Role of 3D chromatin conformation in gene regulation at the pupal stage and in the adult brain of Drosophila melanogaster

Marion Bardou

▶ **To cite this version:**

Marion Bardou. Role of 3D chromatin conformation in gene regulation at the pupal stage and in the adult brain of Drosophila melanogaster. Genomics [q-bio.GN]. Université de Montpellier, 2024. English. NNT : . tel-04718178

HAL Id: tel-04718178
https://hal.science/tel-04718178v1

Submitted on 2 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biologie Santé

École doctorale Sciences Chimiques et Biologiques pour la Santé, CBS2

Unité de recherche Centre de Biologie Structurale (CBS), INSERM U1054

Role of 3D chromatin conformation in gene regulation at the pupal stage and in the adult brain of *Drosophila melanogaster*

Présentée par Marion BARDOU

Le 20 septembre 2024

Sous la direction de Marcelo NOLLMANN et Jean-Charles WALTER

Devant le jury composé de

Giacomo CAVALLI, Directeur de Recherche, IGH, Montpellier	Président du Jury
Maria Cristina GAMBETTA, Cheffe d'équipe, CIG, Lausanne	Rapportrice
Yad GHAVI-HELM, Chargée de Recherche, ENS de Lyon, Lyon	Rapportrice
Emmanuelle FABRE, Directrice de Recherche, Laboratoire Génomes, Biologie Cellulaire et Thérapeutiques, Paris	Examinatrice
Jean-Charles WALTER, Chargé de Recherche, L2C, Montpellier	Directeur de thèse
Marcelo NOLLMANN, Directeur de Recherche, CBS, Montpellier	Directeur de thèse

UNIVERSITÉ DE MONTPELLIER

Acknowledgements

First and foremost, I would like to extend my deepest gratitude to the jury members and the thesis committee for dedicating some of their valuable time to evaluate the work that I have done over the past four years. Your expertise and feedback have been, and will be, invaluable.

I am profoundly grateful to Marcelo Nollmann, the LabMuse EpiGenMed, and the Fondation pour la Recherhe Médicale (FRM) for providing me with the opportunity to pursue this fascinating project. My recruitment through the LabMUSE EpiGenMed collaborative PhD program allowed me to collaborate closely with Giacomo Cavalli and his PhD student Nazli Akilli, as well as Andrea Parmeggiani and his PhD student Loucif Remini. Working with all of you has been a true pleasure. Our extensive discussions on genetics, epigenetics, chromatin conformation, polymer physics, and modelling have significantly enhanced my scientific acumen. I would like to particularly thank Giacomo for his feedback on my work over the years. Your willingness to share your insights has been greatly appreciated. Andrea, your unwavering support, kindness, and curiosity about my project and me have been immensely encouraging. Loucif, you have been an exceptional teammate during the final months of my thesis. Thank you for patiently explaining how your work could be applied to my data, countless times. Together, we achieved remarkable results from my *Drosophila* brains, which would not have been possible without your skills, support and kindness.

I am also extremely thankful to Jean-Maurice Dura, Emmanuel Périsse and Ana Boulanger for taking the time to answer my questions about the adult *Drosophila* brain, behaviour, and the mushroom bodies, as well as for sharing the fly lines I needed for my project. I would like to extend special thanks to Emmanuel Périsse and Ana Boulanger for training me in the dissection of the adult *Drosophila* brain.

I would like to express my gratitude to my colleagues from my host lab: Anna, Antoine, Christel, David and Gautham. The meals we shared, and our lengthy conversations have been memorable. Special thanks to Marie, JB and Maxime for their assistance with my project; to Christophe for always being available to help me in the biochemistry lab and for his kindness; to Xavier for helping me troubleshoot pyHi-M bugs, for our deep conversations about ecology and politics, and for becoming a great friend. I am also thankful to former PhD students, Julian and Sara, for their tremendous support and for being the best colleagues one could hope for. I miss you guys. Last but certainly not least, Olivier, your knowledge and assistance have been pivotal to my thesis. Despite my limited background in genetics and epigenetics, you patiently answered my numerous questions. I am deeply grateful for your help and genuine kindness. I am confident you will have a well-deserved and successful academic career.

I extend my thanks to everyone at CBS, especially those from the 60, particularly Caro, Christine and Manouk, for the lunches, their kindness and the stimulating conversations. To all the current and former PhD students and postdocs from the 60 — Ari, Davide, Emilie, Enrico, Gosha, Lisa and both Matteos — thank you for the coffees, drinks and long conversations. Thanks to Andreas and Andy who were truly there for me. During the long hours at the lab, you made it feel like a second home. I am lucky to have met all of you.

Victoria and Ana-Maria, thank you for the long conversations about everything and anything while we took care of the flies. Meeting you ladies has been a pleasure.

I would also like to thank everyone who has contributed in one way or the other to my love for science and research. Specifically, I would like to thank Raphaël Kribich, who was my first internship supervisor and gave me my first glimpse at academic research and the passion for it; Sandrine Lévêque-Fort and Nicolas Bourg for giving me the opportunity to use advanced microscopy techniques to study biological samples and for playing a significant role in shaping me into a real biophysicist; and finally, Clément Cabriel, who, without intending to, inspired me to pursue a PhD after spending almost two years together. Seeing him thrive as a PhD student allowed me to understand the day-to-day life of a passionate scientist.

I am also grateful to the Women in Science network, especially Lise Chauveau, for being an incredible mentor. You encouraged me to ask questions at conferences and step out of my comfort zone to engage with other scientists. Your guidance over the year we passed together helped me understand what is truly important in Science and in my professional life. Thank you for everything, including the shared beers.

Now, I would like to thank the most important people in my life: my friends and family. Your constant support, concern for my mental health, curiosity about my flies, and the many shared meals and drinks have meant the world to me. I may not mention each of you by name, but you know how much you mean to me.

Finally, Charles, without you, none of this would have been possible. You have been my rock, always believing in me and my abilities even when I doubted myself. Our shared love of science and our long conversations about my work have been incredibly helpful. Thank you for being my greatest love and my best friend.

In conclusion, I hope that the academic environment evolves for future generations of PhD students. Working in science is both a passion and a privilege, but it is disheartening that mental health is not more valued. One cannot perform at their best if their work and time in the lab are not appreciated and recognised. Passion alone is not always enough. To all the other PhD students continuing in Academia and Science, you are achieving what I could not. I wish you the best of luck and thank you for your contributions to Society.

Résumé en français

Le génome des cellules eucaryotes présente une organisation tridimensionnelle hautement hiérarchique. Parmi ces niveaux de compartimentation, les domaines d'association topologique (TADs) sont des régions riches en interactions chromatiniennes et sont impliqués dans la régulation des gènes. Ces domaines doivent rester intacts pour maintenir la proximité physique entre les éléments régulateurs des gènes tels que les enhancers et les promoteurs, ce qui est essentiel pour une expression génique spécifique. La perturbation de la communication entre enhancers et promoteurs peut entraîner une expression génique aberrante, contribuant à des maladies, y compris des cancers ou des troubles génétiques. Ainsi, l'interaction entre les enhancers et les promoteurs est cruciale pour le maintien des fonctions cellulaires normales et pour répondre aux signaux développementaux et environnementaux.

L'objectif initial de mon projet de thèse est d'étudier les mécanismes complexes sous-jacents aux interactions enhancer-promoteur, à la conformation de la chromatine et aux profils d'expression génique dans différents types cellulaires. Pour aborder ces questions, les deux études présentées dans ce manuscrit ont été réalisées sur deux types échantillons biologiques différents : la patte de la pupe et les cerveaux adultes de la *Drosophile*. Une partie de mon projet de thèse s'est concentré sur le développement et l'application de Hi-M (une technique d'imagerie reconstruisant la conformation de la chromatine à l'échelle de la cellule unique) afin d'étudier la régulation de l'expression génique par les interactions entre promoteurs et enhancers.

Le chapitre des résultats de ce manuscrit se divise en un projet secondaire et un projet principal. Ces deux études examinent le rôle de la structure chromatinienne dans la régulation de l'expression génique au sein de différents tissus de *Drosophila melanogaster*. Dans l'étude de *Denaud et al. (2024)*, nous explorons comment une boucle entre deux éléments de réponse Polycomb (PRE), localisée au niveau du gène *dac*, influence les interactions entre enhancers et promoteurs et la transcription. Nous avons pu démontrer que la boucle PRE au locus *dac* agit comme un élément structurel qui restreint et spécifie la communication entre enhancers et promoteurs, mettant en lumière les mécanismes régulateurs complexes impliqués dans le contrôle de l'expression génique.

D'autre part, mon projet principal explore le lien entre conformation du génome et expression génique dans le cerveau adulte de la *Drosophile*. Durant ma thèse, j'ai développé des outils pour détecter la conformation du génome et identifier les interactions enhancer-promoteur spécifiques à différents types de neurones dans le cerveau adulte de la *Drosophile*. Je me suis particulièrement concentré sur l'étude de la conformation de chromatine de régions génomiques incluant des gènes impliqués dans la formation de la mémoire.

En examinant les interactions entre les enhancers et les promoteurs et les motifs de repliement de la chromatine dans les cellules de Kenyon, mon travail complète les résultats du papier *Denaud et al. (2024)* en fournissant des informations sur les réseaux régulateurs plus largement impliqués dans un tissu différencié. L'exploration des interactions différentielles entre enhancer et promoteur ainsi que l'étude de la conformation de la chromatine, que nous avons réalisées, devraient participer à une meilleure compréhension de la manière dont la conformation de la chromatine, telle que la boucle PRE situé dans le locus de *dac*, peuvent être impliquées dans l'expression de certains gènes dans des tissus complexes.

En conclusion, l'article de recherche de *Denaud et al. (2024)* et mon projet principal participent collectivement à une meilleure compréhension des mécanismes régulateurs médiés par les structures chromatiniennes dans le contrôle de l'expression génique. Avec ce travail, nous espérons que nos études

fourniront des informations précieuses à la communauté, bien que des investigations supplémentaires sur les processus moléculaires soient nécessaires pour expliquer le rôle de la conformation de la chromatine dans des mécanismes complexes tels que l'expression génique ou la formation de la mémoire chez la *Drosophile*.

English summary

The genome of eukaryotic cells exhibits a highly hierarchical three-dimensional organisation. Among these levels of compartmentalisation, topologically associating domains (TADs) are regions rich in chromatin interactions and are involved in gene regulation. These domains must remain intact to maintain the physical proximity between gene regulatory elements such as enhancers and promoters, which is essential for precise gene expression. Disruption of enhancer-promoter communication can lead to aberrant gene expression, contributing to diseases including cancers and genetic disorders. Thus, the enhancer-promoter interaction is critical for sustaining normal cellular functions and for responding to developmental and environmental cues.

The initial objective of my thesis project is to study the complex mechanisms underlying enhancer-promoter interactions, chromatin conformation, and gene expression profiles in different cell types. To address these questions, the two studies presented in this manuscript were conducted on two different types of biological samples: the pupal leg and the adult brains of *Drosophila melanogaster*. Part of my thesis project focused on the development and application of Hi-M (an imaging technique that reconstructs chromatin conformation at the single-cell level) to study the regulation of gene expression by enhancer-promoter interactions.

The results chapter of this manuscript is divided into a secondary project and a primary project. Both studies examine the role of chromatin structure in gene expression regulation within different tissues of *Drosophila melanogaster*. In the study by *Denaud et al.* (2024), we explore how a loop between two Polycomb response elements (PRE), located at the *dac* gene, influences enhancer-promoter interactions and transcription. We demonstrated that the PRE loop at the *dac* locus acts as a structural element that restricts and specifies communication between enhancers and promoters, shedding light on the complex regulatory mechanisms involved in gene expression control.

On the other hand, my primary project explores the link between genome conformation and gene expression in the adult *Drosophila* brain. During my thesis, I developed tools to detect genome conformation and identify enhancer-promoter interactions specific to different types of neurons in the adult *Drosophila* brain. I focused particularly on studying the chromatin conformation of genomic regions that include genes involved in memory formation.

By examining enhancer-promoter interactions and the chromatin folding patterns in Kenyon cells, my work complements the results of *Denaud et al.* (2024) by providing insights into the broader regulatory networks involved in a differentiated tissue. The exploration of differential enhancer-promoter interactions and the study of chromatin conformation that we conducted should contribute to a better understanding of how chromatin conformation, such as the PRE loop located at the *dac* locus, can be involved in the expression of certain genes in complex tissues.

In conclusion, the research article by *Denaud et al.* (2024) and my primary project collectively contribute to a better understanding of the regulatory mechanisms mediated by chromatin structures in gene expression control. With this work, we hope that our studies will provide valuable information to the scientific community, although further investigations into the molecular processes are needed to explain the role of chromatin conformation in complex mechanisms such as gene expression or memory formation in *Drosophila*.

Résumé long en français

Le génome des cellules eucaryotes présente une organisation tridimensionnelle hautement hiérarchique. Parmi ces niveaux de compartimentation, les domaines d'association topologique (TADs) sont des régions riches en interactions chromatiniennes et sont impliqués dans la régulation des gènes. Ces domaines doivent rester intacts pour maintenir la proximité physique entre les éléments régulateurs des gènes tels que les enhancers et les promoteurs, ce qui est essentiel pour une expression génique spécifique. La perturbation de la communication entre enhancers et promoteurs peut entraîner une expression génique aberrante, contribuant à des maladies, y compris des cancers ou des troubles génétiques. Ainsi, l'interaction enhancer-promoteur est cruciale pour le maintien des fonctions cellulaires normales et pour répondre aux signaux développementaux et environnementaux.

Au cours de la dernière décennie, de nombreuses études ont cherché à comprendre le rôle des interactions enhancer-promoteur (E-P) dans la régulation de la transcription dans des cellules ou des tissus prédifférenciés et différenciés [Ghavi-Helm et al., 2014][Bonev et al., 2017][Freire-Pritchett et al., 2017][Espinola, Götz et al., 2021][Ing-Simmons et al., 2021][Pollex et al., 2024]. Chez la *Drosophile*, deux modèles ont été proposés sur la base de ces études. Le premier modèle suggère un repliement du génome "conservé" à travers les différents types cellulaires que les gènes soient actifs ou non [Ghavi-Helm et al., 2014][Espinola, Götz et al., 2021][Ing-Simmons et al., 2021]. Le deuxième modèle, connu sous le nom de modèle de conformation "acquise", postule que la conformation de la chromatine évolue avec la différenciation et l'expression des gènes [Pollex et al., 2024].

L'objectif initial de cette étude est de comprendre les mécanismes complexes sous-jacents aux interactions enhancer-promoteur, à la conformation de la chromatine et aux profils d'expression génique dans différents types cellulaires. Pour aborder ces questions, les deux études présentées dans le chapitre des résultats ont étudié deux échantillons biologiques différents : la patte de la pupe et les cerveaux adultes de la *Drosophile*. Sur la base de ces données, nous avons cherché à déterminer quel modèle de conformation décrit le mieux le repliement du génome dans ces différents tissus (Figure 1). Les questions à la base de ce manuscrit de thèse sont les suivantes :

- La communication entre enhancer et promoteur est-elle instructive pour l'activation transcriptionnelle ?
- Les interactions enhancer-promoteur et la conformation de la chromatine varient-elles dans différents types cellulaires d'un tissu différencié ?
- Quels autres mécanismes pourraient être impliqués dans l'activation de la transcription de gènes spécifiques à un type cellulaire ?

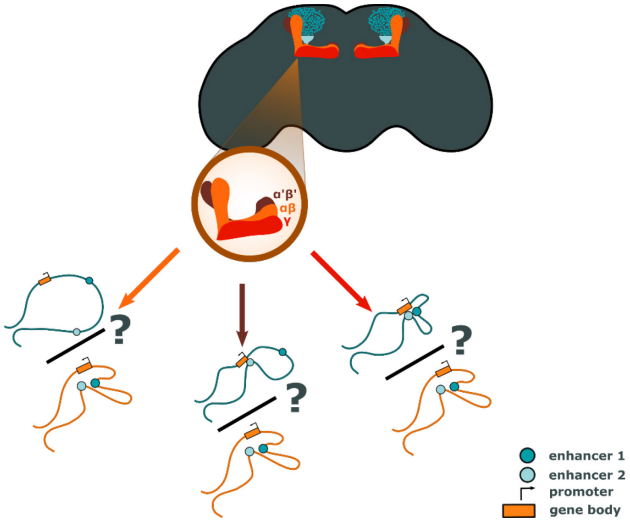


Figure 1 : Modèles de conformation de la chromatine dans les corps pédonculés adultes de la *Drosophile* - Les corps pédonculés sont composés de cellules de Kenyon, qui sont divisées en trois sous-types : γ (rouge), α'β' (marron) et αβ (orange). Représentation schématique des modèles de conformation de la chromatine : "acquis" (bleu) ou "conservé" (orange clair).

Afin de pouvoir répondre à ces questions, mon projet de thèse se concentre sur le développement et l'application de Hi-M (une technique d'imagerie reconstruisant la conformation de la chromatine à l'échelle de la cellule unique) afin d'étudier la régulation de l'expression génique par les interactions enhancer-promoteur (Figure 2). Le chapitre des résultats de ce manuscrit présente deux études qui examinent le rôle de la structure de la chromatine dans la régulation de ces interactions et de l'expression génique dans différents tissus de la *Drosophile*.

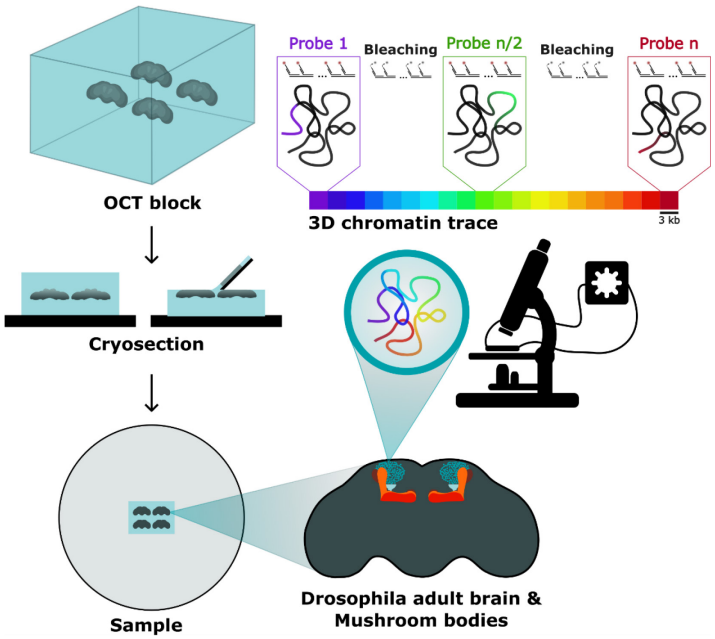


Figure 2 : Représentation schématique d'une acquisition Hi-M et reconstruction de la conformation de la chromatine dans le cerveau adulte de la *Drosophile* - Schéma illustrant la stratégie basée sur l'imagerie pour étudier la conformation des chromosomes au niveau de la cellule unique dans les tissus cérébraux cryo-coupés de la *Drosophile* (Hi-M). Chaque couleur correspond à une sous-région de 3 kb. Toutes les couleurs ensemble forment la région génomique d'intérêt. Chaque sous-région, également appelée sonde, est imagée puis photo-blanchie pour permettre l'acquisition de l'ensemble du locus par ADN FISH séquentiel.

Dans l'étude "A PRE loop at the *dac* locus acts as a topological chromatin structure that restricts and specifies enhancer-promoter communication" de *Denaud et al. (2024)*, nous explorons comment une boucle entre deux éléments de réponse Polycomb (PRE), localisée au niveau du gène *dac*, influence les interactions enhancer-promoteur, la conformation globale du locus ainsi que la transcription. Dans cette étude, nous avons examiné la conformation de la chromatine des différents segments tarsiens du disque des pattes de la pupe de la *Drosophile*. Dans les conditions de type sauvage, seul le premier segment exprime le gène *dac*, tandis que les cellules des autres segments ne l'expriment pas. Malgré l'utilisation de Hi-M, nous n'avons pas pu observer d'augmentation significative de la proximité entre l'enhancer spécifique au type cellulaire et le promoteur du gène d'intérêt dans le premier segment tarsien. Comme dans les études précédentes réalisées lors de l'embryogenèse précoce de la *Drosophile* [Espinola, Götz et al., 2021][Ing-Simmons et al., 2021], nos observations suggèrent que la modulation de la proximité physique entre un enhancer et un promoteur n'est pas nécessairement un déterminant majeur de l'activation des gènes puisqu'aucune interaction forte n'a été observée entre le promoteur de *dac* et le "ring enhancer" (RE) à travers les différents segments (Figure 3). Ces résultats sont en accord avec le modèle de "conservation" et du modèle d'"action à distance" pour la régulation du gène cible par son enhancer [Yang et al., 2024].

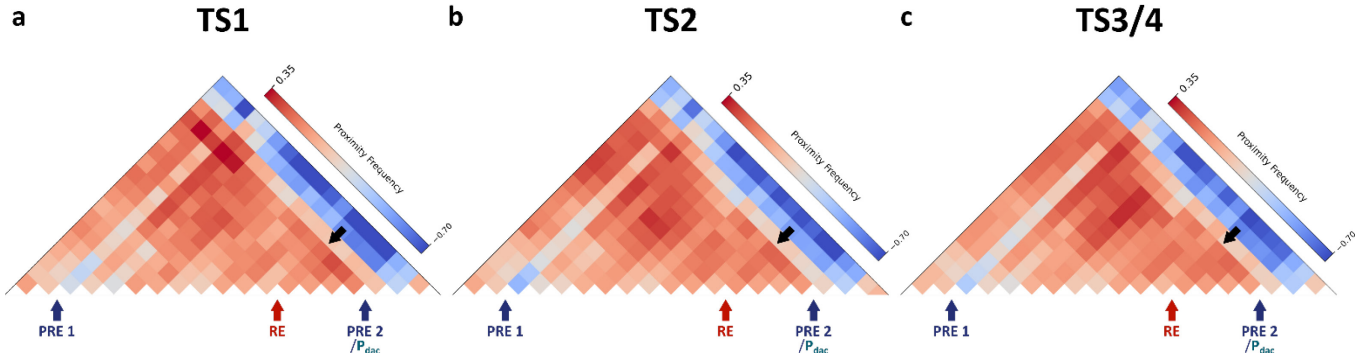


Figure 3 : Conformation de la chromatine du locus *dac* dans les différents segments tarsaux de la patte de la pupe - Matrices de proximité normalisées par la distance génomique des segments tarsaux de la patte de la pupe. (a) Matrice du premier segment tarsal (TS1), (b) matrice du second segment tarsal (TS2) et (c) matrice moyenne des segments tarsaux trois et quatre (TS3/4). Les flèches bleues indiquent la position des PREs ainsi que la position du promoteur de *dac*. La flèche rouge indique la position du ring enhancer RE. La flèche noire indique la zone de la matrice où l'interaction entre le promoteur et l'enhancer devrait être visible.

D'autre part, mon projet de recherche principal explore le lien entre conformation du génome et expression génique dans le cerveau adulte de la *Drosophile*, en particulier dans les corps pédonculés et les cellules de Kenyon. Ce type neuronal est divisé en trois sous-types cellulaires : les cellules de Kenyon γ, αβ and α'β' (Figure 1). Chacun de ces sous-types a une fonction et un programme transcriptionnel qui lui est propre, ce qui fait des cellules de Kenyon un excellent modèle pour étudier les implications physiologiques de la conformation du génome. Conformément aux résultats récemment publiés par *Pollex et al.* (2024), nous avons observé dans le cerveau adulte de la *Drosophile* que de nombreuses proximités entre enhancer-promoteur sont spécifiques à chaque type cellulaire (Figure 4a). Ces résultats soutiennent le modèle instructif des interactions enhancer-promoteur [De Laat et al., 2013]. De plus, il est intéressant de noter que des interactions enhancer-promoteur ont été identifiées comme "communes" dans tous les sous-types de cellules de Kenyon au sein des deux loci, ce qui pourrait indiquer que des topologies instructives entre E-P pourraient se former sur des paysages chromatiniens préexistants [Pollex et al., 2024].

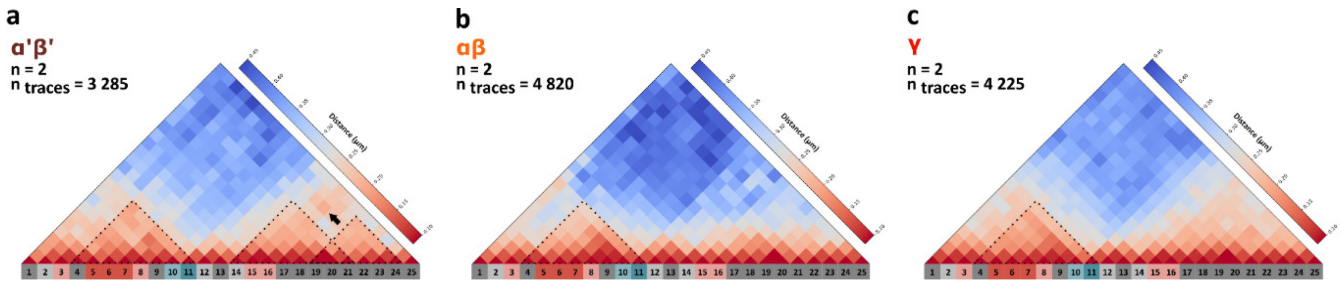


Figure 4 : Conformation de la chromatine du locus *rut* dans les différents sous-types de cellules de Kenyon - Matrice médiane de distance du locus *rut* dans les différents sous-types de cellules de Kenyon (a) cellules de Kenyon α'β', (b) cellules de Kenyon αβ et (c) cellules de Kenyon γ. Les lignes noires en pointillé mettent en évidence les différentes sous-structures de la région. La flèche noire indique une interaction spécifique aux cellules de Kenyon α'β' entre un enhancer prédit et un promoteur. Les sondes de la librairie Hi-M sont colorées de la manière suivante : promoteur (sarcelle), enhancers testés (rouge), enhancers prédits par [Janssens et al., 2022] (rose), régions accessibles (gris clair) et régions inaccessibles (gris foncé).

Les deux études soulignent l'importance des structures de la chromatine dans la modulation de l'expression génique et des interactions régulatrices au sein de loci génomiques spécifiques. En examinant les interactions enhancer-promoteur et les motifs de repliement de la chromatine dans les cellules de Kenyon, les observations faites dans mon projet principal complètent les résultats du papier *Denaud et al. (2024)* en fournissant des informations sur les réseaux régulateurs plus largement impliqués dans un tissu différencié. L'exploration des interactions différentielles entre enhancer et promoteur ainsi que l'étude de la conformation de la chromatine, que nous avons réalisées, devraient participer à une meilleure compréhension de la manière dont la conformation de la chromatine, telle que la boucle PRE situé dans le locus du gène *dac*, peuvent être impliquées dans l'expression de certains gènes dans des tissus complexes. Avec ce travail, nous espérons que nos études fourniront des informations précieuses à la communauté, bien que des investigations supplémentaires sur les processus moléculaires soient nécessaires pour expliquer le rôle de la conformation de la chromatine dans des mécanismes complexes tels que l'expression génique ou la formation de la mémoire chez la *Drosophile*.

Table of contents

List of Figures

Chapter 1 : Introduction

1.1. Chromatin

For DNA to fit (around two meters in human cells) into a nucleus of a few microns, DNA is compacted and folded into a highly organised structure known as chromatin. Chromatin consists of B-form DNA wrapped around an octamer of histones. Histones (H2A, H2B, H3, and H4) are positively charged, allowing them to interact effectively with the negatively charged DNA polymer [Zlatanova et al., 1998]. These histones contain two distinct domains : a structured domain and an unstructured N- and C-terminal domain called "tails". The 145-147 base pairs of DNA wound around each histone octamer (consisting of two copies of each histone) form a structure called the nucleosome [Luger et al., 1997][Richmond et al., 2003].

Chromatin represents the first level of genome folding. Within the nucleus, the various levels of chromatin organisation play a crucial role in the regulation of transcription, although some mechanisms remain poorly understood [Oudelaar and Higgs, 2021]. In this chapter, I will first introduce transcription regulation and its associated regulatory elements, followed by a detailed overview of the different layers of genome organisation. Subsequently, I will present optical microscopy techniques employed to study chromatin organisation. Finally, the introduction will conclude with the selected model system we chose to explore the role of chromatin folding in gene regulation within differentiated tissues.

1.2. Transcription regulation and regulatory elements

1.2.1. Overview of transcription regulation

Despite carrying the same genotype, cells are highly diverse with distinct phenotypes and function to form a complex living organism [Raj et al., 2008][Choi et al., 2019]. This complexity is partly dependent on cellular differentiation during development and throughout the individual's life [Arendt et al., 2016]. In all Metazoan, the early steps of development are crucial for differentiation and for cell types to start setting their own specific transcriptional programs. During development, maternally deposited mRNAs and proteins can generate localised patterns for gene activity highly controlled in space and time, such as the striped gene expression pattern in *Drosophila* [Pankratz et al., 1990][Blair et al., 2008].

To achieve differential gene expression, complex regulatory networks are in place within cells. Transcription is the first step of gene expression. For transcription to be initiated RNA polymerase II (Pol II) first binds to a precise region upstream of the gene body called a promoter. Once this initial step is completed, RNA Pol II reads one of the accessible and unwounded DNA strands and catalyses the synthesis of the complementary pre-messenger RNA (mRNA). Transcription ends when RNA Pol II meets a stop sequence within the gene. Thus, completing the creation of the mRNA molecule [Orphanides et al., 2002].

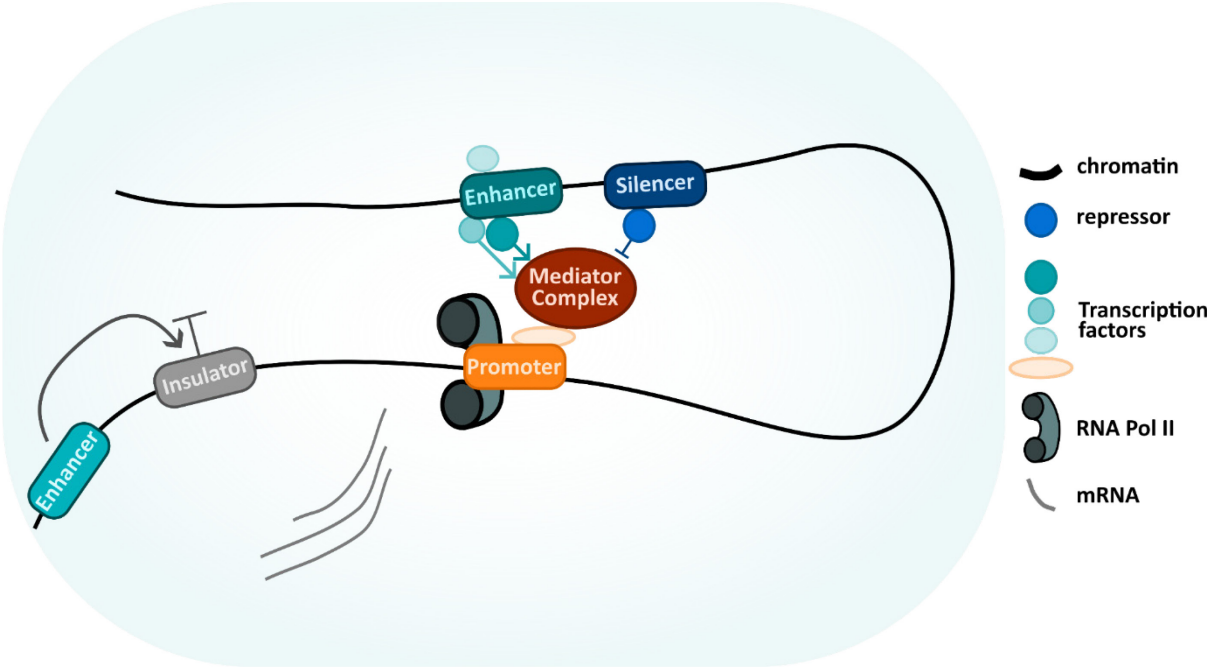


Figure 5 : Transcription regulatory network

1.2.2. Histone Post-Translational Modifications

Histone tails can undergo a variety of Post-Translational Modifications (PTMs), that will impact their interactions with DNA. These modifications can participate to the regulation of chromatin accessibility which play a crucial role in transcription regulation. Two types of PTMs will be principally mentioned throughout this manuscript, histone acetylation and methylation [Tollerney and Lunyak, 2012]. Histone acetylation is mostly associated with transcription activation while lysine methylations are involved in both activation and repression. Three main methylation sites are involved in the activation of transcription: H3 Lys4 (H3K4), H3K36, and H3K79, whereas H3K9 and H3K27 are associated with gene repression [Millan-Zambrano et al., 2022].

