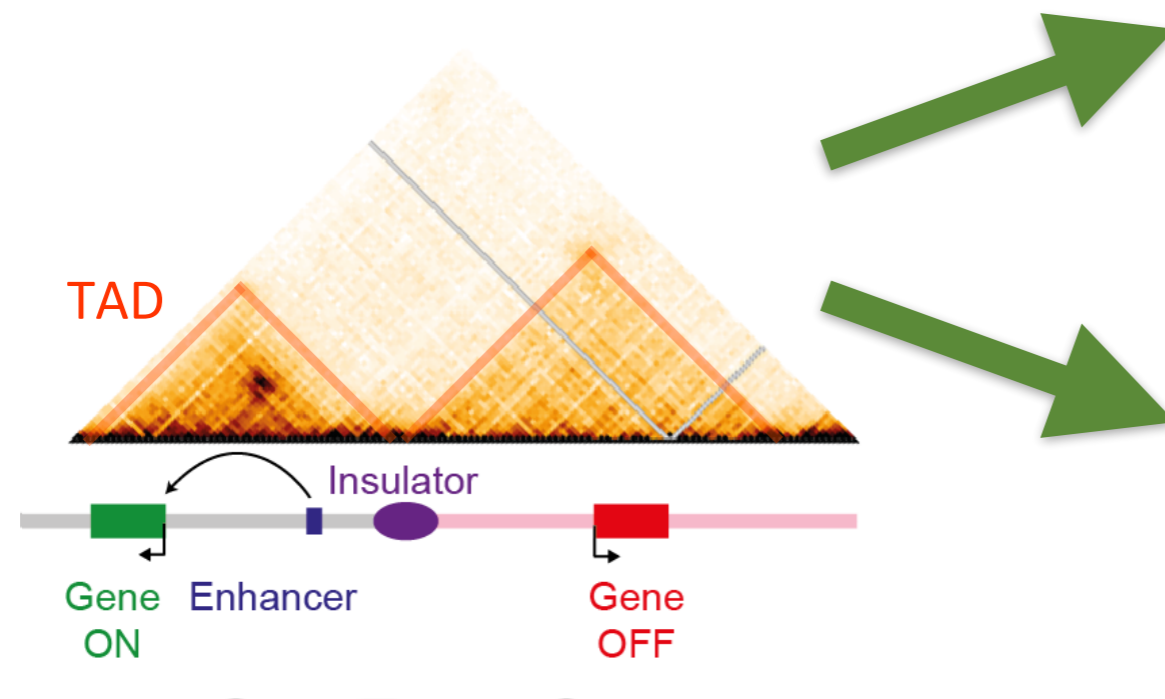


Topologically Associating Domains (TADs) are the units of spatial chromatin organization

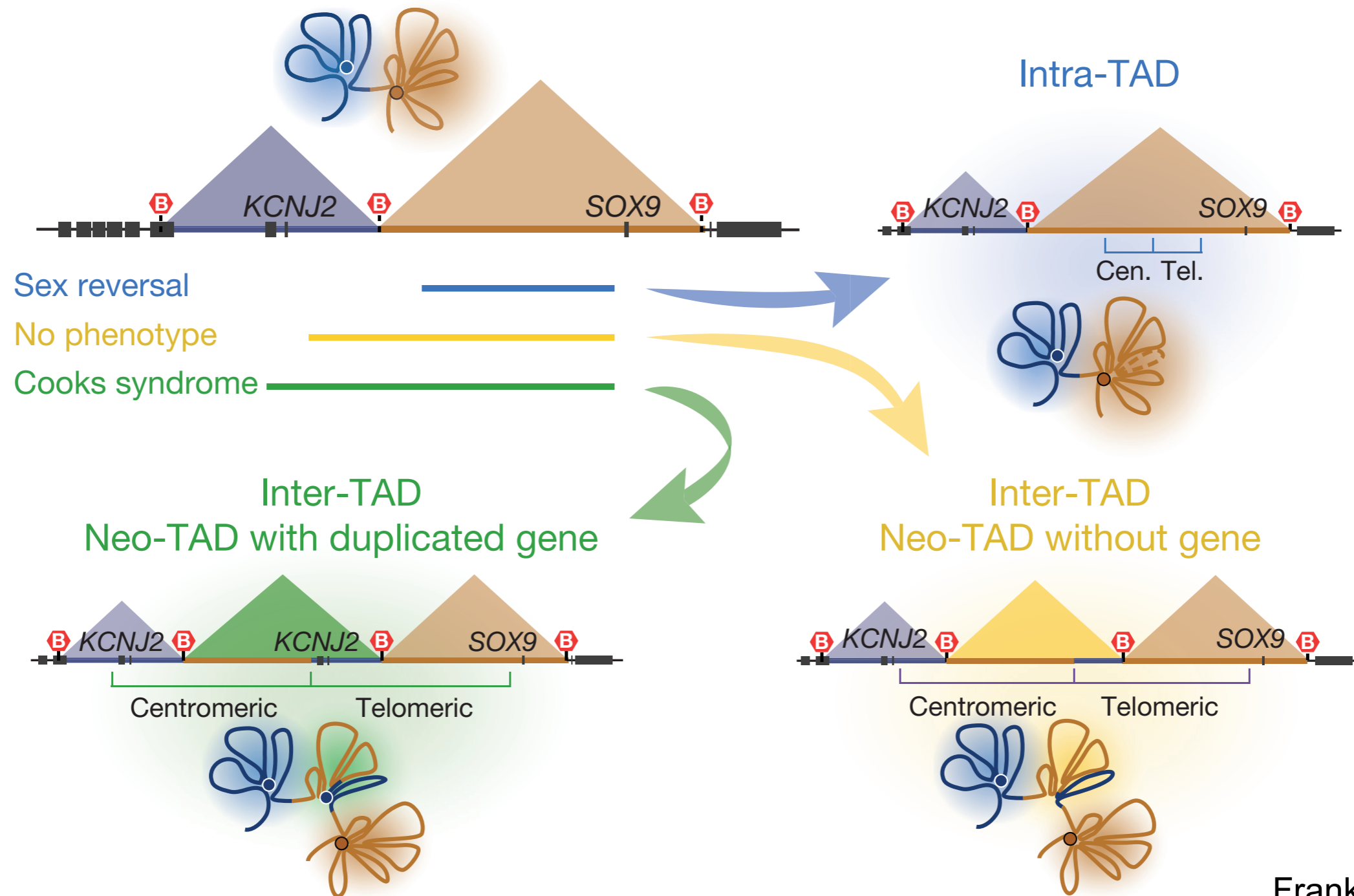


Nora *et al.* Nature 2012
Dixon *et al.* Nature 2012
Sexton *et al.* Cell 2012

Chromatin organization gets disrupted after perturbing TAD boundaries or insulator binding sites



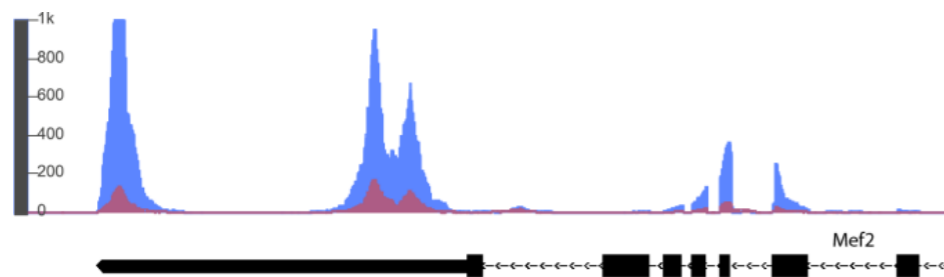
Changes in gene expression at disease loci are associated with genomic rearrangements



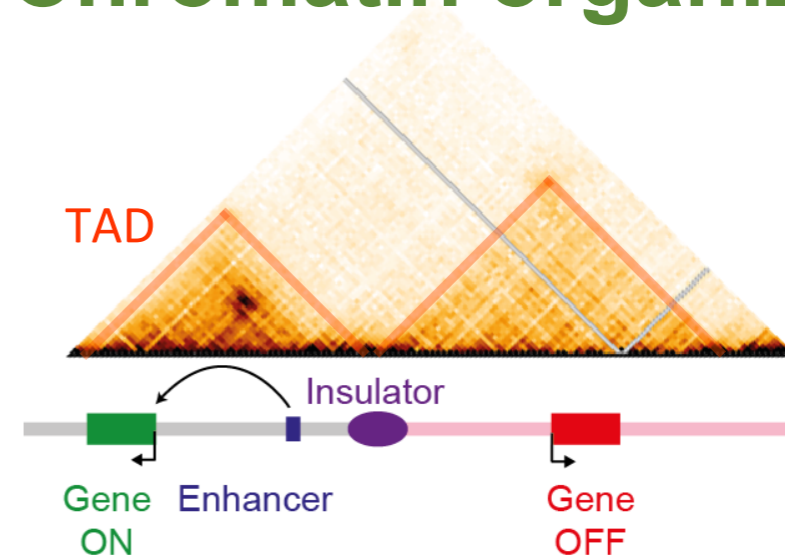
Franke et al. 2016

Our aim: systematically assess the relationship between gene expression and chromatin organization

Gene expression



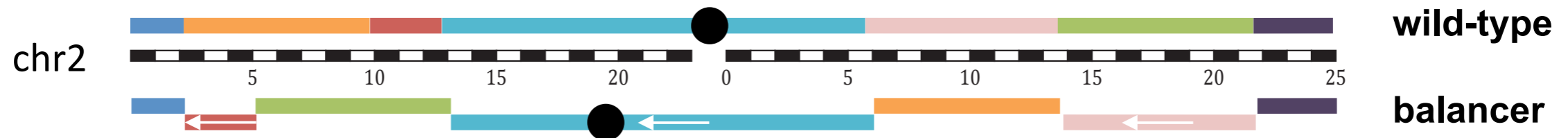
Chromatin organization



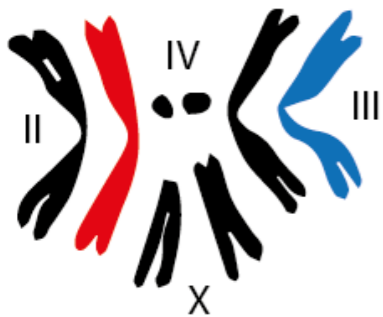
What do we need?

- Both big and small rearrangements.
- Measure the effect in *cis* (both alleles from the same nucleus).
- Excellent tool that we could use: balancer chromosomes.

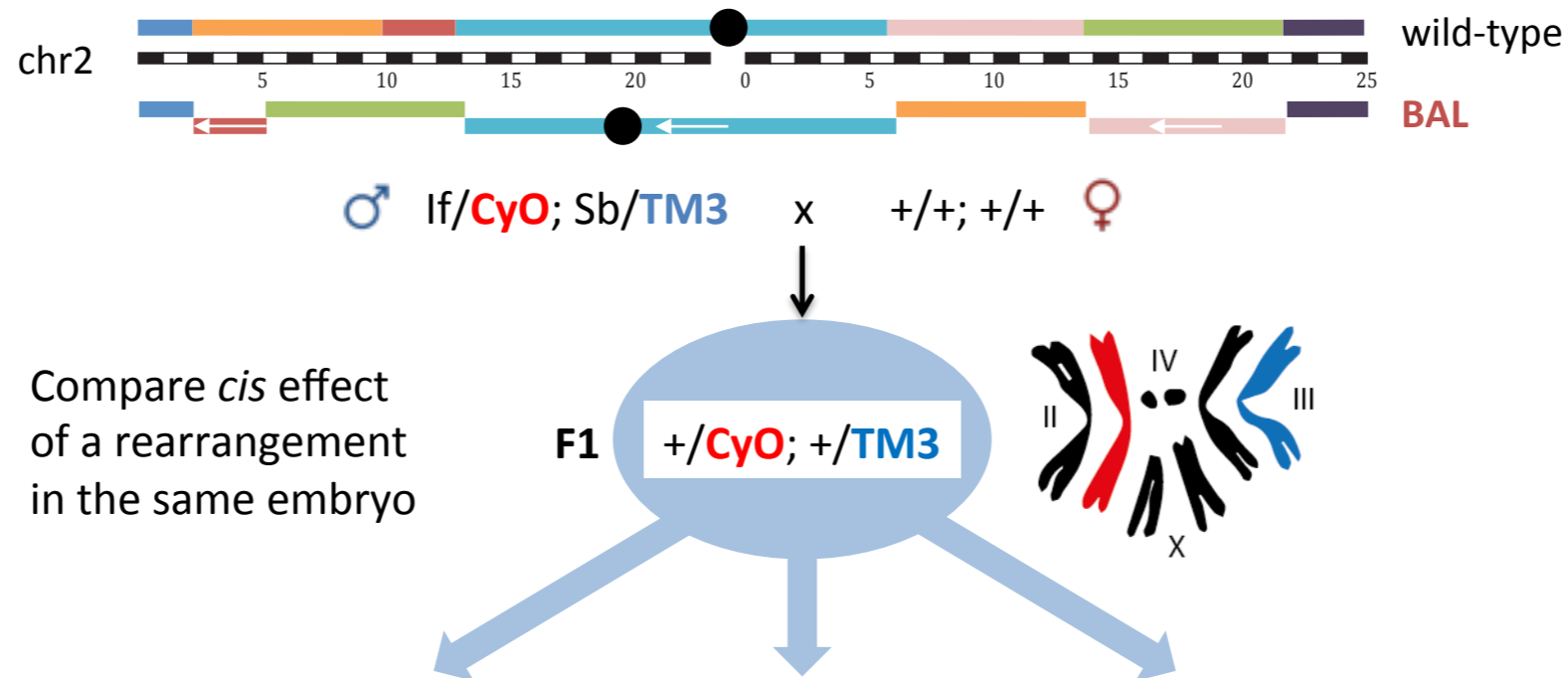
Balancer chromosomes serve as an excellent tool to rearrange the genome



- Balancer chromosomes are highly scrambled homologs of wild-type chromosomes, frequently used in *Drosophila* genetics.
- They were originally obtained through X-ray mutagenesis.
- We used two balancers (CyO and TM3), covering 76% of the fly genome.

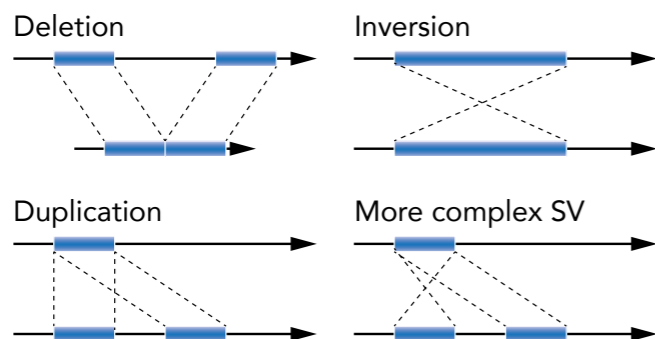


Functional impact of genomic rearrangements on gene expression and chromatin organization



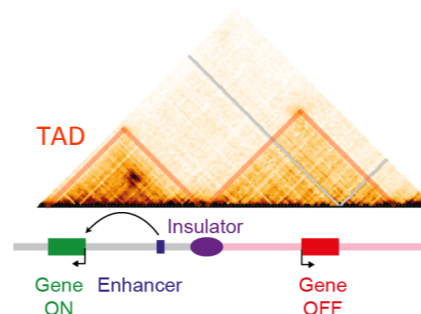
Characterize the genomic rearrangements in both haplotypes

- Structural Variants (SVs)
- Single Nucleotide Variants (SNVs)



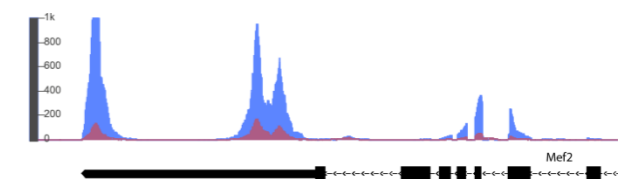
Quantify differences in chromatin organization

