



› AI FOR 3D GENOME

Seeking the genome determinants of 3D organisation

➤ Studied questions

- ▶ Testing *in silico* KO to detect determinant elements
- ▶ Testing *Orca* model robustness
 - ▶ How good is the prediction for other species?
 - ▶ What is the prediction over random sequences?

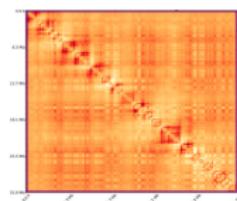
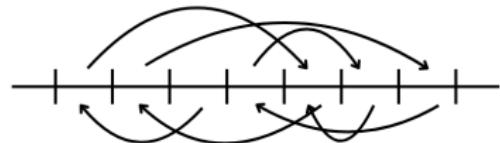
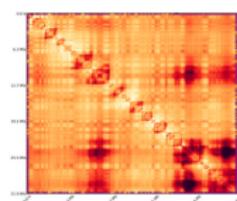
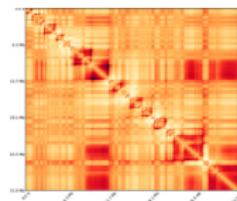


INRAE

AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

➤ Experimental procedure

Get DNA sequence → Mutate (*in silico*) → Predict (Orca) → Analyse

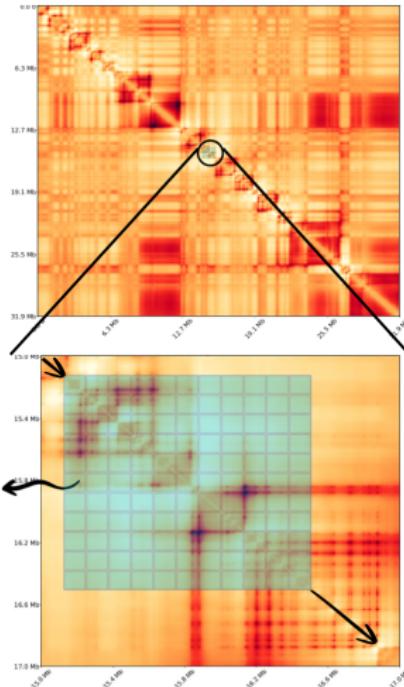


➤ Studied indicators

2 scores :

- ▶ Insulation score (IS)
- ▶ PC1

$$IS_i = \sum \frac{\text{NB}}{\text{NB}}$$



➤ Functional elements studied

- ▶ CTCF binding sites : Protein implied in genome regulation and gene expression
- ▶ H3K4me3 binding sites : Fixation site for promotors
- ▶ H3K9me3 binding sites : Role in constitutive heterochromatin formation
- ▶ CDS : Coding sequences
- ▶ PhastCons : Conserved elements (here of size 50pb) excluding CDS