Chromatin in eukaryotic cells can be broadly classified into two distinct environments: euchromatin and heterochromatin. Euchromatin is typically associated to actively transcribed genes and is characterized by high levels of histone acetylation (ac) or methylations (me) such as H3K9ac, H4K16ac, H3K27ac, H3K4me1 and H3K4me2-3 [Talbert and Henikoff, 2017].

In contrast, heterochromatin is defined by the presence of compact, tightly packed chromatin and is classified into two categories: constitutive and facultative heterochromatin. Constitutive heterochromatin is characterized by the presence of H3K9me3, which is responsible for the formation of stable heterochromatin at pericentromeric regions and at telomeres. Facultative heterochromatin, on the other hand, is characterized by the presence of Polycomb repressive marks such as H3K27me3 [Talbert and Henikoff, 2017].

1.2.3. Trans-acting factors

To switch genes on and off, trans-acting factors are proteins that directly or through an intermediate protein bind to specific sequences of the genome to participate in gene transcription regulation (Figure 5). Transcription factors (TFs) are a classic example of trans-acting factors [Reuveni et al., 2018]. Once bound to specific DNA sequences called cis-regulatory sequences or cis-regulatory elements (CREs), TFs can

participate in transcriptional control by activating or repressing RNA polymerase II (Pol II) through their DNA binding domain [Latcham, 1997][Robert, 2000][Garvie and Wolberger, 2001]. Transcription factors can either be ubiquitous or cell-type-specific to modulate the activity or the strength of gene expression. Ubiquitous or general transcription factors are more often involved in expression regulation of housekeeping genes [O'Connor et al., 2016]. These genes have been defined as being essential for cellular existence regardless of their specific function in the tissue or organism and are generally expressed independently of tissue type, developmental stage, cell cycle state, or external signal [Joshi et al., 2022]. On the other hand, transcription factors can also bind in a cell-type-specific manner to cis-regulatory elements to help initiate and maintain transcriptional programs [Brivanlou and Darnell, 2002].

In addition to sequence-specific DNA-binding domains, TFs possess separate activation domains to interact with other transcription regulators such as the mediator complex (or mediator of RNA polymerase II transcription) [Allen and Taatjes, 2015]. The mediator is a multiprotein complex with a transcriptional coactivator function that can be found in all eukaryotes. The mediator is thought to regulate gene expression through the "bridge model" allowing the conversion of biological input (binding of TFs to enhancers/silencers) into physiological response (gene expression) by regulating RNAPII initiation and elongation [Borggrefe and Yue, 2011]. Thus, mediator complexes are involved in the regulation of transcription, including transcription initiation, transcription elongation, chromatin architecture and CRE-CRE looping [Allen and Taatjes, 2015].

1.2.4. Cis-regulatory elements

Cis-regulatory elements are non-coding-regions located within short and generally nucleosome-free regions of the genome consistent with their recognition by TFs in a sequence-specific manner [Noonan, McCallion 2010][Gilchrist et al., 2010]. CREs include promoters, enhancers, insulators and silencers. Promoters are regions where RNA Pol II is recruited to start gene transcription [Levine et al., 2014]. Enhancers are short DNA regions to which TFs bind to activate the transcription of the genes they regulate [Banerji et al., 1981]. Insulators, or boundary elements, are genetic elements that prevent the spreading of heterochromatin and euchromatin, or block interactions between enhancers and promoters [Udvardy et al., 1958][Kellum and Schedl., 1991]. Silencers are non-coding sequences that negatively regulate gene expression [Huang et al., 2019][Gisselbrecht et al., 2020].

In the following sections, I will provide an overview of different CREs involved in the studies presented in the results chapter.

1.2.4.1. Promoters

Generally, transcription initiates at the transcription start site (TSS) located at the 5' end of a gene. The TSS is embedded within a core promoter, which is a short sequence that can extend upstream and downstream for 30 to 40 base pairs (bp) of the TSS [Smale et al., 2003]. The core promoter is composed of binding sites for the transcription machinery, which comprises Pol II and general TFs. The first core promoter sequence discovered is called the TATA box [Lifton et al., 1978], named for its composition: TATAAA. This sequence is generally located 25 to 35 bp upstream from the gene TSS. The TATA box defines the direction of transcription and which DNA strand is transcribed and is also where specific transcription binding factors, such as TBP (TATA-binding protein), bind [Struhl et al., 1998]. Since then, many different core promoter sequences have been identified, but the most common in mammals is a GC-rich region located within CpG islands [Cooper et al., 2006]. The general size of this core promoter is around 200 nt. The full promoter (core + upstream regions) can extend up to 1,000 bp [Kamath et al., 2008].

For transcription to be initiated, polymerases must first gain access to promoter regions at the beginning of genes [Fuda et al., 2009]. If promoters are in nucleosome dense regions, nucleosomes can inhibit initiation and should be remodelled for transcription to occur [Cramer, 2019]. Thus, active promoters are found in nucleosome-depleted regions [Schones et al., 2008]. Nucleosomes downstream of an active promoter tend to present precise histone modifications such as H3K4me3 and H3K27ac [Haberle and Stark, 2019].

1.2.4.2. Enhancers

The notion of "enhancer" was introduced over forty years ago with the first description of a 72 bp repeat sequence enhancing the activity of the β-Globin gene independently of its position or orientation [Banerji et al., 1981]. Following this study, enhancers have been defined as short DNA sequences containing multiple transcription factor binding sites that can activate the transcription of a gene independently of their position, distance, or orientation relative to the promoter they regulate [Banerji et al., 1981]. Enhancers can be located within the gene body, in close proximity or hundreds of kilobases away from the promoter of their target gene [Noonan, McCallion 2010][Panigrahi et al., 2018][Balasubramanian et al., 2024]. Enhancers contain several highly tuned TFs binding sites that control the spatiotemporal patterns of gene expression, cell type-specific gene expression and responses to external stimuli [Heinz et al., 2015]. To ensure that genes are expressed at the proper developmental stage and in a cell-type-specific manner, the same promoter can be regulated by multiple enhancers [Levo et al., 2022].

Enhancer can be broadly divided into three different categories: primed, poised, active and inactive [Ernst, Kellis 2010]. Primed enhancers are characterised by open chromatin but do not participate in transcription [Heinz et al., 2015]. Poised enhancers can be defined as primed enhancers that also contain repressive epigenetic chromatin marks [Calo and Wysocka, 2013]. In their active state, enhancers are bound by general transcription factors and Pol II [Natoli and Andrau, 2012]. Inactive enhancers are located within compact chromatin and typically lack transcription factor binding and histone modifications.

Active enhancers display enrichment in histone mono or dimethylation of H3 Lys4 (H3K4me1 and H3K4me2, respectively), H3K27ac and are often in the presence of actively transcribing Pol II. Enhancers can generally be discriminated from promoters by being devoid of H3K4me3 histone mark [Heintzman et al., 2009]. Primed enhancer regions are marked with H3K4me1 and H3K4me2, while they lack histone acetylation. Enhancers marked by the repressive mark H3K27me3 are considered has poised [Creyghton et al., 2010].

To assess the activity of enhancers, multiple in vitro and in vivo methods have been developed, including FACS-seq [Gisselbrecht et al., 2013], CRE-seq [Mogno et al., 2013], MPRA [Melnikov et al., 2012], MPFD [Patwardhan et al., 2012], and STARR-seq [Arnold et al., 2013]. For example, STARR-seq (Self-Transcribing Active Regulatory Region Sequencing) enables genome-wide assessment and quantification of enhancer activity [Arnold et al., 2013]. Such techniques, combined with enhancer assays, have identified over 24,000 distinct *Drosophila* enhancers [Halfon et al., 2008].

In recent years, deep learning and transfer learning, combined with synthetic biology, have been employed to decipher the precise logic of enhancers and their role in controlling cell-type-specific gene expression in *Drosophila* [Taskiran et al., 2024][de Almeida et al., 2024]. In *Taskiran et al.* (2024), Stein Aerts' lab demonstrated that deep learning models could be successfully applied to design synthetic, cell-type-specific enhancers, allowing for the analysis of enhancer properties at single-nucleotide resolution. The authors created fully synthetic enhancers active in the neural cells of transgenic flies for over 75% of the

time. In addition, using similar methods and sequence-to-activity models, *de Almeida et al.* (2024) designed 40 synthetic enhancers for different *Drosophila* tissues, with 78% being active and 68% capable of driving tissue-specific expression.

A recent study using single-cell technologies aimed to assess the role and dynamics of enhancers in gene expression at different stages of *Drosophila* embryogenesis [Calderon et al., 2022]. This study revealed that enhancer accessibility and usage vary temporally and spatially during embryonic stages, correlating with specific gene expression patterns and tissue differentiation. For example, enhancers involved in anterior-posterior axis patterning showed differential accessibility, with a pronounced anterior skew at the even-skipped (eve) locus. The high temporal and spatial resolution of their data enabled precise mapping of enhancer activity, providing insight into the complex regulatory mechanisms controlling gene expression during early development.

This research, along with the previously cited studies, underscores the critical role of enhancers in orchestrating the complex processes of gene expression, cellular differentiation, and development.

1.2.4.3. Insulators

In 1990, a study by *Gyutkovics et al.* introduced one of the first observations of insulators in *Drosophila*. Within the Bithorax complex, they identified that the *Abd-b* (*Abdominal-B*) gene was bordered by insulators to protect it from the influence of its chromosomal environment. The authors predicted that in case of weakening or removal of these boundary elements, the *Abd-b* gene would lose its parasegment-specific expression in the adult *Drosophila* [Gyutkovics et al., 1990]. These predictions were further explored by studying the effect of the transposition of the gypsy element in between a gene and its tissue-specific distal enhancer. For example, Cai and Levine analysed the impact of gypsy insulator positioning on the interaction between the promoter and two enhancer elements of the eve gene. During Drosophila embryonic development, the spatial expression of the eve gene is controlled by these enhancers, which establish two expression stripes (stripes 2 and 3) along the anterior-posterior axis. The study demonstrated that placing the stripe 2 enhancer upstream of a 340 bp fragment of the gypsy insulator effectively blocked stripe 2 expression, while stripe 3 expression remained unaffected. Similarly, reversing the order of the stripe 2 and stripe 3 enhancers resulted in the opposite effect, with stripe 2 expression being restored and stripe 3 expression blocked [Cai and Levine, 1995]. Thus, the transposition of the gypsy element between enhancers and promoters can induce a loss of cell-type-specific expression, highlighting one of the key features of insulators that is to block interactions between regulatory elements [Özdemir and Gambetta 2019]. In transgenic reporter assays, the gypsy insulator can also behave as a chromatin border that can block spreading of H3K27me3 [Comet et al., 2006].

Indeed, since the late 80s' to early 90s', numerous studies have highlighted the role of insulators as barrier elements to prevent the spreading of heterochromatin structures and silencing [Sun and Elgin, 1999]. For example, in *Bowman et al.* (2014), chromatin was purified from single *Drosophila* embryo parasegments and their study revealed a 'stairstep' pattern, across the bithorax complex regulatory domains, between H3K27me3 and H3K27ac regions [Bowman et al., 2014]. They observed that these domains sharp edges were overlapping with *drosophila* CCCTC-binding factor (dCTCF) bound regions. These results suggested that boundaries associated with dCTCF binding might prevent spreading of euchromatin or heterochromatin into neighbouring domains. Over the years, multiple studies have shed light on the cellular mechanisms that are used to maintain the epigenetic characteristics of chromatin domains and the ability of insulators to act as barrier elements [Bushey et al., 2008].

Thus, insulators can be classified into two main categories: enhancer-blocking insulators which can block communication between CREs and barrier insulators which prevent the spread of heterochromatin [Bushey et al., 2008].

There are multiple proteins factors that bind to insulators to modulate their functions, including dCTCF, Centrosomal Protein 190kD (CP190) or Boundary-associated element of 32kD (BEAF-32) [Parkhurst et al., 1986], [Zhao et al., 1995], [Gerasimova et al., 1995], [Scott et al., 1999], [Pai et al., 2004], [Moon et al., 2005]. In *Drosophila*, CTCF facilitates long-range chromatin interactions, is found at most boundaries of regulatory domains and possesses transcriptional repressor activity [Moon et al., 2005][Stadler et al,. 2017]. CP190, along with the cofactor dCTCF, mediates enhancer-blocking and has been shown to bind to other insulator binding proteins in *Drosophila*, thereby contributing to insulator function [Ali et al., 2016]. The BEAF-32 protein was initially identified as a factor interacting with the *scs'* insulator. Genome wide analysis revealed that BEAF preferentially associates with the promoter of active genes, although the exact mechanisms by which it influences gene expression remains unknown [Bushey et al., 2009].

The role of insulators and previously cited proteins in the formation and the regulation of Topologically Associating Domains (TADs) will be discussed in section 1.3.2.

1.2.4.4. Polycomb Response Elements

Polycomb group (PcG) proteins are evolutionary conserved epigenetic factors that mediate the inheritance of silent chromatin states throughout development [Schuettengruber and Cavalli, 2009]. Initially discovered in *Drosophila*, PcG proteins were identified as key factors in maintaining the repression of Hox genes, a critical set of genes that determine cell identity along the head-tail axis [Lewis, 1978]. PcG proteins associate to repress their target genes and modify local chromatin architecture by compacting nucleosomes which render chromatin inaccessible to the transcription machinery [Francis et al., 2004].

In *Drosophila*, specific DNA sequences known as Polycomb Response Elements (PREs) serve as nucleation sites for PcG proteins, leading to the formation of Polycomb domains. Within these domains, PREs recruit two main Polycomb repressive complexes: PRC1 and PRC2. PRC1 catalyses the mono-ubiquitination of lysine 119 on histone H2A (H2AK119Ub), while PRC2 is responsible for the deposition of the H3K27me3 mark [Schuettengruber and Cavalli, 2009]. Notably, PREs can form specific chromatin loops within Polycomb domains, which are essential for locking genes in a repressed state during *Drosophila* development [Ogiyama et al., 2018].

The role of PcG and PREs in regulating chromatin architecture highlights a broader aspect of genome organisation. Beyond local chromatin compaction, CREs are intricately linked to the spatial arrangement of the genome within the nucleus. This organisation is crucial for ensuring that genes are properly regulated in a three-dimensional context, allowing for coordinated gene expression patterns across different developmental stages and cell types.

1.3. The organisation of the genome

In the nucleus of a cell, the genome has several levels of organisation that can contribute to gene expression regulation.

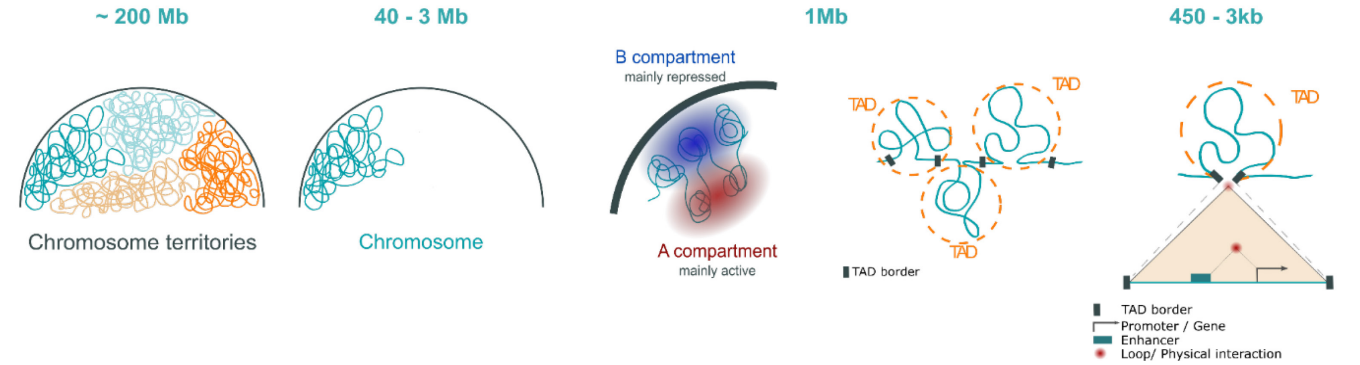


Figure 6 : Schematic representation of genome folding

1.3.1. Chromosome territories and compartments

To increase the degree of packing of DNA inside the nucleus, chromatin undergoes successive folding to create higher order structures and ultimately chromosomes [Oudelaar and Higgs, 2021]. The notion of territorial organisation was first proposed in 1885 by Carl Rabl. The term Chromosome Territories (CTs) (Figure 6) was introduced decades later by Theodor Boveri in 1909 [Cremer et al., 2010]. Advances in fluorescence microscopy and the popularisation of In Situ Fluorescence Microscopy (FISH) provided evidence of CTs' existence. Various studies showed that chromosomes are localised in distinct volumes of the nucleus and rarely intermix [Lichter et al., 1988][Pinkel et al., 1988][Cremer and Cremer 2001]. More recently, chromatin conformation capture techniques confirmed that interactions are more frequent within the same chromosome (*cis*) than between different chromosomes (*trans*) [Lieberman-Aiden et al., 2009].

Chromosomes are formed of two mutually excluded types of chromatins identified as “A” and “B” compartments (Figure 6) [Lieberman-Aiden et al., 2009][Rao et al., 2014][Schrader et al., 2016]. Within chromosomes these compartments exhibit preferential *cis* long-range interactions than trans. The A compartment is generally more accessible and transcriptionally active, while the B compartment is enriched in closed chromatin and correlates more with inactive genes. Each compartment is globally enriched in either euchromatin (A) or heterochromatin (B) and tends to be segregated in space [Lieberman-Aiden et al., 2009][Oudelaar and Higgs 2021]. The A compartment is usually located toward the centre of the nucleus, while the B compartment is often found at its periphery, interacting with the nuclear lamina [Boyle et al., 2001][Bickmore et al., 2013]. In *Drosophila*, regions enriched in inactive chromatin tend to be located close to the nuclear envelope [Pickersgill et al., 2006][van Steensel and Belmont, 2017].

As presented in the previous paragraph, chromatin is divided into two main classes : heterochromatin and euchromatin. In addition, chromatin can be classified in five types of epigenetic domains : two types of heterochromatins, repressive chromatin and two types of euchromatin [Filion et al., 2010]. The first type of heterochromatin wears the H3K27me3 mark and associates with polycomb group proteins that can remodel chromatin to induce gene silencing [Sparmann and van Lohuizen, 2006]. The second one is marked by H3K9me3 and overrepresentation of heterochromatin protein 1 (HP1) binding [Hediger and Gasser, 2006]. HP1 is predominantly associated with pericentric heterochromatin, but it is also found at many sites along chromosome arms. Depending on the chromosomal context and on the presence of other proteins, HP1 can either promote repression or activation of transcription [Hediger and Gasser, 2006]. The repressive chromatin type lacks classic heterochromatin markers, is generally gene poor and presents low to no transcription activity [Filion et al., 2010][Kharchenko et al., 2011][Schwartz et al.,2010]. The two

types of transcriptionally active domains are associated with histone marks such as H3K4me3, H3K36me3, and hyperacetylation (H3K27ac, H4K8ac, and H4K16ac) [Nègre et al., 2011][Sexton et al., 2012][Talbert et al., 2017].

1.3.2. Topologically associating domains (TADs)

In 2012, a new level of chromatin organisation called Topologically Associating Domains (TADs) was described [Dixon et al., 2012][Nora et al., 2012][Sexton et al., 2012]. TADs are sub-megabase structures presenting increased self-association that include genes with similar expression patterns and epigenetic states (Figure 6) [Dixon et al., 2012][Nora et al., 2012][Sexton et al., 2012][Hou et al., 2012][Ulianov et al., 2016][Schauer et al., 2017]. TADs' genomic locations are generally conserved throughout development and across tissue types in many different species, indicating that they may represent a conserved feature of genome organisation [Dixon et al., 2012][Nora et al., 2012][Sexton et al., 2012][Dong et al., 2017][Kaaij et al., 2018]. In addition, these structures may facilitate the communication between genes and their CREs [Symmons et al., 2014][Robson et al., 2019]. An important feature of TADs is their insulating boundaries, which play a crucial role in maintaining the self-associating property of TADs and are characterised by decompacted chromatin regions bound by insulator proteins [Stadler et al., 2017].

In vertebrates, TAD boundaries often coincide with a great number of DNA binding proteins, such as CTCF or with the presence of the structural maintenance of chromosomes (SMC) complex (cohesin and condensin), which play a major role in specifying TAD boundaries [Dixon et al., 2012][Phillips-Cremins et al., 2013][Vietri Rudan et al., 2015][Rao et al., 2017]. The loop extrusion model has been proposed as a mechanism underlying the formation of loops and TADs, in mammals. In this model, proteins such as cohesins bind to the chromatin and slide along the chromatin fibre until when two loop extruding factors meet or when they encounter CTCF binding sites (properly oriented) [Fudenberg et al., 2016].

In *Drosophila*, the homolog version of CTCF is mildly enriched at TAD boundaries [Matthews and White, 2019]. During early *Drosophila* embryonic development, dCTCF proteins do not seem essential for cell viability [Gambetta et al., 2018]. However, complete ablation of CTCF in vivo, can affect small amounts of boundaries (less than 10% of the total number of boundaries) and impact the formation of only a few specific domains [Kaushal et al., 2021]. Instead, other insulator proteins (including BEAF-32 and CP190) were shown to be more enriched at TAD boundaries than dCTCF [Sexton et al., 2012][Hou et al., 2012][Ulianov et al., 2016][Hug et al., 2017][Ramirez et al., 2018]. Interestingly, dCTCF is only found in ~8% of the domain boundaries in *Drosophila* larvae and its mutation has little effect on the overall chromatin structure[Kaushal et al., 2021].

Messina et al. (2022) have recently highlighted that chromatin regions displaying high interactions are preferentially bound by insulators and that TADs border although they do interact with each other, they represent only a small fraction of self-interacting chromatin regions detected.

In addition, *Cavalheiro et al*. (2023) have highlighted that TADs are able to form without BEAF-32, CTCF or CP190 although they present lower insulation. In addition, depletions of these insulator binding proteins (IBPs) do not drastically perturb global gene expression during Zygotic Genome Activation (ZGA) with only hundreds of *Drosophila* genes being downregulated. Misexpressed genes did not seem to be correlated with changes in genome topology. Then, the authors focused on one TAD and observed that removal of the active promoter (and thereby transcription) had a greater impact on TAD structure than removal of the insulator-bound region itself. These results suggest that in some conditions an active promoter/transcription may have more impact on domain insulation than specific IBPs during embryogenesis.

TAD folding has raised significant questions regarding their role in regulating transcription. Insulators might help encourage transcription by stabilising interactions to allow a dynamic binding of transcription factors, or by participating in transcription regulation with the mediation of 3D chromatin conformation [Bhattacharya et al., 2024].

1.3.3. Enhancer-Promoter interactions

Enhancer-promoter (E-P) communication has a critical role in transcription regulation and their interaction can occur over large genomic distances [Deng, Jim and Lim 2022]. Despite many years of research, the principles and mechanisms by which enhancers communicate with their target genes remain still largely controversial. Four main models were identified to regulate their communication. The first two could be considered as 1D genome models, as they posit that E-P communication could be established along the chromatin, without necessitating spatial proximity of enhancers and promoters in the 3D space [Furlong and Levine, 2018][Yang and Hansen, 2024]. The first 1D model, referred to as the tracking model, proposes that enhancers recruit activators or Pol II that track along chromatin towards the promoter while pulling the enhancer, thus stimulating transcription when the enhancer reaches the promoter [Moreau et al., 1981][Kong et al., 1997]. The second 1D model relies on formation of large protein chains connecting E-P pairs located several kilobases to tens of kilobases away from each other [Morcillo et al., 1997]. This 1D model is identified as the "linking model" or the "spreading model". Although these models can explain some properties of enhancer-promoter communication, some researchers believe that the tracking model would be over reliant on motor proteins and that it is unlikely that large chains of proteins would be employed genome-wide as described in the linking model [Panigrahi and O'Malley, 2021].

The "looping model" is currently widely considered as the main model to explain E-P interactions in the 3D space. According to this model, enhancers are thought to physically interact with the promoter region of the target genes to initiate transcription. This model is based on a large body of evidence showing that distal enhancers and promoters form chromatin loops to induce transcription [Carter et al., 2002][Nolis et al., 2009][Chen et al., 2018][Panigrahi et al., 2018]. In *Deng et al.* (2012), authors forced the formation of a chromatin loop between an enhancer and a promoter and induced strong transcriptional activation. In this study, the mouse β-globin (Hbb) promoter and its enhancer were forced to loop, causing the activation of the Hbb gene, highlighting that E-P contacts can lead to transcription activation. If enhancers are linked to target promoters through structural bridging of transcription factors and/or mediator complex, E-P distances should be within ~10–30 nm according to protein diameters [Nollmann and Gregor, 2020][Kawasaki and Fukaya, 2024].

Previous studies from the Furlong lab [Ghavi-Helm et al., 2014], and more recent work from the Vaquerizas' lab and my host lab have shown that E–P interactions are similar in different cell types during early *Drosophila* embryogenesis [Ing-Simmons et al., 2021][Espinola, Götz et al., 2021]. Moreover, E-P loops are independent of transcriptional status [Espinola, Götz et al., 2021]. Other concurrent work showed that enhancers and promoters are not necessarily within short distances even in cells where the promoter can be activated by the enhancer [Alexander et al., 2019][Benabdallah et al., 2019][Gomez Acuna et al., 2024]. One striking example of this contact-independent activation was reported for the mouse sonic hedgehog (Shh) gene [Benabdallah et al., 2019]. Surprisingly, 3D DNA-FISH revealed that the distance between the Shh gene and its distal enhancers increased from ~200–300 nm to ~400 nm upon transcription activation in the developing mouse brain.

This E-P decreased proximity could also be consistent with newly proposed 'kiss-and-kick' or 'hit-and-run' models. These models rely on the relatively dynamic interactions of E-Ps observed during live

imaging of chromatin where E–P pairs are transiently in contact, and then diffuse away from each other or are separated due to transcriptional bursting. These models describe E-P communication as highly stochastic and dynamic in individual living cells. In these non-linear models, enhancers may need to contact promoters many times to activate transcription [Xiao, Hafner and Boettiger, 2021][Mach and Giorgetti, 2023].

The lack of correlation between E-P distance and transcription suggest that the textbook looping model cannot describe every type of E-P interactions involved in transcription regulation and gave rise to a fourth model for E-P communication. In this model, so-called "action-at-a-distance", enhancers manage to activate promoters without physically contacting promoters [Yang and Hansen, 2024]. This was hypothesised to occur by two possible mechanisms. The first mechanism would be driven by transcription condensates surrounding E-Ps to mediate communication. These condensates would form through liquid-liquid phase separated TFs, co-activators and RNA Pol II [Hnisz et al., 2017][Cho et al., 2018][Sabari et al., 2018][Wang, Cairns and Yan, 2019]. The second mechanism could involve the formation of transcription factor activity gradients (TAG) [Karr et al., 2022]. In this mechanism, enhancers bound by co-activators act as a source of a concentration gradient of transcription factors, encouraging interactions between TFs and proximal promoters [Yang and Hansen, 2024].

From a development and differentiation point of view, E-P communication can take different forms. During development, cell types or pre-differentiated tissues can show different gene expression patterns across cell-types but similar chromatin organisations [Ghavi-Helm et al., 2014][Ing-Simmons et al., 2021][Espinola, Götz et al., 2021][Rubin et al., 2017]. In this scenario, specific interactions between CREs can be established before target genes start to be activated, with chromatin conformation acting as a scaffold mediating gene expression for when enhancers get activated. This fits with a "permissive" model, where E–P interactions are temporally and/or spatially separated from transcription activation [De Laat et al., 2013].

More recently, the question of whether this observation holds during later stages of development and cell differentiation was addressed in *Drosophila* [Pollex et al., 2024]. The article of *Pollex et al.* (2024) shows that during terminal tissue differentiation, E-P interactions switch from a "permissive" to a more "instructive" mode, where new E–P loops start to emerge to possibly enable plasticity and diversification during tissue differentiation. Although this study is highly informative and offers a new perspective on how chromatin organisation evolves in different cell types of *Drosophila*, the question of whether and how chromatin conformation contributes to cell-type-specific transcription remains unanswered and difficult to address. Thus, the simultaneous detection of cell identity, genome folding and transcription is critical to deepen our understanding of how interactions between promoters and enhancers regulate gene expression in differentiated tissues. One way to extract such information is to use microscopy-based chromosome conformation capture methods.

1.4. Microscopy based chromosome conformation capture methods

1.4.1. DNA-FISH

Over the last decades, Fluorescent In Situ Hybridization (FISH) has been widely employed to image and probe chromatin at different scales. This technique allows detection and localisation of specific genomic regions by hybridizing single strands of DNA (DNA probes) that are complementary to the sequence of interest. In the late 60s, FISH was developed on *Xenopus* cells with radioactive molecules bound to DNA probes [Pardue and Gall, 1969]. Years later, radioactive labels were replaced by fluorescently

labelled DNA probes that were, obviously, less hazardous, more stable and easier to use [Rudkin and Stollar, 1977][Bauman et al., 1980]. Historically, probes were synthesised from cloned sequences of bacterial artificial chromosome libraries or from polymerase chain reaction (PCR) of the targeted regions and followed by labelling the DNA probes with fluorescent dyes [Wiegant et al., 1991]. In recent years, new tools paved the way for more flexible, rapid and efficient FISH labelling approaches such as Oligopaint. Oligopaint is a technique that uses short, custom-designed oligonucleotides (oligos) as probes for specific DNA sequences and are much shorter (20-60 nucleotides) than traditional DNA-FISH probes. These oligos are synthesized to hybridize precisely to the target region. Since its implementation, complex libraries of single-stranded DNA oligos were applied to label genomic regions of interest from few kilobases to full chromosomes [Beliveau et al., 2012][Beliveau et al., 2015][Beliveau et al., 2018]. Now a standard protocol for conventional DNA-FISH relies on four main steps: gentle fixation using paraformaldehyde (PFA), permeabilization, denaturation of the genomic DNA (around 80°C), and hybridization of the fluorescent probes to the genomic DNA. These steps preserve the nuclear structure, at least up to a spatial resolution of ~120 nm [Solovei et al., 2002][Markaki et al., 2012]. Since then, FISH has been undoubtedly the most popular imaging-based technique to investigate genome organisation of two or three loci of interest. This small number of loci probed is constrained by the availability of different fluorescent dyes spectrally separated, thus limiting the number of loci that can be detected simultaneously. In addition, resolving genomic regions that are under 250 nm in the 3D space has been extremely hard as this technique is usually combined with widefield microscopy, due to intrinsic spatial resolution limits. To overcome these limitations, super-resolution microscopy principles were applied to DNA-FISH.