- Allele-specific Hi-C
- Allele-specific Capture-C

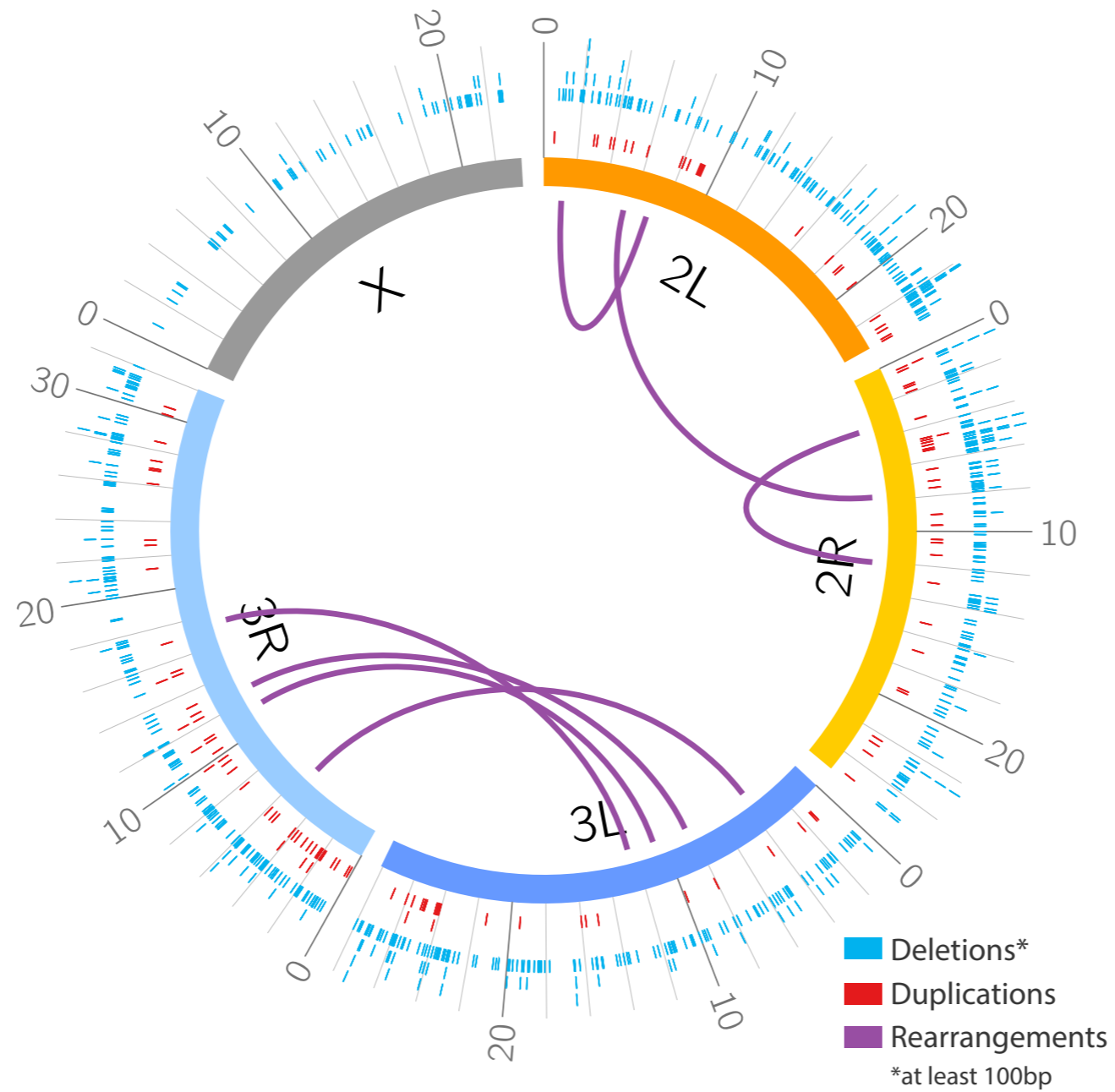


Quantify differences in gene expression

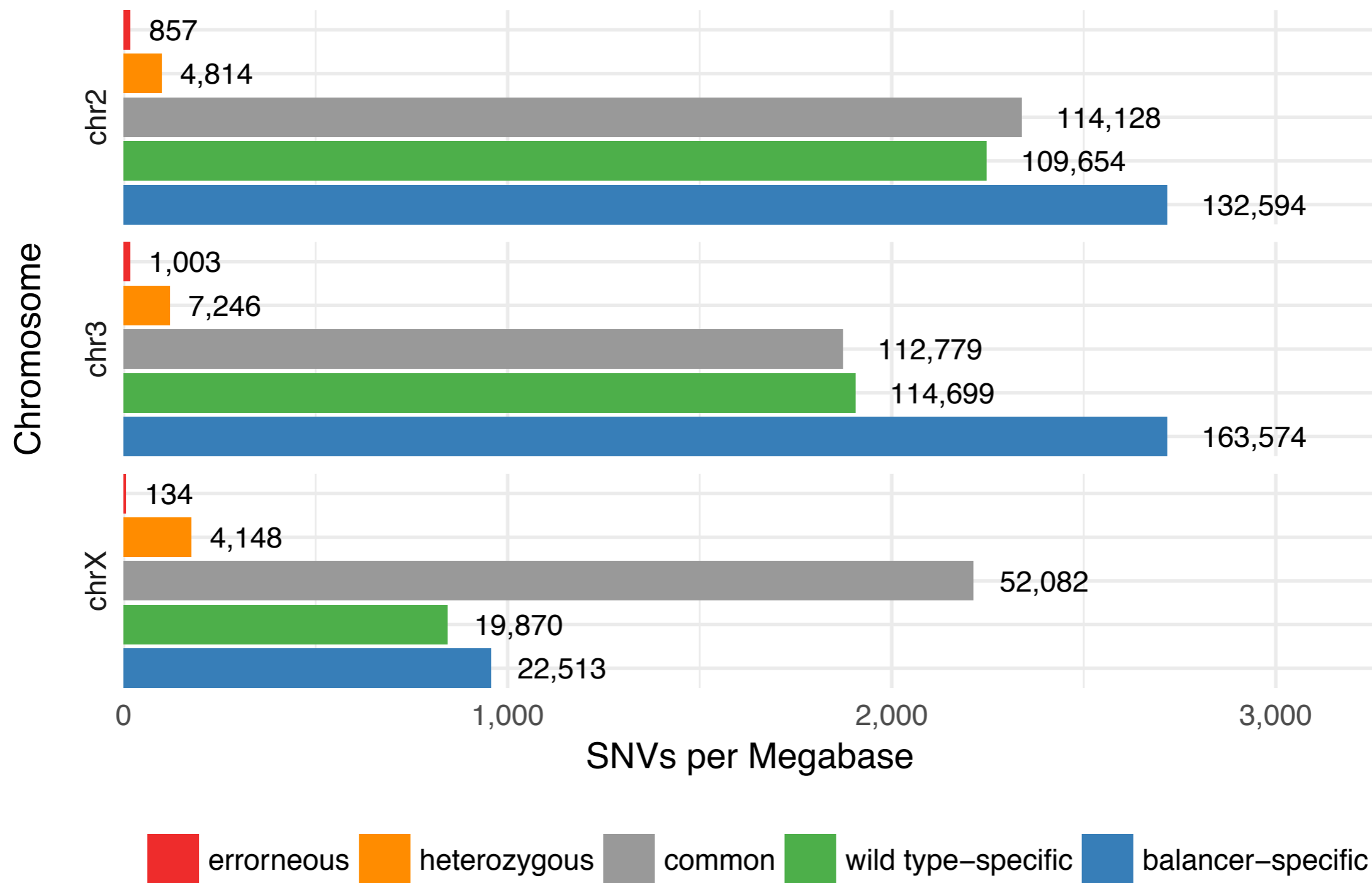
- Allele-specific RNA-seq



Summary of structural variants



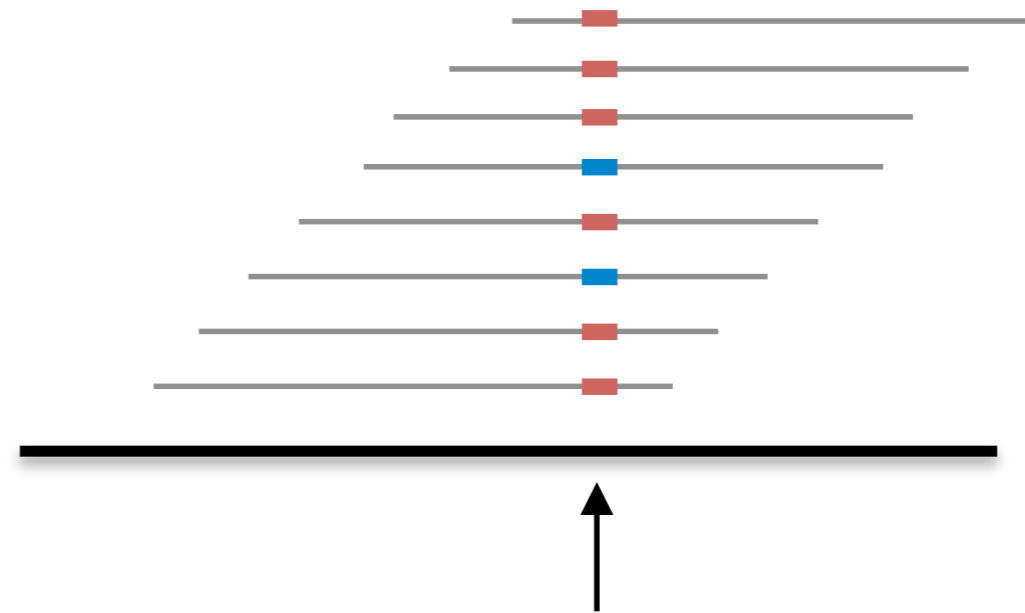
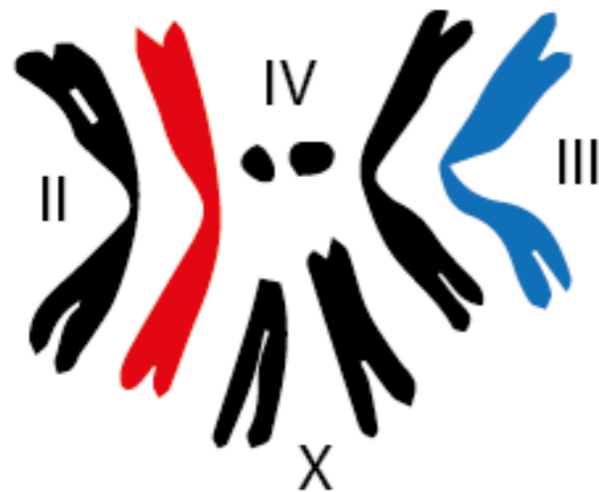
Accumulation of Single Nucleotide Variants in balancer chromosomes



Separation of reads into alleles

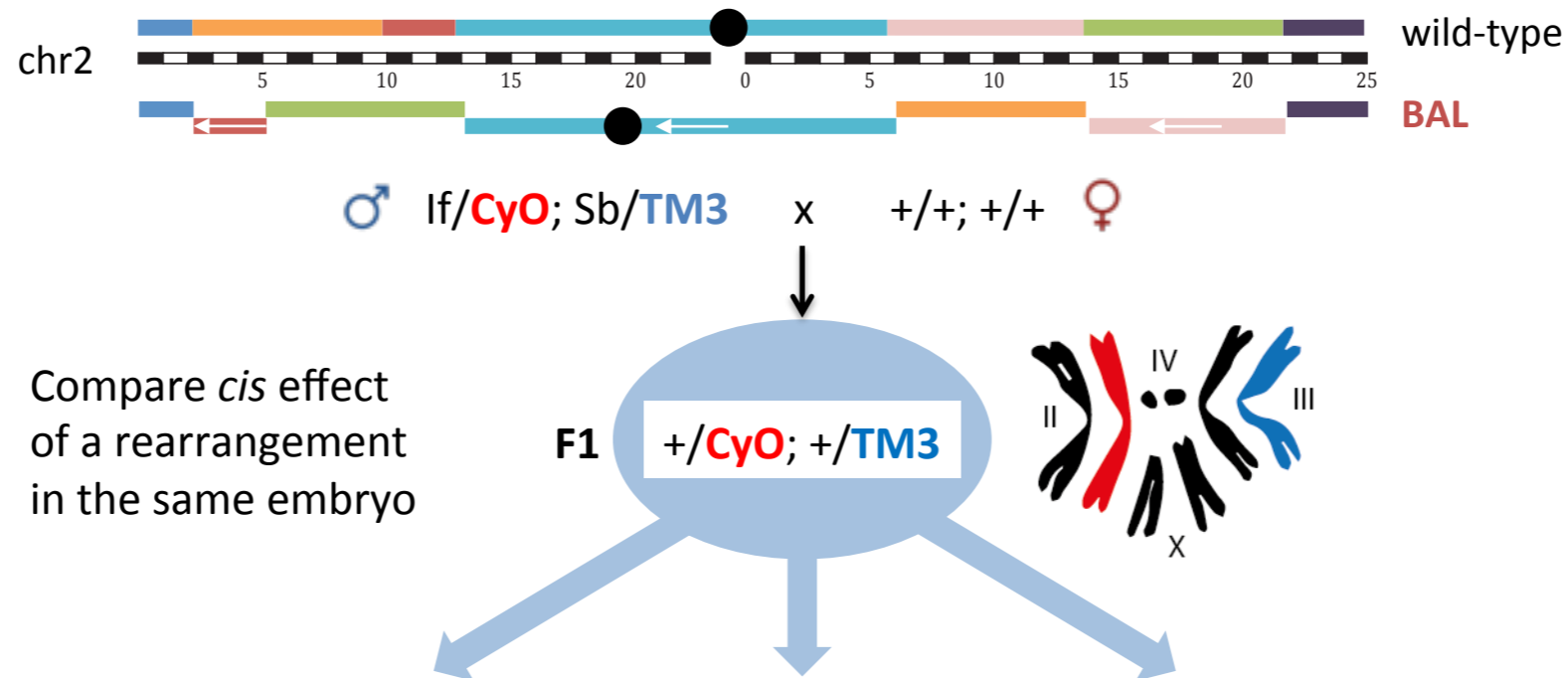
Separate reads from the wild-type and balancer alleles

F1 **+/CyO; +/TM3**



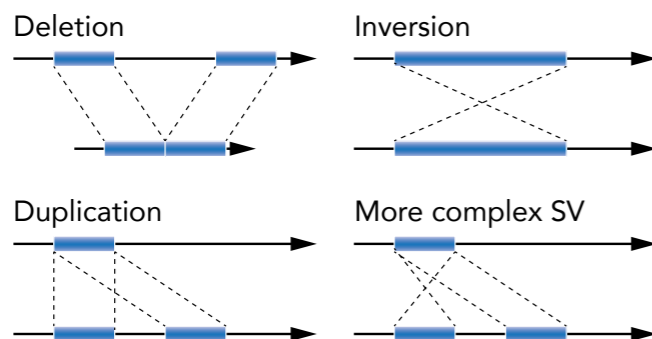
SNV distinguishing
balancer from wild-type

Functional impact of genomic rearrangements on gene expression and chromatin organization



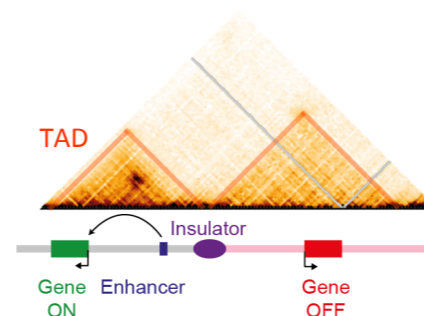
Characterize the genomic rearrangements in both haplotypes

- Structural Variants (SVs)
- Single Nucleotide Variants (SNVs)



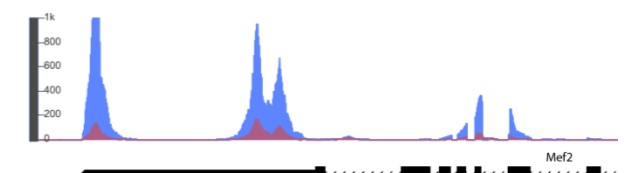
Quantify differences in chromatin organization