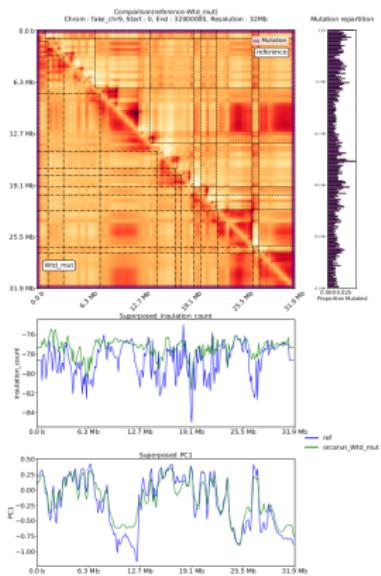
INRAE

AI for 3D genome

Florian FORGÉ JSGA July 1st 2025

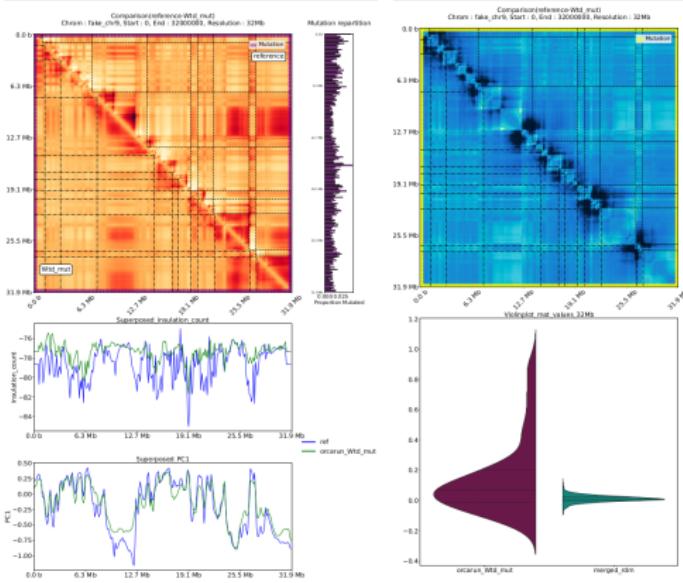
▶ Controls - Mutating CTCF sites

Shuffling each CTCF binding site sequence



➤ Controls - Mutating CTCF sites

Shuffling each CTCF binding site sequence

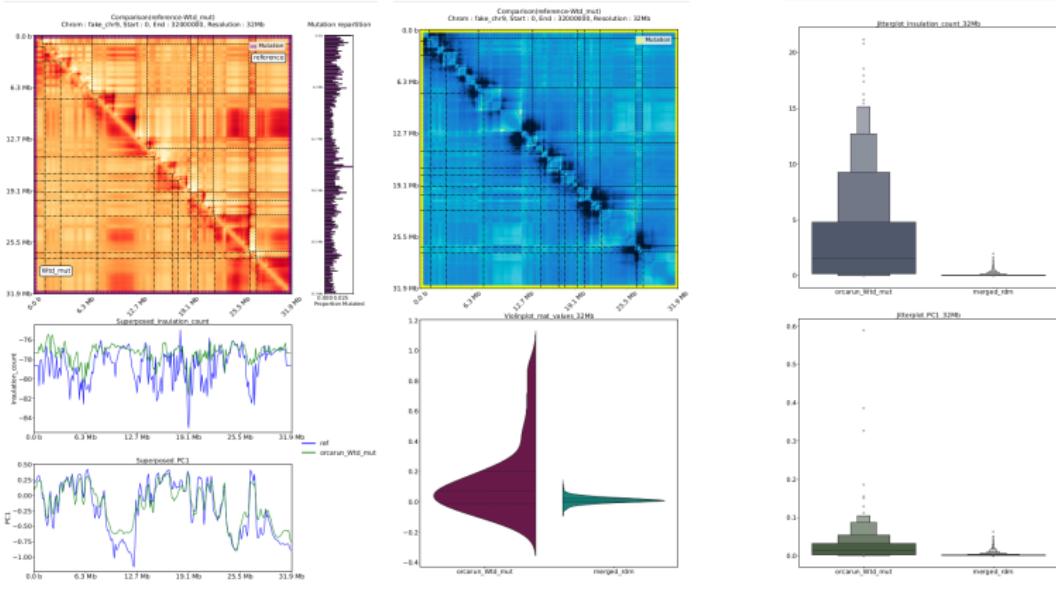


INRAE

AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

➤ Controls - Mutating CTCF sites

Shuffling each CTCF binding site sequence



➤ Ways to produce random mutations



Mutation reference (H3K9me3)



Globally shuffled mutated intervals
(from one of the random test)

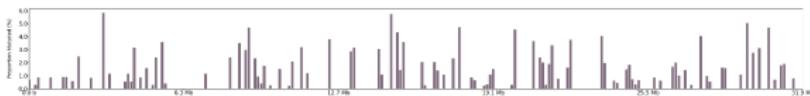
➤ Ways to produce random mutations



Mutation reference (H3K9me3)



Globally shuffled mutated intervals
(from one of the random test)



Mutated intervals shuffled per bin
(from one of the random test)



INRAE

AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

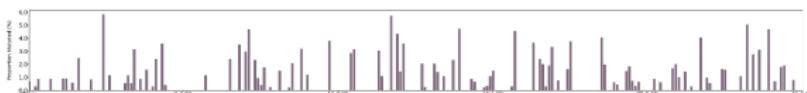
➤ Ways to produce random mutations



Mutation reference (H3K9me3)



Globally shuffled mutated intervals
(from one of the random test)



Mutated intervals shuffled per bin
(from one of the random test)



Mutated intervals shifted per bin
(from one of the random test)