1.4.2. Diffraction limit, super-resolution microscopy and principle of localisation

Due to diffraction, the image of a single fluorescent molecule, or source point, through an optical microscope is known as the Point Spread Function (PSF), which can be theoretically described by an Airy pattern. This spot's size is dependent on the wavelength of the source and on the numerical aperture of the system. For a circular aperture, the size of the spot is called the Airy radius:

$$r_{Airy} = 0.61 \times \frac{\lambda_{fluo}}{NA}$$

where r_{Airy} is the radius of the airy pattern, λ_{fluo} is the fluorescence wavelength and NA is the numerical aperture of the system.

Thus, spatial resolution in fluorescence microscopy is usually generally limited by the numerical aperture of the objective lens of the microscope. Two fluorescent molecules, imaged with a microscope, that are closer to each other than the Airy radius, cannot be distinguished in an image. This last criterion empirically defines the resolution limit in optical microscopy (Rayleigh criterion). In conventional widefield microscopy, the lateral resolution (directly correlated to the Airy radius) is around 250 nm, while the axial resolution is around 500 nm.

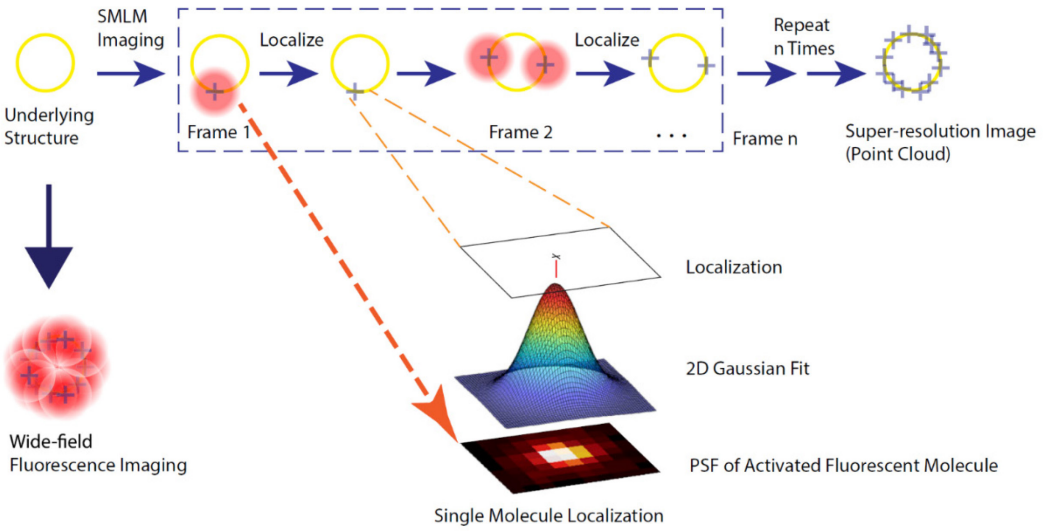


Figure 7 : Principle of Single Molecule Localisation Microscopy - from [Khater et al., 2020]

Multiple super-resolution methods were developed to overcome the resolution limit [Khater et al., 2020]. Amongst the many different methods, Single Molecule Localization Microscopy (SMLM) provides the highest spatial resolution (Figure 7). From a simplified point of view, the principle of SMLM is to excite, image and localise one fluorescent source at a time instead of imaging multiple spatially close sources simultaneously. Different strategies were applied to constrain the number of single emitters that become excitable at any given time [Lelek et al., 2021]. Once all the fluorescent molecules have been sequentially excited and imaged by the camera, the images can be registered (to compensate for drift that occurred during the acquisition) and the centre of each PSF can be precisely extracted through image post-processing. Several methods are available to extract the centre of a PSF, such as detection of the centre of mass by fitting the PSF using a 2D Gaussian model. Thus, the ability to distinguish two Airy patterns no longer relies on the optical system's resolution but on the uncertainty in the estimated position of the PSFs which can reach 10 nm and depends on the brightness of the fluorophore, the microscope stability, and the camera pixel size. This principle of localisation developed for SMLM was also applied to DNA-FISH to increase localisation precision of fluorescent probes.

1.4.3. Sequential DNA-FISH

Since 2016, a set of new microscopy-based conformation capture methods have been developed to reconstruct 3D chromatin folding at kilobases scale using sequential DNA-FISH [Wang et al., 2016][Bintu et al., 2018][Cardozo Gizzi et al., 2019][Mateo et al., 2019]. While conventional DNA-FISH relies on imagining simultaneously different probes labelled with different colours, with sequential DNA-FISH tens to hundreds of genomic loci can be localised using sequential labelling and imaging. These chromatin tracing methods depend on the 3D centroid localization of probes, similar to SMLM. Instead of relying on blinking [Rust et al., 2006][Heilemann et al., 2008] or photoactivatable [Betzig et al., 2006][Hess et al., 2006] dyes, sequential DNA-FISH images one genomic locus of interest at a time. Using image post-processing, the 3D chromatin tracing of each single cell in the field of view can be reconstructed.

Similarly to sequencing-based chromatin conformation capture methods, sequential DNA-FISH can be used to reconstruct matrices of distances or contact frequencies of genomic loci. One of the strengths of these methods is that it can simultaneously detect transcription activity by RNA-FISH and genome folding by DNA-FISH within the same cell [Mateo et al., 2019][Cardozo Gizzi et al., 2019]. The combination

of both FISH methods has allowed the assignment of given chromatin domain structures to specific transcription activities in *Drosophila* embryos [Cardozzo et al., 2019][Mateo et al., 2019]. These studies revealed that genome folding can vary according to the epigenome and the transcriptional activity.

1.4.4. Hi-M

In this thesis, I used a sequential DNA-FISH method known as Hi-M (microscopy-based chromosome conformation capture), which was developed in my host lab alongside the aforementioned methods. Hi-M, similarly to other sequential DNA-FISH methods, employs a microfluidics device coupled with a widefield microscope to deliver and incubate various solutions onto previously stained samples, enabling the acquisition of 3D images of fluorescent DNA probes. To reconstruct the chromatin folding of the genomic region of interest, the 3D positions of tens of probes across thousands of single cells are recorded and subsequently processed using a pipeline called pyHi-M, a software developed by my host lab [Devos, Fiche et al., 2024]. This pipeline can reconstruct thousands of chromatin traces, and median pairwise distance (PWD) matrices can be generated to visualise the results (Figure 8 - for a detailed explanation of the Hi-M steps, see Chapter 4 : Materials and Methods).

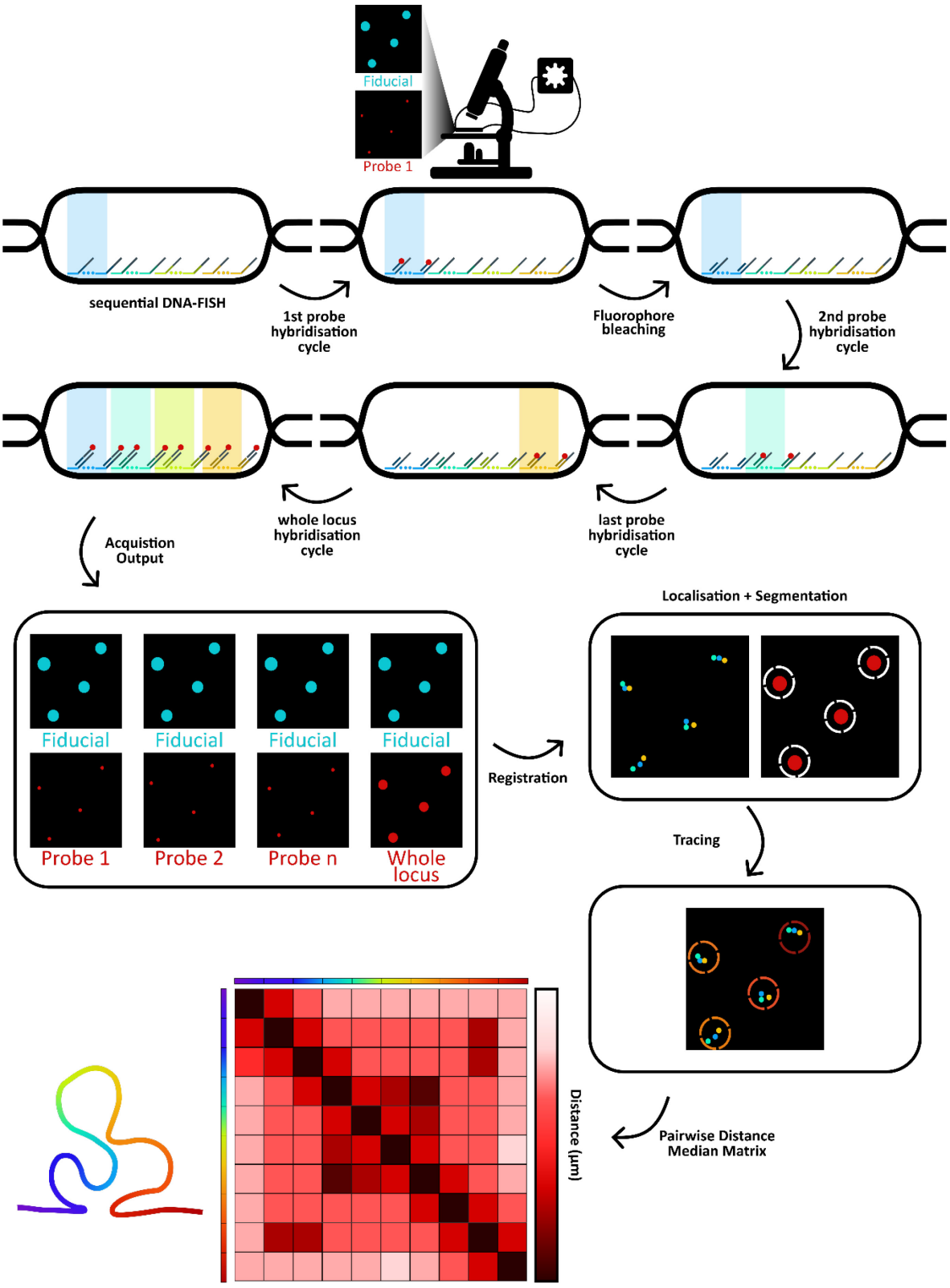


Figure 8 : Schematic representation of a Hi-M experiment and the pyHi-M post processing - Overview of the strategy used for a sequential DNA-FISH experiment. First, oligopaint probes are hybridised to genomic DNA. Then, using a microfluidics device coupled with a widefield microscope, locus-specific readout sequences, fiducial and the whole locus are imaged sequentially. The acquisition, of all loci of the region of interest, relies on sequential imaging and bleaching cycles. Once the Hi-M acquisition is complete images are registered and segmented with the analysis pipeline, pyHi-M. Probes can then be localised and affiliated to a chromatin trace. Based on these traces, pairwise distance (PWD) matrices are reconstructed.

Over the last few years, Hi-M was applied in different studies to understand how chromatin structure evolves depending on the cell type, transcription status or interactions with specific proteins. In *Espinola, Götz et al.* (2021), Hi-M was applied to *Drosophila* embryos using a 3kb resolution to explore chromatin architecture and gene regulation. This study revealed similar TAD organisation between different cell types (mesoderm, neuroectoderm, dorsal ectoderm) at the *snail* and *dorsocross* loci. In the the *dorsocross* locus (that includes *doc1*, *doc2* and *doc3* genes), they observed that loops between CREs were independent of gene activation and cell fate during early *Drosophil*a development. These results were made possible by using Hi-M and its ability to trace chromatin and detect RNAs in single cells. In *Götz et al.* (2022), the authors used single-trace analysis to show that chromatin organisation of expressing and non-expressing cells were similar in single cells within nuclear-cycle 14 (nc14) *Drosophila* embryos. In *Messina et al.* (2023), the authors investigated the role of IBPs in 3D chromatin organisation in *Drosophila*, particularly focusing on how these proteins influence chromatin interactions and transcriptional regulation during development. By combining Hi-M and bioinformatics analysis, they were able to highlight that while IBPs play a role in chromatin organisation, their influence is to fine-tune interactions rather than establishing stable structural domains. In *Gurgo et al.* (2024), the authors studied the role of Polycomb proteins in chromatin organisation in the 3D space of the nucleus at early (nc14) and late stages (S15–S16) of *Drosophila* embryo development. Their results reveal that three-dimensional proximity between Polycomb (Pc) targets is rare in *Drosophila* embryos. Their finding suggests that coalescence of multiple distant Pc domains to form large repressive compartments is uncommon. Their results suggest that long-range interactions between Pc genes are infrequent but are temporally and spatially regulated by their transcription status.

1.5. *Drosophila* adult brain, behaviour and the mushroom bodies

1.5.1. *Drosophila* adult brain and behaviour

Model organisms, such as *Drosophila*, have been especially useful for the genetic dissection of developmental and anatomical traits [Nusslein-Volhard et al., 1987][Lasko, 2020]. The fact that many genes found in flies have structural or functional homologues in vertebrates, including humans, implies that genetic discoveries in the fruit fly can contribute to our general understanding of evolutionarily conserved developmental and physiological processes [Tolwinski, 2017][Ogienko et al., 2022]. While mechanisms that underlie mammalian behaviour are more complex than those in the fly, the basic components of such mechanisms are often conserved [Takahashi et al., 2009]. The fruit fly can exhibit a wide range of complex behaviours, including foraging [Amrein et al., 2005], courtship [Siegel et al., 1979], and learning [Quinn et al., 1974][Tempel et al., 1983][Sokolowski, 2001][Mollá-Albaladejo and Sánchez-Alcañiz, 2021][Davis, 2023]. Thus, by studying molecular, cellular and evolutionary bases of behaviours in *Drosophila*, *Drosophila* behavioural geneticists have been able to connect, compare and identify some of the basic characteristics of mammalian behaviour [Dubnau and Tully 1998][Adel and Griffith, 2021][Davis, 2023].

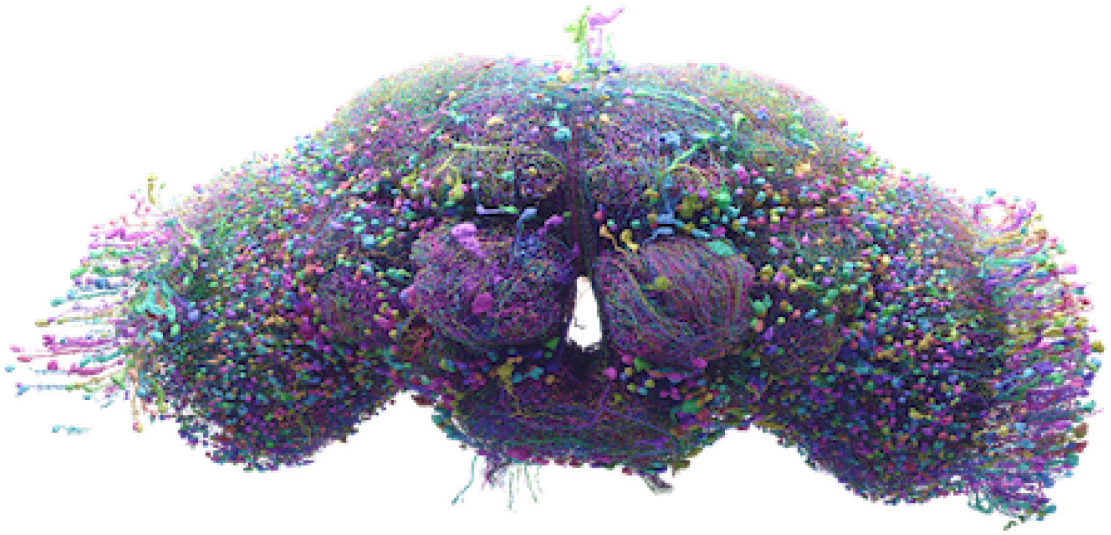


Figure 9 : Whole *Drosophila* adult brain connectome - from FlyWire (flywire.ai)

During recent years, the *Drosophila* brain has also been extensively studied. Now, a vast literature grants access to data, including connectome (Figure 9), transcriptome or chromatin accessibility atlases for different stages of brain and Central Nervous System (CNS) development including the adult brain [Costa et al., 2016][Crocker et al., 2016][Croset et al., 2017][Li et al., 2017][Konstantinides et al., 2018][Zheng et al., 2018][Davie et al., 2018][Dorkenwald et al., 2022][Janssens et al., 2022][Scheffer et al., 2022][Winding et al., 2023]. These studies have revealed an immense diversity of neuronal and glial cell types that underlie an array of functional and behavioural traits in the fly. In the following section, the behavioural trait of *Drosophila* that will be described is learning and memory formation driven by a specific structure of the brain called the mushroom bodies.

1.5.2. Olfactory memory and the mushroom bodies

The fruit fly possesses various forms of memory that support different types of learning such as visual learning, taste learning, motor learning, spatial orientation learning, courtship learning, olfactory associative learning and others [Margulies et al., 2005][Davis, 2023]. Notably, there are multiple types of olfactory associative memory, which vary in terms of duration, neuronal type and molecular mechanisms.

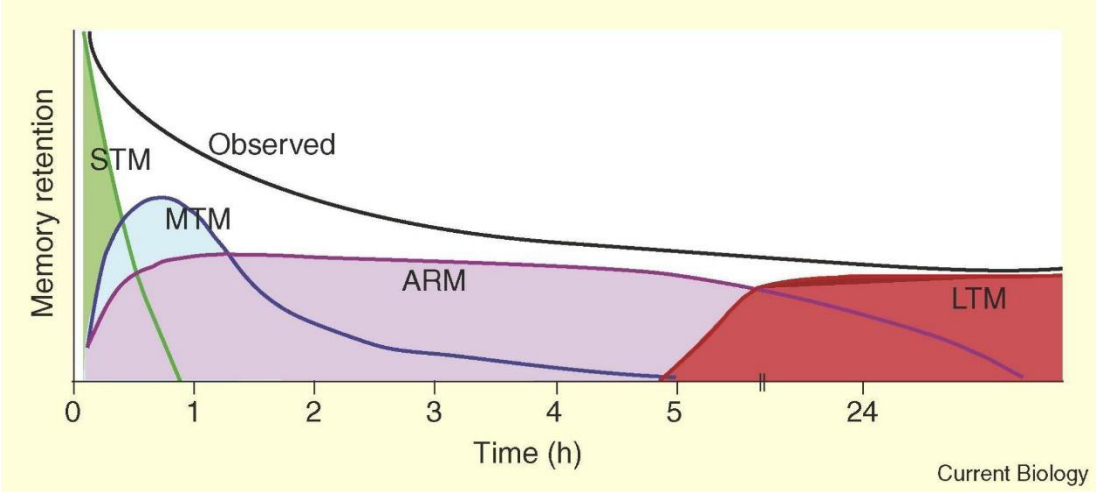


Figure 10 : Different phases of olfactory memory formation in the adult brain - from [Margulies et al., 2005]

In *Drosophila*, associative olfactory memory can be categorised into four types: short-term memory (STM), medium-term memory (MTM), anaesthesia-resistant memory (ARM), and long-term memory (LTM) (Figure 10). Various paradigms can be employed to study each type of memory. For instance, the appetitive memory paradigm associates a scent with a reward (e.g., sugar), while the aversive memory paradigm links it to a punishment (e.g., electric shock) [Quinn et al., 1974][Tempel et al., 1983].

Various studies have demonstrated the role of a brain structure called the mushroom bodies in olfactory memory and learning [Heisenberg et al., 1985][De Belle et al., 1994][McBride et al., 1999]. When functionally comparing the mushroom bodies to structures in the mammalian brain, they are most similar to the hippocampus and/or the cerebellum [Strausfeld et al., 1998].

In the *Drosophila* adult brain, mushroom bodies are symmetrical structures located in the central brain. The neurons constituting each mushroom body originate from the division of eight neuroblasts [Fahrbach et al., 2006]. Throughout development, these precursor cells give rise to 5,000 intrinsic neurons called Kenyon cells (KCs), representing 3% of the neurons in the *Drosophila* adult brain. Their cell bodies are clustered in the dorsal posterior part of the brain surface [Ito et al., 1997]. Their dendrites are grouped in a region just below the cell bodies, called the calyx. The axons project from the calyx into the anterior portion of the brain via a dense structure called the peduncle, giving rise to five distinct lobes: two vertical lobes (α and α') and three horizontal lobes (γ, β, and β') (Figure 11). Kenyon cells are divided into three neuronal types based on the branching patterns of their axonal projections: α/β neurons, α'/β' neurons, and γ neurons [Crittenden et al., 1998].

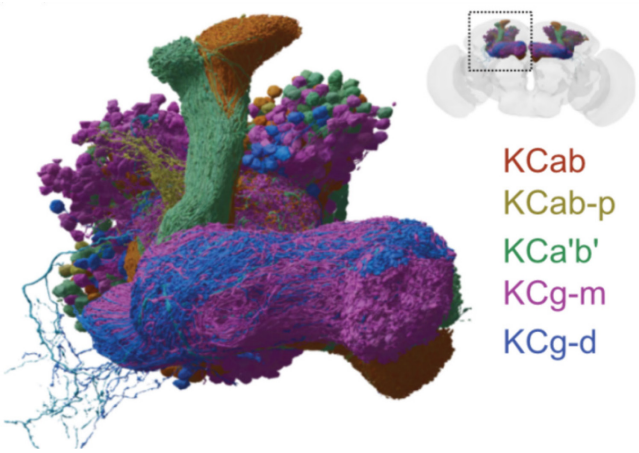


Figure 11 : 3D rendering of Mushroom Bodies and Kenyon cells - from [Schlegel et al., 2023]

With the improved Split-GAL4 technique developed by Gerald Rubin's lab [Pfeiffer et al., 2010], the anatomy of the mushroom bodies has been extensively described [Aso et al., 2014]. After generating 7,000 different fly lines, the Kenyon cells have been categorised into seven subtypes based on their morphology. The γ lobe is divided into two layers: main (γm - 600 neurons) and dorsal (γd - 70 neurons). The α'β' KCs have been classified into middle (α'β'm - 170 neurons) and anterior-posterior (α'β'ap - 200 neurons) lobes. Finally, the αβ lobe is composed of αβ posterior (αβp - 90 neurons), αβ core (αβc - 400 neurons), and αβ surface (αβs - 600 neurons) KCs [Aso et al., 2014].

The three main KC subtypes all have specific roles in olfactory memory formation. For instance, γ KCs are involved in STM, αβ KCs participate in LTM, and α'β' KCs play a part in memory consolidation [Yu et al., 2006][Krashes et al., 2007][Blum et al., 2009][Tomchik et al., 2009][Trannoy et al., 2011].

1.5.3. Kenyon cells and cell-type specificities

The differences between Kenyon cells subtypes are not only quantifiable by their roles in memory formation but also by their molecular changes [Henry et al., 2012][Crocker et al., 2016][Croset et al., 2017][Davie et al., 2018][Janssens et al., 2022].

At the cell type level, Kenyon cells compared to other neurons or glial cell types of the brain exhibit cell-type-specific gene expression and histone modifications (H3K4me3, H3K27ac, and H3K27me3)[Henry et al., 2012]. Using the INTACT method (a simple method for gene expression and chromatin profiling of individual cell types within a tissue [Steiner et al., 2012]) combined with a OK107-GAL4xUAS-GFP fly line, *Henry et al.* (2012) were able to highlight the transcriptional regulatory networks that underlie Kenyon cells identity. For example, the locus of the eyeless (ey) gene (a well-known transcription factor of the Kenyon cells) was globally active and lacked repression specifically in Kenyon cells.

Studies exploring the expression of the adult brain at the single cell level showed that Kenyon cells subtypes had cell-type-specific transcription signatures [Croset et al., 2017][Davie et al., 2018]. Using t-SNE clustering, the authors of *Croset et al.* (2017) were able to identify 26 genes differentially expressed in KC subtypes involved in gene regulation, signal transduction or synapse function. This study identified three clusters using a combination of markers for α/β (*sNPF*), α'/β' (trio), and γ (*sNPF*, trio) neurons. These two markers were identified based on previously published papers [Awasaki et al., 2000][Johard et al., 2008]. In their study of the ageing *Drosophila* brain, *Davie et al. (2018)* have highlighted various gene expression differences between the KCs subtypes that were related, for example, to mitochondrial activity. These differences in transcriptional programs between the three major classes of Kenyon cells indicate possible mechanistic differences.

Following the publication of *Davie* et al. (2018), Stein Aerts' lab explored the link between neuronal and glial cell type diversity and gene regulatory networks. In *Janssens* et al. (2022), they identified more than 95,000 regulatory regions that are used in different neuronal cell types, of which 70,000 were linked to developmental trajectories involving neurogenesis, reprogramming and maturation. Using a combination of motif discovery, deep learning and scATAC-seq data (Assay for Transposase-Accessible Chromatin with high-throughput sequencing), they applied a software package (DeepFlyBrain) to identify enhancer architectures and to predict cell type specific enhancers in 40 different cell types including the α/β, α'/β', and γ Kenyon cells.

A study exploring the relationship between mushroom bodies and memory formation has exhibited that mushroom body neurons present distinct patterns of gene expression following the induction of long-term memory [Crocker et al., 2016]. This cell-type-specific transcriptome analysis revealed memory-related changes for 23 genes, indicating that transcriptional changes in specific neuron types are associated with memory formation [Crocker et al., 2016].

Given that memory formation, transcription and chromatin accessibility can be cell type-specific in the mushroom bodies, it is possible that chromatin conformation might also differ between KCs subtypes to regulate transcription levels. This theory is supported by studies in mice that have established a close relationship between chromatin architecture and behaviour, including formation of long-distance enhancer-promoter interactions after motor-learning circuit activation [Yamada et al., 2019][Ito et al., 2014]. In *Drosophila*, many studies have also highlighted the importance of epigenetic gene regulation in fine-tuning behavioural responses [Anreiter et al., 2019] highlighting another level of complexity in behavioural phenotypes.

1.6. Concluding remarks

The genome of eukaryotic cells exhibits a highly hierarchical three-dimensional organisation. Among the various levels of compartmentalisation, topologically associating domains (TADs) are regions rich in chromatin interactions and play a crucial role in gene regulation. These domains must remain intact to maintain the physical proximity between gene regulatory elements, such as enhancers and promoters, which is essential for precise gene expression. Disruption of enhancer-promoter communication can lead to aberrant gene expression. Thus, the interplay between enhancers and promoters is critical for sustaining normal cellular functions and for responding to developmental and environmental cues.

Building on this understanding of chromatin structure, recent advancements in optical microscopy have enabled the simultaneous reconstruction of chromatin conformation, quantification of gene expression, and identification of cell types with single-cell resolution, while preserving tissue architecture. Despite these technological strides, the question of whether and how chromatin conformation contributes to cell-type-specific gene expression remains unresolved and challenging to address. As highlighted earlier, multiple factors and mechanisms intricately regulate gene expression and transcriptional programs unique to each cell type. Therefore, it is essential to measure how physical interactions between gene regulatory elements influence expression and how these interactions are modified during differentiation to maintain the different transcriptional programs of a differentiated tissue such as the *Drosophila* adult brain.

1.7. Aims of the thesis

The aim of my thesis manuscript is to explore the intricate processes underlying enhancer-promoter interactions, chromatin conformation and gene expression across different tissues in *Drosophila*. To achieve this, my objectives were as follows :

- Apply Hi-M to study enhancer-promoter interactions and chromatin conformation in different cell types in *Drosophila*
- Investigate whether E-P interactions and chromatin conformation vary across different cell types
- Examine whether E-P communication is instructive for transcription activation
- Identify types of E-P interactions

To approach these objectives, the first study presented in this thesis explores how topological chromatin structures restrict and specify communication between enhancers and promoters. This collaborative work aims to provide insights into the mechanisms underlying gene expression control and chromatin organisation at a specific genomic locus.

In the second project, I attempted to address all the previously mentioned objectives. Although, I am particularly interested in exploring the potential role of E-P topologies in memory formation. To effectively investigate such relationship, it is essential first to reconstruct genome folding in the *Drosophila* adult brain. Recently, *Mohana et al. (2023)* produced the first Hi-C matrices for the *Drosophila* adult Central Nervous System (CNS). However, despite this advancement, there is currently no established protocol for sequential DNA-FISH or RNA-FISH in the *Drosophila* adult brain, while such protocols are necessary to simultaneously detect chromatin conformation and transcription status at the single cell level within a preserved tissue. Thus, I aimed on developing tools to detect genome conformation and identify enhancer-promoter interactions specific to different types of neurons in the adult *Drosophila* brain for genes involved in memory formation.

Chapter 2 : Results

During my PhD, my work focused primarily on studying 3D structural organisation in single cells using the Hi-M technique to understand the complex mechanisms underlying enhancer-promoter communication and cell-type-specific gene expression.

In the first study, "A PRE loop at the *dac* locus acts as a topological chromatin structure that restricts and specifies enhancer-promoter communication" by *Denaud et al. (2024)*, we explored how a loop between two PREs, located at the *dac* locus, influences enhancer-promoter interactions, the overall conformation of the locus and transcription. This work is presented in section 2.1. in the form of a research article which is currently accepted for publication in the journal Nature Structural & Molecular Biology.

In the second part of this chapter, I investigated the link between genome conformation and gene expression in the adult *Drosophila* brain, particularly in the mushroom bodies and the Kenyon cells. This neuronal type is divided into three subtypes : γ Kenyon cells, αβ Kenyon cells, and α'β' Kenyon cells. Each of these subtypes has its own function and transcriptional program, making Kenyon cells an excellent model for studying the physiological implications of genome conformation. The objectives of my main project were as follows:

- Develop tools to simultaneously detect genome conformation in different cell types in the adult *Drosophila* brain
- Identify differentially expressed genes and their potential enhancers in various cell types
- Investigate variations in interactions between gene regulatory regions in Kenyon cells to establish a link between chromatin conformation and transcription

The results of this ongoing work are detailed in section 2.2.

2.1. Research article: A PRE loop at the *dac* locus acts as a topological chromatin structure that restricts and specifies enhancer promoter communication

The aim of this paper is to investigate the role of a PRE loop at the *dac* locus in regulating enhancer-promoter interactions. This collaborative project spanned over the four years of my PhD, starting in September 2020. Initially, I was involved in testing the *dac* locus library in embryos. After Marcelo Nollmann improved the initial design of the *dac* library, Sandrine Denaud and I progressed on adapting Hi-M for imaging pupae with the support of Giacomo Cavalli and Marcelo Nollmann. *Drosophila* pupae had never previously been tested in our team (Marcelo Nollmann's team) thus a few developments had to be made. Subsequently, I acquired multiple replicates and started the analysis using pyHi-M. Post-processing of the data began in December 2022 and continued intermittently until the end of April 2024. The Hi-M results I generated were discussed with Sandrine Denaud, Bernd Schuttengruber, Giacomo Cavalli and Marcelo Nollmann. This paper has been accepted by Nature Structural & Molecular Biology and should be in press shortly.

A PRE loop at the *dac* locus acts as a topological chromatin structure that restricts and specifies enhancer promoter communication

Sandrine Denaud[1], Marion Bardou[2+], Giorgio-Lucio Papadopoulos[1+], Stefan Grob[3], Marco Di Stefano[1], Gonzalo Sabarís[1], Marcelo Nollmann[2], Bernd Schuettengruber[1*], Giacomo Cavalli[1*]

[1] Institute of Human Genetics, UMR9002 CNRS and University of Montpellier, 141 Rue de la Cardonille, 34396, Montpellier Cedex 5, France

[2] Centre de Biologie Structurale, University of Montpellier, CNRS UMR 5048, INSERM U1054, Montpellier, France

[3] Department of Plant and Microbial Biology and Zurich-Basel Plant Science Center, University of Zurich, Zurich, Switzerland.