- Allele-specific Hi-C
- Allele-specific Capture-C

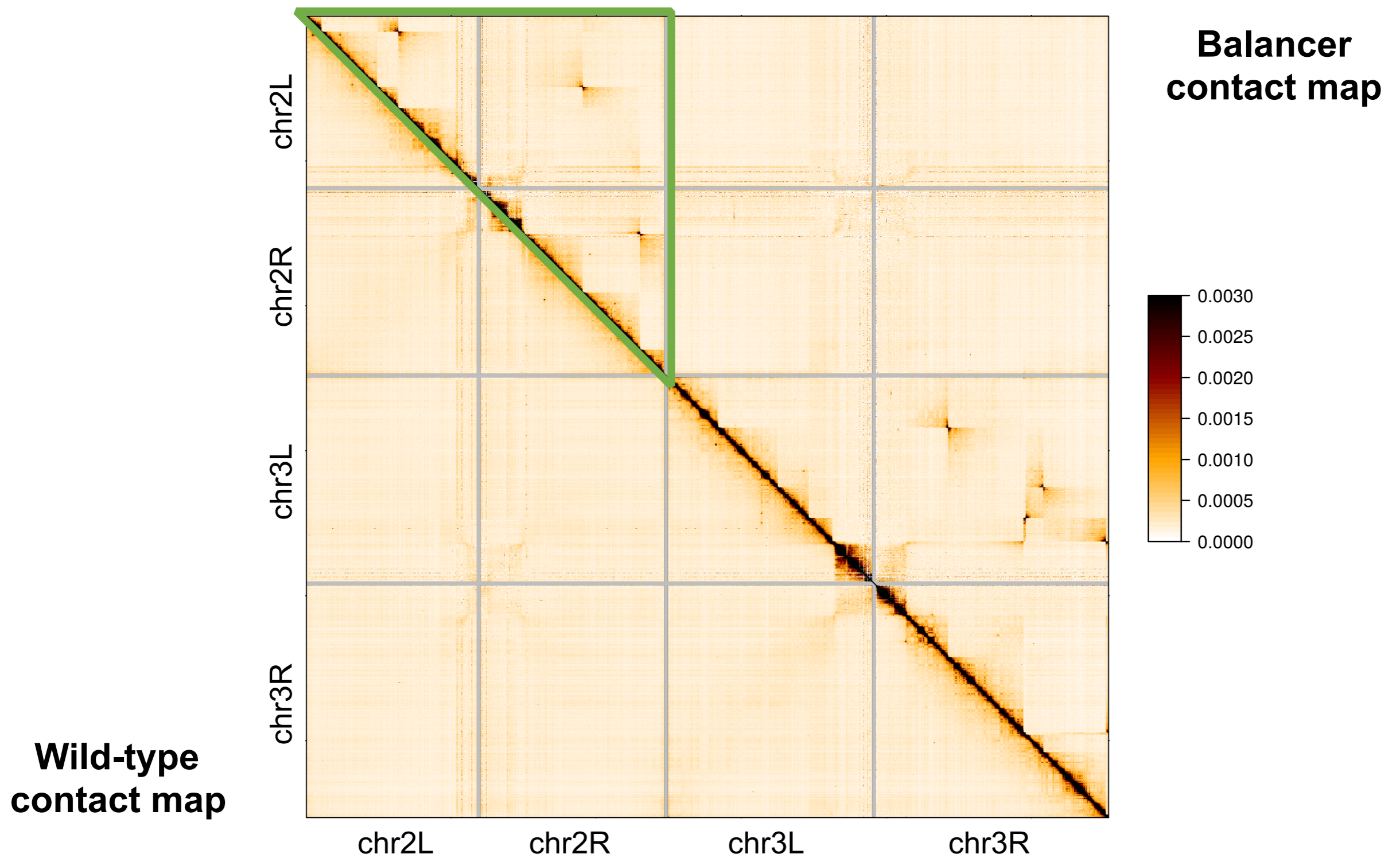


Quantify differences in gene expression

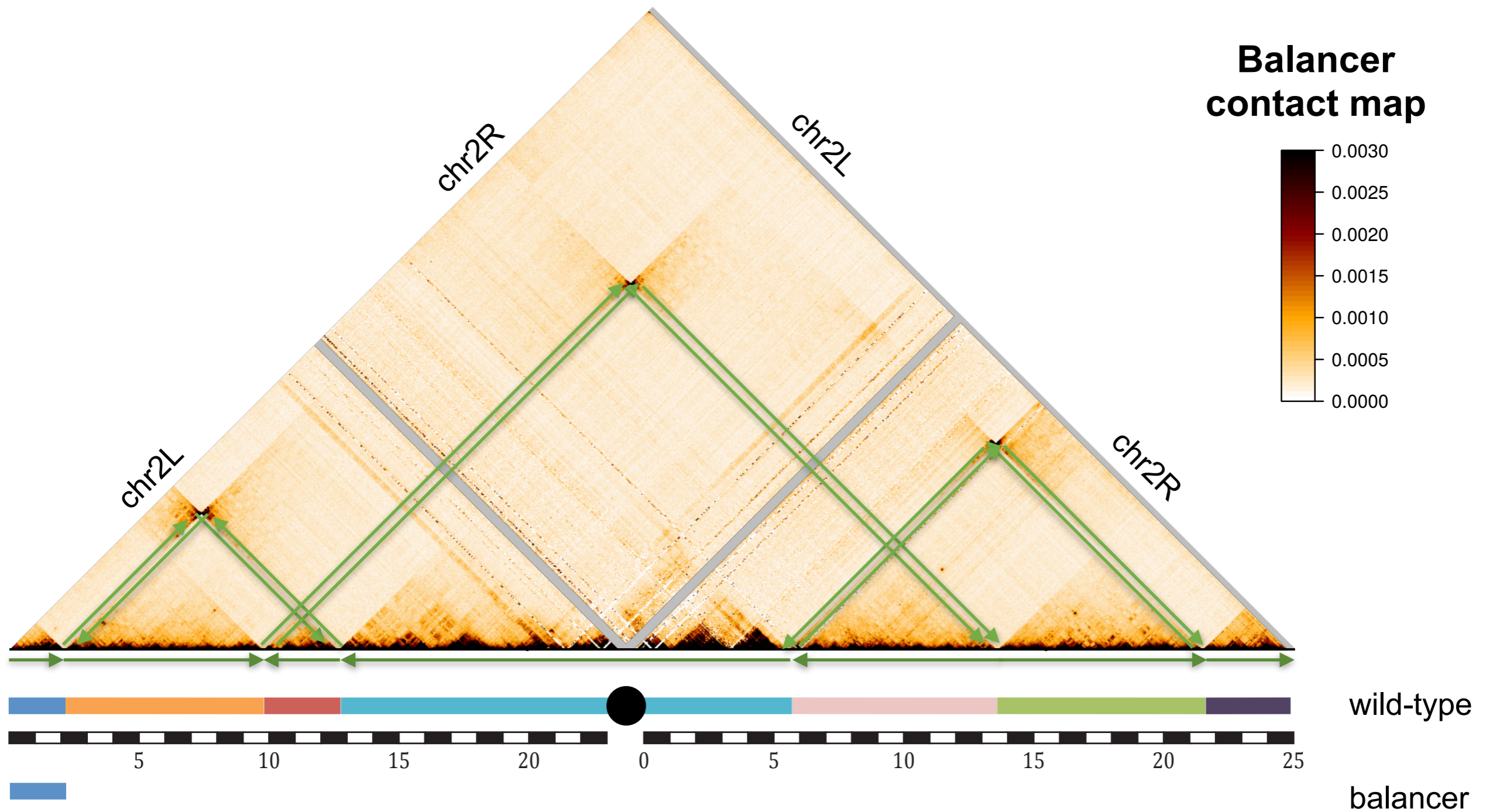
- Allele-specific RNA-seq



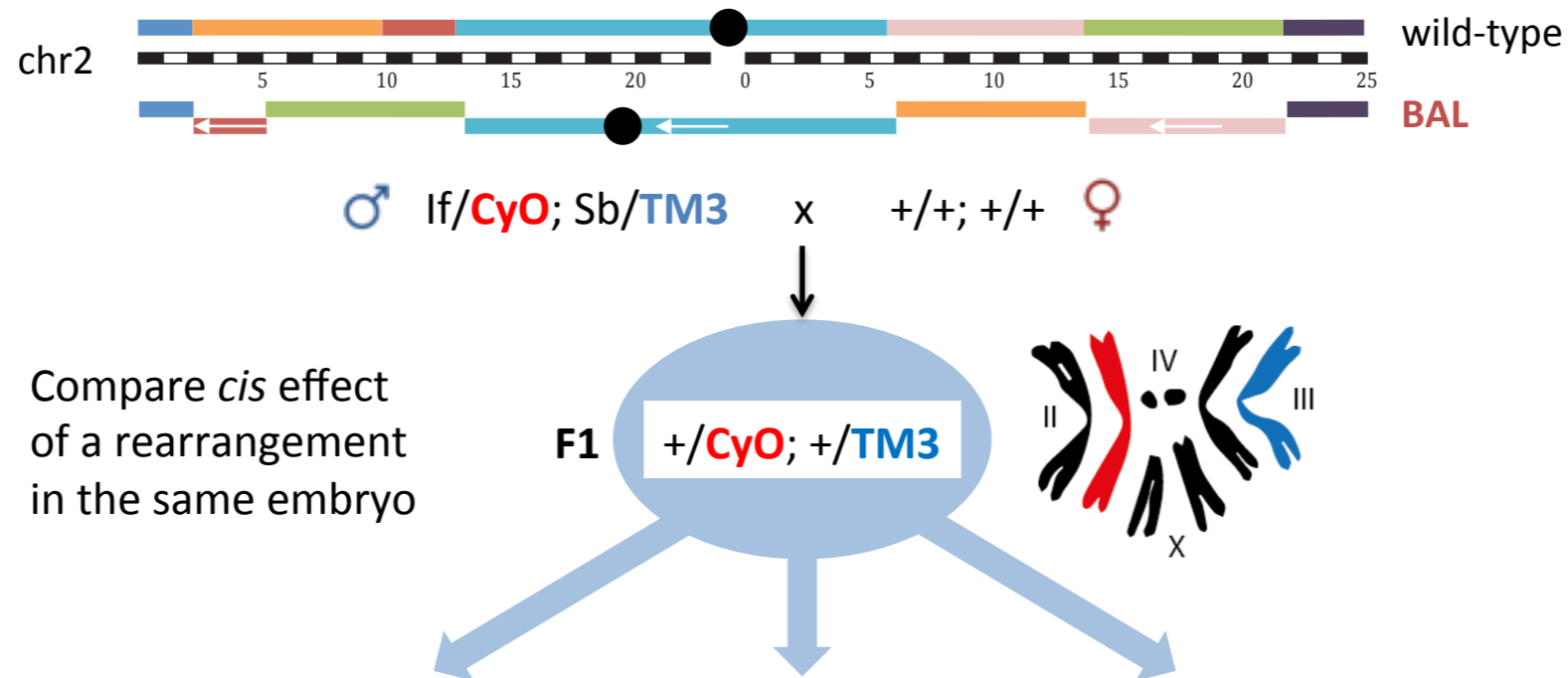
Comparing allele-specific Hi-C contact maps



Hi-C data recapitulate balancer genome structure

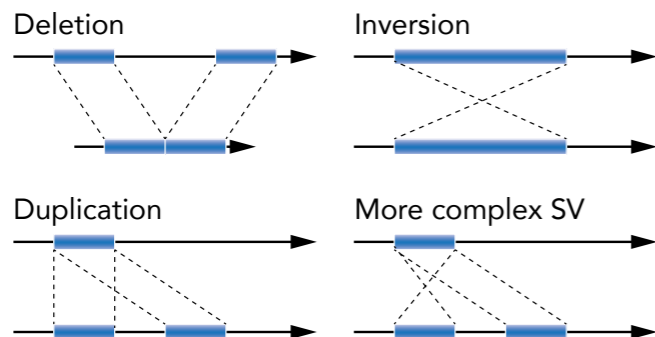


Functional impact of genomic rearrangements on gene expression and chromatin organization



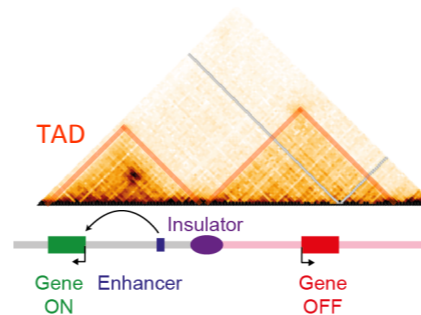
Characterize the genomic rearrangements in both haplotypes

- Structural Variants (SVs)
- Single Nucleotide Variants (SNVs)



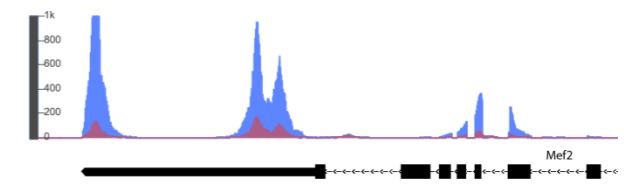
Quantify differences in chromatin organization

- Allele-specific Hi-C
- Allele-specific Capture-C

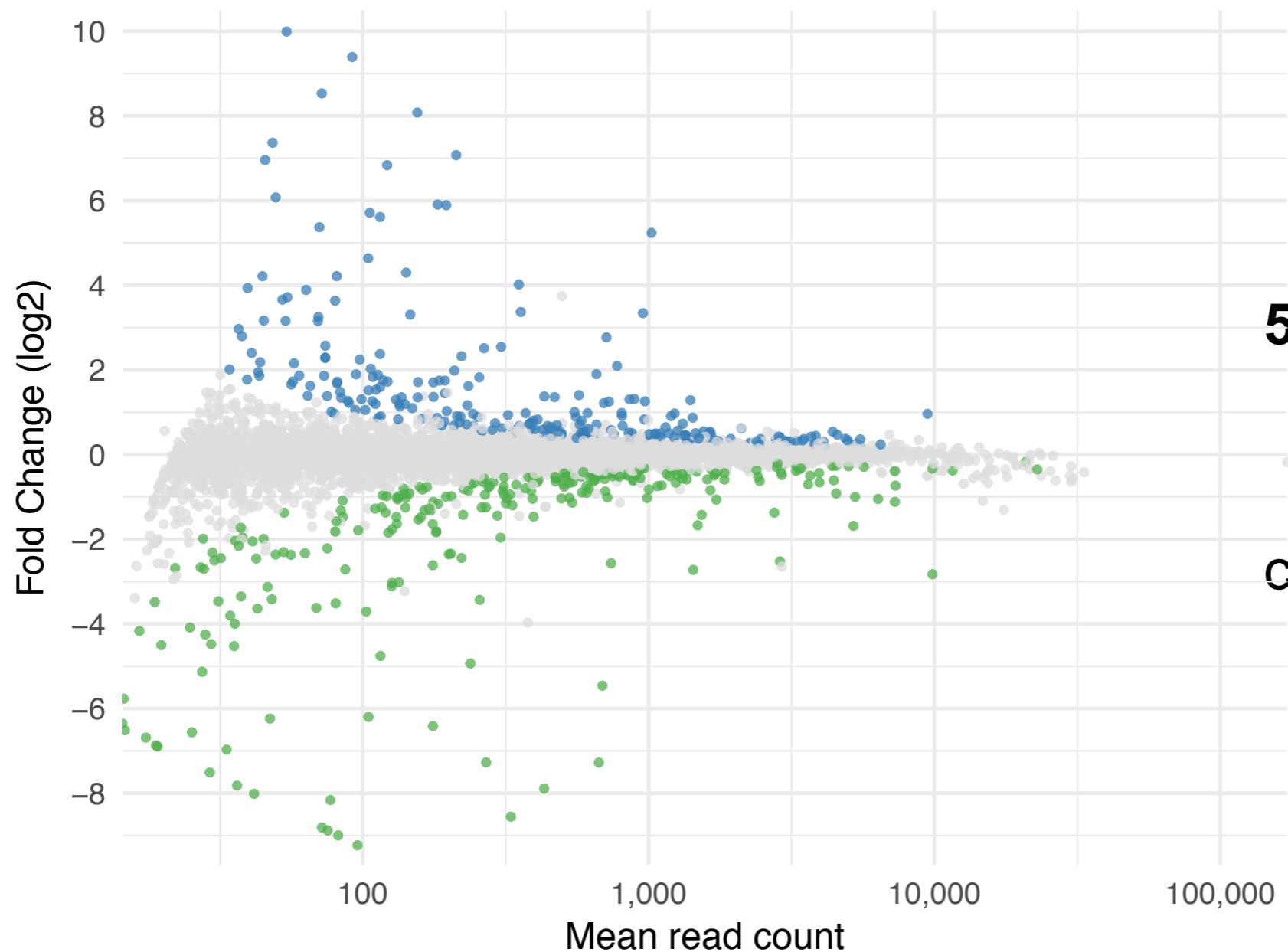


Quantify differences in gene expression

- Allele-specific RNA-seq



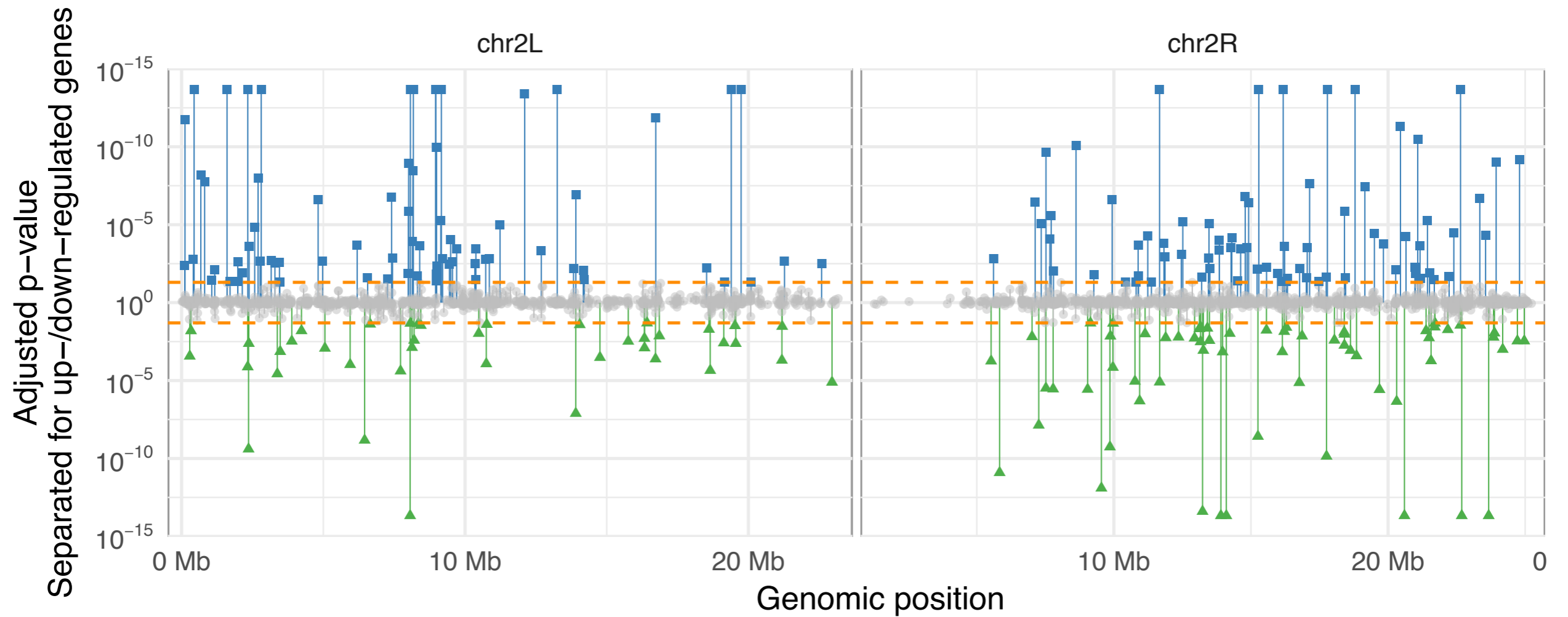
Identification of differentially expressed genes between the alleles