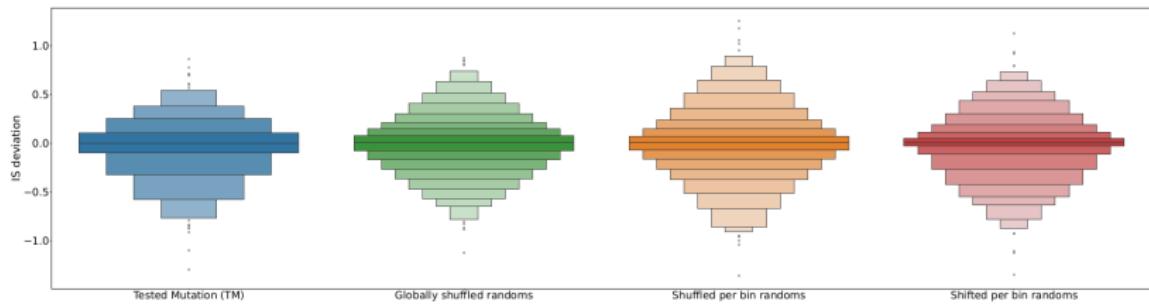


INRAE

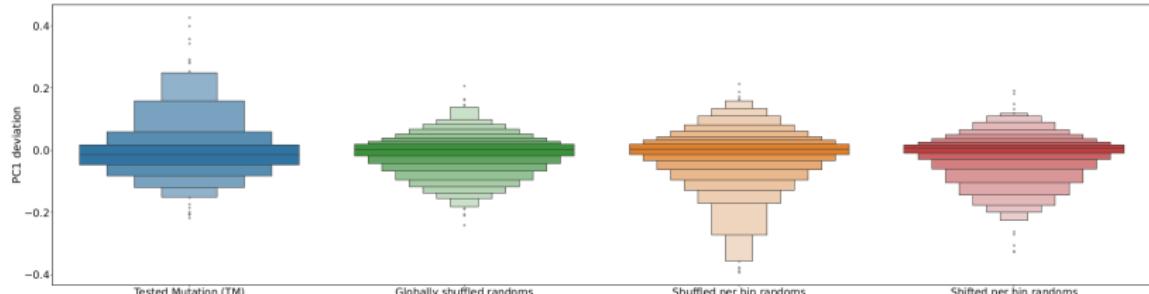
AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

➤ Impact of the random procedure

Compared IS deviation for H3K9me3

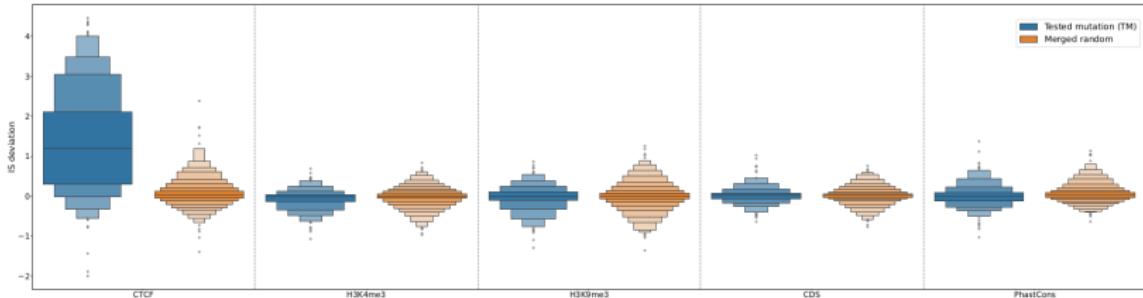


Compared PC1 deviation for H3K9me3

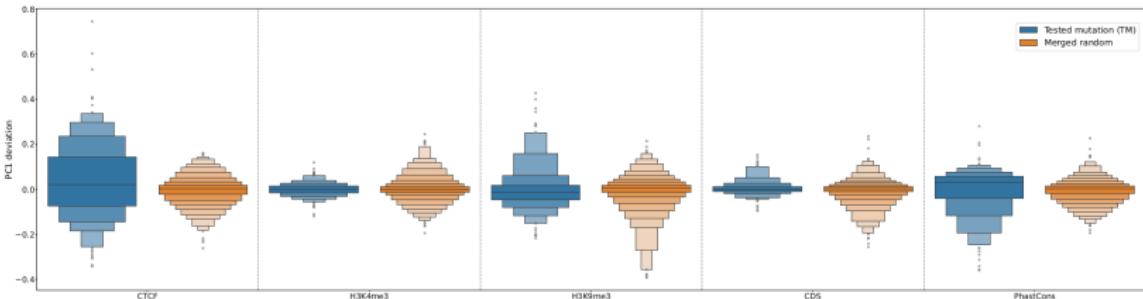


➤ Functional element classes

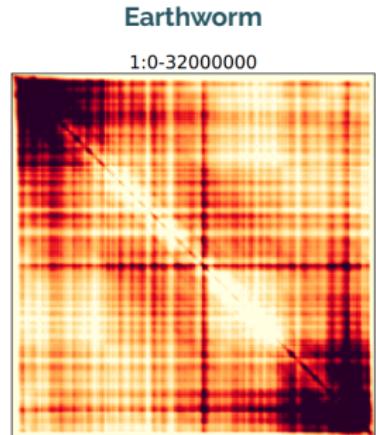
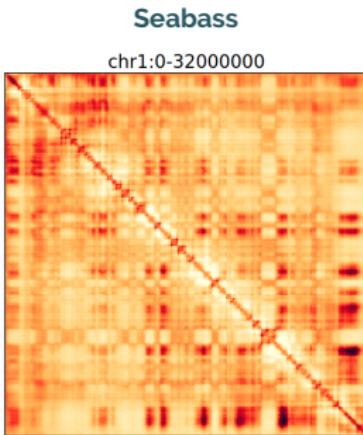
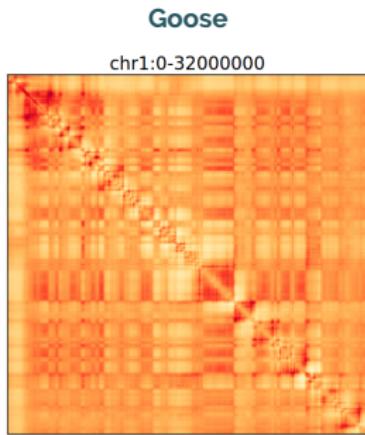
Compared IS deviation