[+] These authors contributed equally

[*] Contact Information

giacomo.cavalli@igh.cnrs.fr

bernd.schuttengruber@igh.cnrs.fr

Abstract

3D genome folding plays a fundamental role for the regulation of developmental genes by facilitating or constraining chromatin interactions between cis-regulatory elements (CREs). Polycomb response elements (PREs) are a specific kind of CREs involved in the memory of transcriptional states in *Drosophila melanogaster*. PREs act as nucleation sites for Polycomb group (PcG) proteins, which deposit the repressive histone mark H3K27me3, leading to the formation of a class of topologically associating domain (TADs) called Polycomb domains. PREs can establish looping contacts that stabilize gene repression of key developmental genes during development. However, the mechanism by which PRE loops fine tune gene expression is unknown. Using CRISPR/Cas9 genome engineering, we specifically perturbed PRE contacts or enhancer function and used complementary approaches including 4C-seq, Hi-C, and Hi-M to analyze how chromatin architecture perturbation affects gene expression. Our results suggest that the PRE loop at the *dac* gene locus acts as a constitutive 3D chromatin scaffold during *Drosophila* development that forms independently of gene expression states and has a versatile function: it restricts enhancer promoter communication and contributes to enhancer specificity.

Introduction

Eukaryotic genomes are highly organized within the 3D nuclear space. The development of chromatin conformation capture (3C)-based methods and advanced microscopy approaches highlighted the importance of the 3D chromatin topology of cis-regulatory elements (CREs) in gene regulation [1]. Regulatory interactions of CREs define, maintain and change the expression pattern of key developmental genes to ensure proper development. Loss of these control mechanisms is a frequent feature of cancer and disease [2,3].

In interphase nuclei, chromosomes are hierarchically organized into topologically associating domains (TADs) [4-6] that modulate gene regulation, although their exact function is debated. Extensive rearrangement of chromosomes and TADs in the *Drosophila* genome does not correlate with changes in gene expression [7] and removal of proteins involved in TAD border formation have no dramatic effect on gene expression [8]. On the other hand, genomic rearrangements of TADs can cause gene misexpression and diseases [9-13]. TADs facilitate local promoter-enhancer interactions and prevent inappropriate interactions between different TADs j[4,9,14]. Genome organization can shape transcription dynamics by two complementary mechanisms: “tethering elements” (TEs) within TADs foster interactions between CREs, whereas insulators and/or TAD boundaries prevent inappropriate interactions between enhancers and promoters [15].

In *Drosophila*, a specific class of TADs is characterized by the presence of H3K27me3 Polycomb mark, which can cover several hundreds of kilobases (kb), including key developmental regulatory genes [16,17]. Within these Polycomb domains, epigenetic regulatory sequences named Polycomb response elements (PREs), act as nucleation sites for the recruitment of the Polycomb repressive complexes 1 and 2 (PRC1 and PRC2), which are responsible for the deposition and spreading of H3K27me3 and H2AK118ub, respectively (reviewed in [18]). Many Polycomb domains contain multiple PREs and genes, which tend to be coregulated or involved in related developmental pathways [16,17]. PREs can participate in the maintenance of both active and repressed gene expression states throughout development (reviewed in [19]) and a subset of PREs can engage in specific chromatin contacts within Polycomb domains, forming chromatin loops (PRE loops) [20-22]. PRC1 is likely to play an important role in PRE looping, since PRC1 is critical for chromatin condensation of Polycomb domains and for the establishment of their long-range interactions in mammals (reviewed in [18]). Looping interactions might be mediated via oligomerization of the SAM domain of Polyhomeotic (PH), a subunit of PRC1, which is crucial for the condensation of individual Polycomb domains [23] as well as for mediating long-range Polycomb domain interactions [24,25].

Using the Polycomb domain associated with the leg patterning gene *dachshund* (*dac*) as a paradigm, we previously showed that the loss of PRE contacts induced a very specific gain of

function phenotype in the adult fly leg, without affecting gene expression during early development [22]. While this suggests that PRE loops have repressive functions and contribute to stabilize gene silencing during development, another PRE at the HOX gene locus that coincides with a loop anchor of tethering elements (TE) is involved in gene activation at the embryonic stage [15]. This suggests that chromatin loops involving PREs can mediate both, gene activation and repression, but many questions concerning PRE loop formation and their functions remain unsolved. It has not been defined whether PRE loops are spatially or temporally regulated and the consequences of loss of PRE contacts on the global 3D chromatin architecture of Polycomb TADs is unknown. Furthermore, the nature, timing and underlying molecular mechanisms of gene misexpression upon loss of a PRE loop remain to be studied.

In order to address these questions, we exploited the *dac* Polycomb domain by generating mutant fly lines in which we specifically interfere with PRE contacts and/or enhancer function and analyzed consequences on TAD architecture and gene expression during *Drosophila* development. Intriguingly, both PREs at the *dac* locus have been identified as tethering elements (TEs) in early embryogenesis[15]. Here, we show that the PRE loop constitutes a 3D chromatin scaffold of the Polycomb domain that forms independently of gene expression states and is present at all stages during fly development. Insertion of an insulator at various positions generates local insulation within Polycomb domains without creating a new TAD boundary. Intriguingly, the insulator blocks PRE looping, whereas an essential leg- enhancer can bypass the insulator to activate the *dac* gene. Loss of PRE looping does not lead to widespread gene activation, but correlates with the activation of the *dac* gene at a precise developmental stage and tissue. Importantly, the modulation of physical proximity between the leg enhancer and promoter is not a major determinant for gene activation. Finally, we show that reduced PRE looping results in loss of enhancer specificity, leading to ectopic activation of a neighboring gene. We propose that PRE loops form a topological scaffold structure within Polycomb domains that have a dual function: they restrict promoter enhancer communication, while they also contribute to enhancer promoter specificity.

Results

Dac expression during development upon loss of PRE looping

To analyze the importance of PRE looping for the 3D organization of Polycomb domains and gene expression, we created three classes of CRISPR/Cas9 mutant flies associated with the *dac* gene locus (**Fig. 1a**). The first class comprises PRE deletion lines that affect both PcG recruitment and PRE looping [22]. The second category corresponds to the deletion of an enhancer driving *dac* expression in the leg. Finally, the third class consists of inserting a gypsy

insulator sequence between two PREs at various positions up or downstream the leg enhancer.

We previously showed that the disruption of PRE looping does not change the embryonic expression pattern of the *dac* gene. However, a very specific gain of function phenotype was observed in the adult fly: the appearance of extra sex comb bristles (ESC) on the tarsal segment 2 (TS2) of male flies on the first legs [22]. To determine the nature and timing of *dac* misexpression inducing this phenotype, we analyzed *dac* expression patterns during leg development in PRE deletion fly lines by performing immunostaining of 3rd instar larval imaginal leg discs and early pupal imaginal leg discs, when leg segmentation takes place. In larval leg discs *dac* is expressed in a ring like shape corresponding to the medial leg structures (trochanter, femur, tibia and first tarsal segment) and no significant changes in *dac* expression pattern can be observed at the larval stage upon mutation of PRE sequences (**Fig. 1b**).

At the early pupal stage, *dac* is normally expressed in the proximal leg part and the first tarsal segment (TS1), whereas expression in the second tarsal segment (TS2) is low and it is completely absent in the more distal segments (TS3-5) (**Fig. 1c**). Importantly, expression of *dac* is significantly increased, specifically in the second tarsal segment (TS2) of the developing leg in PRE mutant fly lines, whereas *dac* remains repressed in tarsal segments 3-5 (**Fig. 1c,d**). No significant difference in global mRNA levels were observed by RT-qPCR analysis in pupal and larval imaginal leg discs (**Extended Data Fig. 1a**). This shows that PRE deletion leads to derepression of *dac* specifically in cells of the TS2 segment of pupal leg discs. Increased *dac* expression at this precise developmental stage and tissue is likely involved in transforming the segment TS2 into segment TS1 identity, consistent with the induction of ectopic sex combs - normally found exclusively on the more proximal segment TS1 - in the more distal segment TS2 (**Extended Data Fig. 1b**).

The PRE loop acts as a constitutive 3D chromatin scaffold

We hypothesized that changes in DAC expression, leading to the transformation of leg segment identity, may involve alterations in 3D chromatin architecture in mutant flies. To test this, we performed Hi-C experiments of 3rd instar larval imaginal leg discs, and early pupal leg discs and compared them to Hi-C data in embryos (**Fig. 2a**). Interestingly, *dac* intraTAD interactions decreases during development (**Fig. 2b, left**). This might reflect global decondensation of the domain correlating with an increase in *dac* expressing cells during the larval and pupal stages (approximately 20% *dac* expressing cells in late embryos versus 50% *dac* expressing cells in leg discs). Importantly, the local contact enrichment of the PRE loop is present at similar levels from embryos to pupae (**Fig. 2b, right**). Furthermore, no additional chromatin loop or major change in intra-TAD structure or TAD borders were observed during these stages of development. This indicates that the PRE loop is a prominent chromatin feature

of the *dac* TAD that persists at all stages of *Drosophila* development. Since the proportion of *dac* expressing cells significantly increases during development, we hypothesized that the PRE loop within the *dac* TAD may not be the key element determining cell-type specific expression.

To test this model, we determined the conformation of the *dac* TAD in the different leg segments expressing or not *dac*. For this, we turned to Hi-M, a multiplexed DNA-FISH technology that captures chromatin conformations in single cells [26]. We designed and amplified an oligopaint library tiling most of the *dac* TAD with different barcodes, reaching a mean resolution of 4 kb (**Supplementary Table 1**). This library was hybridized, and each barcode was sequentially imaged in early pupal leg discs (see Methods). Ensemble Hi-M maps obtained from pooled nuclei comprising segments 1-4 (TS1-4) confirmed the presence of a prominent long-range chromatin interaction between barcodes including and adjacent to the two PRE sequences (**Fig. 2c**). Hi-M maps obtained from different replicates were highly correlated (**Extended Data Fig. 2b**). As in Hi-C maps, the PRE loop includes additional neighboring chromatin regions, indicating that it involves an extensive set of chromatin interactions in the whole regions surrounding the PREs. Next, we analyzed the chromatin organization of the *dac* TAD in the different tarsal segments, where the *dac* gene is either highly expressed (TS1), weakly expressed (TS2) or repressed (TS3/4). Interestingly, PRE interactions can be observed with similar frequency in all these segments (**Fig. 2d**). Taken together, these results show that the *dac* PRE loop forms with similar frequencies at different stages of development, and within leg segments displaying different levels of *dac* expression. Thus, we conclude that the regulation of *dac* expression does not require changes in PRE looping.

A ring enhancer is necessary and sufficient for *dac* expression

Since loss of PRE function induces activation of the *dac* gene exclusively during leg development, we searched for putative regulatory regions driving expression of the *dac* gene specifically in the leg. Interestingly, a well conserved 567 bp sequence has been previously shown to recapitulate *dac* expression in a ring like shape in larval leg disc in transgenic reporter gene assays [27]. Therefore, this regulatory region, 20kb downstream to the *dac* promoter (**Fig. 1a**) was called ring enhancer (RE).

Knowing that the RE is sufficient to recapitulate *dac* expression in reporter assays, we asked whether the RE is essential for regulating the correct *dac* expression pattern at the endogenous *dac* TAD. Therefore, we created a mutant fly line carrying a 1 kb deletion encompassing the ring enhancer sequence (**Fig. 1a**). Heterozygous flies for the RE deletion are viable and do not show any morphological phenotype. In contrast, homozygous flies display short and deformed legs with fused segments, characteristic of *dac* loss of function (**Fig. 3a**). Strong loss of *dac*

expression was observed by immunostaining experiments in imaginal leg discs and early pupal discs (**Fig. 3b**). Although *dac* expression is absent in the large majority of cells, few random spots of cells expressing *dac* can be observed and morphological perturbations of the imaginal disc appear at early pupal stage during metamorphosis (**Fig. 3b**).

Together, these results show that the RE is the major cis-regulatory element driving the expression of the *dac* gene in larval and early pupal leg discs and is both necessary and sufficient for *dac* activation in the leg disc.

Loss of *dac* expression and PRE looping

RE deletion and concomitant loss of *dac* expression affects the formation and topology of the repressive H3K27me3 domain and in particular PRE looping.

Therefore, we first performed Hi-C experiments in wild type or ΔRE larval imaginal leg discs in order to determine the consequences of RE deletion on the PRE loop and global TAD architecture (**Fig. 3c**). Analysis of the Hi-C data and quantification of the interaction frequency of the PRE loop revealed that overall TAD structure, boundaries and PRE looping are not significantly affected upon deletion of the RE (**Fig. 3d,e**).

Next, we performed CUT&RUN experiments using H3K27me3 antibodies, in wild type or ΔRE leg imaginal discs (**Fig. 3f**). We did not observe any major changes in the distribution of the repressive H3K27me3 mark across the *dac* locus and quantitative analysis of H3K27me3 levels confirmed that the deposition of H3K27me3 across the TAD is not significantly changed upon deletion of the RE. This indicates that the RE does not play a role in PRE function and the deposition of H3K27me3. Overall, these experiments show that the PRE loop is formed independently of *dac* expression states and that the ring enhancer has no effect on PRE function and global TAD architecture.

Gypsy insertion induces insulation and reduces PRE looping

The identification and characterization of the RE as an essential cis-regulatory element driving *dac* expression in the leg strongly suggests that derepression of *dac* in TS2 upon loss of PRE function is mediated by the RE. To disentangle the functional relationship of PRE looping, RE-promoter communication and *dac* expression, we generated a panel of mutant fly lines, where we inserted a gypsy insulator between the RE and the *dac* promoter and/or between the two PREs (**Fig. 1a**). The gypsy insulator element has the ability to reduce enhancer-promoter or PRE-PRE interactions when placed between these elements [28,29]. In a previously characterized fly line (gypsy 1) the insulator is inserted upstream of the RE and is not located between the RE and *dac* promoter (gypsy 1 line, **Fig. 1a**). This insulator is not expected to interfere with the interaction of the RE and the *dac* promoter. We therefore created two lines where we inserted the gypsy insulator at different positions between the RE and the

dac TSS (gypsy 2 and gypsy 3 line, **Fig. 1a**). We hypothesized that in these lines the gypsy element should interfere with both, PRE looping and RE-promoter communication.

qChIP experiments confirmed that all three gypsy insertions generate an ectopic Su(Hw) binding site, indicating that the insulator sequence is functional (see **Extended Data Fig. 3a** for gypsy 2 and gypsy3 lines, Su(Hw) binding to gypsy 1 was analyzed in [22]). Additionally, CUT&RUN experiments in larval imaginal leg discs showed that PRE-mediated deposition of H3K27me3 is not significantly changed upon the insertion of the gypsy insulator in all three gypsy lines, confirming that recruitment of PcG complexes to PREs is not affected (**Extended Data Fig. 3b, c**).

To analyze the physical insulation activity of the gypsy element and its impact on PRE looping, we performed Hi-C experiments in larval imaginal leg discs of the three gypsy insertion lines (**Fig. 4**). Visual inspection of Hi-C contact maps revealed that no new TAD borders are formed around the gypsy insertion sites (**Fig. 4a**). A moderate insulation activity of the gypsy sequences can be visualized by plotting the maps of differential score enrichments of interactions in the WT line versus each of the gypsy mutant lines (**Fig. 4b**). Gypsy 1 insertion has the weakest effect on local insulation, whereas in comparison to gypsy 2 and gypsy 3 lines the WT had stronger contact enrichment in the region bypassing the gypsy insertion sites, indicating an insulation activity of gypsy on the surrounding chromatin. Accordingly, quantification of the insulation scores at the gypsy insertion sites showed a significant increase in insulation at the gypsy 2 and gypsy 3 insertion sites, whereas no significant statistical differences between the insulation profiles of the different conditions are observed in the region upstream of the PRE1 site (**Extended Data Fig. 4a,b**). Although gypsy 1 insertion showed no significant increase in local insulation in the Hi-C approach (**Fig. 4c,d**), we do observe an insulation of the surrounding chromatin at the gypsy 1 insertion site insulation activities by 4C-seq experiments using viewpoints up or downstream the gypsy insertion (**Extended Data Fig. 5a,b**. This suggests that, even though with different strength, all three gypsy elements have insulation activity.

Next, we analyzed how the physical insulation activity in the different gypsy lines affects the PRE loop. Quantification of the looping interactions between the two PRE regions revealed a significant reduction of PRE contacts upon gypsy insertions in each of the three mutant lines (**Fig. 4e**). Quantitative differences were still observed between the gypsy lines, with gypsy 1 insertion weakly reducing PRE contacts, whereas gypsy 2 and gypsy 3 insertions have stronger effects on PRE looping. Reduced PRE looping of the gypsy 1 line was further confirmed by 4C-seq experiments, using the *dac* promoter as a viewpoint (**Extended Data Fig. 5b**).

Finally, we analyzed changes in gypsy insulation activity and its consequences on PRE looping during *Drosophila* development. For this, we exclusively used the gypsy 2 line and performed

4C-seq experiments at different developmental stages (embryos, larval leg discs and pupal leg discs). Using the *dac* promoter (PRE2) (**Extended Data Fig. 6a**) and the RE (**Extended Data Fig. 6b**) as viewpoints we observed that gypsy insertion results in a similar physical insulation activity at all developmental stages and PRE looping is reduced at all developmental stages investigated (**Extended Data Fig. 6a,b**).

Altogether, these experiments indicate that gypsy insertion interferes with PRE looping during all stages of fly development investigated, although the position of insertion of the gypsy insulator affects the strength of the insulation effect.

Gypsy insulator does not block RE function

Next, we asked whether gypsy insertions affect enhancer-promoter interactions and change *dac* expression. We predicted that insertion of the gypsy insulator upstream the RE (gypsy 1), which interferes with PRE looping, might induce *dac* expression in the TS2, as it is the case in the PRE deletion lines (**Fig. 1**). On the other hand, gypsy 2 and gypsy 3 insulators, which are inserted between the RE and the *dac* promoter might block enhancer-promoter communication.

Intriguingly, insertion of the gypsy insulators between the RE and the *dac* promoter (gypsy 2 and gypsy 3) resulted in a significant increase in *dac* expression in the second tarsal segment of pupal leg discs (**Fig. 5a,b**), whereas *dac* activation upon gypsy insertion downstream the RE (gypsy 1) is much weaker. Moreover, activation of *dac* in TS2 in the gypsy 2 and gypsy 3 lines induced the *dac* gain of function phenotype (ESC) on adult male fly legs (**Fig. 5c**). The penetrance of ESC in the gypsy 2 and 3 lines (about 25% and 50% of all male flies, respectively) was significantly stronger than the penetrance observed in the gypsy 1 line. Notably, there is a good correlation between the penetrance of ESC, and the levels of overexpression of *dac* in TS2 (**Fig. 5b,c**), which in turn inversely correlate with the reduction of PRE contacts (**Fig. 4e**), suggesting that loss of PRE looping and *dac* activation in TS2 are functionally linked.

The absence of enhancer-blocking activity when the insulator is inserted between RE and promoter is surprising, given the previously reported enhancer blocking function of the gypsy element[30]. One possible explanation could be that gypsy blocks the RE but, in this case, a shadow enhancer might take over and begin to induce *dac* expression.

In order to test whether the overexpression of *dac* upon gypsy insertions is induced by the RE, we deleted the RE in the presence of the gypsy 2 insertion (gypsy2+ΔRE line, see **Fig. 1a**). Indeed, if the gypsy insertion induces *dac* overexpression independently of the RE, we would expect at least a partial rescue of the loss of function phenotype upon deletion of the RE (**Fig. 3a**). Instead, we observed that homozygous flies carrying the gyspy 2 insertion together with a deleted RE (gypsy2+ΔRE) display the same strong crippled leg phenotype than the deletion

of the RE alone, with a complete penetrance (**Fig. 5d**). Furthermore, RT-qPCR analysis showed that *dac* expression is reduced to the same extent in the gypsy2+ΔRE line compared to the ΔRE alone (**Fig. 5e**).

Altogether, these results indicate that the RE can bypass the gypsy insulator to activate *dac* expression irrespective of the genomic location of the gypsy insulator. Moreover, activation of the *dac* gene upon gypsy insertion and reduced PRE looping is strictly dependent on the presence of RE.

Unchanged enhancer promoter proximity upon *dac* gene activation

Next, we tested whether the bypass of the insulator and activation of the *dac* gene involves changes in the physical proximity between enhancer-promoter sequences. For this, we performed Hi-M experiments in pupal leg discs in the gypsy 2 mutant, which allows us to analyze RE-promoter distances specifically in the tarsal segment where *dac* is activated. Comparing the interaction profiles of gypsy 2 mutant cells to WT cells confirmed that the PRE loop is reduced in all tarsal segments (TS1, TS2 or TS3/4) independent of the *dac* expression status (**Extended Data Fig. 7a**). Furthermore, we observed increased short-range interactions up or downstream the gypsy insulator, whereas long range interactions between sequences up and downstream the insulator insertion are reduced, consistent with the insulator activity of the gypsy sequence (**Extended Data Fig. 7a,b**).

Next, we compared the distances between the RE and the *dac* promoter in the tarsal segments where *dac* is active (TS1), moderately expressed (TS2), or inactive (TS3/4) in WT or gyspy 2 mutant flies. Therefore, we created virtual 4C plots derived from Hi-M experiments using the *dac* promoter as a viewpoint (**Fig. 6a**). Remarkably, despite the upregulation of *dac* in the gypsy mutant, the RE-*dac* promoter distance was not decreased, but displayed a very modest increase between WT and gypsy mutant in TS2 (≤ 10 nm). A similar weak increase in RE-promoter distances is observed in tarsal segments, where *dac* remains repressed (TS3/4), indicating that this increase in RE-promoter distance is not sufficient for gene activation. We note that the RE-*dac* promoter distance changes were statistically significant for these segments (Wilcoxon two-sided rank test, **Fig. 6a**), but were in all cases comparable to the measurement error in the median distance as estimated by bootstrapping analysis (~8-25 nm, see Methods). We therefore concluded that activation of *dac* expression in TS2 upon loss of PRE looping is not the consequence of large-scale changes in RE-promoter distances. Likewise, RE-*dac* promoter distances exhibited no significant changes when we compared TS1 segment (where *dac* is active) to segments TS3/4 (where *dac* is completely repressed) in WT pupal leg discs (**Fig. 6b**). Importantly, the distance distributions across individual cells between PREs or enhancer-promoter elements show no evidence for bimodality in all tarsal segments including TS2, which comprises a mixed population of *dac* expressed and repressed

cells (**Extended data Fig. 8a,b**). This argues against the existence of different chromatin conformations within the examined tarsal segments.

Together, these results indicate that increased physical proximity between the RE and *dac* promoter does not appear to be a major mechanism involved in *dac* gene activation.

The PRE loop contributes to enhancer-promoter specificity

Since the major structural effect of gypsy insertion is to reduce PRE looping at the *dac* domain, we asked whether other genes within the domain are affected by the loss of looping interactions. Therefore, we first performed RT-qPCR analysis at different developmental stages in mutant fly lines, where we interfered with either PRE function (Double, ΔPRE2), PRE looping (gypsy2) or enhancer function (ΔRE) (**Fig. 7a**). We found that the *CG5888* gene is induced already at the larval stage, upon loss of PRE contacts. The *Idgf1* gene is also activated, whereas two other genes, *Idgf2* and *Idgf3*, stay repressed at all developmental stages analyzed (**Extended Data Fig. 9a**). RNA FISH analysis confirmed the transcriptional activation of *CG5888* upon reduction of PRE looping in both the gypsy 2 and gypsy 3 larval leg discs (**Fig. 7b, Extended Data Fig 9b,c**). Interestingly, *CG5888* is mainly overexpressed in cells that also express *dac* and where the RE is active (**Fig. 7c**), resulting in a similar "ring like" shaped expression. Importantly, expression of the *CG5888* gene is not significantly changed upon deletion of the RE alone (**Fig. 7b,d,e**), indicating that *CG5888* expression is not controlled by the RE in WT conditions. However, the ectopic expression of the *CG5888* gene upon PRE looping reduction depends on the presence of the RE, since *CG5888* activation was lost when we deleted the RE in the presence of the gypsy2 insertion (ΔRE +gypsy2) (**Fig. 7b,d,e**). Importantly, the perturbation of the PRE-loop does not lead to the activation of all genes within the *dac* TAD (**Extended Data Fig. 9a**), and is restricted to the region where the RE is active during leg development. This indicates that the loss of the PRE loop does not create a global permissive environment facilitating transcription *per se*, but rather regulates RE specificity towards permissive promoters, like the *CG5888* gene. Altogether, these results suggest that, in addition to TADs, which can restrict enhancer-promoter communication across TAD borders, PRE loops within TADs constitute an additional gene regulatory layer by contributing to intra-TAD enhancer specificity (**Extended Data Fig. 10a**).

Discussion

We showed that the *dac* PRE loop constitutes a topological chromatin structure that has a versatile function: on one hand, it can regulate enhancer-promoter communication in a developmental stage and tissue specific manner. On the other hand, the PRE loop contributes to enhancer-promoter specificity by restricting the enhancer activity to its specific target

promoter (the *dac* gene), as shown by the RE-dependent illegitimate activation of the *CG5888* gene upon insertion of a gypsy insulator between the two PREs.

Gypsy insulator reduces PRE looping but not enhancer function

Insulator elements are DNA sequences that act as chromatin boundaries and regulate interactions between genomic regulatory elements. The gypsy element is one of the best-characterized insulators. It contains three core components Su(Hw), CP190 and Mod(mdg4) and the zinc finger protein CLAMP that promotes gypsy enhancer blocking activity [31].

At the *dac* TAD, the gypsy insulator sequence interferes with the interaction of the two PREs when inserted between them. However, gypsy insertion between the RE and the *dac* promoter does not block enhancer-promoter communication. Although surprising, these results are consistent with previous work showing that the Su(Hw) protein binds at thousands of sites throughout the genome, yet is not systematically associated with the physical boundaries of gene units [5]. Although more than 20 enhancers have been shown to be blocked by the gypsy insulator (see [30] and references therein), this element does not necessarily establish an impermeable chromatin barrier [32,33]. In all the enhancer and promoter pairs that were analyzed in these studies, each element of the pair was located close to the gypsy element and insulation might perhaps depend on the ability of the gypsy element to reduce short-range chromatin contacts. Furthermore, a recent genome-wide analysis suggests that gypsy-binding proteins do not generally act by blocking loop formation but rather induce local insulation at their binding sites [34].

In transgenic reporter assays, the gypsy insulator behaves as a chromatin border that is able to block spreading of H3K27me3 and prohibit contacts between a PRE and a distal promoter [29,35]. In contrast, at the *dac* TAD, insertion of the same gypsy insulator between the two PREs does not interfere with the formation of the repressive H3K27me3 domain. This is in agreement with the presence of endogenous Su(Hw) binding sites within the *dac* TAD that, at this genomic locus, do not act as classical chromatin domain borders and do not interfere with the deposition of H3K27me3. This indicates that the function and effect of the gypsy insulator on the physical interaction between cis-regulatory elements or the partitioning of chromatin domains is context dependent.

PRE contacts constitute a specific form of TE loops

A recent paper used the Micro-C technique to demonstrate that the genome is organized by insulator elements and so called “tethering elements” (TEs) in early *Drosophila* embryos[15]. These TEs correspond to “organizational elements” that form chromatin loops in order to facilitate specific enhancer-promoter contacts for rapid gene activation.

Intriguingly TEs and PREs show several similarities. First, both PRE loop and TE loop disruptions have little effect on the overall structure of TADs, whereas they have an impact on enhancer-promoter communication. Secondly, both elements are frequently bound by GAGA factor, which has been proposed to mediate chromatin loops [36]. Finally, and most compellingly, we observed that 48% (109 out of 225) of PREs identified in embryos coincides with TEs (**Extended Data Fig. 10b**), notably including the two PREs of the *dac* gene locus. We therefore propose that PRE loops actually constitute a specific form of TE loops.

In contrast to the *dac* PRE loop, which restricts E-P communication in pupal leg discs, a chromatin loop between TEs (that also corresponds to PREs) at the *scr* locus has been shown to promote enhancer-promoter contacts [15]. Another example of a TE loop that correspond to a PRE loop involves the regulation of the cut (*ct*) gene. Importantly, a natural insertion of a gypsy retrotransposon between the two PREs/TEs of the *ct* gene locus leads to *ct* down-regulation and loss of enhancer-promoter communication [28]. So clearly, chromatin loops involving PREs/TEs can mediate, both, gene activation and repression. Intriguingly deletion of both looping anchor points of the *dac* PRE loop results in reduced *dac* gene expression in embryos [22]. Although this reduced expression does not result in developmental defects, this suggests that even the same PRE/TE loop can have different functions in gene activation and repression depending on the developmental stage or tissue. These antipodal functions of PRE/TE loops might be mediated by the presence of developmental or stage specific factors. Interestingly, a recent study analyzed chromatin loops of paralogous gene pairs, proposing an additional function of chromatin loops in the fine-tuning of coordinated expression levels of genes with related function [37]. As it is the case for the *dac* PRE loop, the same chromatin loop can have different functions in the regulation of gene expression [37].

Together, these data suggest that PRE/TE loops constitute a chromatin scaffold structure that is used to regulate enhancer-promoter communication positively or negatively, depending on the gene locus, the developmental stage and on the specific tissue.

PRE looping as a regulator of enhancer-promoter communication

Loss of PcG binding to the PREs is unlikely to be the major mechanism leading to *dac* gene activation upon loss of PRE looping. We previously observed that deletion of both PREs, and concomitant loss of PcG binding, is not sufficient to globally activate *dac* expression, but only results in the same tissue specific activation of the *dac* gene than the insertion of the gypsy insulator. Moreover, PcG proteins have been shown to remain associated with PREs when their target genes are active [38,39]. In contrast, activation of the *dac* gene in the 2nd tarsal segment upon loss of PRE looping is strictly dependent on the ring enhancer, indicating that gene activation involves changes in enhancer-promoter communication.

A popular model of enhancer-promoter communication proposes chromatin looping as a mechanism to bring the two regulatory regions in close proximity [40] and experimentally forced enhancer-promoter looping contributes to gene activation [41]. By applying Hi-M to detect chromatin 3D organization in single nuclei [42], including loops at distances similar to the one separating RE from the *dac* TSS, we were unable to detect increased contact frequency between RE and the *dac* promoter compared to surrounding regions. In addition, no significant differences in contact frequencies of the *dac* promoter and RE was observed in repressed versus active cells, or upon loss of PRE looping, suggesting that no stable contact is needed for *dac* transcription. This agrees with a previous report, showing no differences between enhancer-promoter pairs in active or inactive transcriptional states during early *Drosophila* development [42]. In addition, live cell imaging analyses challenge the idea of stable loops as a general mechanism for all enhancer-promoter communication [43,44]. Finally, increased enhancer-promoter distance has been reported to accompany Shh gene activation during neural differentiation [45], supporting evidence that enhancers can act at a distance.

Several “action-at-a-distance models” have been proposed to explain functional interaction of distant enhancers with their promoters [46], including liquid-liquid phase separation, where enhancers function as binding surfaces that concentrate transcriptional activators in ‘‘transcriptional condensates,’’ allowing them to induce gene expression even when promoters are not in absolute proximity. Alternatively, in the transcription factor activity gradient model, enhancer-bound co-activators can activate target promoters within a “permissive range”. It is conceivable that the PRE-loop might act as crucial scaffold to create a particular chromatin environment or a “permissive range”, thereby regulating activation of the *dac* gene and enhancer specificity.

It is worth mentioning that the ring enhancer at the *dac* locus is not associated with the previously well-established molecular signatures of active enhancers (like H3K4me1 and H3K27Ac enrichment) in embryos or larval disc tissues [47]. This surprising lack of histone marks at an essential enhancer might indicate that other such enhancers with an important function but undetectable epigenomic signature exist in the genome. In the future, it will be interesting to study whether chromatin looping or contact-independent mechanisms are differently involved in gene activation for different classes of enhancer elements.