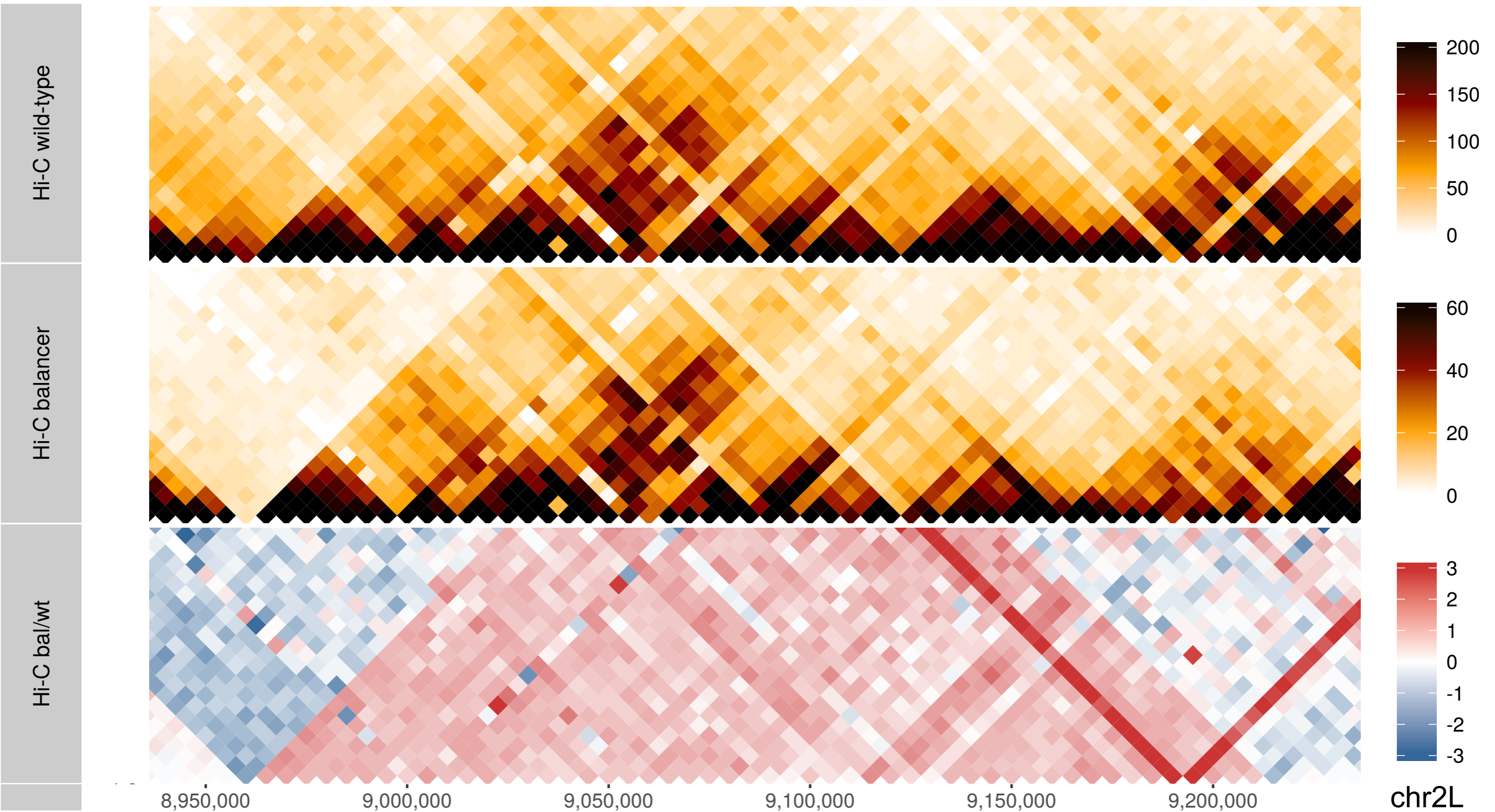
5,357 testable genes
512 DE genes (5% FDR)
343 DE genes with fold change of at least 1.5

Significant genes (5% FDR) ● increased in balancer ● increased in wild type ● not significant

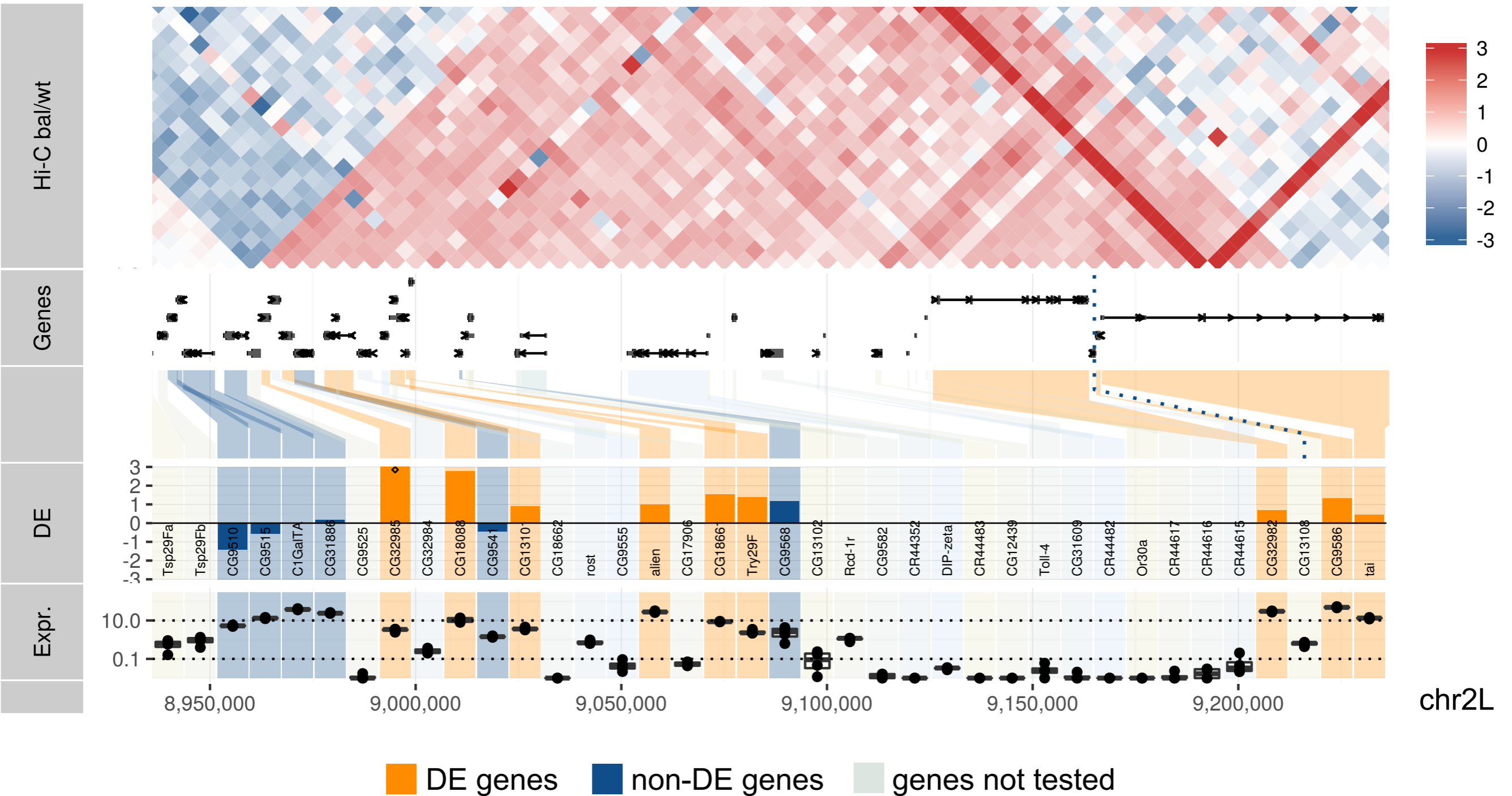
Differentially expressed genes are not clustered along the chromosomes



Hi-C data recapitulate duplications and other copy number variation events

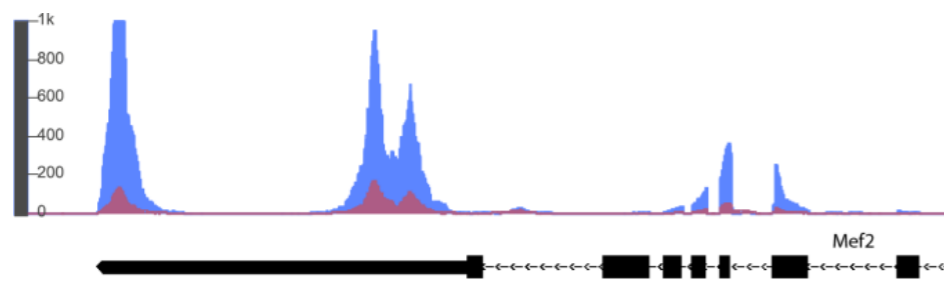


Duplications and other copy number variation events partially explain the differential gene expression

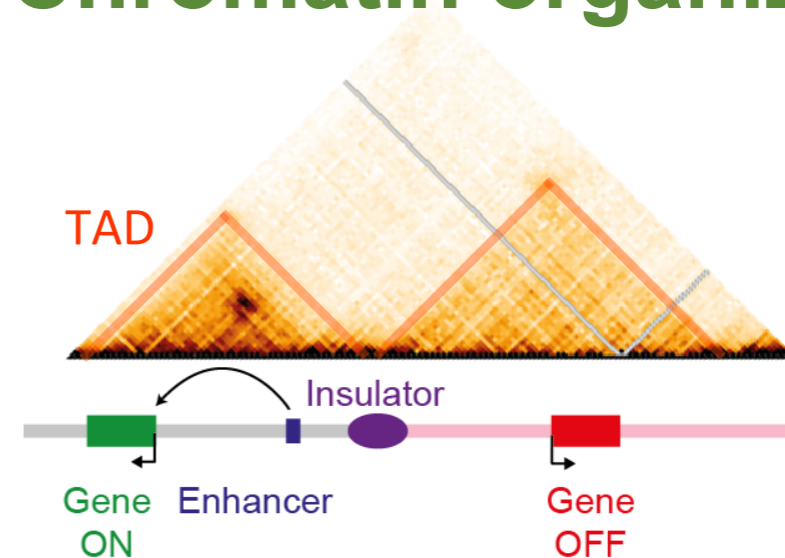


We aim to systematically assess the relationship between gene expression and chromatin organization

Gene expression

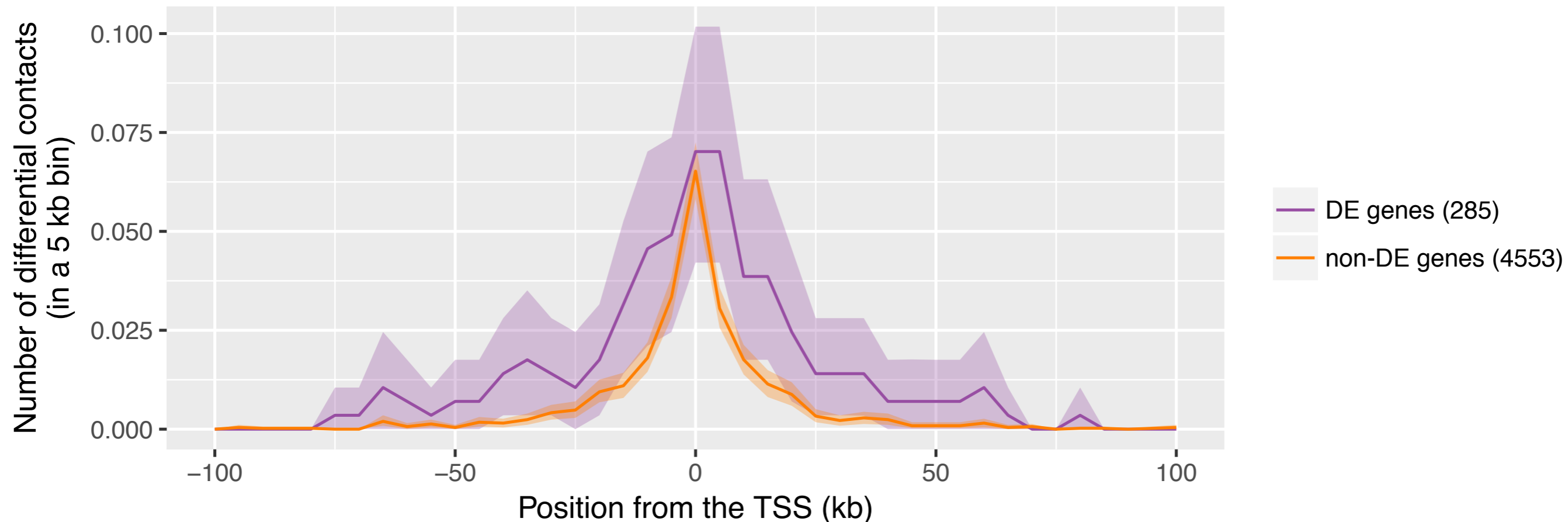


Chromatin organization



Differential gene expression is correlated with local changes in genome topology

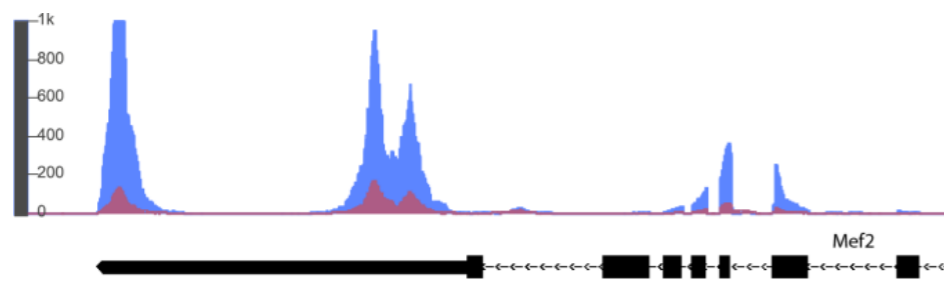
Differential Hi-C contacts from gene promoters



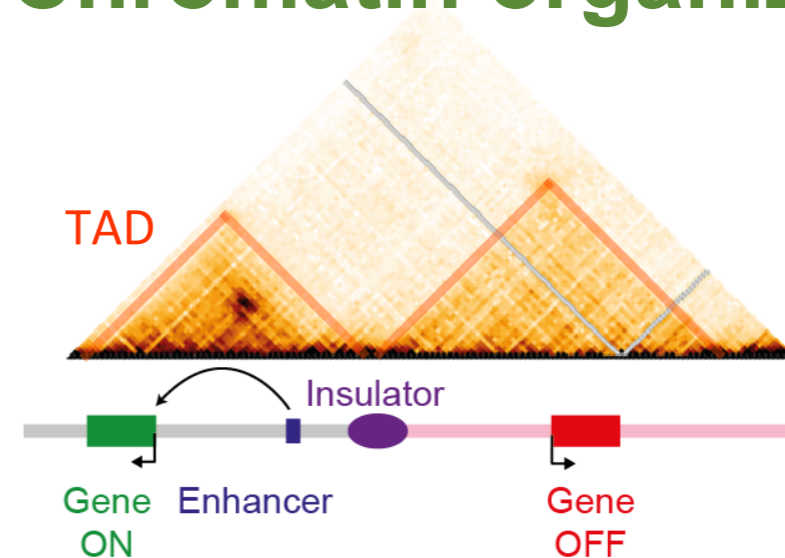
DE genes have on average 0.57 differential Hi-C contacts, non-DE genes have 0.24.
 $p = 5.99 \cdot 10^{-6}$, Wilcoxon test with a triangular kernel. 95% CI obtained by bootstrapping.