Compared PC1 deviation

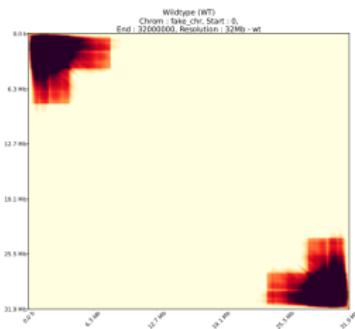


➤ Predictions for other species



► Predictions for random sequences

Random sequence

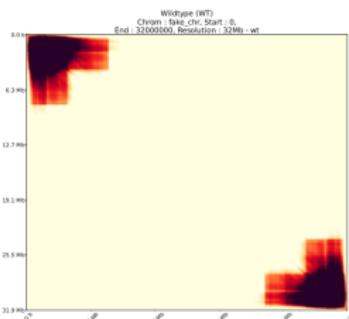


INRAE

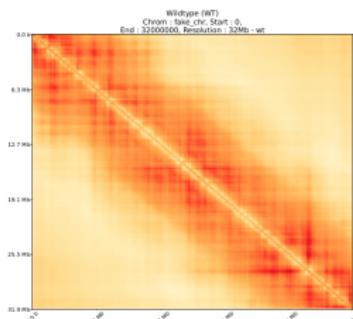
AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

► Predictions for random sequences

Random sequence



Random sequence (variable probability)

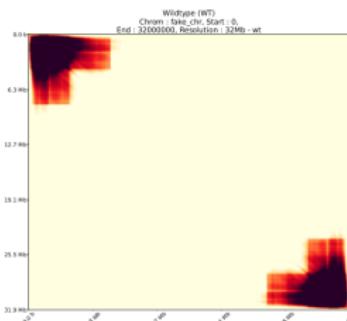


INRAE

AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

► Predictions for random sequences

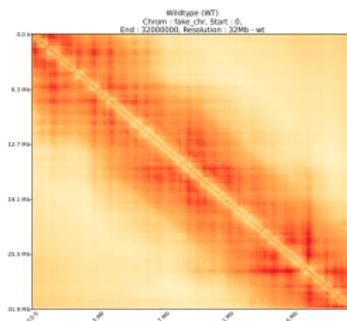
Random sequence



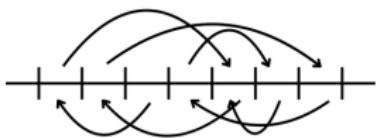
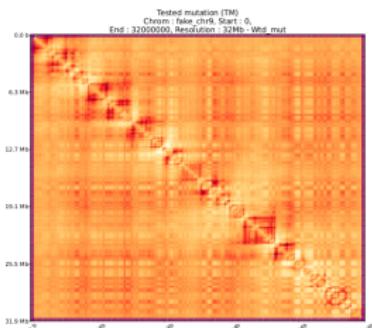
INRAE

AI for 3D genome
Florian FORGÉ JSGA July 1st 2025

Random sequence (variable probability)



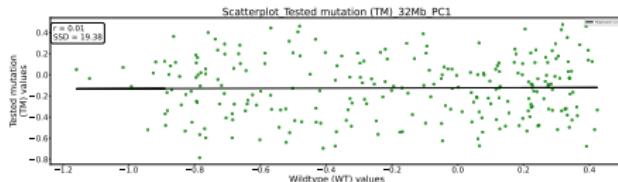
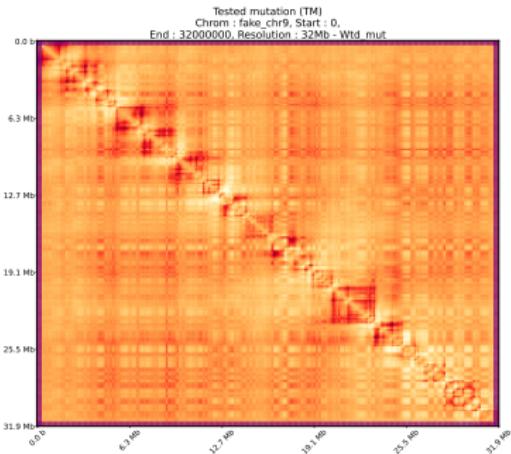
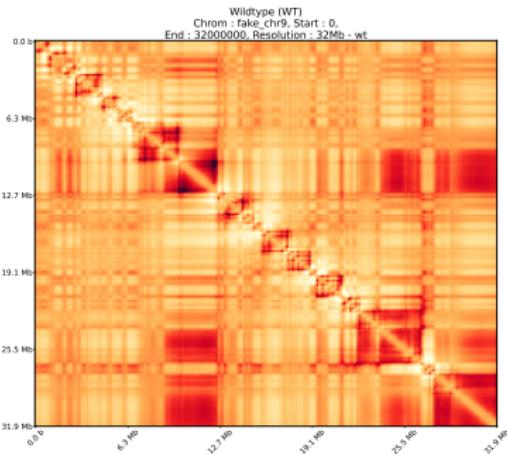
Permutation of all the bins
- shuffling 250 portions of 128kb -