Acknowledgments

We are grateful to Victor Corces (Department of Human Genetics, Emory University School of Medicine, Atlanta) for Su(Hw) antibody. We would like to thank MRI and Drosophila facilities (BioCampus Montpellier), CNRS, INSERM, and University de Montpellier. This work was supported by grants from the European Research Council (Advanced Grant 3DEpi, G.C), the European CHROMDESIGN ITN project (Marie Skłodowska-Curie grant agreement No

813327, G.C.), the European E-RARE NEURO DISEASES grant “IMPACT”, by the Agence Nationale de la Recherche (PLASMADIFF3D, grant N. ANR-18-CE15-0010, LIVCHROM, grant N. ANR-21-CE45-0011, G.C.), by the Fondation ARC (EpiMM3D, G.C.), the Fondation pour la Recherche Médicale (EQU202303016280, G.C.), by the MSD Avenir Foundation (Project GENE-IGH), and by the French National Cancer Institute (INCa, PIT-MM grant N. INCA-PLBIO18-362, G.C.). This project was also funded by the European Union’s Horizon 2020 Research and Innovation Program (Grant ID 724429) (M.N.), the Bettencourt-Schueller Foundation (prize ‘Coup d'élan pour la recherche Française’, M.N.), France-BioImaging (grant ID ANR-10-INBS-04, ‘‘Investments for the Future’’, M.N.), and the Drosophila facility (BioCampus Montpellier, CNRS, INSERM, Université Montpellier, France). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author Contributions

GC and BS conceived this study. SD, MB and BS performed experiments. SG performed bioinformatic analysis of 4C experiments. MN and MB conceived and designed Hi-M experiments and MB performed image acquisition and data analysis. GP and MDS performed bioinformatic analysis of Hi-C experiments. GS performed analysis of C&R data. BS, GC wrote the manuscript. BS, CG and MN edited the manuscript.

Competing interests

The authors declare no competing interests.

Figure Legends:

Fig. 1: Expression of *dac* gene during *Drosophila* development upon loss of PRE function and/or PRE looping

(**a**) Top: Schematic representation of the *dac* TAD (WT). Gray shadow represents H3K27me3 domain.

Bottom: CRISPR/Cas9 mutant fly lines used in this study. Orange flash indicates mutated PRE deletion lines and Gypsy 1 have been described in [22]. Enhancer deletion lines and gypsy2 and gypsy 3 lines have been generated in this study.

(**b**) DAC immunostaining analysis of 3rd instar larval imaginal leg discs of WT and PRE deletion lines. White bars indicate 30 micrometers.

(**c**) DAC immunostaining analysis in early pupal imaginal leg discs (4-5h after pupation). Tarsal segments 1 to 5 (TS1-5) are indicated. White bars indicate 30 micrometers. At this pupal stage,

the imaginal disc is composed of a single layer of cell that connect in 3D in a “tubal” shape. The images correspond to one side of the pupal leg disc imaged by confocal microscopy.

(**d**) Quantification of DAC immunostaining signals. The average signal ratio between TS2/TS1 segments is plotted for WT versus mutant lines. A minimum of 6 pupal discs (WT: n=13, Double: n=6, ΔPRE2 n=9, ΔPRE1 n=7) were scored. Error bars indicate s.d. *** indicates p-value < 0.001, ** indicates p-value < 0.01, * indicates p-value < 0.1 (two-sided unpaired t-test).

Fig. 2: PRE loop during *Drosophila* development and *dac* expressing or repressed cells

(**a**) Hi-C score maps (see Methods) of a 200 kb region at 3kb resolution on chromosome 2L including the *dac* gene locus in whole late embryos (left), 3rd instar leg imaginal discs (middle) and early pupal stage (4-5 hours after pupation) (right). Black circle indicates the position of the *dac* PRE loop. Violet bars indicate position of PREs. Black arrows indicate Gene promoters of the *dac* and the CG5888 genes.

(**b**) Left: Distributions of the Log2 ratios of the frequencies of observed Hi-C contacts within Polycomb-associated TADs in chr2L in embryo over the larval and pupal leg discs (n=25 values in each distribution). The contact frequency for each condition is computed against the total number of valid-pairs of the corresponding condition. The ratio is computed as the contact frequency in embryo over the equivalent quantity in larvae or pupae. Hence, a positive value of the Log2 for the *dac* domain indicates a decrease of contact frequency during the analyzed developmental stages. Boxplots show median (central line), Q1=25th and Q3=75th percentiles (box limits), and Q1+1.5×IQR to Q3+1.5×IQR (whiskers), where IQR is the interquartile range. Outliers are not shown. The unpaired two-sided Wilcoxon statistical test was used to estimate the reported *p-values*.

Right: Quantification of Hi-C scores the *dac* PRE loop in embryos (n=195), larval leg discs (n=172) and early pupal leg discs (n=183). Reported *p-values* (top) result from comparing the embryos with the larval and the pupal leg-discs distributions, respectively. The number of points per distribution (middle) is reported (see Methods). Boxplots show median (central line), Q1=75th and Q3=25th percentiles (box limits), and Q1+1.5×IQR to Q3+1.5×IQR (wiskers), where IQR is the interquartile range. The unpaired two-sided Wilcoxon statistical test was used to estimate the reported *p-values*.

(**c**) Ensemble Hi-M proximity matrices generated from all tarsal segments (TS1-5) for early pupal leg discs (4-5 h after pupation). On the scale bar, red and blue represent, respectively, high proximity and low proximity frequencies. The proximity frequency has been normalized by the genomic distance. Black circle indicates the position of the PRE1-PRE2 loop within the *dac* TAD. Barcodes 29 (PRE1) and 48 (PRE2) are highlighted in purple.

(**d**) Ensemble Hi-M proximity maps in early pupal stage tarsal segments (TS1, *dac* active), (TS2, *dac* weakly active) and (TS3/4, *dac* repressed). The proximity frequency has been

normalized by the genomic distance. Matrices were generated from 6535 traces for TS1, 7363 traces for TS2 and 9993 traces for TS3/4. Black circles indicate the position of the PRE1-PRE2 barcodes. Barcodes 29 (PRE1) and 48 (PRE2) are highlighted in purple.

Fig. 3: Phenotypic, transcriptional and 3D chromatin conformational consequences of the deletion of the ring enhancer (ΔRE) sequence.

(**a**) Representative picture of homozygous adult flies and legs homozygous mutant for the ring enhancer (ΔRE). Tibia (ti) and tarsal segments (TS) are fused.

(**b**) DAC immunostaining analysis in WT and ΔRE mutant larval 3rd instar imaginal leg disc (left) or early pupal imaginal leg discs (4-5h after pupation) (right). White scale bar indicates 30 micrometers.

(**c**) Hi-C score (Methods) maps of a 200 kb region at 3kb resolution on chromosome 2L at *dac* gene locus in 3rd instar imaginal leg disc in WT or ΔRE flies. Black circle indicates the position of the *dac* PRE loop. Violet bars indicate position of PREs. Black arrows indicate Gene promoters of the *dac* gene and the CG5888 genes. Red square indicates Ring enhancer (RE).

(**d**) Quantification of the *dac* PRE loop interaction scores. Hi-C interaction score in WT larval leg discs (n=172) and discs containing a deletion in the ring enhancer (∆RE, n=73). Reported p-values result from comparing the WT and ∆RE distributions. The number of points per distribution is reported (see Methods). Boxplots show median (central line), Q1=25th and Q3=75th percentiles (box limits), and Q1+1.5×IQR to Q3+1.5×IQR (wiskers), where IQR is the interquartile range. The unpaired two-sided Wilcoxon statistical test was used to estimate the reported *p-values*.

(**e**) Insulation profile shown at 3kb resolution along the *dac* 200kb region in larval WT and ∆RE mutant leg discs is shown as the mean value (line) +/- the standard deviation (shaded area) over the insulation scores (IS) computed using 5 different values of the window parameter (w=100, 150, 200, 250, and 300 kb, see Methods).

(**f**) CUT&RUN profile for H3K27me3 mark at the dac domain performed in 3rd instar imaginal leg disc in WT or ΔRE flies. Red bar indicates position of the RE. Violet bars indicate positions of PREs. Grey bar below H3K27me3 tracks demarcates the PcG TAD [22]. The scatter plot on the right shows the H3K27me3 enrichment in the 131 Drosophila Polycomb domains in ΔRE as a function of WT flies (log2 scale). The *dac* Polycomb domain is highlighted in blue.

Fig. 4: Consequences of gypsy insertions for the *dac* Polycomb domain architecture analyzed by Hi-C.

(**a**) Hi-C score (see Methods) maps of a 200 kb region at 3kb resolution on chromosome 2L including the *dac* gene locus in 3rd instar imaginal leg disc carrying the indicated gypsy insertions. Black circle indicates the position of the *dac* PRE loop. Violet bars indicate position

of PREs. Black arrows indicate Gene promoters of the *dac* gene and the CG5888 genes. Red square indicates Ring enhancer (RE). Green bar indicates the position of the gypsy insulator.
(**b**) Differential Hi-C scores maps (WT *vs* gypsy 1, gypsy 2 or gypsy 3 mutants) of a 150 kb region of the *dac* gene in 3rd instar imaginal leg disc (see Methods). Black dashed lines within differential Hi-C maps indicate position of local insulation. Black arrows indicate gene promoters of the *dac* gene and the *CG5888* genes. Red square indicates Ring enhancer (RE). Green bar indicates the position of the gypsy insulator.
(**c**) Insulation profile shown at 3kb resolution along the *dac* 200kb region in larval WT and gypsy1, gypsy 2 and gypsy 3 mutant leg discs is shown as the mean value (line) +/- the standard deviation (shaded area) over the insulation scores (IS) computed using 5 different values of the window parameter (w=100, 150, 200, 250, and 300 kb, see Methods). Gypsy insulator insertions (gypsy 1, gypsy 2, gypsy 3) are indicated.
(**d**) P-values from the comparisons of insulation scores (IS) at gypsy 1, gypsy 2 and gypsy 3 insertion sites between the WT and the corresponding fly line (see Methods). The *p-values* resulted from a two-sided Welch t-test between the WT condition and each of the gypsy mutant at the corresponding locus.
(**e**) Quantification of the *dac* PRE loop Hi-C interaction score in WT (n=172), gypsy 1 (n=106), gypsy 2 (n=106), and gypsy 3 (n=111) mutant flies. Reported p-values result from comparing the WT and gypsy 1, gypsy 2 and gypsy 3 distributions. The number of points per boxplot (middle) is reported (see Methods). Boxplots show median (central line), Q1=25th and Q3=75th percentiles (box limits), and Q1+1.5×IQR to Q3+1.5×IQR (wiskers), where IQR is the interquartile range. The unpaired two-sided Wilcoxon statistical test was used to estimate the reported *p-values*.

Fig. 5: Transcriptional and phenotypic consequences of gypsy insulator insertions at the *dac* Polycomb domain.

(**a**) RNA FISH images of WT and gypsy insertions in early pupal imaginal leg discs (4-5h after pupation) (*dac* gene: violet). White bars indicate 30 micrometers. Tarsal segments 2 is encircled by a dashed line. These images are z projection of one side of the pupal leg disc representing a single layer of cell.
(**b**) Quantification of FISH signals of the *dac* gene in TS2 vs TS1 is shown. For each segment of the pupal leg disc, the number of *dac* FISH spots has been quantified and normalized by the surface of the segment. Y axis represents the average intensity ratio of TS2/TS1 values of 4 imaginal discs (n=4). Data are presented as the mean values ± s.d (error bars) of replicates. ** indicates p-value < 0.01; *** indicates p-value < 0.001 (two-sided unpaired t-test).
(**c**) Quantification of the extra sex comb (ESC) phenotype in the indicated fly lines grown at 25°C. A minimum of 50 male flies were scored. ESC phenotype in gypsy1 flies can only be

observed in a sensitized mutant background, where levels of the PcG protein PH are reduced (PH^{410} mutant background) [22].

(**d**) Representative example of adult flies carrying a deletion of the RE and/or an insertion of the gypsy insulator (gypsy 2).

(**e**) RT-qPCR analysis of 3rd instar imaginal leg discs in the indicated fly lines. The fold change between the indicated mutant and wild type (control) line is shown. Three independent experiments have been performed. Data are presented as the mean values ± s.d (error bars) of replicates. *** indicates p-value < 0.001, ** indicates p-value < 0.01 (two-sided unpaired t-test).

Fig. 6: RE-*dac* promoter distances in Tarsal segments (TS) of early pupal leg discs in WT versus gypsy 2 flies determined by Hi-M.

(**a**) Virtual 4C plots derived from Hi-M pairwise distances (PWD) difference matrices of WT and gypsy 2 mutant conditions of pupal leg discs in the indicated tarsal segments (TS). Barcode 48 (*dac* promoter) is used as a viewpoint (black cross). Blue dots represent shorter distances for the WT condition and red dots for the gypsy 2 mutant condition. Barcode 43 (salmon bar) indicates the position of the RE. Violet bars highlight the positions of PRE1 (barcode 29) and PRE2 (barcode 48). Green bar indicates the position of the gypsy 2 insulator. Gray bar indicates position of the CG5888 promoter (barcode 37). Asterisks represent p-values obtained by Wilcoxon two-sided rank test: p-value [*] < 0.05.

(**b**) Virtual 4C plots derived from Hi-M PWD difference matrices of TS3/4 segments and TS1 segment in wild type pupal leg discs. Blue dots represent shorter distances for TS3/4 and red dots for TS1 segment. Barcode 43 (salmon bar) indicates the position of the RE. Violet bars (barcodes 29 and 48) highlight the position of PRE1 and PRE2, respectively. Green bar indicates the position of the gypsy insulator. Gray bar indicates the position of the CG5888 promoter (barcode 37).

Fig. 7: Consequences of gypsy insertions and/or RE deletion on *CG5888* and *dac* gene expression.

(**a**) RT-qPCR analysis at the indicated developmental stages using primers specific for *CG5888* gene. The fold change between the indicated mutant and wild type (control) line is shown. Three independent experiments have been performed. Data are presented as the mean values ± s.d (error bars) of replicates. * indicates p-value < 0.1, ** indicates p-value < 0.01 (two-sided unpaired t-test).

(**b**) RNA FISH images of WT, gypsy2 insertion, RE deletion (ΔRE) and gypsy2 insertion in the context of the deleted RE (ΔRE+gypsy2) in 3rd instar imaginal leg discs (*dac* gene: violet, *CG5888* gene: green). White bar indicates 30 micrometers

(**c**) Quantification of *CG5888* RNA FISH signals observed in *dac* expressing cells in WT and gypsy 2 3rd instar imaginal leg discs. 9 pupal discs (n=9) were scored. Data are presented as the mean values ± s.d (error bars) of replicates. ** indicates p-value < 0.01 (two-sided unpaired t-test).

(**d**) Quantification of *CG5888* (left) or *dac* (right) RNA FISH signals in the indicated fly lines. The number of FISH spots has been quantified and normalized by the total disc area (*CG5888*) or by the number of FISH spots in WT condition (*dac*). Y axis represents the average of 6 imaginal discs values. Error bars indicate s.d. *** indicates p-value < 0.001. ** indicates p-value < 0.01, * indicates p-value < 0.1 (two-sided unpaired t-test).

(**e**) RT-qPCR analysis of WT, gypsy2 insertion, RE deletion (ΔRE) and gypsy2 insertion in the context of the deleted RE (ΔRE+gypsy2) in 3rd instar imaginal leg discs using primers specific for the CG5888 gene. The fold change between the indicated mutant and wild type (control) line is shown. Three independent experiments have been performed. Data are presented as the mean values ± s.d (error bars) of replicates. ** indicates p-value < 0.01 (two-sided unpaired t-test).

References

1. Jerkovic, I. & Cavalli, G. Understanding 3D genome organization by multidisciplinary methods. *Nat Rev Mol Cell Biol* **22**, 511-528 (2021).
2. Peng, A. et al. Regulation of 3D Organization and Its Role in Cancer Biology. *Front Cell Dev Biol* **10**, 879465 (2022).
3. Deng, S., Feng, Y. & Pauklin, S. 3D chromatin architecture and transcription regulation in cancer. *J Hematol Oncol* **15**, 49 (2022).
4. Nora, E.P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-5 (2012).
5. Sexton, T. et al. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458-72 (2012).
6. Dixon, J.R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-80 (2012).
7. Ghavi-Helm, Y. et al. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat Genet* **51**, 1272-1282 (2019).
8. Despang, A. et al. Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat Genet* **51**, 1263-1271 (2019).
9. Lupianez, D.G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-1025 (2015).
10. Franke, M. et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265-269 (2016).
11. Narendra, V. et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017-21 (2015).
12. Guo, Y. et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900-10 (2015).
13. Flavahan, W.A. et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110-4 (2016).

14. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454-1458 (2016).
15. Batut, P.J. et al. Genome organization controls transcriptional dynamics during development. *Science* **375**, 566-570 (2022).
16. Schwartz, Y.B. et al. Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat Genet* **38**, 700-5 (2006).
17. Schuettengruber, B. et al. Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. *PLoS Biol* **7**, e13 (2009).
18. Schuettengruber, B., Bourbon, H.M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34-57 (2017).
19. Schuettengruber, B. & Cavalli, G. Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* **136**, 3531-42 (2009).
20. Eagen, K.P., Aiden, E.L. & Kornberg, R.D. Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proc Natl Acad Sci U S A* **114**, 8764-8769 (2017).
21. Hug, C.B., Grimaldi, A.G., Kruse, K. & Vaquerizas, J.M. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169**, 216-228 e19 (2017).
22. Ogiyama, Y., Schuettengruber, B., Papadopoulos, G.L., Chang, J.M. & Cavalli, G. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during Drosophila Development. *Mol Cell* **71**, 73-88 e5 (2018).
23. Kundu, S. et al. Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Mol Cell* **65**, 432-446 e5 (2017).
24. Isono, K. et al. SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Dev Cell* **26**, 565-77 (2013).
25. Wani, A.H. et al. Chromatin topology is coupled to Polycomb group protein subnuclear organization. *Nat Commun* **7**, 10291 (2016).
26. Cardozo Gizzi, A.M. et al. Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. *Mol Cell* **74**, 212-222 e5 (2019).
27. Giorgianni, M.W. & Mann, R.S. Establishment of medial fates along the proximodistal axis of the Drosophila leg through direct activation of dachshund by Distalless. *Dev Cell* **20**, 455-68 (2011).
28. Jack, J. & DeLotto, Y. Structure and regulation of a complex locus: the cut gene of Drosophila. *Genetics* **139**, 1689-700 (1995).
29. Comet, I., Schuettengruber, B., Sexton, T. & Cavalli, G. A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc Natl Acad Sci U S A* **108**, 2294-9 (2011).
30. Kuhn, E.J., Viering, M.M., Rhodes, K.M. & Geyer, P.K. A test of insulator interactions in Drosophila. *EMBO J* **22**, 2463-71 (2003).
31. Bag, I., Dale, R.K., Palmer, C. & Lei, E.P. The zinc-finger protein CLAMP promotes gypsy chromatin insulator function in Drosophila. *J Cell Sci* **132**(2019).
32. Morris, J.R., Chen, J.L., Geyer, P.K. & Wu, C.T. Two modes of transvection: enhancer action in trans and bypass of a chromatin insulator in cis. *Proc Natl Acad Sci U S A* **95**, 10740-5 (1998).
33. Scott, K.C., Taubman, A.D. & Geyer, P.K. Enhancer blocking by the Drosophila gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* **153**, 787-98 (1999).
34. Kahn, T.G. et al. Topological screen identifies hundreds of Cp190- and CTCF-dependent Drosophila chromatin insulator elements. *Sci Adv* **9**, eade0090 (2023).
35. Comet, I. et al. PRE-mediated bypass of two Su(Hw) insulators targets PcG proteins to a downstream promoter. *Dev Cell* **11**, 117-24 (2006).
36. Li, X. et al. GAGA-associated factor fosters loop formation in the Drosophila genome. *Mol Cell* (2023).

37. Pollex, T. et al. Chromatin gene-gene loops support the cross-regulation of genes with related function. *Mol Cell* **84**, 822-838 e8 (2024).
38. Papp, B. & Muller, J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev* **20**, 2041-54 (2006).
39. Bowman, S.K. et al. H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex. *Elife* **3**, e02833 (2014).
40. Furlong, E.E.M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341-1345 (2018).
41. Petrascheck, M. et al. DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res* **33**, 3743-50 (2005).
42. Espinola, S.M. et al. Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. *Nat Genet* **53**, 477-486 (2021).
43. Fukaya, T., Lim, B. & Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **166**, 358-368 (2016).
44. Alexander, J.M. et al. Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. *Elife* **8**(2019).
45. Benabdallah, N.S. et al. Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Mol Cell* **76**, 473-484 e7 (2019).
46. Yang, J.H. & Hansen, A.S. Enhancer selectivity in space and time: from enhancer-promoter interactions to promoter activation. *Nat Rev Mol Cell Biol* (2024).
47. Loubiere, V., Papadopoulos, G.L., Szabo, Q., Martinez, A.M. & Cavalli, G. Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. *Sci Adv* **6**, eaax4001 (2020).
48. Silicheva, M. et al. Drosophila mini-white model system: new insights into positive position effects and the role of transcriptional terminators and gypsy insulator in transgene shielding. *Nucleic Acids Res* **38**, 39-47 (2010).
49. Olivares-Chauvet, P. et al. Capturing pairwise and multi-way chromosomal conformations using chromosomal walks. *Nature* **540**, 296-300 (2016).
50. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240-4 (2015).
51. Cardozo Gizzi, A.M. et al. Direct and simultaneous observation of transcription and chromosome architecture in single cells with Hi-M. *Nat Protoc* **15**, 840-876 (2020).
52. Messina, O. et al. 3D chromatin interactions involving Drosophila insulators are infrequent but preferential and arise before TADs and transcription. *Nat Commun* **14**, 6678 (2023).
53. Barho, F. et al. Qudi-HiM: an open-source acquisition software package for highly multiplexed sequential and combinatorial optical imaging. *Open Res Eur* **2**, 46 (2022).
54. Gotz, M., Messina, O., Espinola, S., Fiche, J.B. & Nollmann, M. Multiple parameters shape the 3D chromatin structure of single nuclei at the doc locus in Drosophila. *Nat Commun* **13**, 5375 (2022).
55. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
56. Schmid, M.W., Grob, S. & Grossniklaus, U. HiCdat: a fast and easy-to-use Hi-C data analysis tool. *BMC Bioinformatics* **16**, 277 (2015).
57. Schuettengruber, B. et al. Cooperativity, specificity, and evolutionary stability of Polycomb targeting in Drosophila. *Cell Rep* **9**, 219-233 (2014).
58. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-9 (2012).
59. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-4 (2015).
60. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-5 (2016).

61. Lopez-Delisle, L. et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422-423 (2021).

Methods

Fly work and generation of mutant flies by CRISPR/Cas9 genome engineering

All flies were raised on standard corn meal yeast extract medium at 25°C. CRISPR/Cas9 mutant fly lines Double, ΔPRE1, ΔPRE2 and gypsy 1 are described in [22]. Sequences of gRNAs used to create fly lines gypsy2, gypsy3, ΔRE and gypsy2+ ΔRE are described in **Supplementary Table 2**. Sense and antisense oligonucleotides were annealed and phosphorylated by the T4 polynucleotide kinase (NEB#M0201S) before being inserted inside a pCFD3 plasmid (Addgene #49410) previously digested by BbsI (NEB#R0539S). To create the pHD-dsRED donor plasmid (Addgene) containing a removable (floxed) 3XP3-dsRED construct flanked by loxP sites and DNA fragments having homology to the target regions (homology arms) serving as template for homology-directed repair, 1.5 kb genomic DNA fragments were amplified by PCR (**Supplementary Table 2**) and inserted into the pHD-dsRED plasmid using the GIBSON assemble (kit NEBuilder NEB#E2621S).

The gypsy insulator was amplified from the plasmid (Gy)w(Gy) described in [48] and introduced into the donor plasmid cut by SpeI and BglII using GIBSON cloning **Supplementary Table 2)**. To generate mutant fly lines, gRNA-containing pCFD3 and pHD-dsRED donor plasmids were injected into flies expressing Cas9 in the germline (vas-Cas9(X) RFP-; Bloomington stock #55821). Injections and dsRED screening was performed by BestGene (https://www.thebestgene.com/). To remove the dsRED reporter construct, mutant flies were crossed with a fly line expressing CRE recombinase (Bloomington stock #34516). To generate the gypsy2+ΔRE mutant line gRNAs targeting the RE and corresponding donor plasmid were injected into gypsy2 mutant lines previously generated and expressing Cas9 (vas-Cas9(III)). Coordinates and sequences of deleted regions can be found in **Supplementary Table 3**. Genotypes of mutant fly lines were confirmed by PCR genotyping and sequencing analysis of the mutated region.

Immunostaining experiments

For immunostaining, 3rd instar imaginal leg discs were dissected at room temperature in sterile schneider medium. Pupae were selected at the very beginning of pupation that can be recognized by their white color (a pupal stage that lasts 1h) and were dissected 3.5-4 hours later. The discs were then fixed for 20 min in 4% formaldehyde and were permeabilized during 1 hour in PBS+0.5% triton (for larval leg discs) or 0.8% (for pupal leg discs). The samples were then incubated 1 hour in 3% bovine serum albumin (BSA) PBTr (1xPBS + 0.5% tritonX-100).

DAC primary antibody was diluted 1/400 (DSHB mAbdac1-1) in 1%BSA PBTr and incubated over night at 4°C on a rotating wheel. The leg discs were washed in PBTr before adding the secondary antibody at 1/1000 dilution (Thermo Fisher Scientific #A-31571) and incubated for 2 hours at room temperature on a rotating wheel. Finally, the discs were extensively washed in PBTr. The proximal segments of the leg discs were removed by dissection to only keep the tarsal segments that were subsequently mounted on microscope slides using ProLong Gold reagent (Invitrogen # P36930). The different images were acquired on a Zeiss axioimager Z2 Apoptome Leica SP8 confocal microscope using the same settings for all mutant lines and analyzed using Fiji software.

RNA FISH experiments

RNA FISH probes were preparated with a RNA FISH probes kit (Thermofisher F32956) from DNA probes amplified with the primers described in **Supplementary Table 2**. 3rd instar imaginal leg discs were quickly dissected in Schneider medium. Pupae were selected at the very beginning of pupation that can be recognize by their white color (a pupal stage that lasts 1h) and were dissected 3.5-4 hours later. The discs were then fixed with 4% formaldehyde before being permeabilized with 0.5% TritonX-100 in PBS during 4 hours. Subsequently, discs are incubated 10 minutes with 50%PBT (=PBS +1%Triton) / 50% hybridization solution (=50% formamide, 5X saline-sodium citrate (SSC), 100/ml fragmented salmon testes DNA, 50g/ml heparin, 0,2% Tween-20) at R.T. The samples were incubate 45 minutes and then 1 hour in hybridization solution at 55°C. In parallel, a previously tested optimal concentration of labelled probe was diluted in 50µL hybridization solution, heated 2 minutes at 85°C and chilled on ice in order to denaturate RNA secondary structures. The discs were then incubated overnight with 50µL probe solution at 55°C. The day after, the samples were washed 3 times at 55°C with hybridization solution and twice with PBS1%TritonX-100. The proximal segments of the leg discs are removed by dissection to only keep the tarsal segments that are mounted on microscope slides using ProLong Gold reagent (Invitrogen # P36930). Images were acquired on a Zeiss axioimager Z2 APopoteme Leca SP8 confocal microscope using the same settings for all mutant lines and analyzed using Fiji software.

Hi-C experiments

Hi-C experiments were performed using the EpiTect Hi-C Kit (Quiagene#59971). All Hi-C experiment were performed in two or three independent experiments using 50 3rd instar imaginal leg discs or early pupal discs. Briefly, discs were homogenized and fixed in activated Buffer T and 2% Formaldehyde using Tissue Masher tubes (Biomasher II (EOG-sterilized) 320103 Funakoshi). Tissue was digested by adding 25ul Collagenase I and II (40 mg/ml) for 1h at 37°C. Samples were centrifuged and supernatant was carefully aspirated, leaving ~250

µl of solution in the tube. Then 250ul QIAseq Beads equilibrated to room temperature were added to bind nuclei to the beads and all subsequent reactions were performed on the beads according to the manufactures protocol. Libraries were sequenced at BGI (https://www.bgi.com/) PE 150 (approx. 400 million reads per replicate).

Hi-C analysis

Raw data from Hi-C sequencing were processed by using the "shHiC2" pipeline. Sequencing statistics are summarized in **Supplementary Table 4**. Valid interactions were stored in a database using the "misha" R package (https://github.com/msauria/misha-package). Extracting the valid interactions from the *misha* database, the "shaman" R package [https://bitbucket.org/tanaylab/shaman] has been used for computing the Hi-C expected models, Hi-C scores with parameters *k=250* and *k_exp=500* (**Fig.s 2a**, **3c**, and **4a**), and differential Hi-C interaction scores with parameters *k=250* and *k_exp=250* and per each comparison down-sampling the compared datasets to have the same number of valid-pairs in chr2L (**Fig. 4b**). Specifically, Hi-C scores quantify the contact enrichment (positive values) or depletion (negative values) of each bin of the map with respect to a statistical model used to evaluate the expected number of counts. To generate this expected model, we (randomized) shuffled the observed Hi-C contacts using a Markov Chain Monte Carlo-like approach per chromosome [49]. Shuffling is done such that the marginal coverage and decay of the number of observed contacts with the genomic distance are preserved, but any features of genome organization (*e.g*, TADs or loops) are not. These expected maps were generated for each biological replicate separately and contain twice the number of observed cis-contacts. Next, the score for each contact in the observed contact matrix was calculated using *k* nearest neighbors (kNN) strategy [49]. In brief, the distributions of two-dimensional Euclidean distances between the observed contact and its nearest *k_exp* neighbors in the pooled observed and pooled expected (per cell type) data are compared, using Kolmogorov–Smirnov D statistics to visualize positive (higher density in observed data) and negative (lower density in observed data) enrichments. These D-scores are then used for visualization (100 to +100 scale) and are referred to as Hi-C scores in the text. Accordingly, the color-scale of the Hi-C scores comprises both positive and negative values. When computing the Differential Hi-C scores maps of **Fig. 4b** the reference dataset was used as the expected model.