We aim to systematically assess the relationship between gene expression and chromatin organization

Gene expression



Chromatin organization



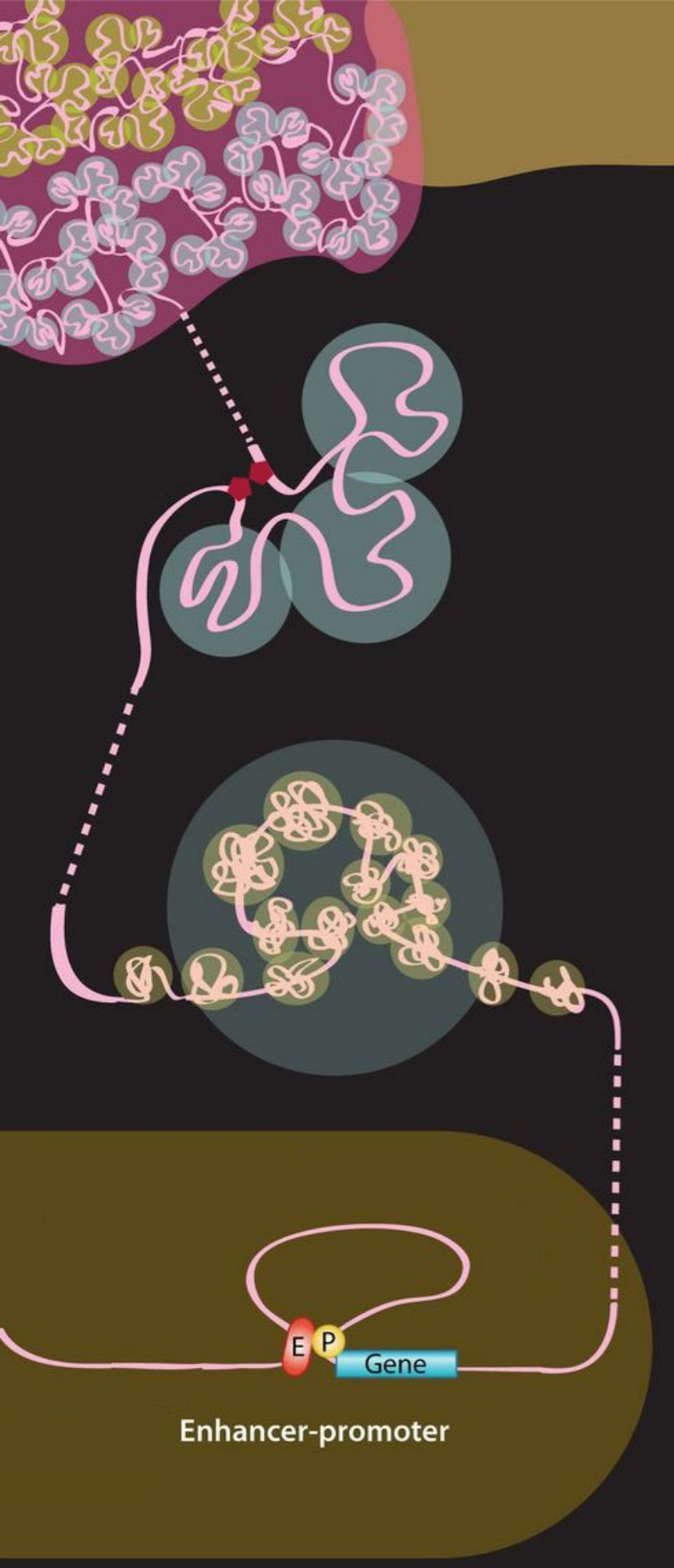
Scales of chromatin organization

Long-range loops (megabase-scale)
disruptions by genomic rearrangements

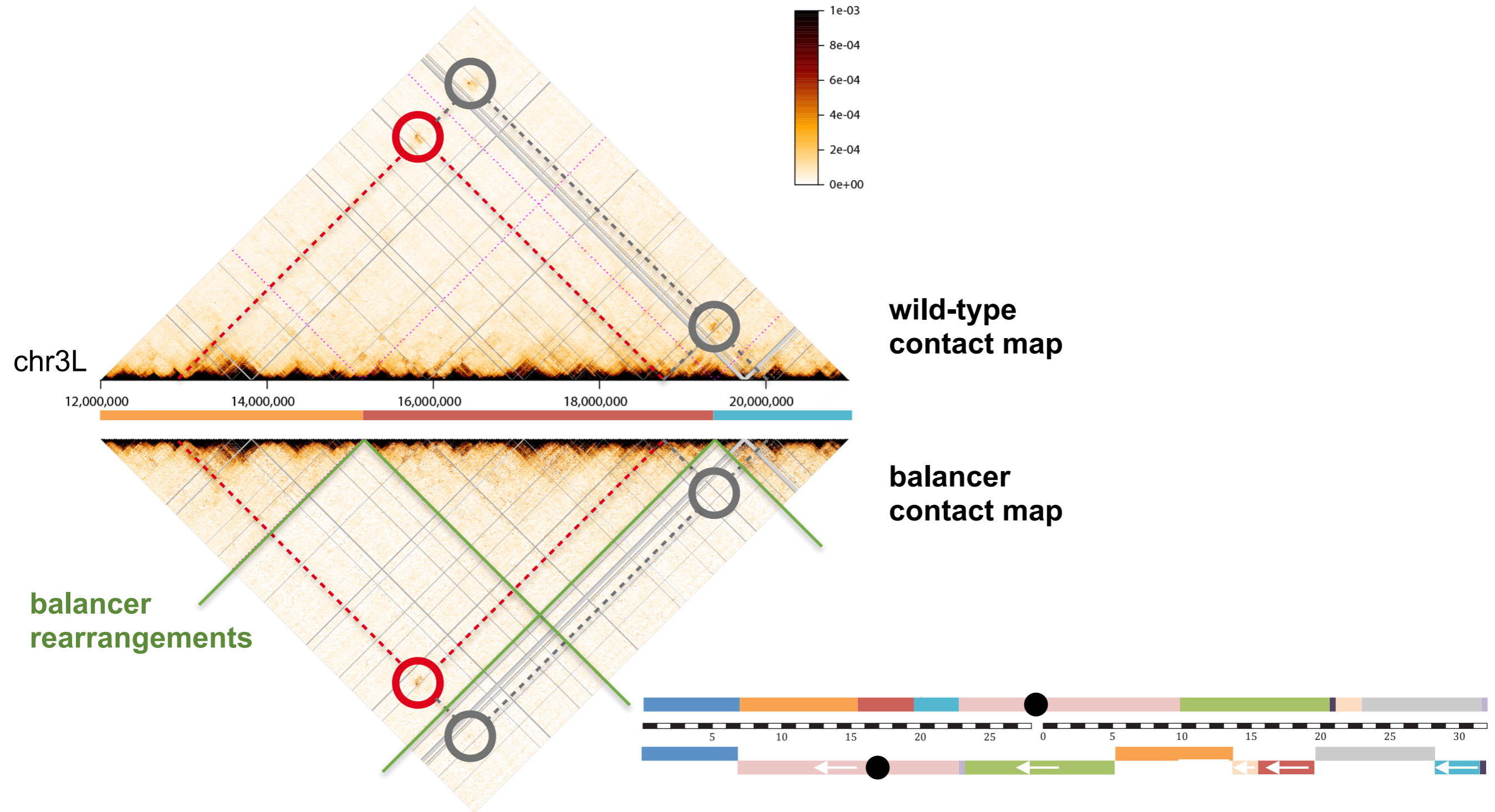
Topologically Associating Domains (TADs)
disruptions by genomic rearrangements
differences in TAD structure between alleles

Contact frequency within TADs

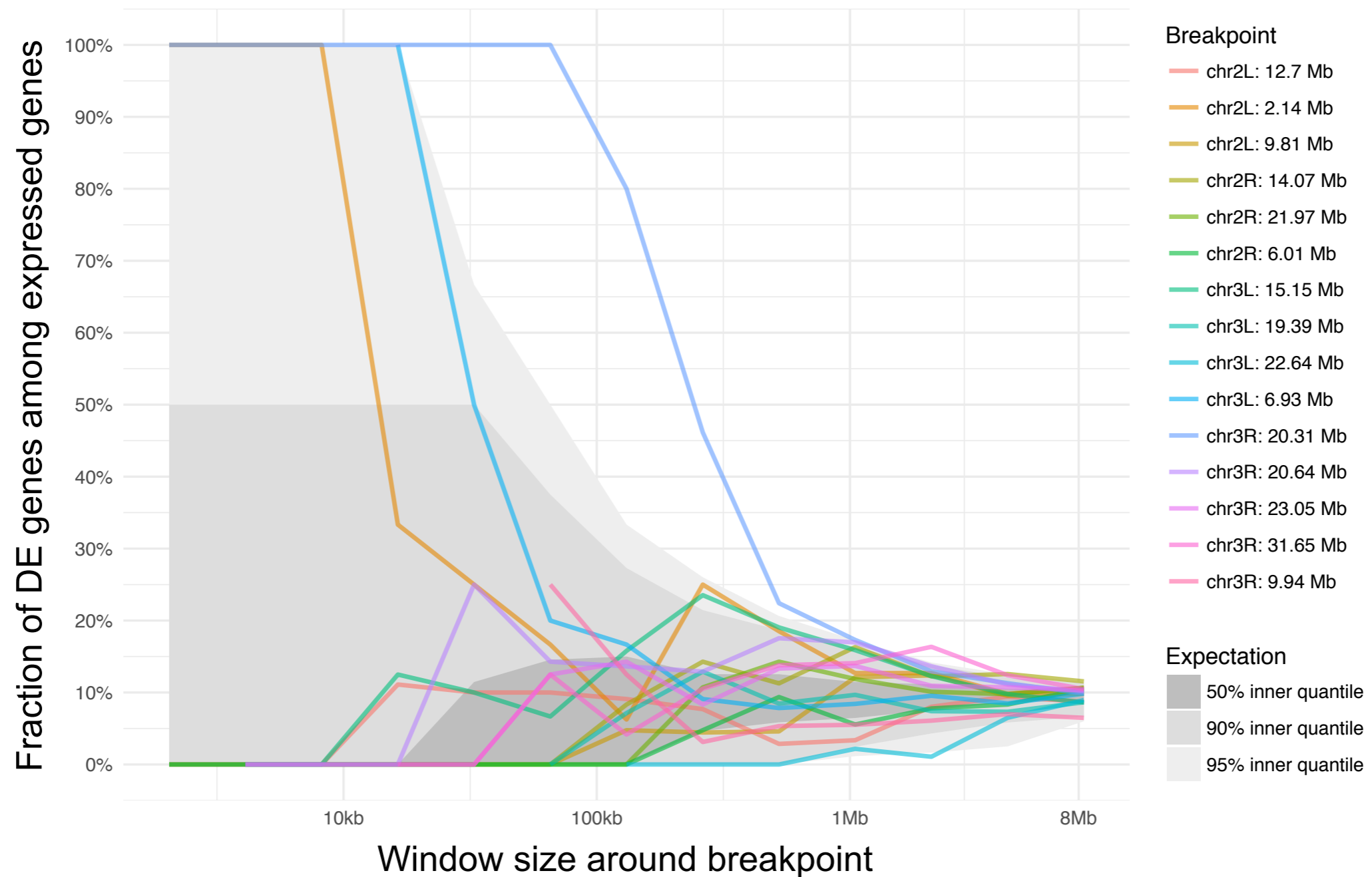
Enhancer-promoter loops



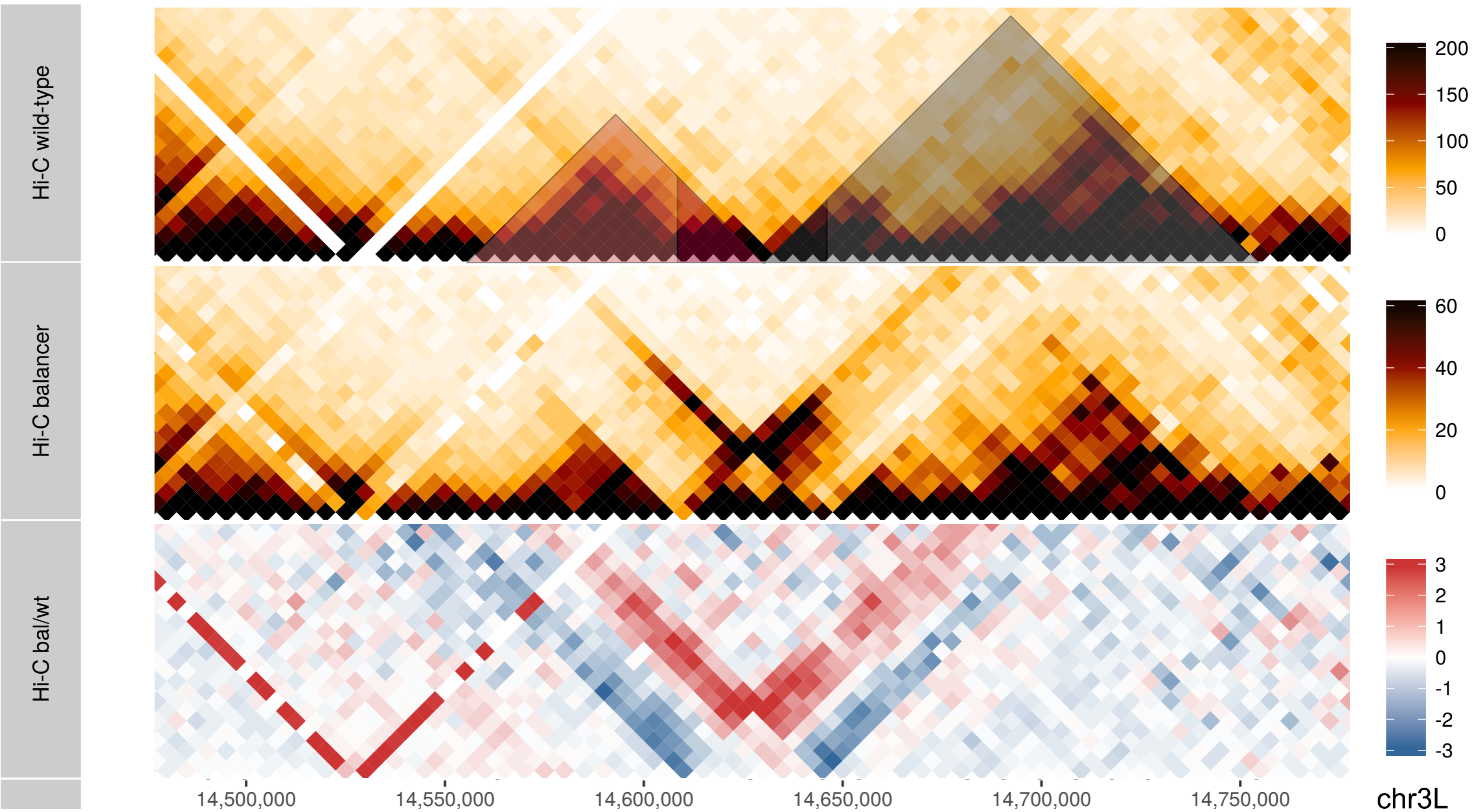
Are long-range loops maintained in the context of genomic rearrangements?