➤ Impact of bin permutations

Shuffled

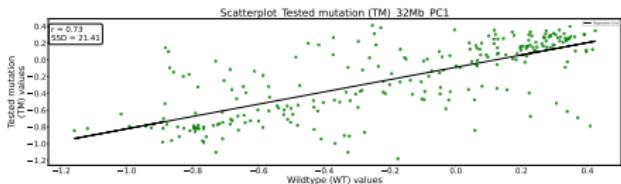
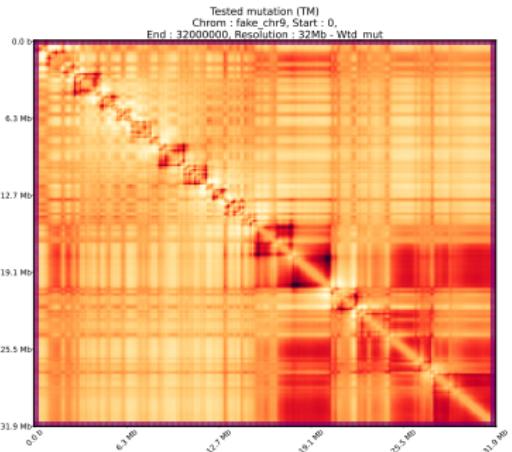
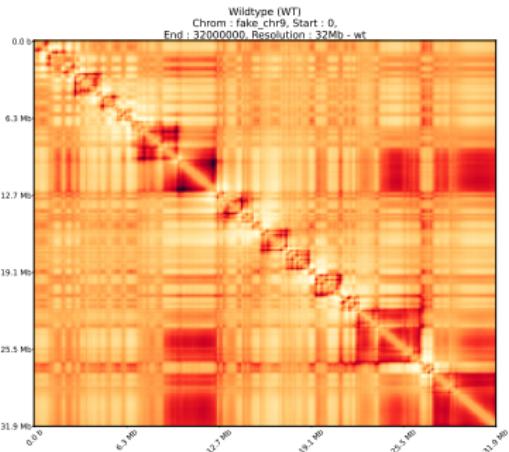
Random permutation of all the bins - shuffling 250 portions of 128kb -



➤ Impact of bin permutations

Segregate compartments A and B

Segregation of compartments - switching portions of 128kb -



➤ Any questions?

**Thank you for your
attention**



INRAE

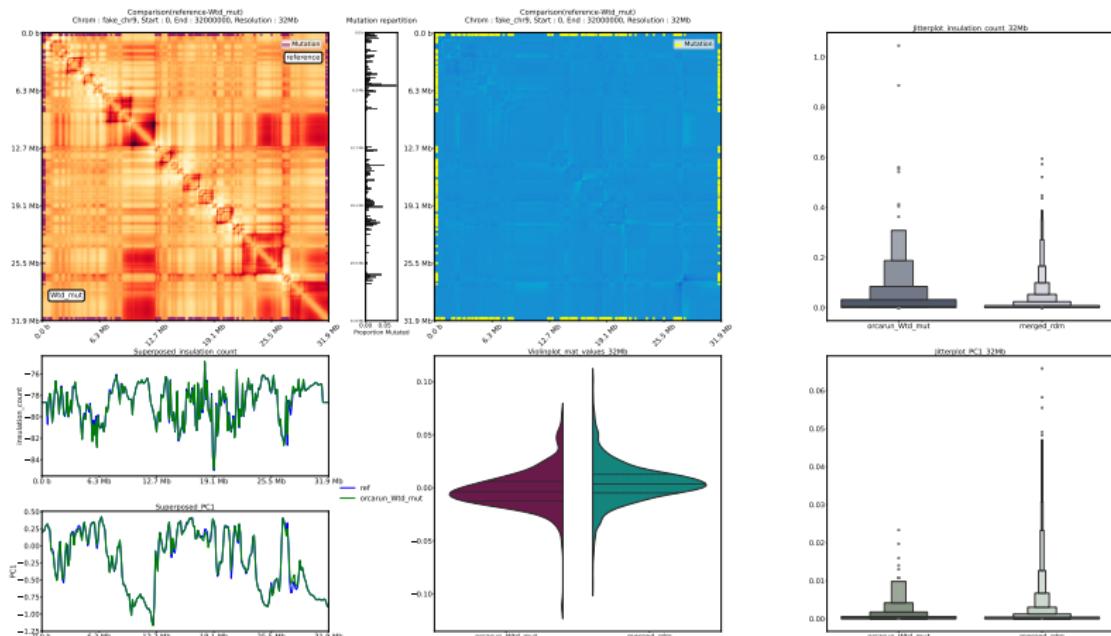
AI for 3D genome

Florian FORGÉ JSGA July 1st 2025

> Annexes

CDS

Coding sequences (187.4kb mutated - 187.4kb locally)