For each condition, the Hi-C interaction quantifications at the *dac* PRE loop (**Fig. 2b**, **3d**, and **4e**) were performed by considering the Hi-C scores between two regions of 6 kb, chr2L:16,419,514-16,425,515bp and chr2L:16,482,929-16,488,930bp), each including the PRE1 and PRE2, respectively (**Supplementary Table 3**). The distributions of Hi-C scores (**Fig. 2b**, **3d**, and **4e**) are represented as boxplots showing: central line, median; box limits, 75th and

25th percentiles; whiskers, 1.5×interquartile range. Each of the comparisons of the Hi-C interaction quantifications at the *dac* PRE loop was performed between a reference conditions - *Embryo* in **Fig. 2b** and *larvae WT* in **Fig. 3d** and **4d**– and each of the other conditions present in the same Figure. An unpaired two-sided Wilcoxon statistical test (H0: true median shift is equal to 0. The two variables are not normally distributed) was used to estimate the reported *p-values*. The annotation of the Polycomb-associated TADs in chr2L in [5] was used to compute the number of Hi-C interactions intra-PcG-TAD, which were then normalized by the total number of valid-pairs at the corresponding developmental stage (embryo, larvae, or pupae). The distributions of these interaction frequencies are shown in the violin- and boxplots of **Fig. 2b** as the Log2 ratios of Embryo over the Larval and Pupal leg discs. The boxplots show: central line, median; box limits, 75th and 25th percentiles; whiskers, 1.5×interquartile range. An unpaired two-sided Wilcoxon statistical test (H0: true median shift is equal to 0. The two variables are not normally distributed) was used to estimate the reported *p-values*. The insulation scores [50] were computed on the observed Hi-C datasets binned at 2kb resolution with windows of 100, 150, 200, 250, and 300kb resulting in five valued per bin and were stored in the *misha* database using an *in-house* R script. The mean and standard deviation per each of the 2kb-bins were computed were used for the plots in **Fig. 3e** and **4c**. The quantification of the insulation scores (IS) at gypsy insertions and R0-12 regions was performed applying the pair-wise statistical comparison of the five IS quantifications per 2kb-bins. The *p-values* in **Fig. 4d** and **Ext. Data Fig. 4a** resulted from a Welch t-test (H0: true difference in means is equal to 0. The variances of the samples are thought not to be equal) between the WT condition and each of the gypsy mutant at the corresponding locus. All plots of Hi-C maps (**Fig. 2a**, **3c**, **4a**, and **4d**), Hi-C interaction scores comparisons (**Fig. 2b**, **3d**, and **4e**), insulation score (IS) profiles (**Fig. 3e** and **4c**), p-values of IS comparisons (**Fig. 4d**) were obtained with *in-house* R scripts (see Code Availability statement).

Hi-M library preparation

The oligopaint library covering the dac region consists of 52-mer sequences with genome homology ordered from CustomArray. These sequences were obtained from the oligopaint public database (http://genetics.med.harvard.edu/oligopaints). From the initial design of the library, we selected 20-mers with an average probe density of 9-17 probes/kb. Each barcode contains 45 probes covering in average 3,8 kb (**Supplementary Table 1**). Each oligo is composed of 5 different regions: (1) a 21 nt forward universal priming region for library amplification, (2) two 20 nt readout regions separated by an A for barcoding, (3) a 42nt genome homology region, (4) a duplication of one 20 nt readout region and (5) a 21 nt reverse universal priming region.

The procedure for oligopaint library amplification was previously described [26,42,51,52]. It consists of seven steps : (1) an emulsion PCR (emPCR) to extract the *dac* library from the oligonucleotide pool using universal primers; (2) a limited-cycle PCR performed on the emPCR to identify the most efficient amplification cycle; (3) a large scale PCR with a T7 promoter on the reverse primer; (4) an in vitro T7 transcription; (5) a reverse transcription to transform RNAs into single-stranded DNA (ssDNA); (6) an alkaline hydrolysis for the removal of the intermediate RNA; and (7) a ssDNA purification and concentration.

Each barcode is unique and specific to an adapter oligo. The adapter oligo serves as a bridge between the readout region and an Alexa Fluor-647-labeled secondary oligonucleotide. The fluorescently labeled part of the secondary probe is attached via a disulfide leakage that can be cleaved (chemical bleaching) during the sequential imaging of FISH probes [51]. For the fiducial, we used an adapter oligo complementary to the reverse primer of the library and specific to a secondary probe bound to a non-cleavable Rhodamine Red fluorophore. Adapters and fluorescently labeled secondary probes were synthesized and purchased from Integrated DNA Technologies (IDT).

Hi-M library hybridization

Pupae were collected at the beginning of pupation (white pupae) and dissected 3.5 - 4h later. The dissected leg discs were fixed with 4% Formaldehyde before being permeabilized with 0.5% Triton in PBS during 4 hours. The discs were then progressively washed in four different concentrations of Triton/pHM (pHM = 2X SSC, NaH 2PO 4 0.1 M pH = 7): 20%, 50%, 80% and 100% pHM for 20 min in each buffer at RT on a rotating wheel. Then, the discs were incubated overnight in 225 pmols of the library diluted in 30µL of FISH Hybridization Buffer (FHB = 50% Formamide, 2X SSC, Salmon Sperm DNA 0.5 mg mL-1, 10% dextran sulfate). The probes and the discs in pHM were heated at 80°C. The incubation of the leg discs in the FHB + probe buffer was performed in a PCR machine from 80°C to 37°C with a temperature decrease of 1°C every 10 minutes. The next day, discs were washed 2 times with 50% formamide, 2x SSC, 0.3% CHAPS and sequentially washed with four different concentrations of formamide/PBT: 40%, 30%, 20%, 10% formamide during 20 min per buffer on a rotating wheel. Finally, the discs were washed with PBS/1%TritonX-100, fixed with 4% formaldehyde in PBS, washed with PBS and stored at 4°C.

Hi-M imaging system

Hi-M experiments were performed with a homemade widefield and epifluorescence microscope. This setup includes a Rapid Automated Modular Microscope (RAMM) (Applied Scientific Instrumentation) coupled with a microfluidic device previously described [26,42]. The microscope and fluidics system were controlled using Qudi-HiM (our homemade hardware

control package) [53]. The fluidics system permitted the automated and sequential hybridizations of the probes. The solutions were delivered to the sample by a combination of three eight-way valves (HVXM 8-5, Hamilton), a negative pressure pump (MFCS-EZ, Fluigent) and a FCS2 flow chamber (Bioptechs). The excitation was performed by three different lasers : 405 nm (Obis 405, 100 mW, Coherent), 561 nm (Sapphire 561 LP, 150 mW, Coherent) and 642 nm (VFL-0-1000-642-OEM1, 1W, MPB communications Inc.). The fluorescence was collected through a Nikon APO x60 1.2 NA water immersion objective lens mounted on a closed-loop piezoelectric stage (Nano-F100, Mad City Labs Inc.). Images were acquired using a sCMOS camera (ORCA Flash 4.0 V3, Hamamatsu) with an effective optical pixel size of 106 nm. To correct axial drift in real time, we used a homemade autofocus system composed of a 785 nm laser (OBIS 785, 100 mW, Coherent) and an infrared sensitive camera (CMOS - DCC1545M, Thorlabs).

Acquisition of Hi-M datasets

The proximal part of the pupal leg discs was removed by dissection in order to only keep the tarsal segments. About 15-20 tarsal segments were aligned on a 2% agar:PBS pad, then attached onto a 40 mm round coverslip previously functionalized with trimethoxysilane and 10% poly-L-lysine. The slide was then mounted onto the flow chamber. Pupal leg discs were first incubated with the fiducial adapter (25 nM of the adapter specific to the reverse primer, 2x SSC, 40% v:v formamide) for 20 min and then washed with a washing buffer solution (2x SSC, 40% v:v formamide) for 10 min. To complete the hybridization of the fiducial, we did a second round of incubation with the appropriate secondary oligo (25 nM of Rhodamine-red labeled probe, 2x SSC, 40% v:v formamide) for 20 min and washed again for 10 min with the washing buffer solution. After a 5 min wash with 2x SSC, we proceeded with nuclei staining with 0.5 µg mL-1 of DAPI in PBS for 20 min. After another 5 min wash with 2x SSC, the imaging buffer (1× PBS, 5% Glucose, 0.5 mg/ml glucose oxidase and 0.05 mg/ml catalase) was injected to limit fiducial photobleaching during the acquisition. An image stack (200µm x 200µm region of interest or ROI) was acquired for each of the 10-15 pupal leg discs. The DAPI and the fiducial were sequentially imaged (using 405 nm and 561 nm lasers) with a z-step size of 250 nm for a total range of 17,5 µm.

Next, adapter oligos and secondary probe were sequentially hybridized, acquired and photobleached to image the whole *dac* oligopaint library. The following steps were performed for each of the 22 barcodes: (1) adapter (40 nM of adapter oligonucleotide, 2x SSC, 40% v:v formamide) injection and incubation for 10 min; (2) imaging probe (40 nM secondary probe, 2x SSC, 40% v:v formamide) injection and incubation for 10 min; (3) 10 min wash with washing buffer solution; (4) 5 min wash with 2x SSC; (5) imaging buffer injection and sequential acquisition of fiducial and barcode with 561 and 642 nm lasers; (6) Chemical bleaching (2 ×

SCC, 50 mM tris(2-carboxyethyl)phosphine (TCEP)) of the imaging probe; (7) 5 min wash with 2x SSC before a new cycle of hybridization.

Image processing and Hi-M analysis

Raw TIFF images were deconvolved using Huygens Professional 21.04 (Scientific Volume Imaging, https://svi.nl). Hi-M analysis was performed using pyHiM, a homemade analysis pipeline (https://pyhim.readthedocs.io/en/latest/) (https://www.biorxiv.org/content/10.1101/2023.09.19.558412v1), as previously described [54]. First, images were z-projected by applying either sum for DAPI channels or maximum intensity for the barcodes and fiducial. For each cycle of hybridization, fiducial images were used to register the corresponding barcode image using global and local registration methods. Barcodes and fiducials were segmented in 3D using a neural network, followed by 3D localisation of the center of each barcode mask (https://www.biorxiv.org/content/10.1101/2023.09.19.558412v1). The fiducial oligo binds to the universal priming regions thus labeling the entire *dac* locus. Therefore, we built chromatin traces by combining the DNA-FISH spots colocalizing within single fiducial masks. DAPI images were used to manually segment the different tarsal segments (TS1, TS2, TS3, TS4 or TS5). Pairwise distances (PWD) matrices were calculated for each single chromatin trace. From a list of pairwise distance maps, we calculated the proximity frequencies as the number of chromatin traces in which pairwise distances were within 250 nm, normalized by the number of chromatin traces containing both barcodes. Hi-M maps of the WT condition were generated from 51 622 total traces from 48 pupal leg discs from two independent biological replicates. Hi-M maps of the Gypsy 2 mutant were produced from 63 458 total traces of 51 pupal leg discs from two independent biological replicates. Hi-M matrices were generated for all the tarsal segments combined (TS1, TS2, TS3, TS4 and TS5), TS1, TS2 and by combining TS3 and TS4 (TS3/4). Each trace contains at least 12% of the barcodes. Virtual 4C figures were obtained by plotting the pairwise distances between the anchored barcode or viewpoint with the rest of the barcodes of an Hi-M matrix.

Wilcoxon two-sided rank tests between the pairwise distance distributions of the barcodes containing RE and the *dac* promoter were performed to test the hypothesis that two independent samples (e.g. WT and gypsy mutant) were drawn from the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).

We estimated the error in the measurement of the median RE-*dac* promoter distance by performing bootstrapping analysis. For this, we performed 1000 bootstrapping cycles drawn from the experimental distribution of pairwise distances to estimate the standard deviation in the determination of the median distance. The errors were between 8 and 25 nm for the WT condition.

4C-seq experiments

For 4C either about 3000 embryos were collected or 300 3rd instar imaginal leg discs were dissected and homogenized and fixed in 2% formaldehyde diluted in nuclear permeabilization (NP) buffer: 15 mM HEPES, pH 7.6, 60 mM KCl, 15 mM NaCl, 4 mM MgCl2, 0.1% Triton-X100, 0.5 mM DTT, 1x protease inhibitors (Roche complete EDTA-free tablets; 11 873 580 001) during 10 minutes at RT. Fixation was stopped by adding 2M glycine during 5 minutes. The samples were then washed once in NP buffer and twice in 1.25xNEB3 buffer and the pellet of fixed cells is frozen in liquid nitrogen and conserved at -80°C.

The chromatin pellet was then layered with 500ul 1.25x DpnII buffer without resuspension and centrifuged. Pellet was resuspended in 250ul 1.25x DpnII buffer. 10ul of 10%SDS was added and incubated for 20min at 65°C and 40min at 37°C. Chromatin was then split into 250ul 1.25x DpnII buffer aliquots of 5-6 10^6 cells and incubated for 1h at 37°C with 3.3% TritonX (final concentration). Samples where the digested with 500U DpnII o/n. The day after, DpnII enzyme was inactivated by heating the samples at 65°C during 20 minutes. The fragments were then ligated during 5 hours at 16°C with T4 ligase 2000U/µL and digested overnight with proteinase K at 65°C. The day after, RNA was degraded by RNAse A solution during 1 hour at 37°C. DNA was purified with Ampure beads without size selection and digested overnight with NlaIII enzyme. The next day, the DNA fragments were circularized by overnight ligation with T4 ligase 2000U/µL in a large buffer volume. Finally, circularized DNA was purified by Ampure beads without size selection. 4C PCR was performed with the primers described in **Supplementary Table 2**. The amplified DNA was purified with Ampure beads. The sequencing libraries were produced with an illumina kit (illumina 20015964). Sequencing (paired-end sequencing 150bp, approx. 4Gb/sample) was performed by Novogene (https://en.novogene.com/).

4C-seq processing and analysis

Using a custom-made python script, fastq sequencing files were split using 4C primer sequences to obtain individual fastq files only containing reads from a single viewpoint per genotype and tissue type. Thereby, the reads were trimmed to remove viewpoint sequences up to the restriction sites. Subsequently, the trimmed reads were aligned against the DM6 reference assembly using bowtie [55] with the parameters -a -v 0 -m 1 (no mismatches and no multiple alignments allowed). The number of successfully aligned reads can be found in **Supplementary Table 5**. The aligned reads were mapped to restriction fragments and genomic bins of 1kb size using HiCdat [56] to obtain tabular files describing the number of reads (i.e. contact frequencies) for a given fragment or genomic bin, respectively. All subsequent analysis steps have been conducted using R. Depending on the 4C samples genotypes and

viewpoints, contact frequencies arising from the viewpoint (±4 bins) and contact frequencies mapping to genotype-specific deletions have been masked by setting them to zero (**Supplementary Table 3**). Then, data from individual samples have been normalized for differing overall library size (counts per million).

To analyze differences between different genotypes, t-test using triplicate data per genotype have been performed for each 1kb genomic bin along the region of interest (chr2L:16,300,00-16,600,000). No multiple testing correction has been performed. Subsequently, the differences of the average of triplicates have been plotted and genomic bins that exhibited p-values < 0.1 have been highlighted.

qRT-PCR experiments

Embryos were collected in a 16h20h developmental time window. 3rd instar imaginal leg discs or early pupal leg discs (3.5-4 hours after pupation) were quickly dissected (<30 min) in Schneider medium and transferred into Trizol. RNA was extracted by using Trizol reagent and purified by RNA clean and concentrator kit (Zymo Research # R1015) following instruction and using the DNAse I from Quiagen (Quiagen#79254). 250 ng of purified RNA was used for the reverse transcription using Maxima First Strand cDNA synthesis Kit for RT-qPCR with dsDNase (Thermos cientific #K1671) following the manufacturer's recommendations. Finally, quantification of the reverse transcription product was performed on LightCycler480 (Roche) with the primers listed in **Supplementary Table 2**. Data analysis were performed on Light cycler software. Expression levels were normalized to the housekeeping gene RP49.

qCHIP experiments

qChIP experiments were performed as described in [57] with minor modifications. Chromatin was sonicated using a Bioruptor Pico (Diagenode) for 7 min (30sec in, 30sec off). Su(HW) antibody was diluted 1/100 for the IP. After decrosslinking, DNA was purified using MicroChIP DiaPure columns from Diagenode. Enrichment of DNA fragment was analyzed by real-time PCR Light-cycler 480 (Roche). Primers used are indicated in **Supplementary Table 2**.

CUT&RUN experiments

CUT&RUN experiments were performed as described by Kami Ahmad in protocilas.io (https://dx.doi.org/10.17504/protocols.io.umfeu3n) with minor modifications. 50 eye discs were dissected in Schneider medium, centrifuged for 3 min at 700g and washed twice with wash+ buffer before addition of Concanavalin A-coated beads. MNase digestion (pAG-MNase Enzyme from Cell Signaling) was performed for 30 min on ice. After ProteinaseK digestion, DNA was recovered using SPRIselect beads and eluted in 50ul TE. DNA libraries for sequencing were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina.

Sequencing (paired-end sequencing 150bp, approx. 2Gb/sample) was performed by Novogene (https://en.novogene.com/). H3K27me3 antibody (Active motif, 39155) was diluted 1:100. IgG antibody (1:100, Cell Signaling Technology, 2729S) was used as control.

CUT&RUN analysis

The quality of the reads was assessed using FastQC. Fastq files were aligned to the D. melanogaster reference genome dm6 using Bowtie 2 (v 2.4.2) [58] with the following parameters: --local ---very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700. SAM files were compressed into BAM files using samtools (v 1.16.1) and reads with low mapping quality (Phred score <30) were discarded. Duplicate reads were removed using sambamba markdup (v 1.0.0) [59] with the following parameters: -r --hash-table-size 500000 --overflow-list-size 500000. For visualization, replicates were merged using samtools merge with default parameters and reads per kilo base per million mapped reads (RPKM)-normalized bigWig binary files were generated using the bamCoverage (v 3.5.5) function from deepTools2 [60] with the following parameters: --normalizeUsing RPKM --ignoreDuplicates -e 0 -bs 10. Genome browser plots were generated using the pyGenomeTracks package (v 3.8) [61]. The 131 Drosophila Polycomb domains [22] were used for the differential enrichment analysis using the DESeq2 method from the "DiffBind" R package (v 3.12.0). Differential quantification results of H3K27me3 levels within Polycomb domains are summarized in **Supplementary Table 6**.

Data Availability

The accession number for the sequencing data reported in this paper is GEO: GSE247377

Code Availability

All original code was deposited on GitHub (https://github.com/cavallifly/Denaud_et_al_NatStructMolBiol_2024) and is publicly available as of the date of publication. Any additional information required to reanalyze the data reported in this paper is available from the corresponding author upon request.

Methods-only Reference

48. Silicheva, M. et al. Drosophila mini-white model system: new insights into positive position effects and the role of transcriptional terminators and gypsy insulator in transgene shielding. *Nucleic Acids Res* **38**, 39-47 (2010).
49. Olivares-Chauvet, P. et al. Capturing pairwise and multi-way chromosomal conformations using chromosomal walks. *Nature* **540**, 296-300 (2016).
50. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240-4 (2015).

51. Cardozo Gizzi, A.M. et al. Direct and simultaneous observation of transcription and chromosome architecture in single cells with Hi-M. *Nat Protoc* **15**, 840-876 (2020).
52. Messina, O. et al. 3D chromatin interactions involving Drosophila insulators are infrequent but preferential and arise before TADs and transcription. *Nat Commun* **14**, 6678 (2023).
53. Barho, F. et al. Qudi-HiM: an open-source acquisition software package for highly multiplexed sequential and combinatorial optical imaging. *Open Res Eur* **2**, 46 (2022).
54. Gotz, M., Messina, O., Espinola, S., Fiche, J.B. & Nollmann, M. Multiple parameters shape the 3D chromatin structure of single nuclei at the doc locus in Drosophila. *Nat Commun* **13**, 5375 (2022).
55. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
56. Schmid, M.W., Grob, S. & Grossniklaus, U. HiCdat: a fast and easy-to-use Hi-C data analysis tool. *BMC Bioinformatics* **16**, 277 (2015).
57. Schuettengruber, B. et al. Cooperativity, specificity, and evolutionary stability of Polycomb targeting in Drosophila. *Cell Rep* **9**, 219-233 (2014).
58. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-9 (2012).
59. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-4 (2015).
60. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-5 (2016).
61. Lopez-Delisle, L. et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422-423 (2021).

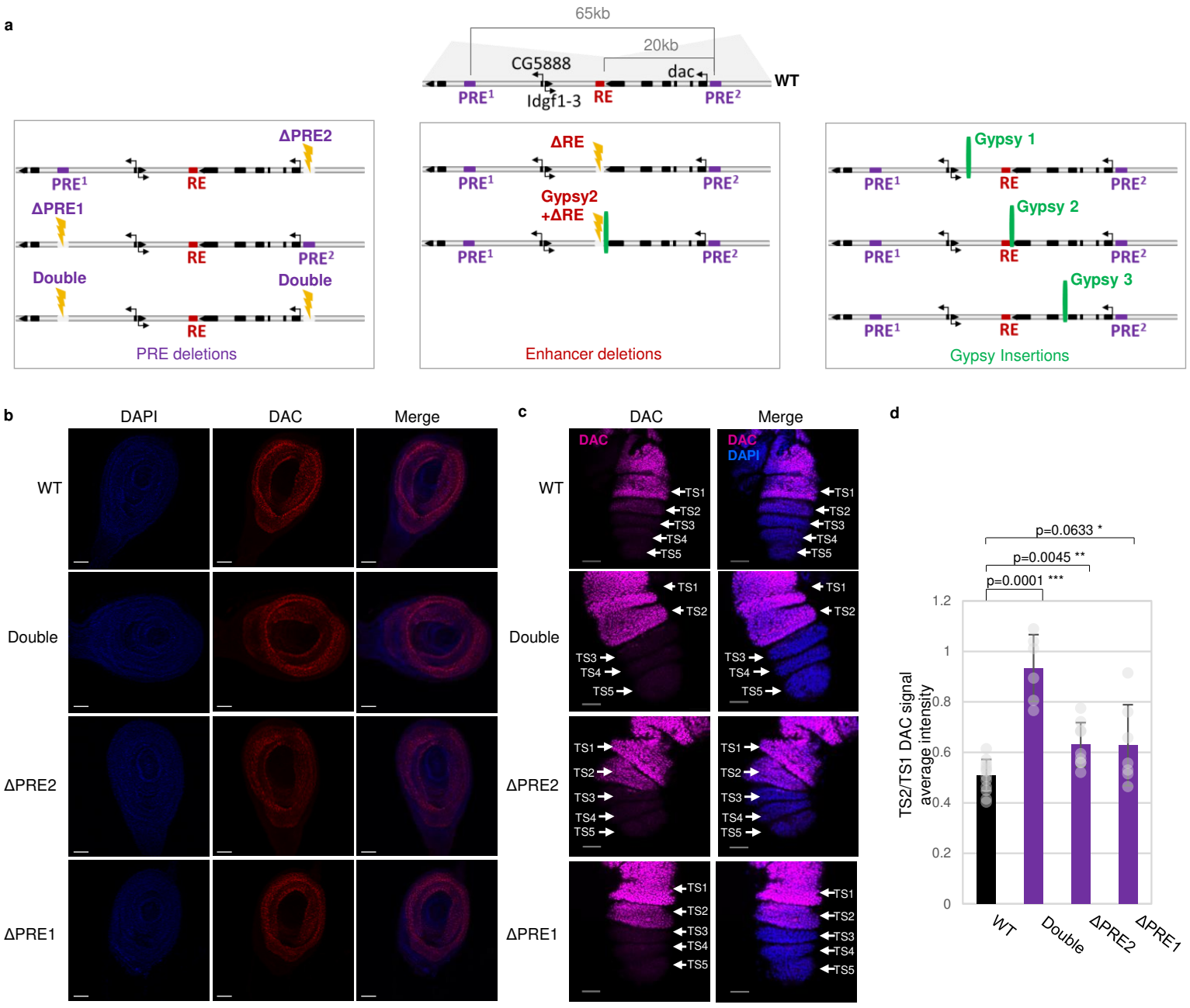

a
65kb
20kb
CG5888
dac
WT
PRE1
Idgf1-3
RE
PRE2
ΔPRE2
ΔPRE1
Double
PRE deletions
ΔRE
Gypsy2
+ΔRE
Enhancer deletions
Gypsy 1
Gypsy 2
Gypsy 3
Gypsy Insertions
b
DAPI
DAC
Merge
WT
Double
ΔPRE2
ΔPRE1
c
DAC
Merge
DAC
DAPI
TS1
TS2
TS3
TS4
TS5
d
p=0.0633 *
p=0.0045 **
p=0.0001 ***
TS2/TS1 DAC signal average intensity
1.2
1
0.8
0.6
0.4
0.2
0
WT
Double
ΔPRE2
ΔPRE1

Figure 1

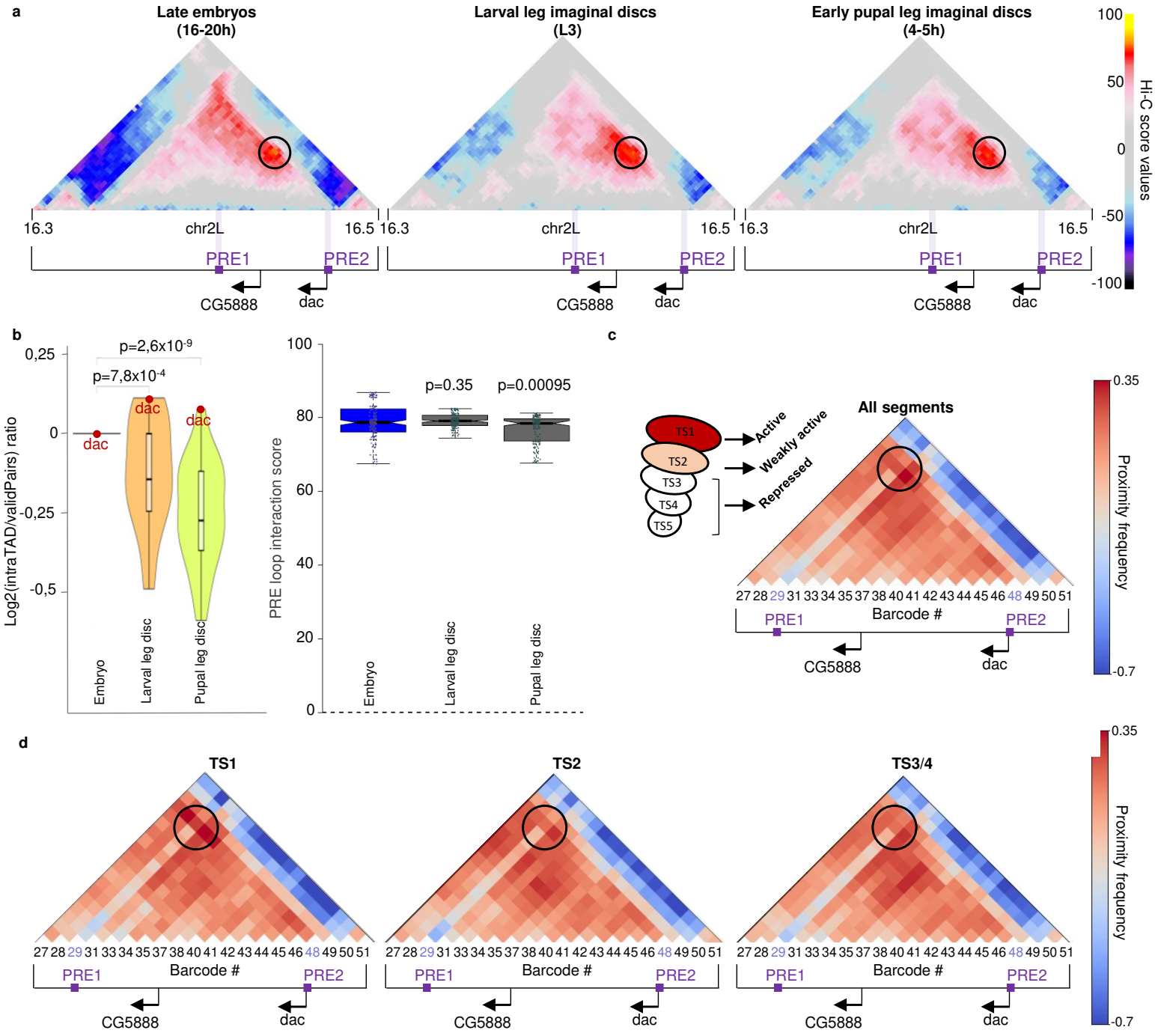

a
Late embryos
(16-20h)
Larval leg imaginal discs
(L3)
Early pupal leg imaginal discs
(4-5h)
Hi-C score values
100
50
0
-50
-100
16.3
chr2L
16.5
PRE1
PRE2
CG5888
dac
b
p=2,6x10-9
p=7,8x10-4
Log2(intraTAD/validPairs) ratio
0,25
0
-0,25
-0,5
Embryo
Larval leg disc
Pupal leg disc
PRE loop interaction score
p=0.35
p=0.00095
c
TS1
TS2
TS3
TS4
TS5
Active
Weakly active
Repressed
All segments
27 28 29 31 33 34 35 37 38 40 41 42 43 44 45 46 48 49 50 51
Barcode #
Proximity frequency
0.35
-0.7
d
TS1
TS2
TS3/4

Figure 2

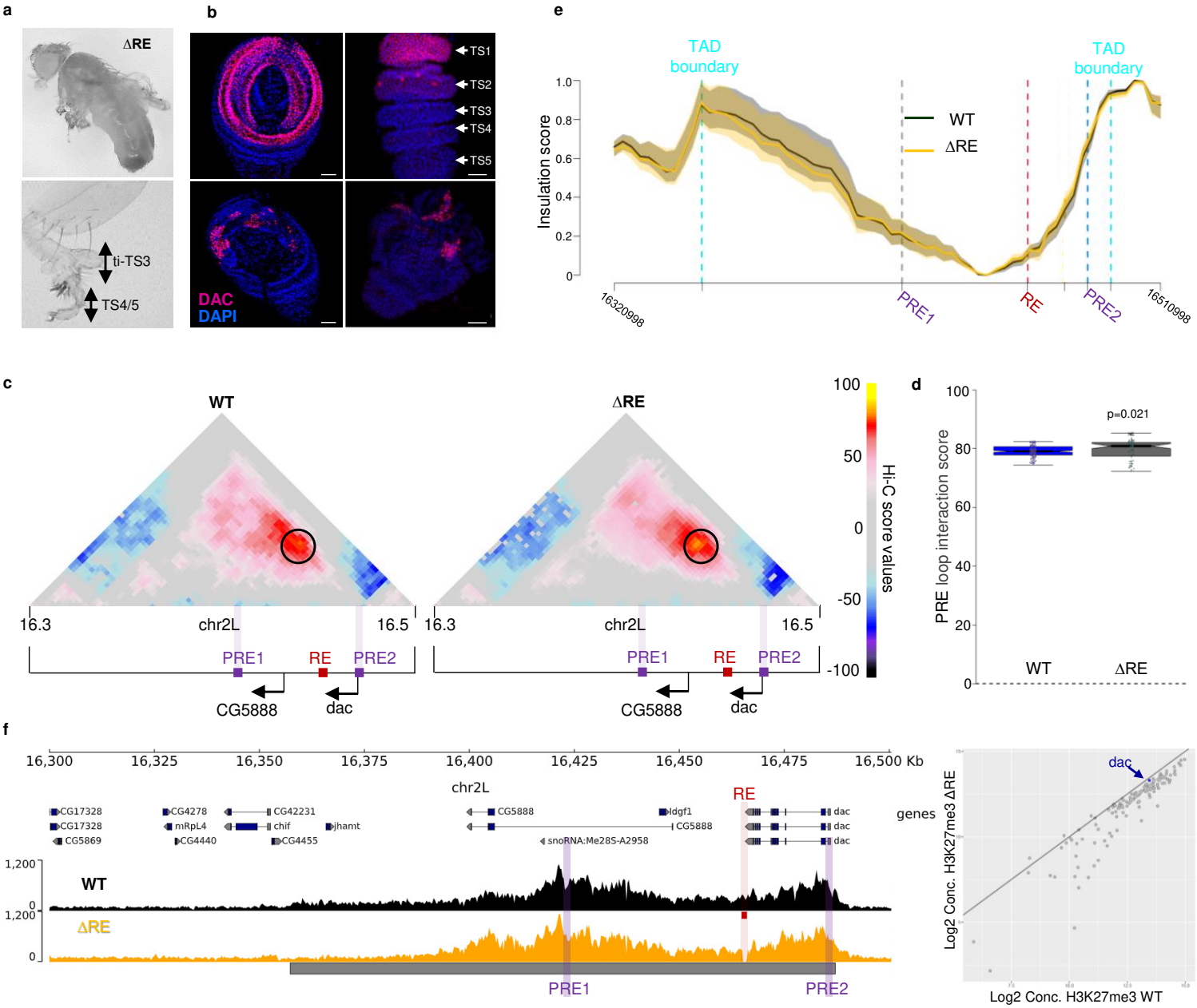

a
ΔRE
ti-TS3
TS4/5
b
TS1
TS2
TS3
TS4
TS5
DAC
DAPI
e
TAD boundary
TAD boundary
Insulation score
1.0
0.8
0.6
0.4
0.2
0
WT
ΔRE
16320998
PRE1
RE
PRE2
16510998
c
WT
ΔRE
100
50
0
-50
-100
Hi-C score values
16.3
chr2L
16.5
PRE1
RE
PRE2
CG5888
dac
d
PRE loop interaction score
p=0.021
WT
ΔRE
f
16,300
16,325
16,350
16,375
16,400
16,425
16,450
16,475
16,500 Kb
chr2L
RE
genes
CG17328
CG17328
CG5869
CG4278
mRpL4
CG4440
CG42231
chif
CG4455
jhamt
CG5888
CG5888
Idgf1
snoRNA:Me28S-A2958
dac
dac
dac
1,200
0
1,200
0
WT
ΔRE
PRE1
PRE2
dac
Log2 Conc. H3K27me3 ΔRE
Log2 Conc. H3K27me3 WT