Around rearrangement breakpoints, differentially expressed genes are not enriched



Hi-C contact maps reveal a 38 kb inversion



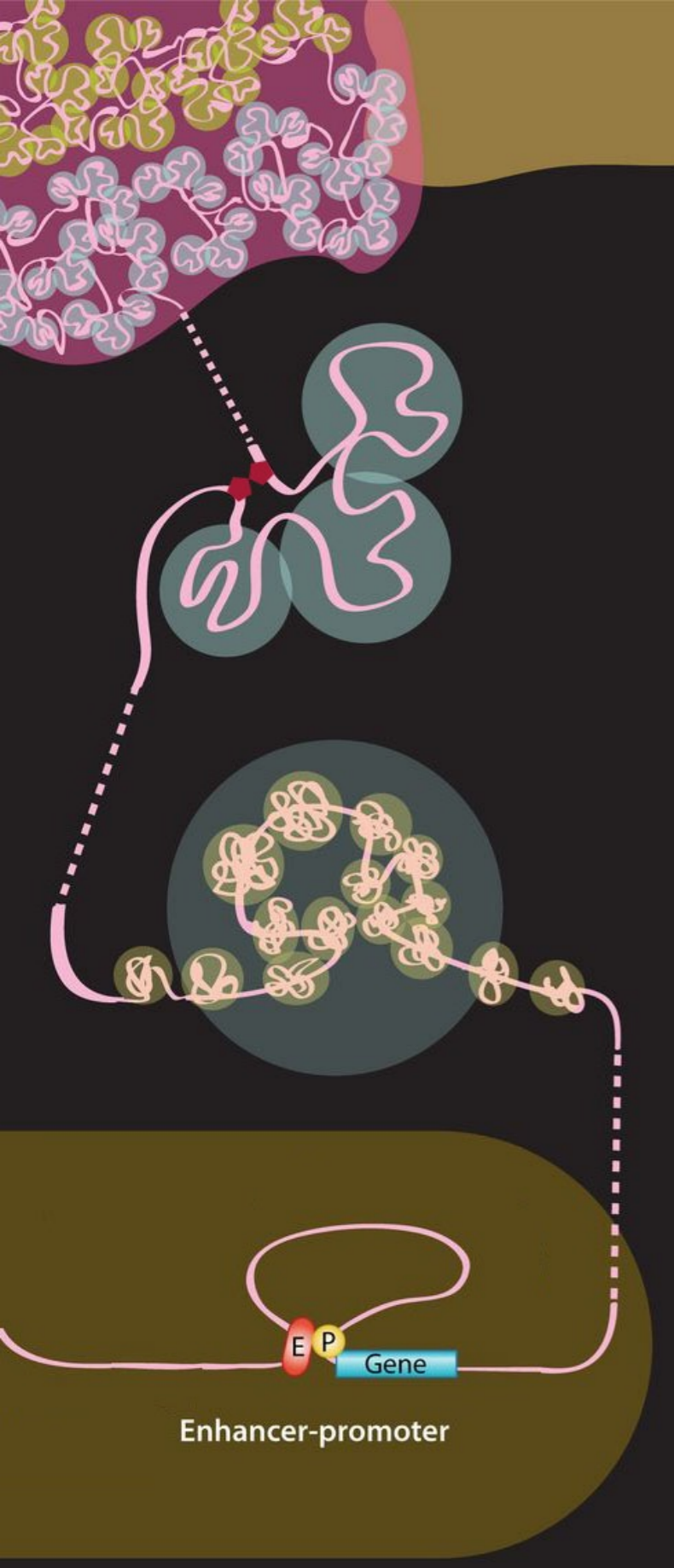
Scales of chromatin organization

Long-range loops (megabase-scale)
disruptions by genomic rearrangements

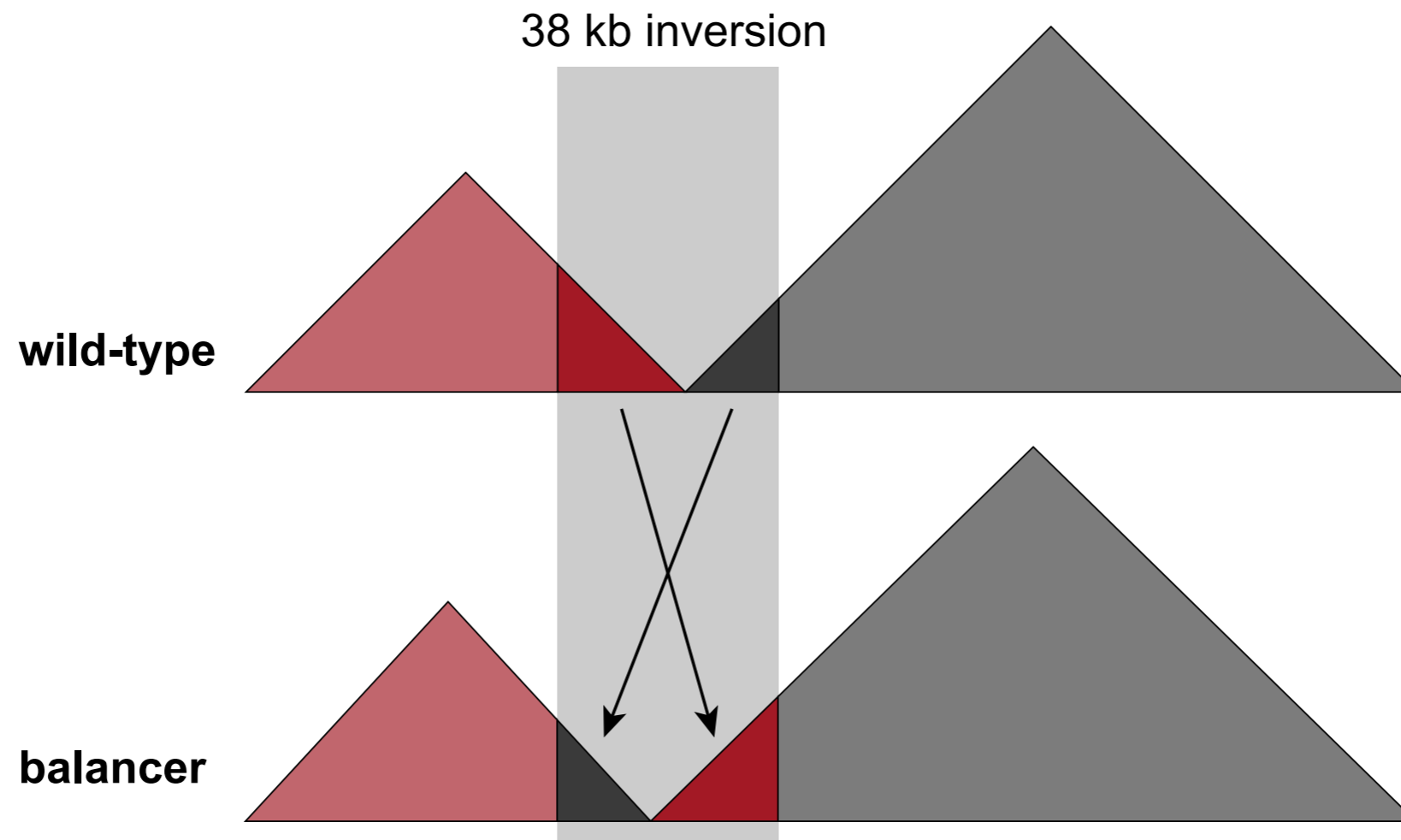
Topologically Associating Domains (TADs)
disruptions by genomic rearrangements
differences in TAD structure between alleles

Contact frequency within TADs

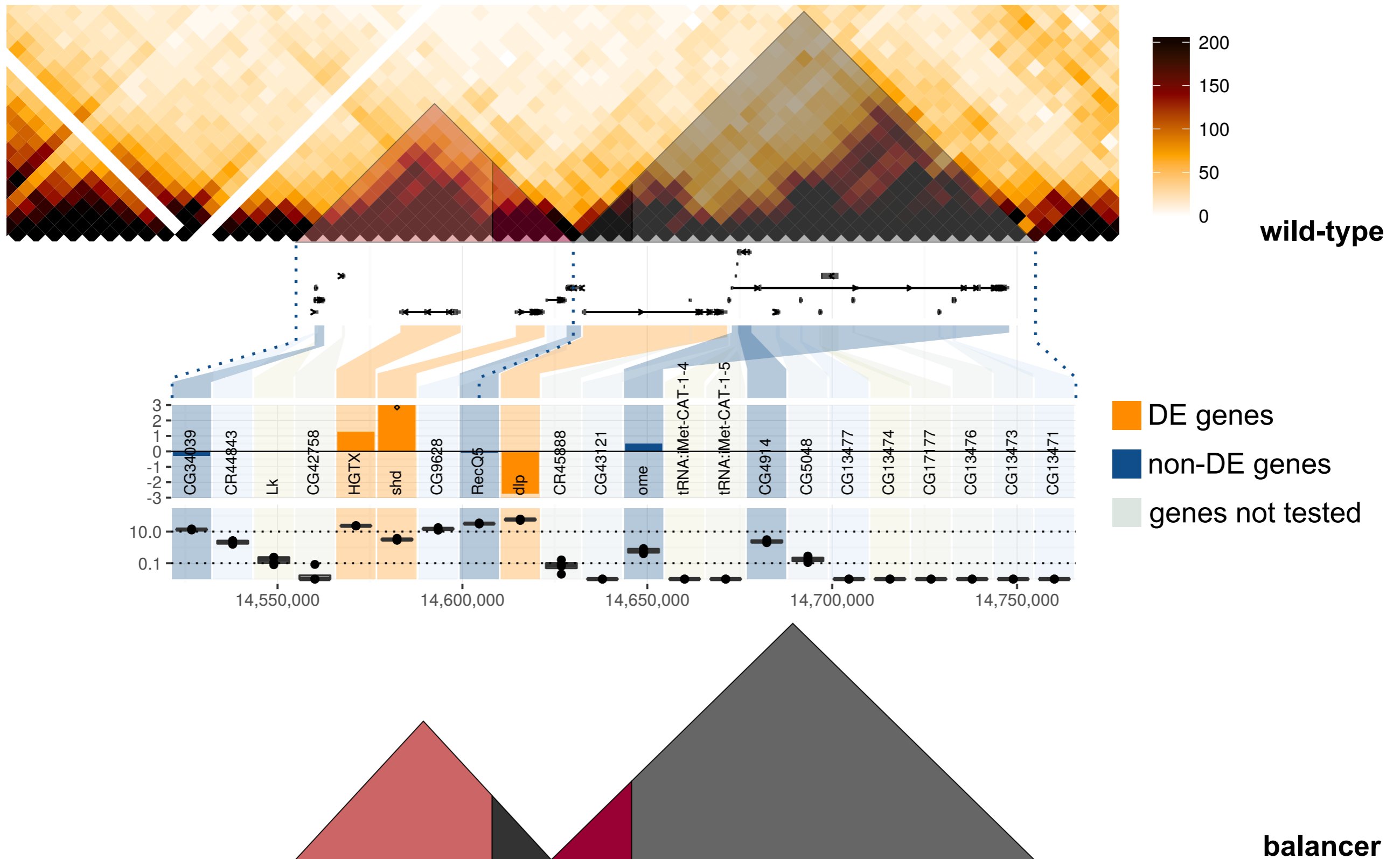
Enhancer-promoter loops



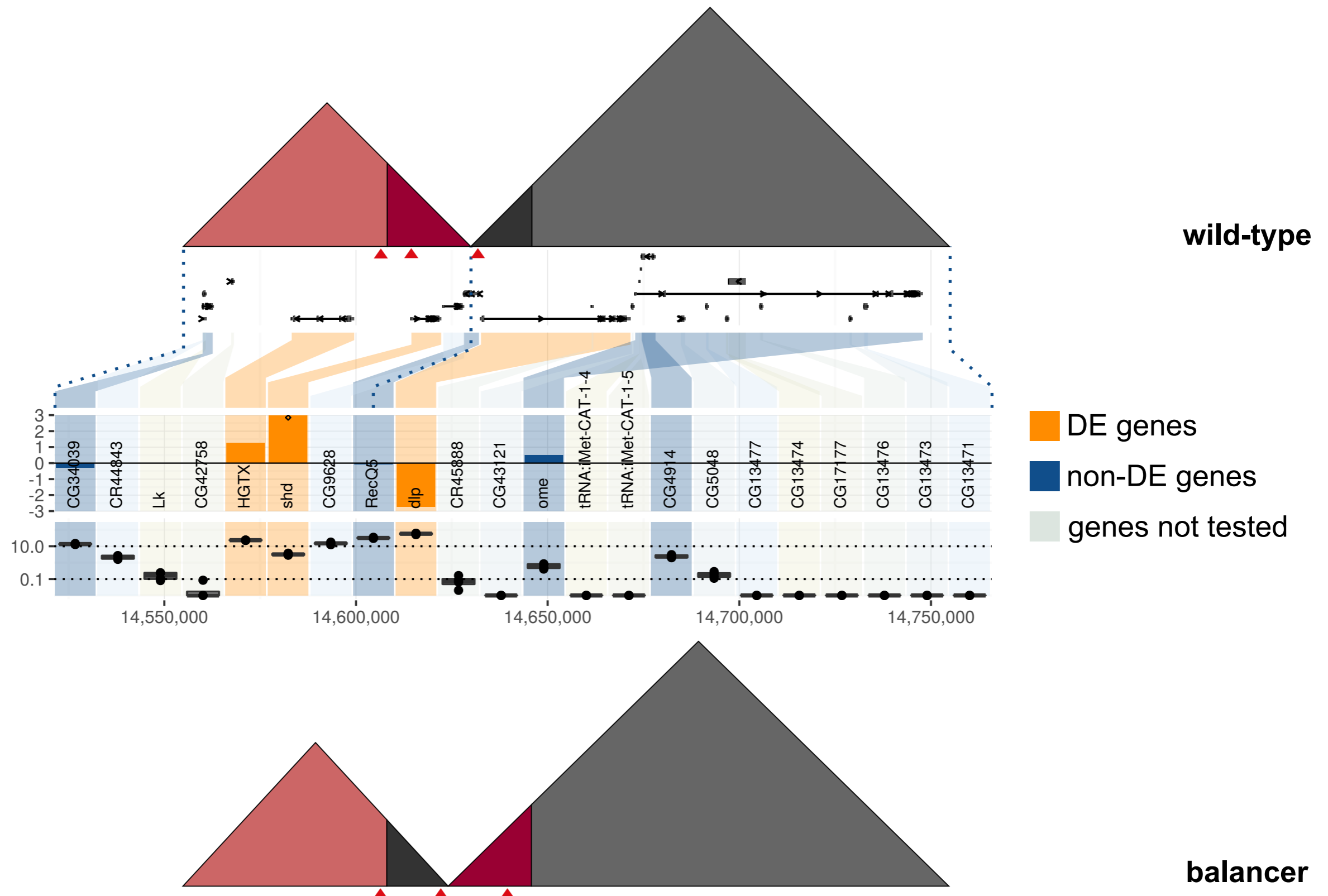
TADs are broken by genomic rearrangements



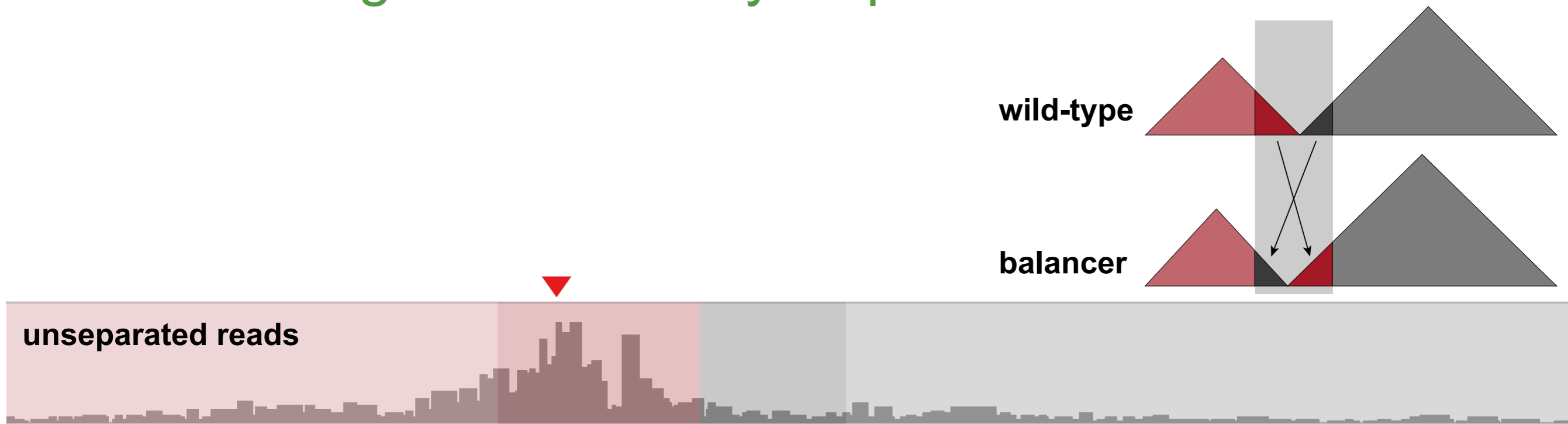
Inversion breaks TADs and affects gene expression