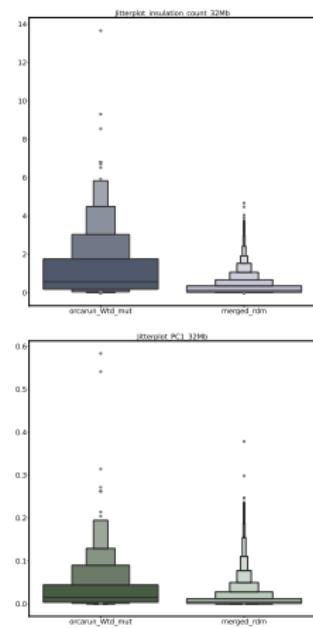
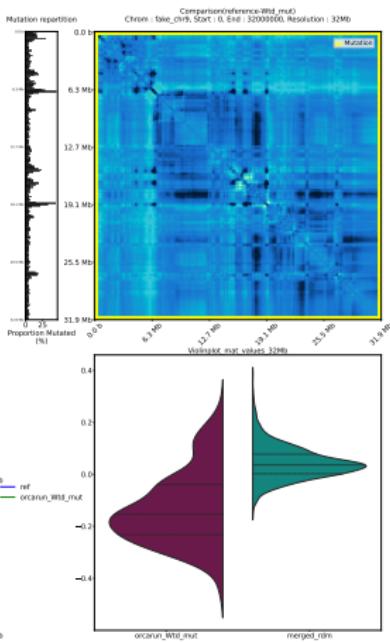
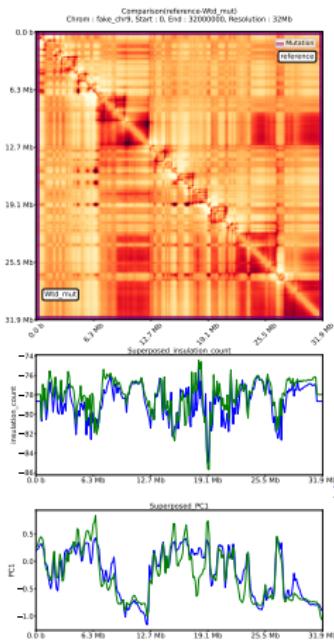


Annexes

SINE

Small Interspersed Nuclear Elements (SINE) (2.2Mb mutated - 2.2Mb locally)

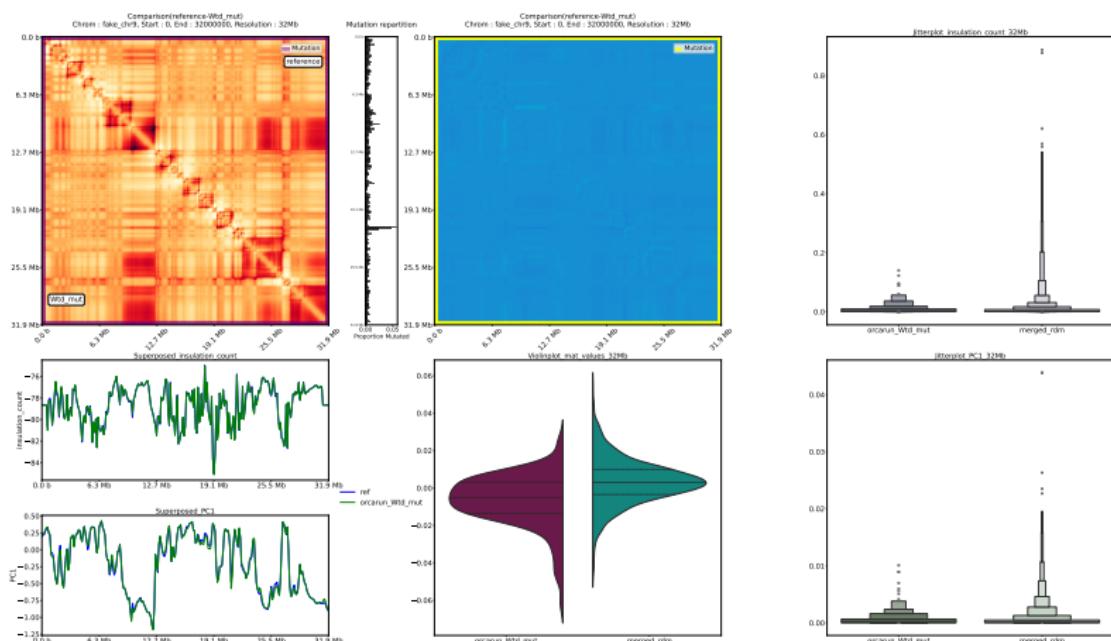
For random mutations :