Figure 3

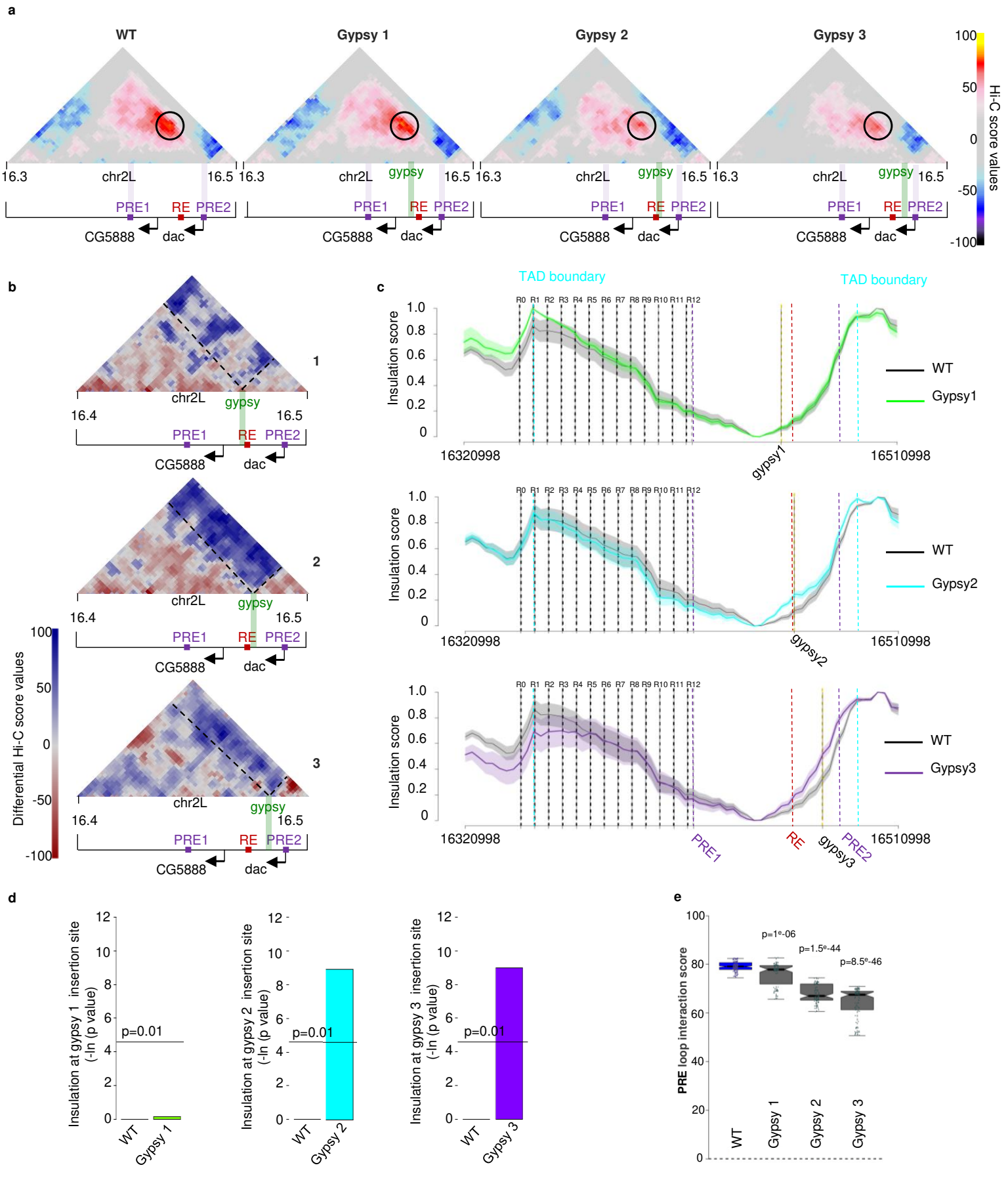

a
WT
Gypsy 1
Gypsy 2
Gypsy 3
Hi-C score values
100
50
0
-50
-100
16.3
chr2L
16.5
gypsy
PRE1
RE
PRE2
CG5888
dac
b
1
2
3
chr2L
gypsy
16.4
16.5
Differential Hi-C score values
c
TAD boundary
TAD boundary
Insulation score
WT
Gypsy1
Gypsy2
Gypsy3
16320998
16510998
gypsy1
gypsy2
gypsy3
d
Insulation at gypsy 1 insertion site (-ln (p value))
Insulation at gypsy 2 insertion site (-ln (p value))
Insulation at gypsy 3 insertion site (-ln (p value))
p=0.01
WT
Gypsy 1
Gypsy 2
Gypsy 3
e
PRE loop interaction score
p=1e-06
p=1.5e-44
p=8.5e-46

Figure 4

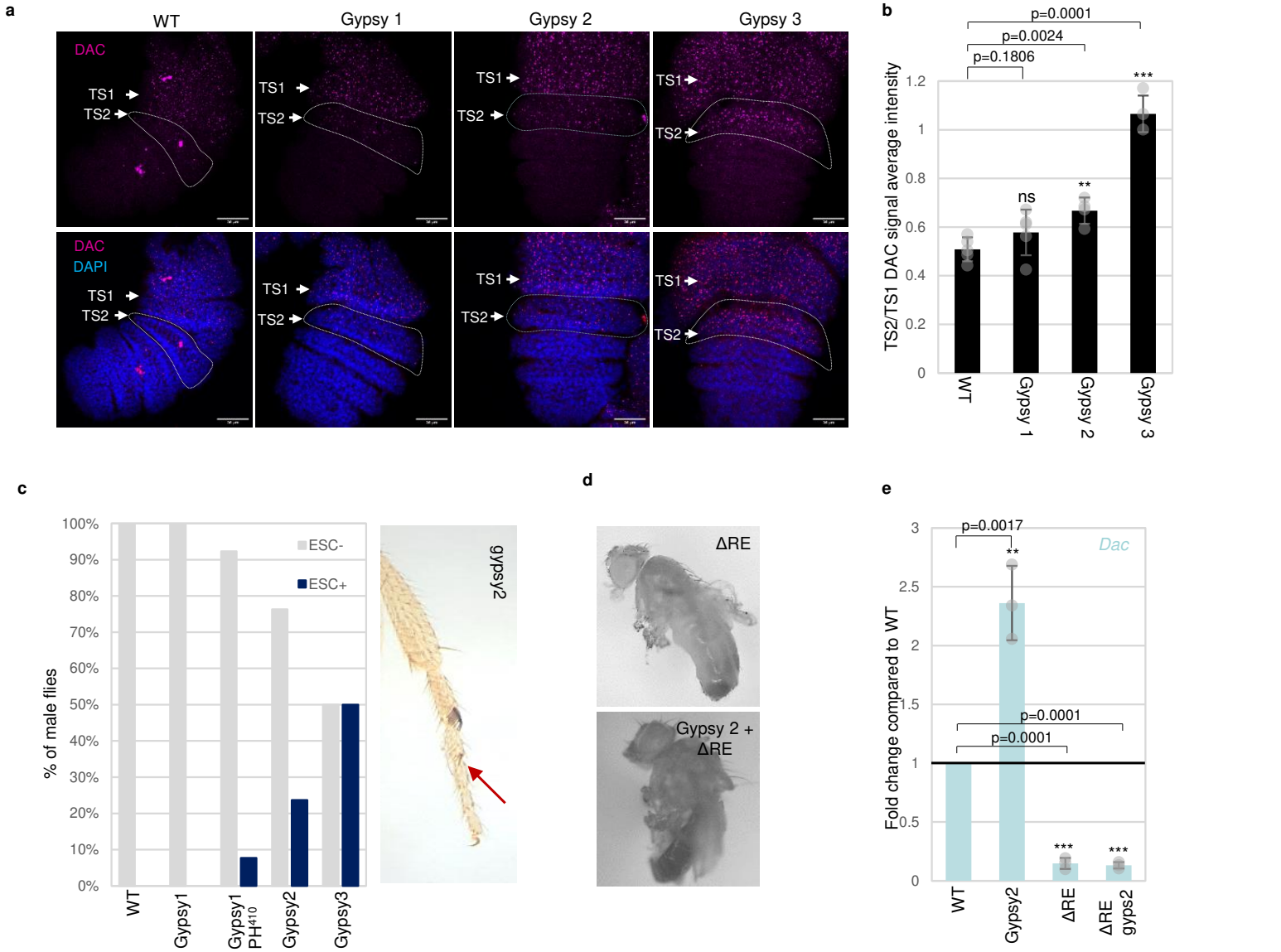

a
WT
Gypsy 1
Gypsy 2
Gypsy 3
DAC
DAPI
TS1
TS2
b
p=0.0001
p=0.0024
p=0.1806
TS2/TS1 DAC signal average intensity
ns
**
***
WT
Gypsy 1
Gypsy 2
Gypsy 3
c
% of male flies
ESC-
ESC+
WT
Gypsy1
Gypsy1 PH410
Gypsy2
Gypsy3
gypsy2
d
ΔRE
Gypsy 2 + ΔRE
e
Dac
Fold change compared to WT
p=0.0017
p=0.0001
p=0.0001
WT
Gypsy2
ΔRE
ΔRE gyps2

Figure 5

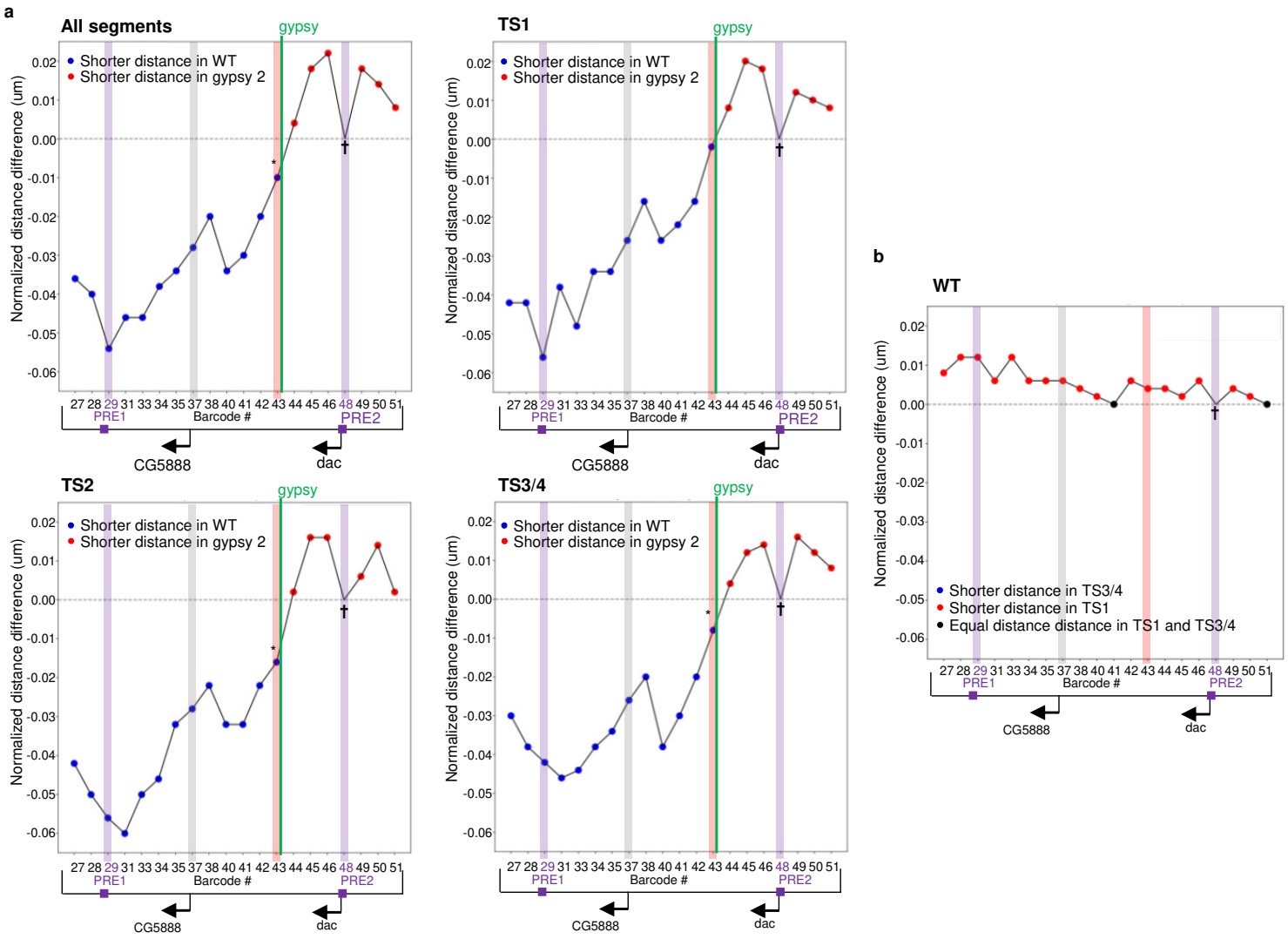

a
All segments
gypsy
Shorter distance in WT
Shorter distance in gypsy 2
Normalized distance difference (um)
0.02
0.01
0.00
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
27 28 29 31 33 34 35 37 38 40 41 42 43 44 45 46 48 49 50 51
PRE1
Barcode #
PRE2
CG5888
dac
TS1
TS2
TS3/4
b
WT
Shorter distance in TS3/4
Shorter distance in TS1
Equal distance distance in TS1 and TS3/4

Figure 6

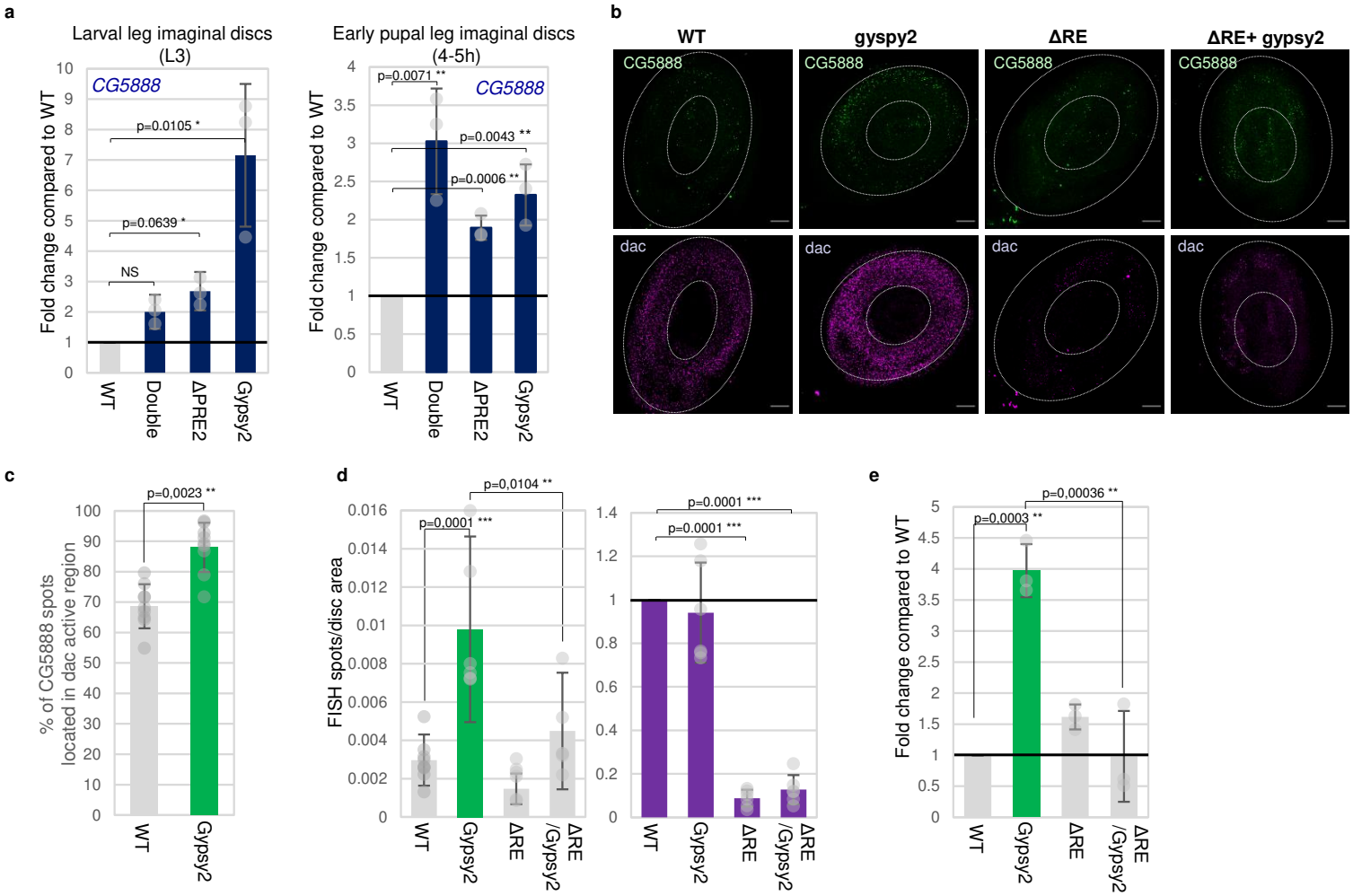
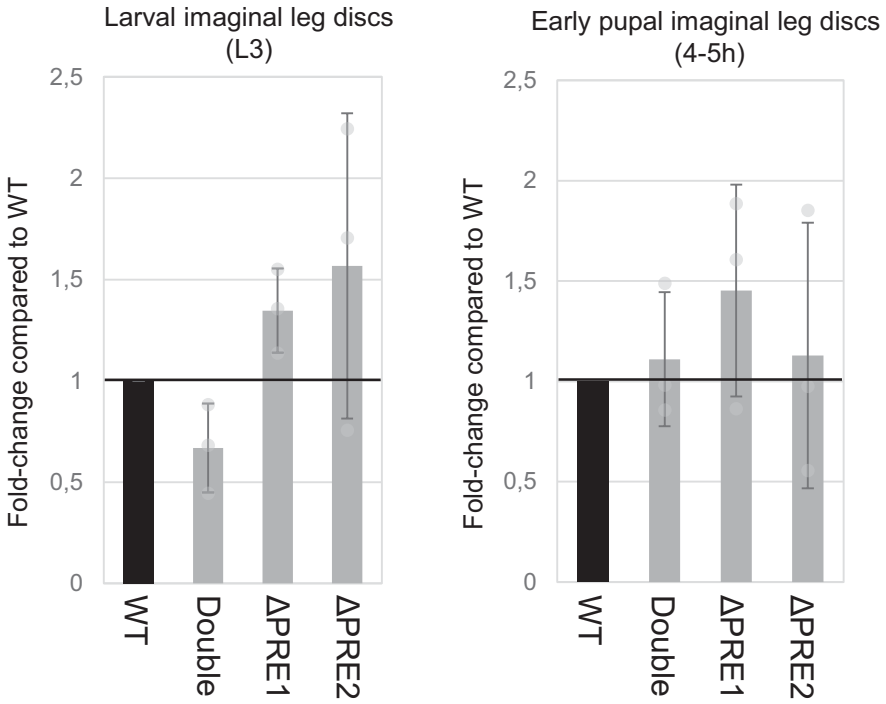
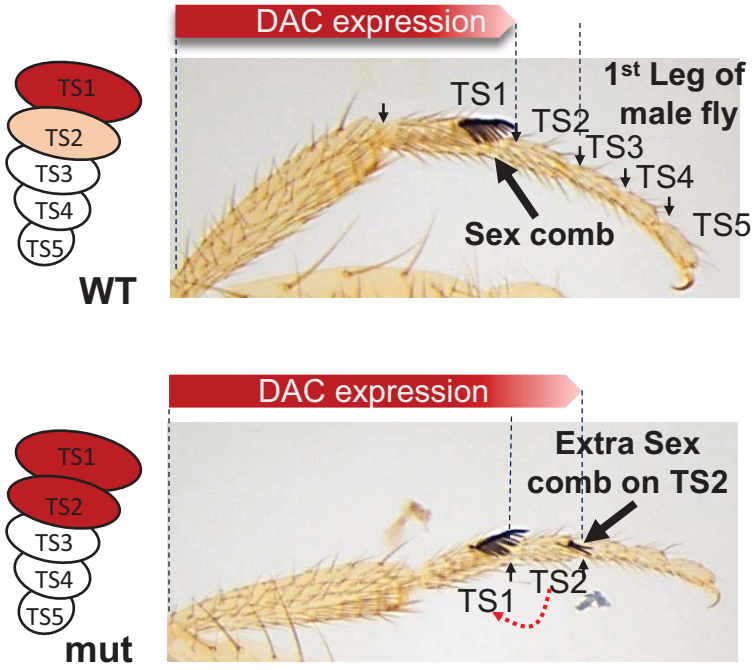

a
Larval leg imaginal discs (L3)
CG5888
Fold change compared to WT
p=0.0105 *
p=0.0639 *
NS
WT
Double
ΔPRE2
Gypsy2
Early pupal leg imaginal discs (4-5h)
CG5888
p=0.0071 **
p=0.0043 **
p=0.0006 **
b
WT
gyspy2
ΔRE
ΔRE+ gypsy2
CG5888
dac
c
% of CG5888 spots located in dac active region
p=0,0023 **
d
FISH spots/disc area
p=0,0104 **
p=0.0001 ***
p=0.0001 ***
p=0.0001 ***
ΔRE /Gypsy2
e
Fold change compared to WT
p=0,00036 **
p=0,0003 **