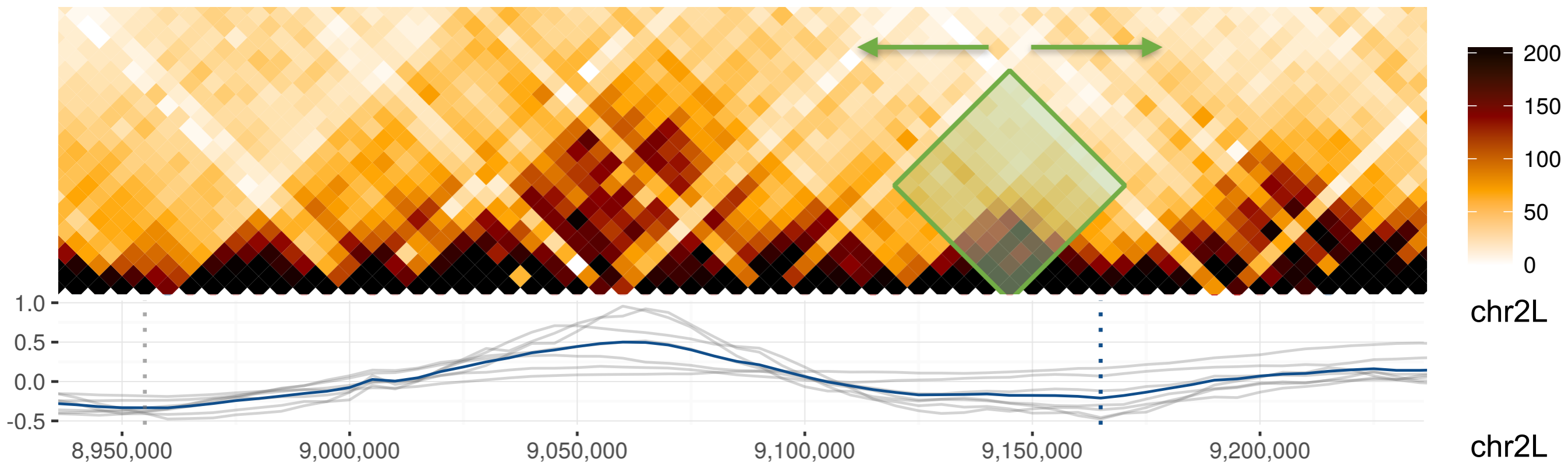
Inversion breaks TADs and affects gene expression



Local chromatin contacts changed by the inversion was investigated further by Capture-C

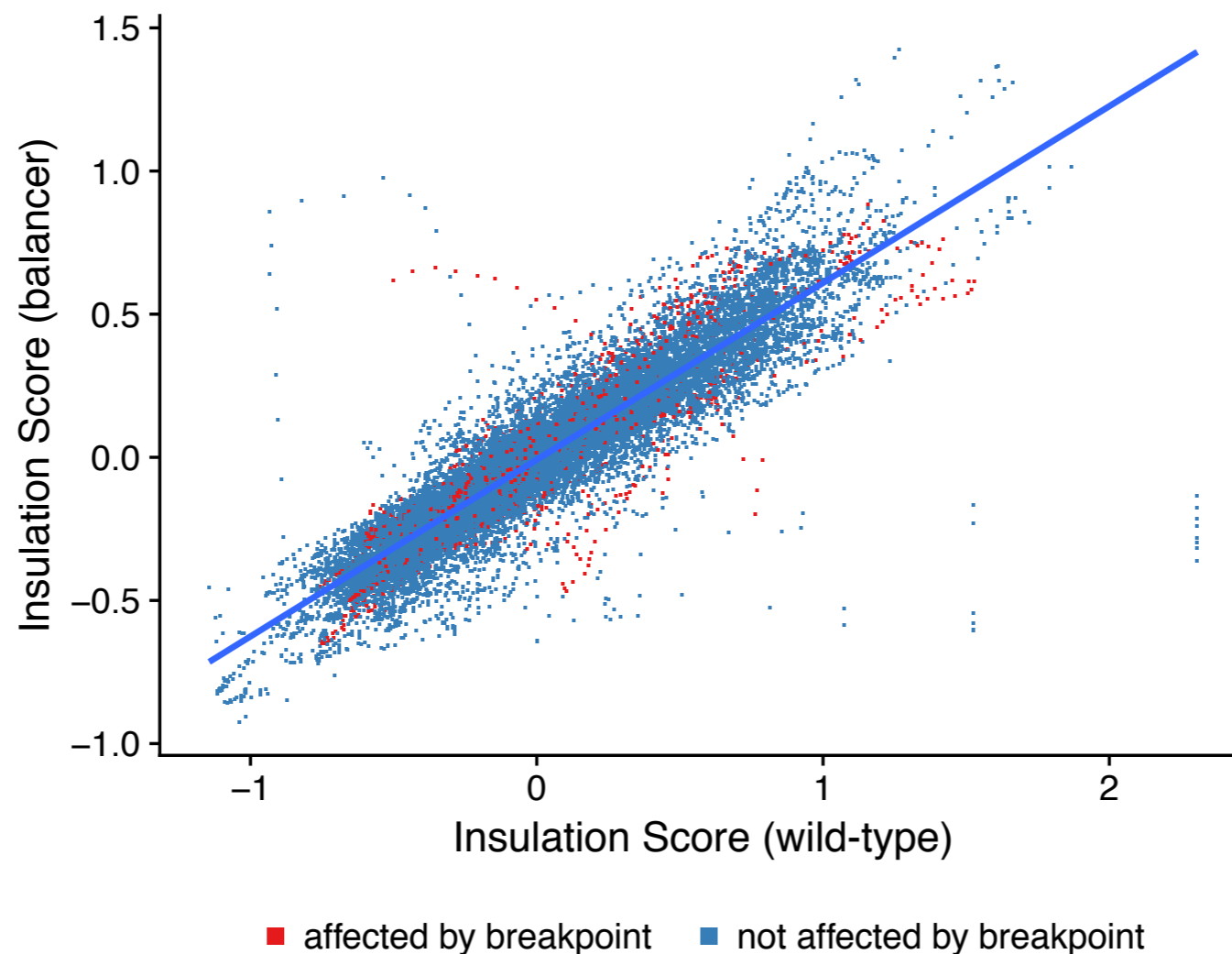


Insulation Score as a metric to identify Topologically Associating Domains (TADs)



TAD boundaries correspond to local minima in the Insulation Score profile.

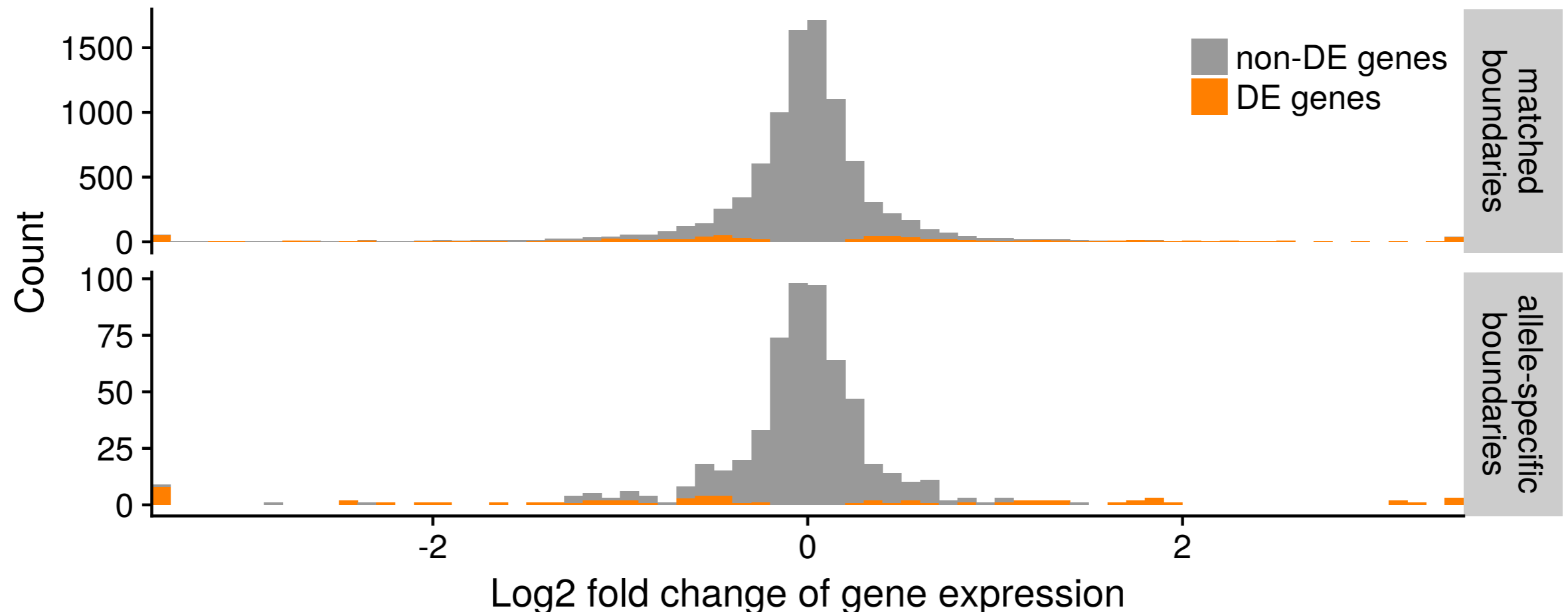
Insulation Scores are highly correlated between alleles — but some boundaries might differ



Pearson's $r = 0.913$

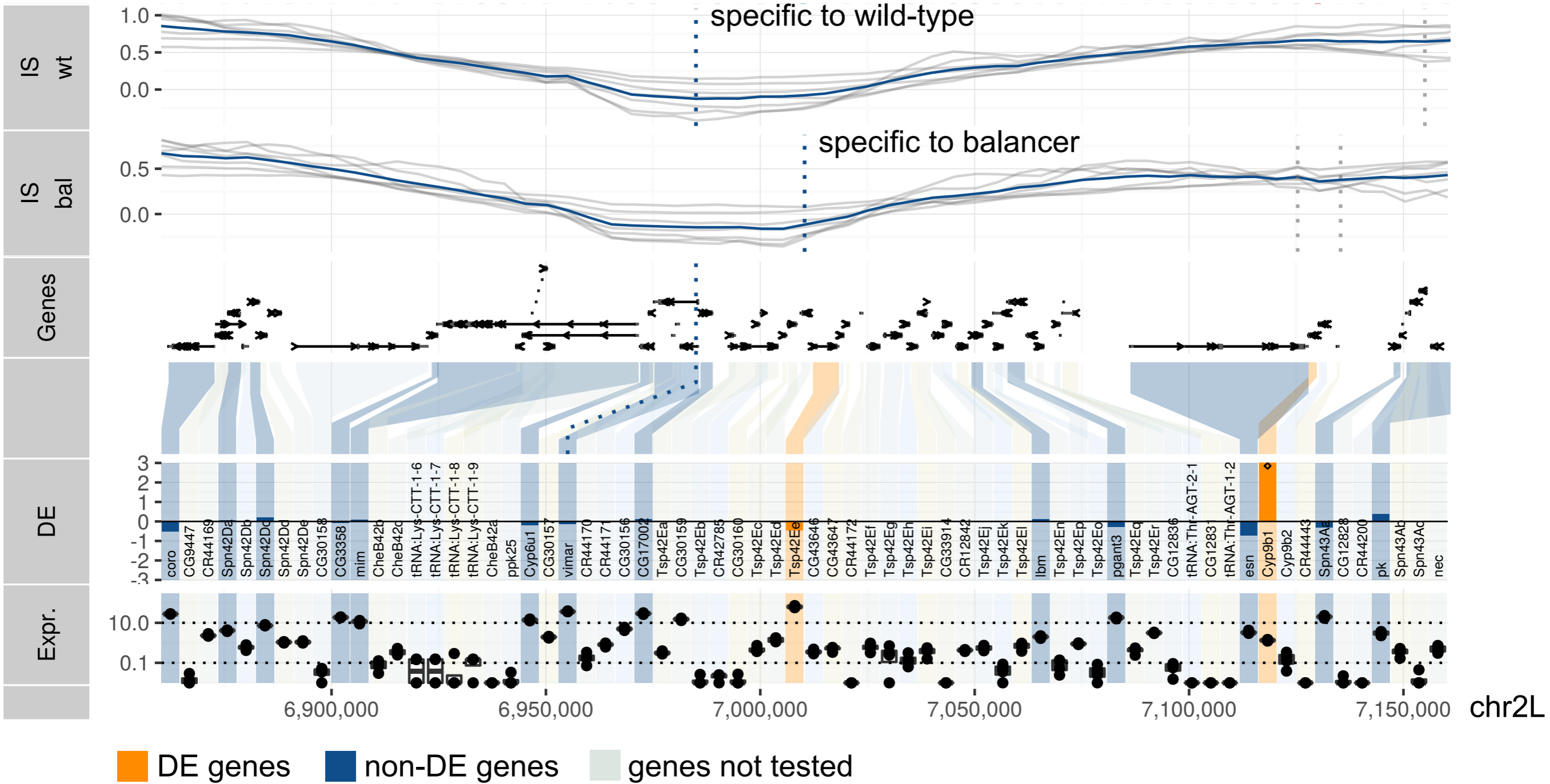
Globally, changes in TAD boundaries do not imply differential gene expression

We identified 441 TAD boundaries matched between the alleles (up to 25 kb shift), as well as 36 wild-type-specific and 22 balancer-specific boundaries.



Genes associated to an allele-specific TAD boundary are more likely to be DE (10.9% vs. 8.5%). The difference is not significant ($p = 0.05$, Fisher's exact test).

Example of allele-specific TAD boundaries



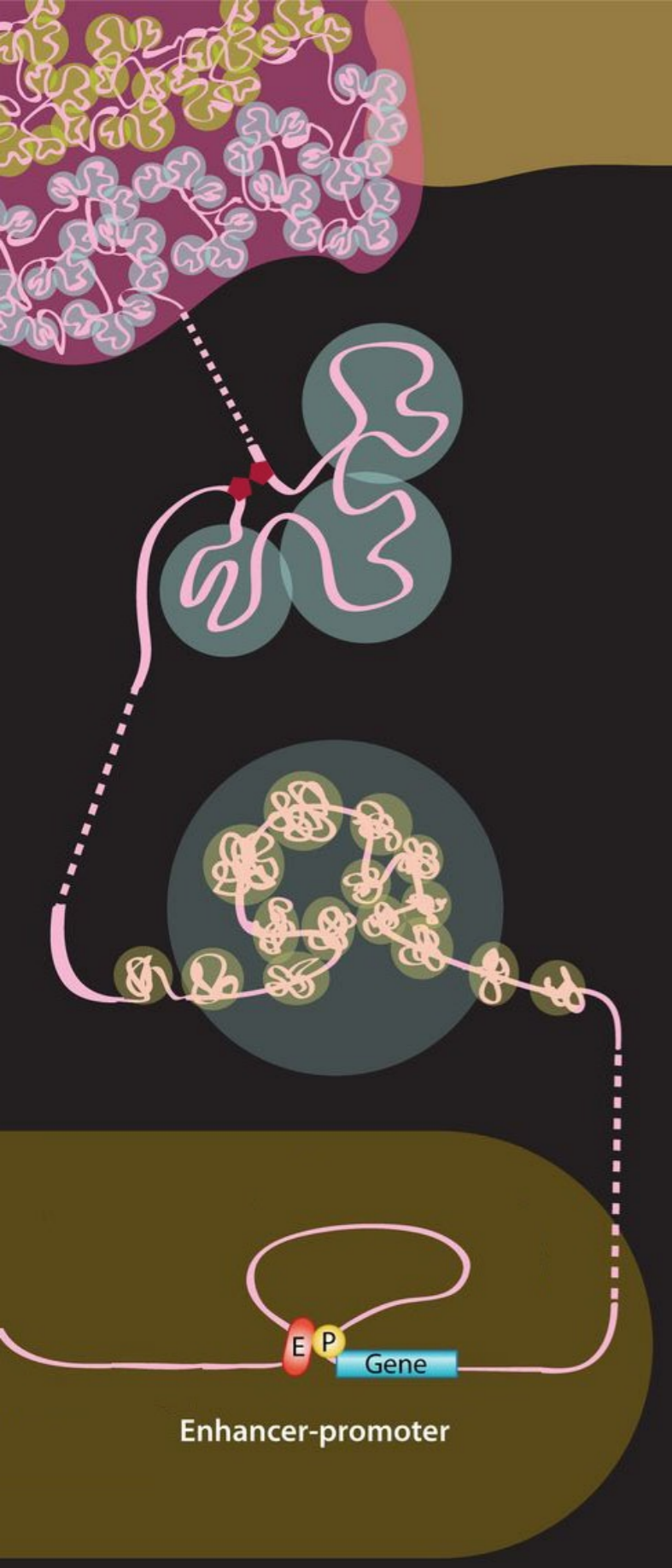
Scales of chromatin organization

Long-range loops (megabase-scale)
disruptions by genomic rearrangements

Topologically Associating Domains (TADs)
disruptions by genomic rearrangements
differences in TAD structure between alleles