➤ Annexes

PhastCons - Less conserved

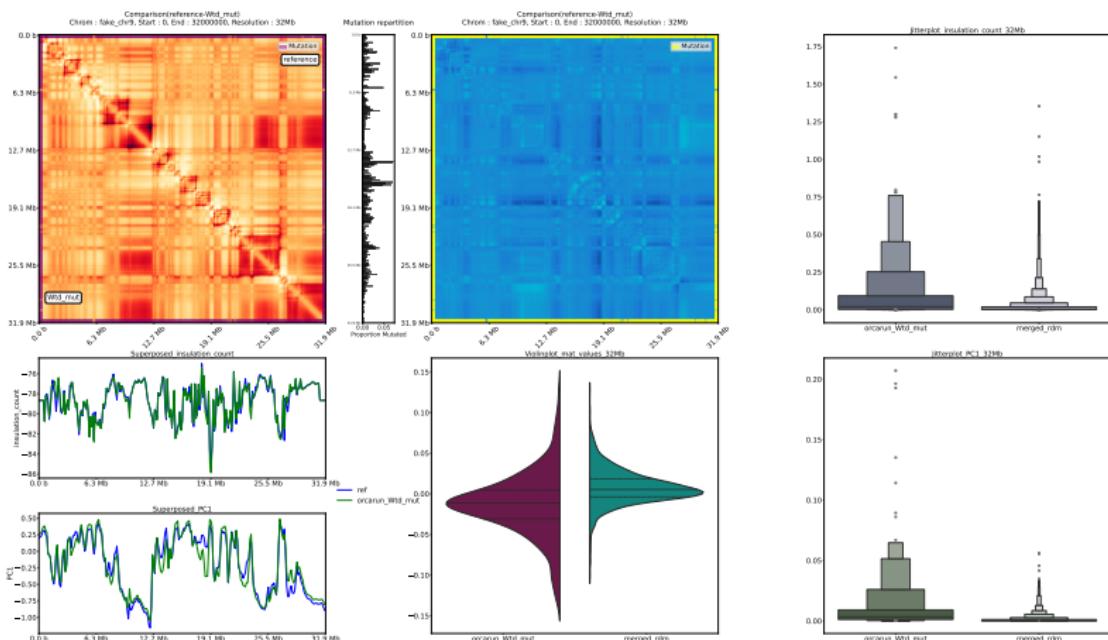
Conserved portions of 50pb -1% less conserved- (189.7kb mutated - 189.7kb locally)



Annexes

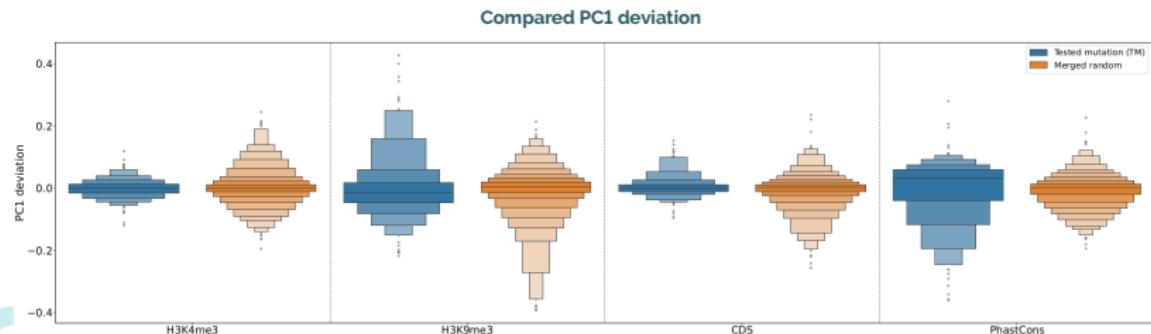
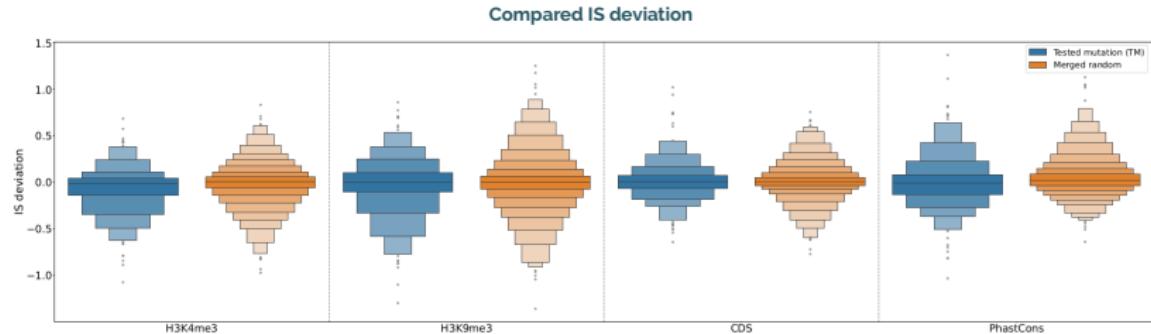
PhastCons - Most conserved

Conserved portions of 50pb -1% most conserved- (372.6kb mutated - 372.6kb locally)



Annexes

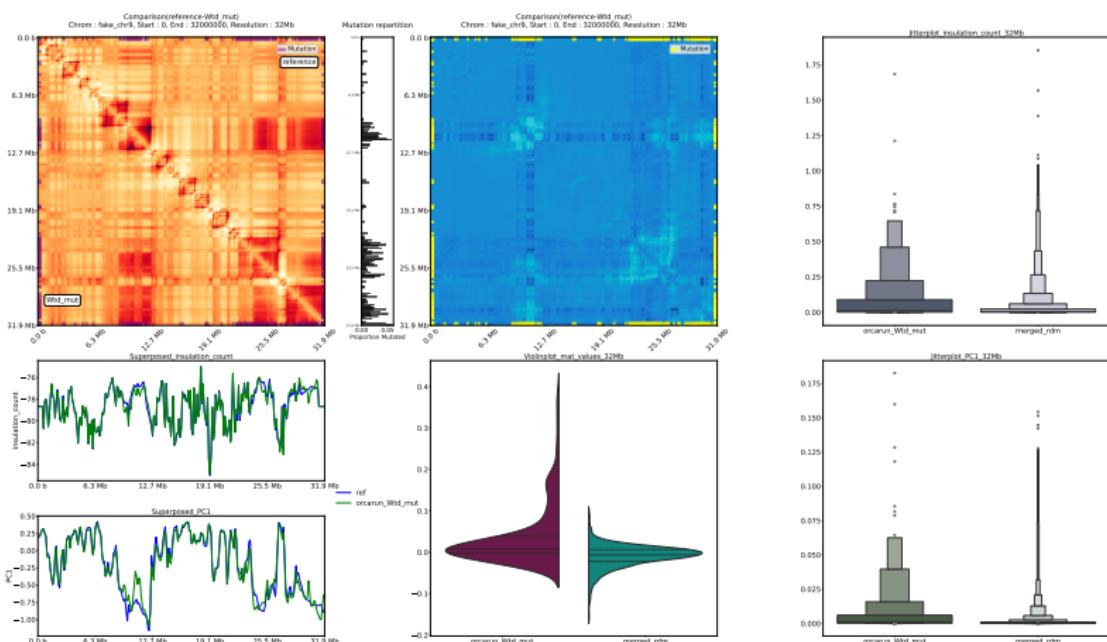
Compared scores without CTCF



Annexes

H3K9me3 with $q - \text{value} \leq 10^{-10}$

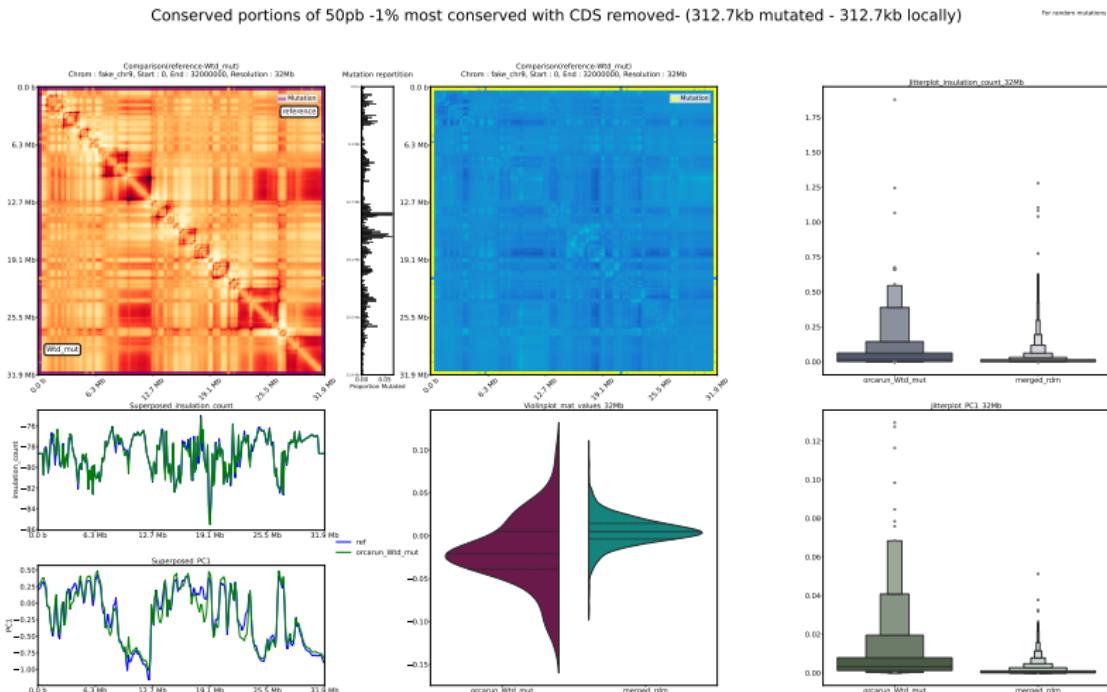
H3K9me3 (Role in constitutive heterochromatin formation) - q -value < 10^{-10} (232.2kb mutated - 232.2kb locally)



Annexes

PhastCons excluding CDS

Conserved portions of 50pb -1% most conserved with CDS removed- (312.7kb mutated - 312.7kb locally)



Annexes

Permute 5 bins (640kb) from different compartments

5 consecutive bins highly dense in SINE permuted with 5 consecutive bins from compartment B (1.3Mb mutated - 1.3Mb locally)