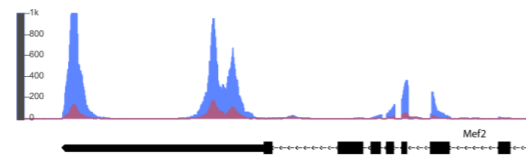
Contact frequency within TADs

Enhancer-promoter loops

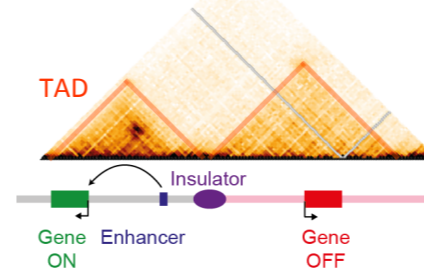


Recap: Differential gene expression is correlated with local changes in genome topology

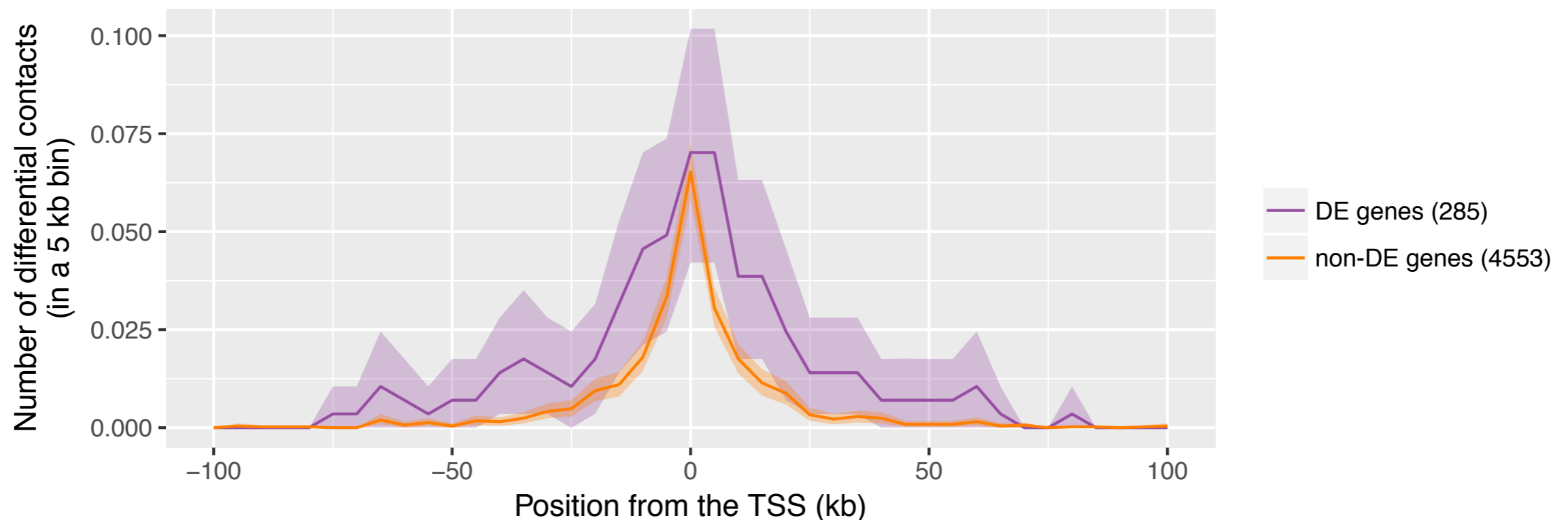
Gene expression



Chromatin organization



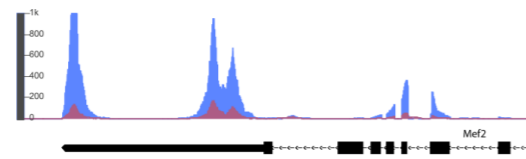
Differential Hi-C contacts from gene promoters



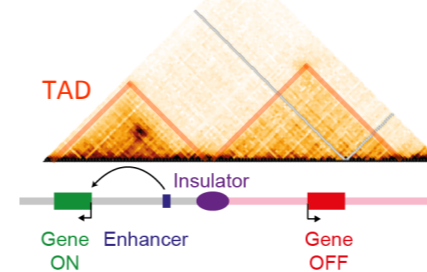
DE genes have on average 0.57 differential Hi-C contacts, non-DE genes have 0.24.
 $p = 5.99 \cdot 10^{-6}$, Wilcoxon test with a triangular kernel. 95% CI obtained by bootstrapping.

Local changes in genome topology do not imply differential gene expression

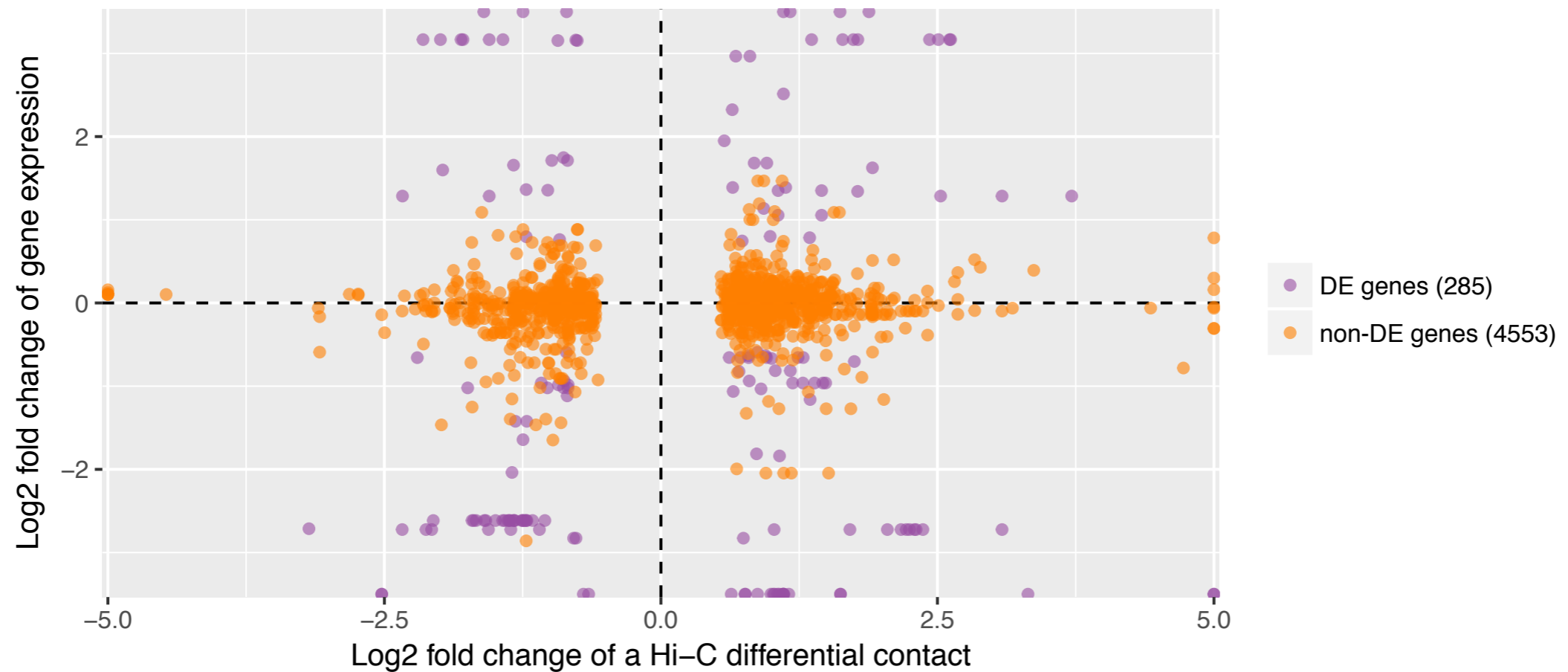
Gene expression



Chromatin organization

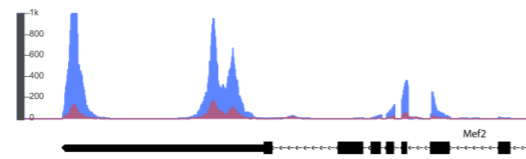


Differential Hi-C contacts from gene promoters

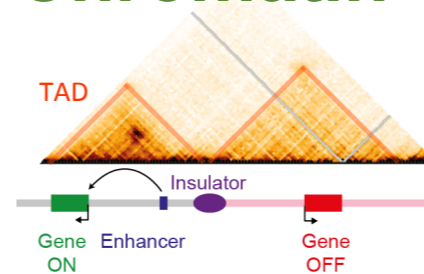


Higher resolution (Capture-C): Differential gene expression is correlated with local changes in genome topology

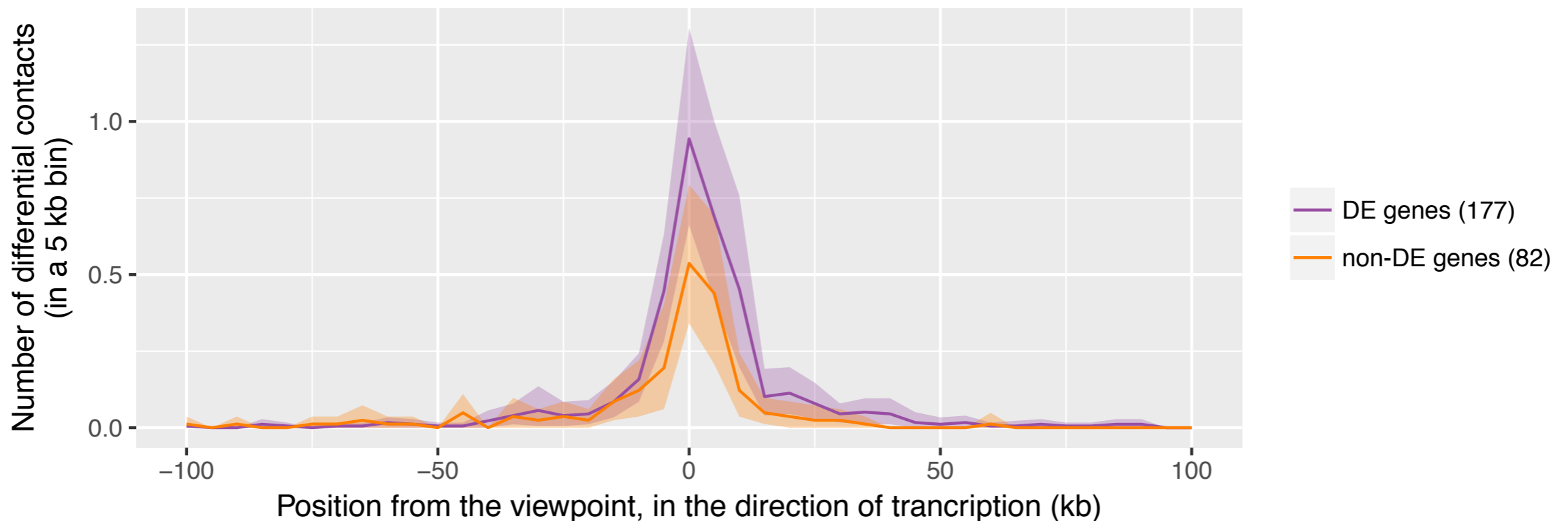
Gene expression



Chromatin organization



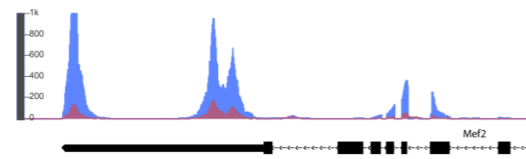
Differential Capture-C contacts from gene promoters



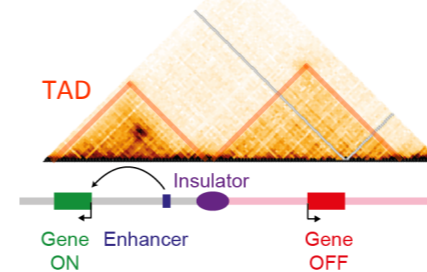
DE genes have on average 3.6 differential Capture-C contacts, non-DE genes have 1.9. $p = 0.085$, Wilcoxon test with a triangular kernel. 95% CI obtained by bootstrapping.

Higher resolution (Capture-C): Local changes in genome topology do not imply differential gene expression

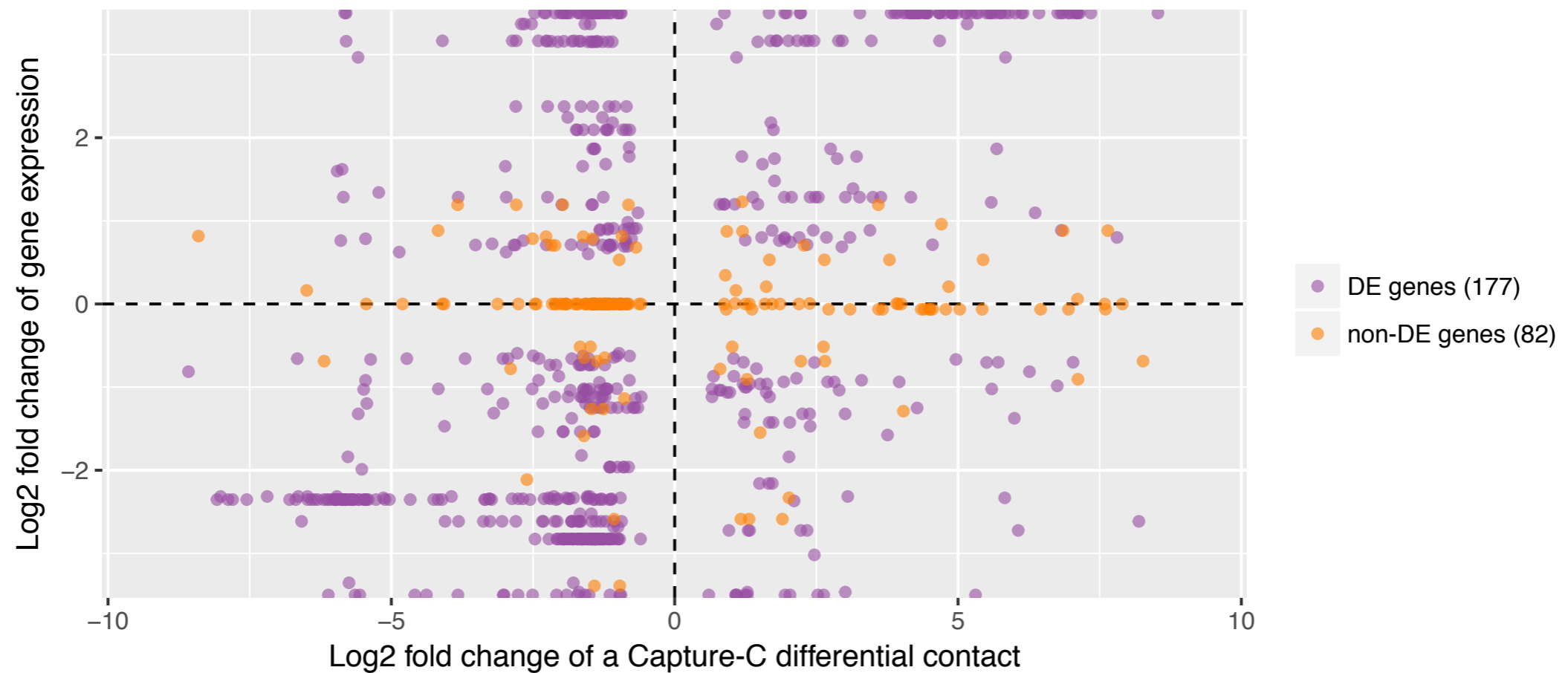
Gene expression



Chromatin organization



Differential Capture-C contacts from gene promoters



Summary

- We use allele-specific Hi-C and RNA-seq on rearranged genomes to dissect the functional impact of genomic rearrangements on gene expression and chromatin organization during embryogenesis in *Drosophila melanogaster*.
- The experimental design obtains a direct readout from the same embryo in an allele-specific manner.
- In line with previous studies, we found that differential gene expression is correlated with local changes in genome topology.
- Surprisingly, we observed that changes in large-scale chromatin organization do not globally correlate with changes in gene expression. There are cases where they do correlate, but in general they do not.
- Overall, our results are indicative of very robust mechanisms buffering genomic variation.

Acknowledgements



Rebecca
Rodríguez Viales

Eileen Furlong

Yad Ghavi-Helm



Sascha Meiers



Jan Korbel