Figure 7

a

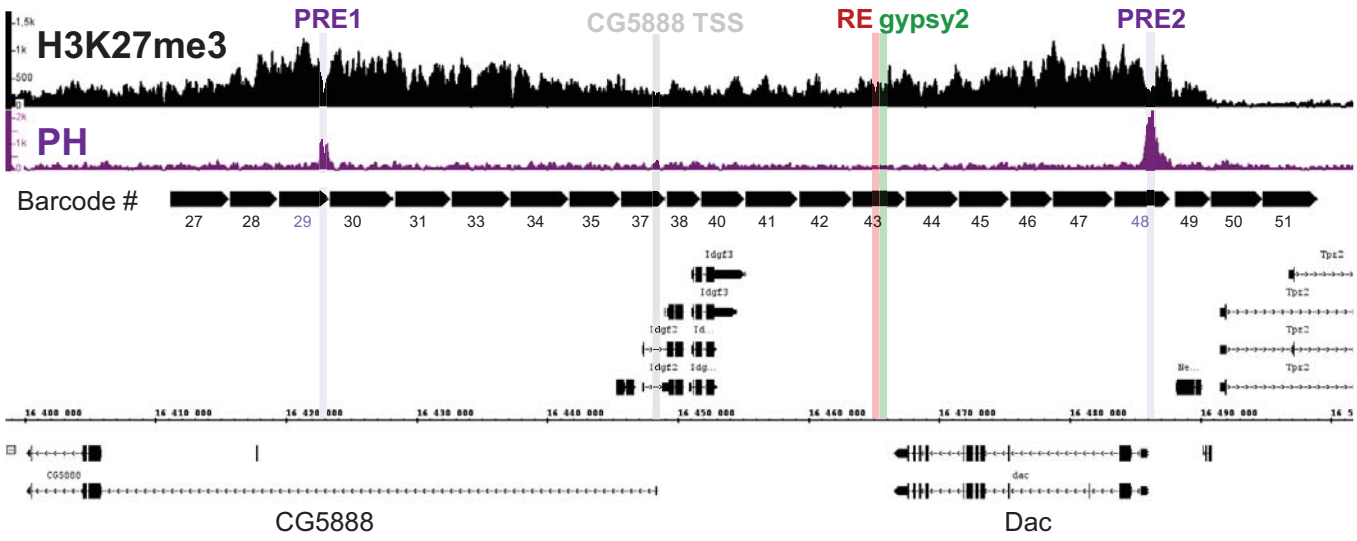


b

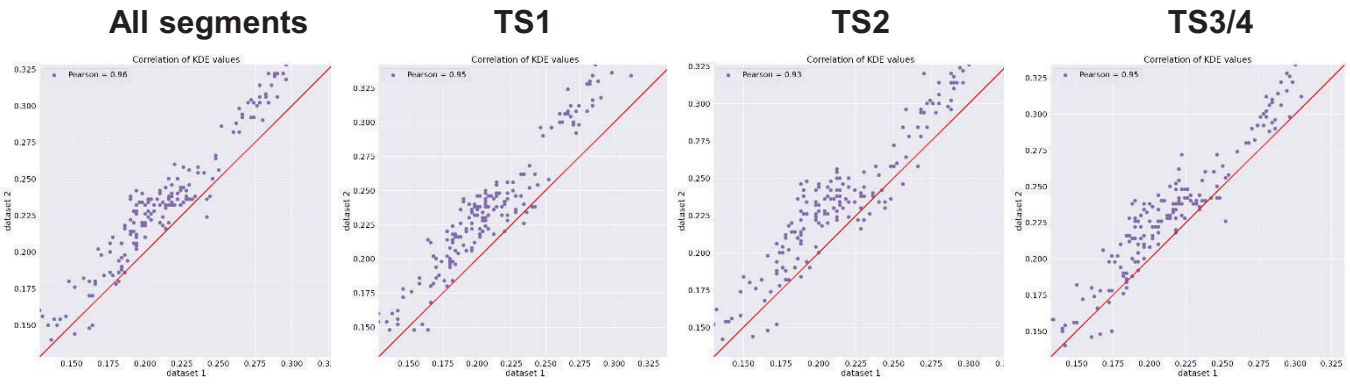


Extended Data Figure 1

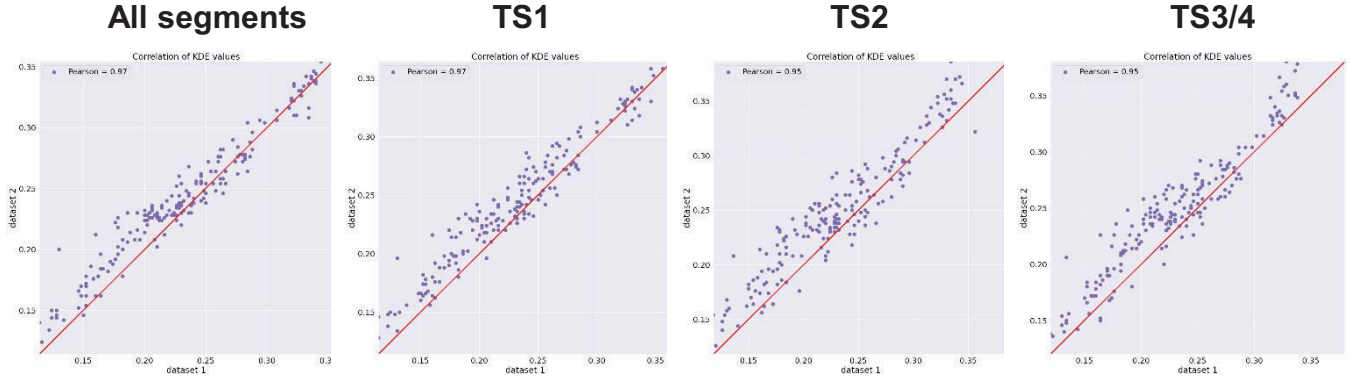
a



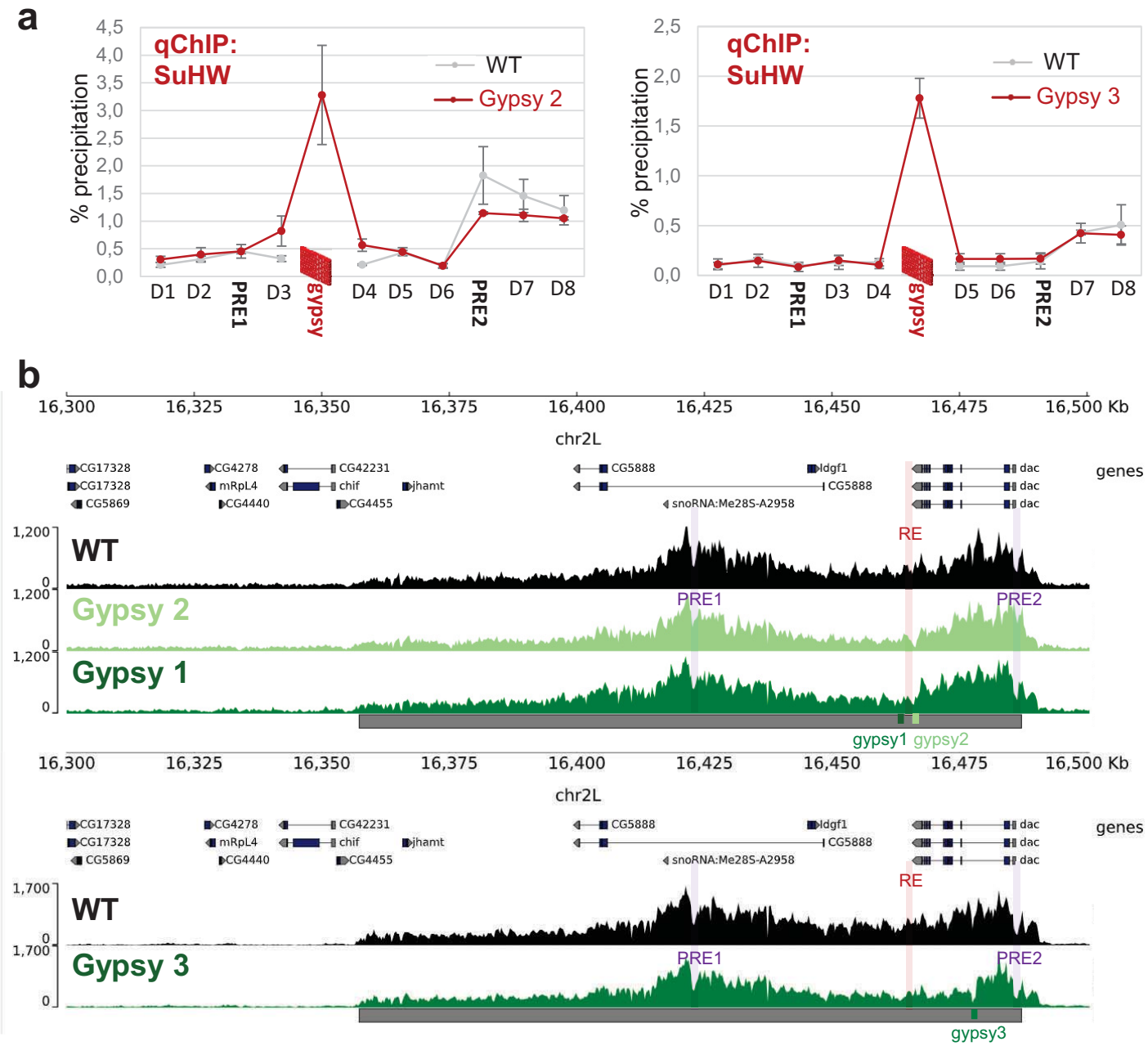
b



c

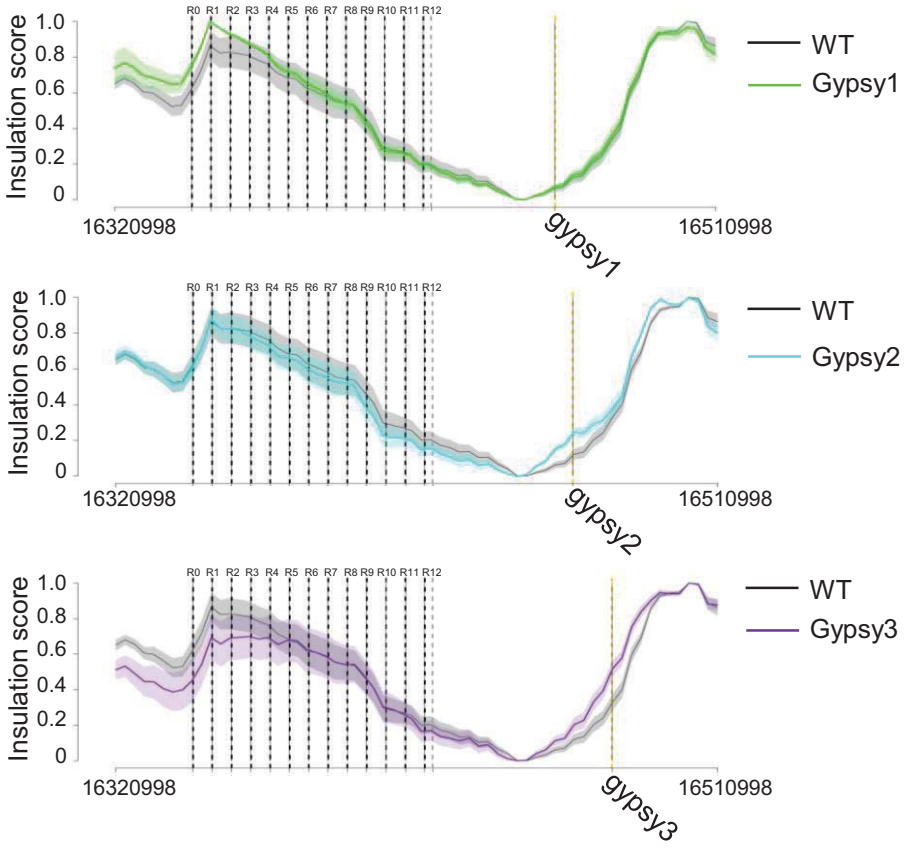


Extended Data Figure 2

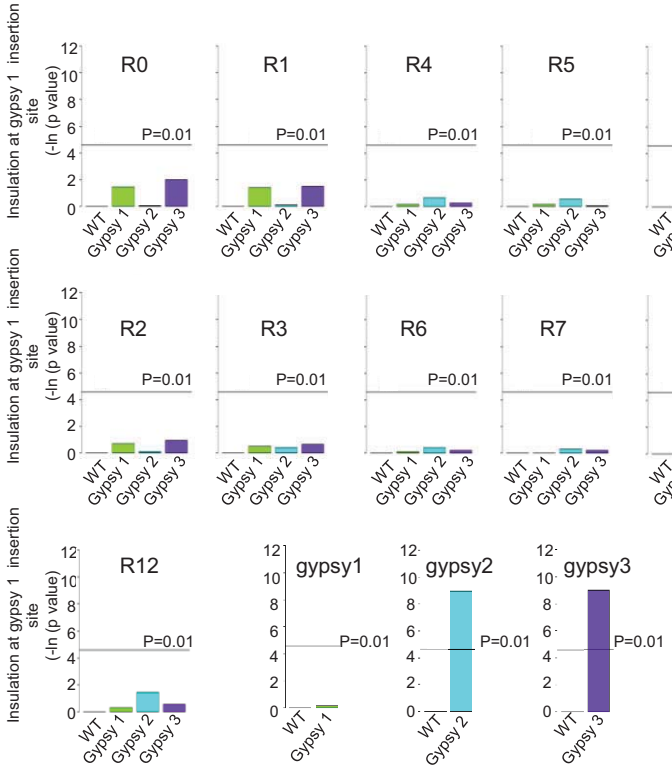


Extended Data Figure 3

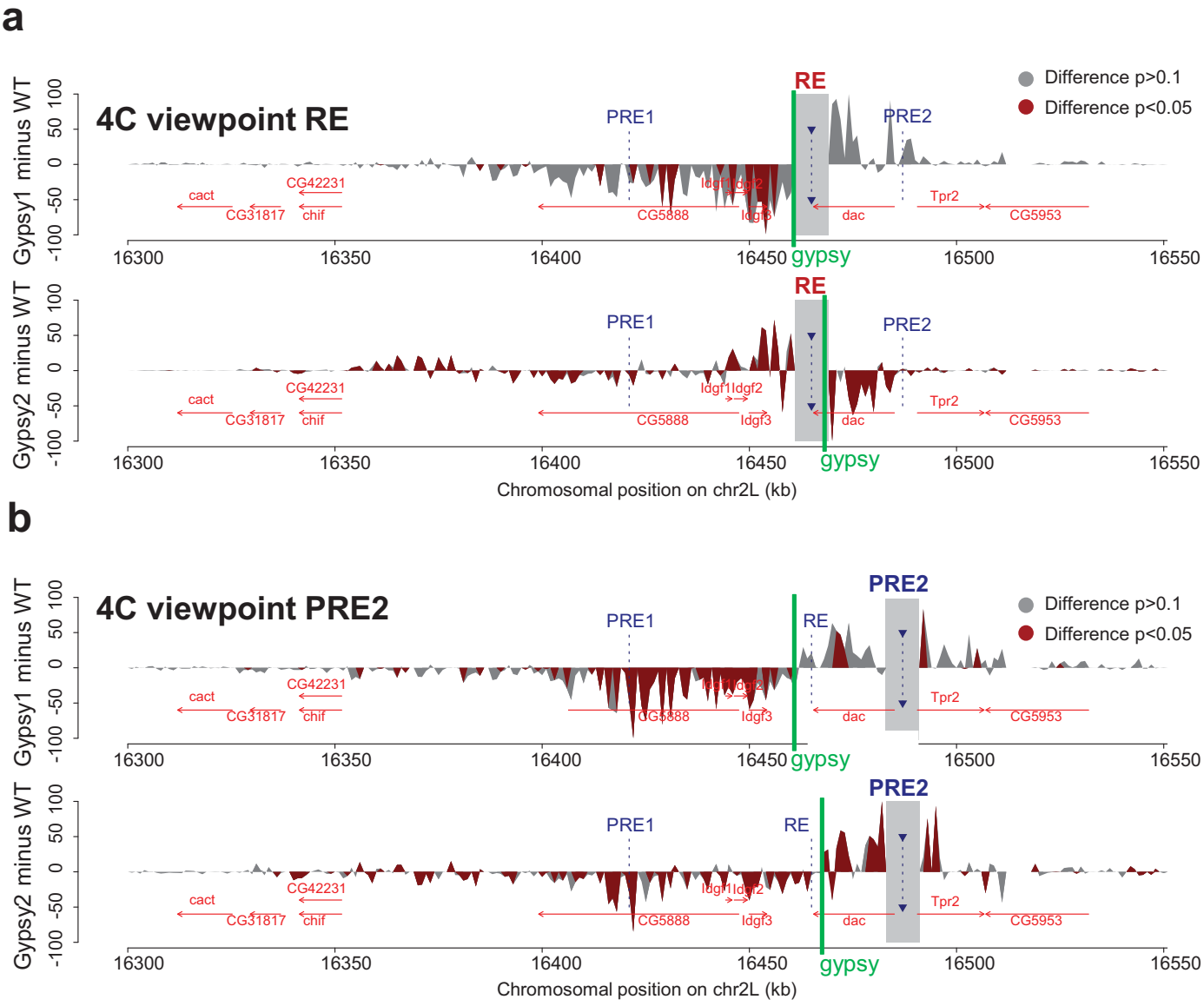
a



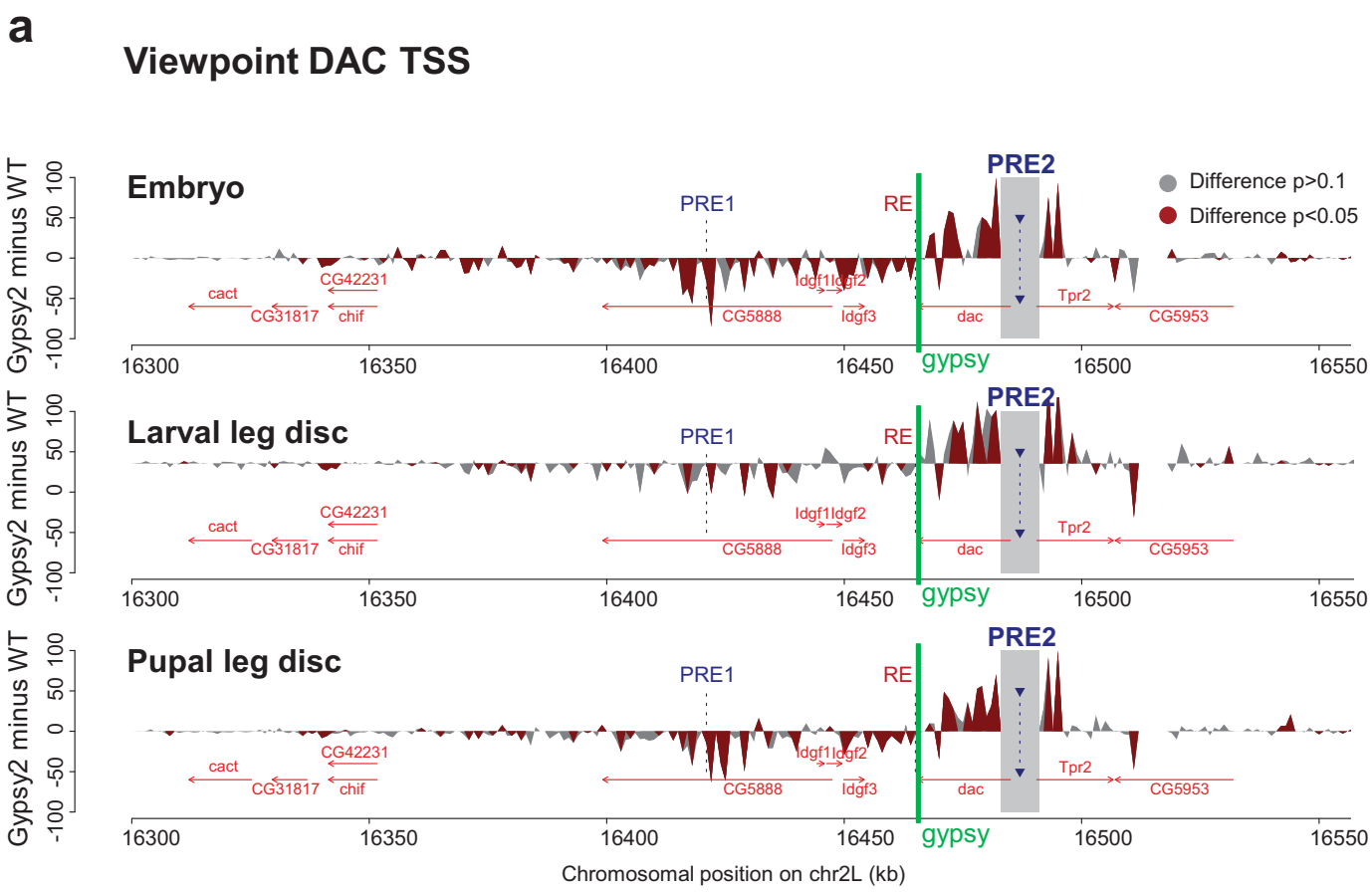
b

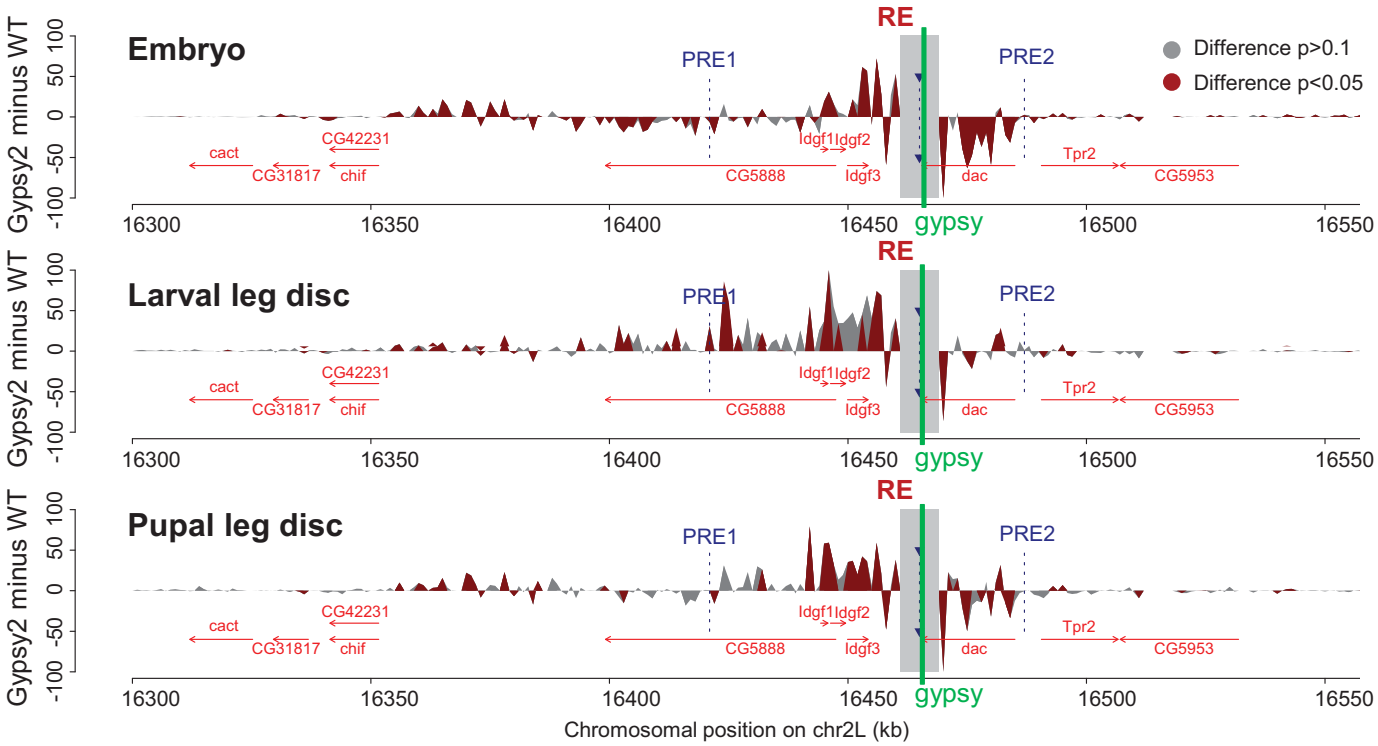


Extended Data Figure 4



Extended Data Figure 5

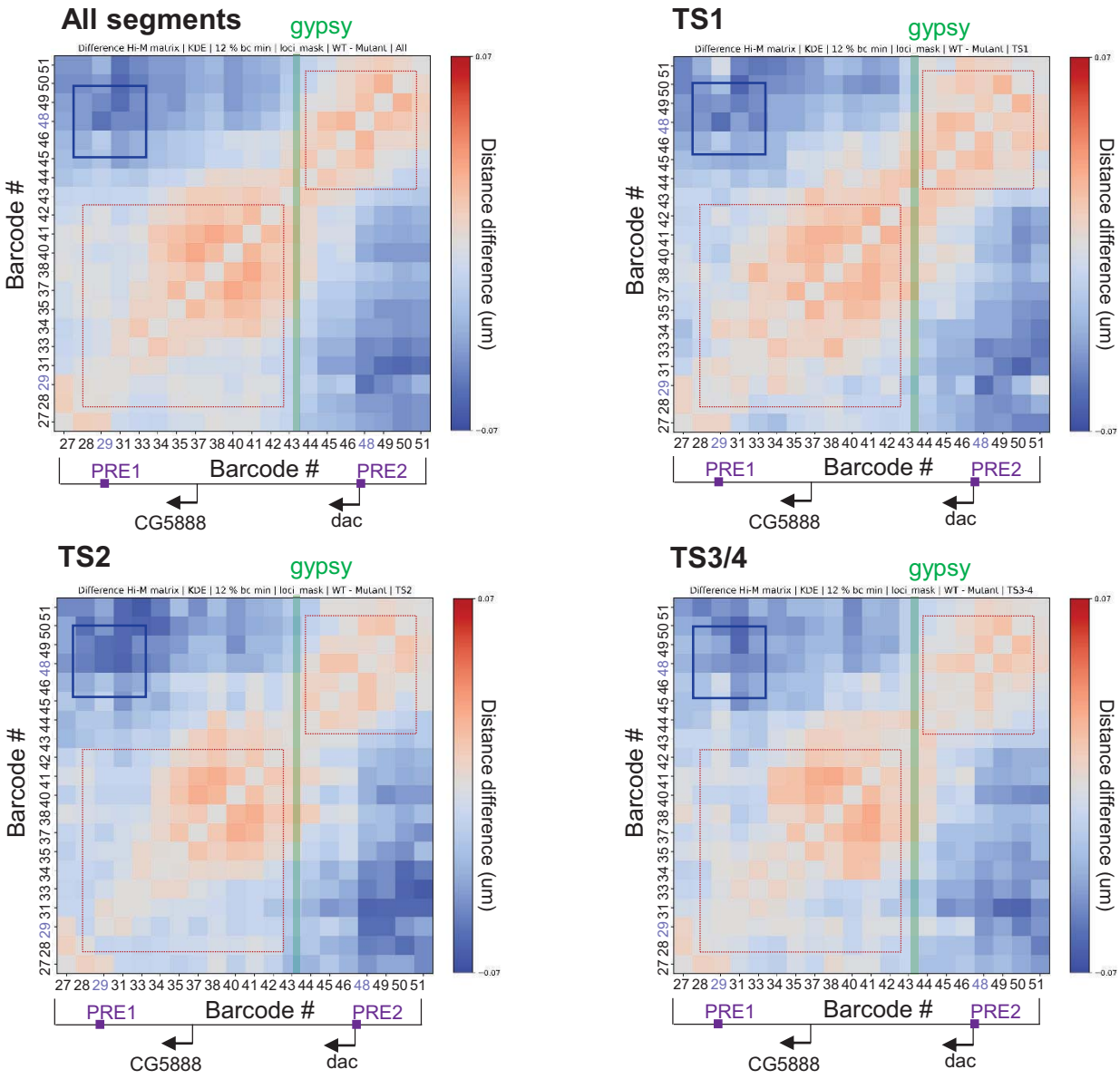




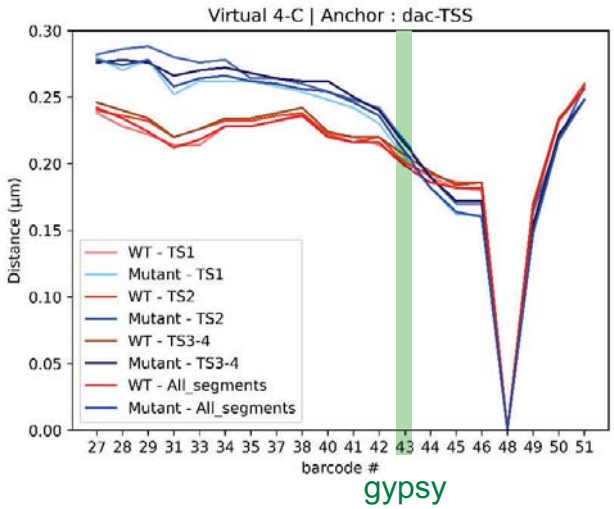
Extended Data Figure 6

a

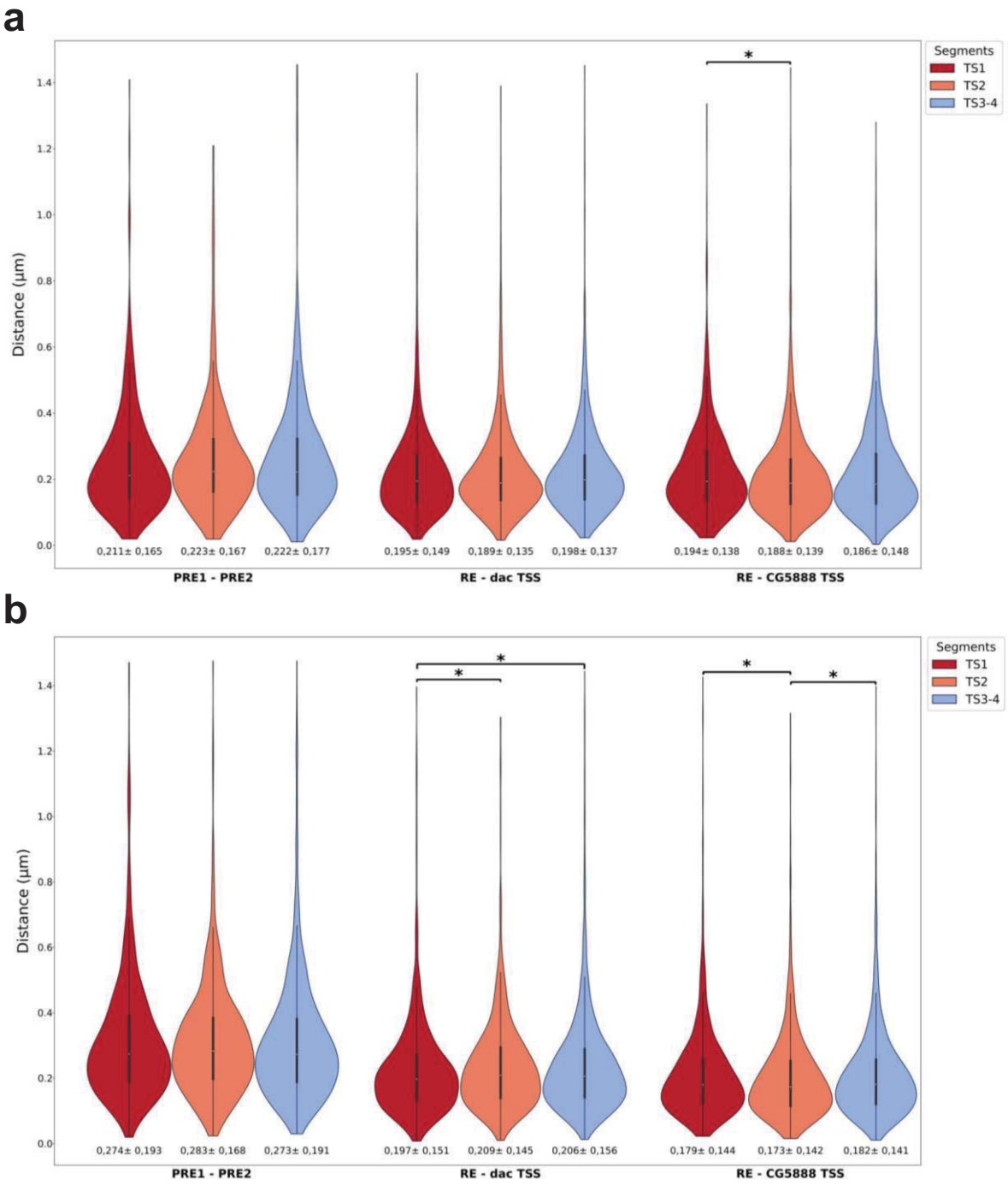
Blue: shorter distance in WT
Red: shorter distance in gypsy 2



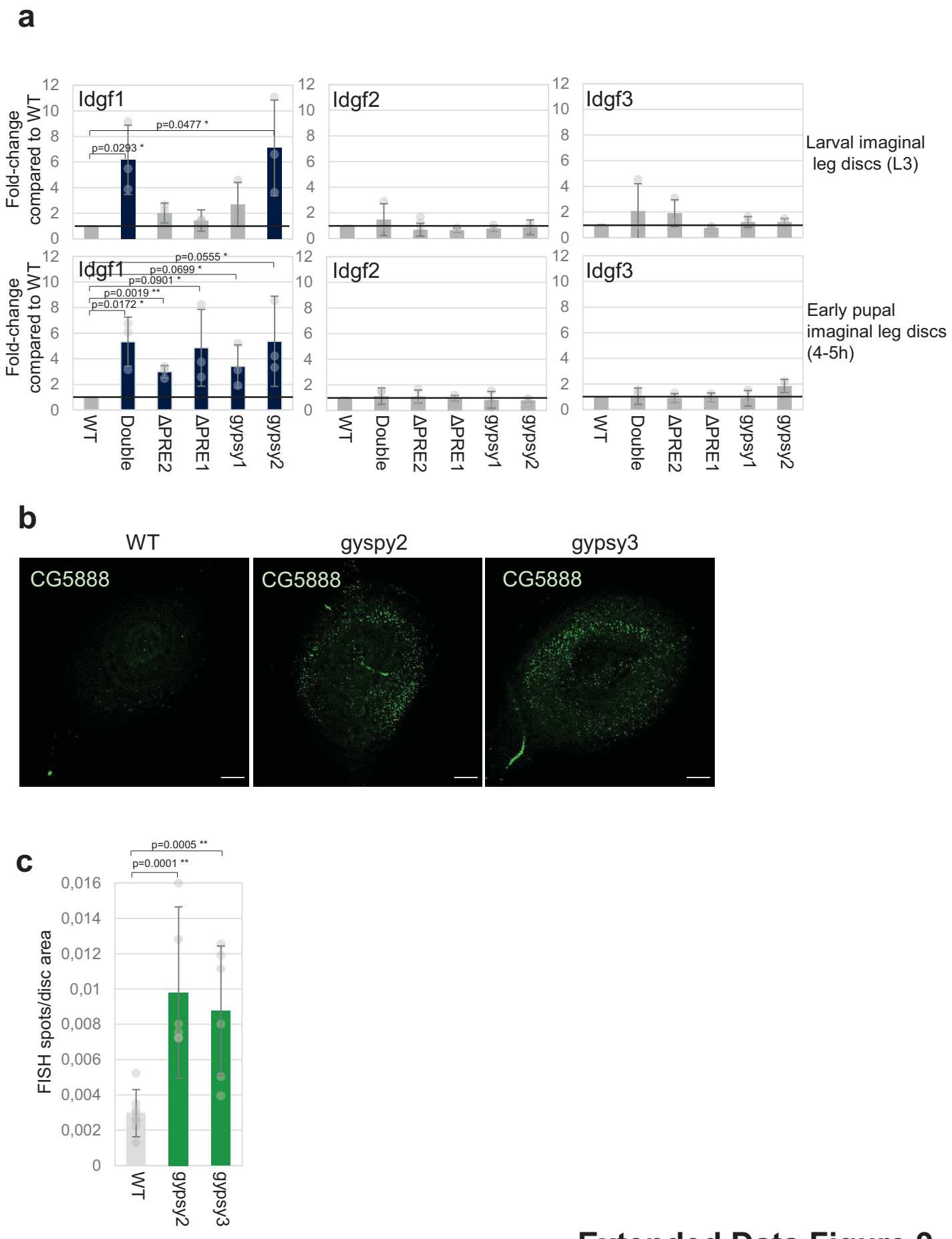
b



Extended Data Figure 7

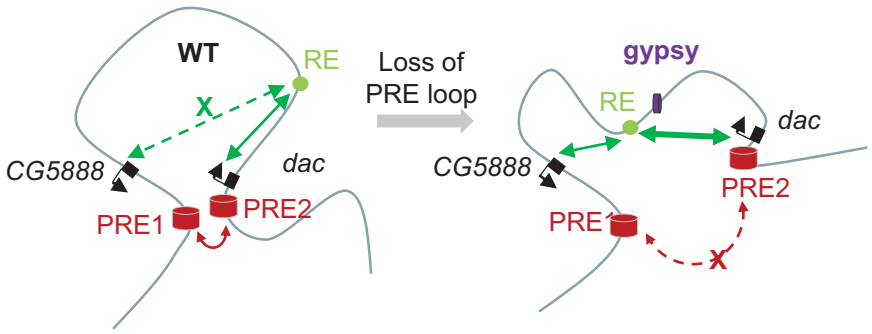


Extended Data Figure 8

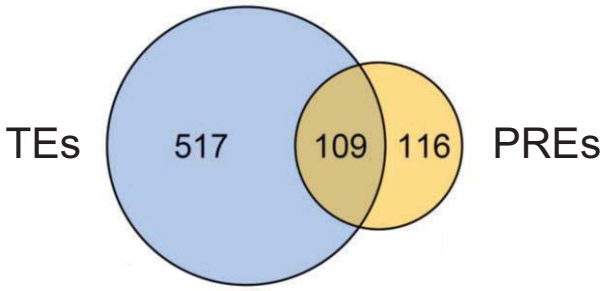


Extended Data Figure 9

a



b



Extended Data Figure 10

2.2. Chromatin organisation in the Kenyon cells for genes related to olfactory memory formation in the *Drosophila melanogaster*'s adult brain

2.2.1. Introduction and Motivation

Although numerous studies have sought to decipher the mechanisms by which genome organisation regulates transcription in both pre-differentiated and differentiated cells or tissues [Ghavi-Helm et al., 2014][Bonev et al., 2017][Freire-Pritchett et al., 2017][Espinola, Götz et al., 2021][Ing-Simmons et al., 2021][Pollex et al., 2024], our understanding of these phenomena remains incomplete. It is therefore crucial to elucidate how physical interactions between gene regulatory elements modulate their expression and how these interactions are affected by differentiation to ensure the transcriptional programs specific to each cell type.

This part of my thesis project aims to develop and apply single-cell imaging technologies to deepen our understanding of how interactions between promoters and enhancers regulate gene expression in specific cell types. I have focused on the *Drosophila* adult brain, particularly the mushroom bodies, which are crucial structures for olfactory memory formation and learning [Heisenberg et al., 1985][de Belle and Heisenberg et al., 1985][Dubnau et al., 2001][McGuire et al., 2001]. Mushroom bodies are composed of Kenyon cells, which are divided into three main neuron subtypes (γ, αβ and α'β'). Each subtype has distinct functions, chromatin accessibility and unique transcriptional programs, making the Kenyon cells an excellent model for studying the physiological aspects of genome organisation.

The objectives of my project are as follows:

- Develop tools to simultaneously detect genome conformation for different cell types in the adult *Drosophila* brain.
- Identify differentially expressed genes and their potential enhancers in various cell types.
- Investigate variations in interactions between gene regulatory regions in Kenyon cells to establish a link between chromatin conformation and transcription.

Currently, there are two models explaining chromatin conformation in pre-differentiated and differentiated cells. The first model suggests "permissive" (Figure 12) of genome folding across different cell types, regardless of whether a gene is active or repressed [Ghavi-Helm et al., 2014][Espinola, Götz et al., 2021][Ing-Simmons et al., 2021]. The second model is known as the "instructive" conformation model (Figure 12) [Bonev et al., 2017][Freire-Pritchett et al., 2017][Winick et al., 2021][Pollex et al., 2024] and posits that chromatin conformation evolves with cell type and gene expression. This part of my project aims to determine which model best describes the adult *Drosophila* brain.

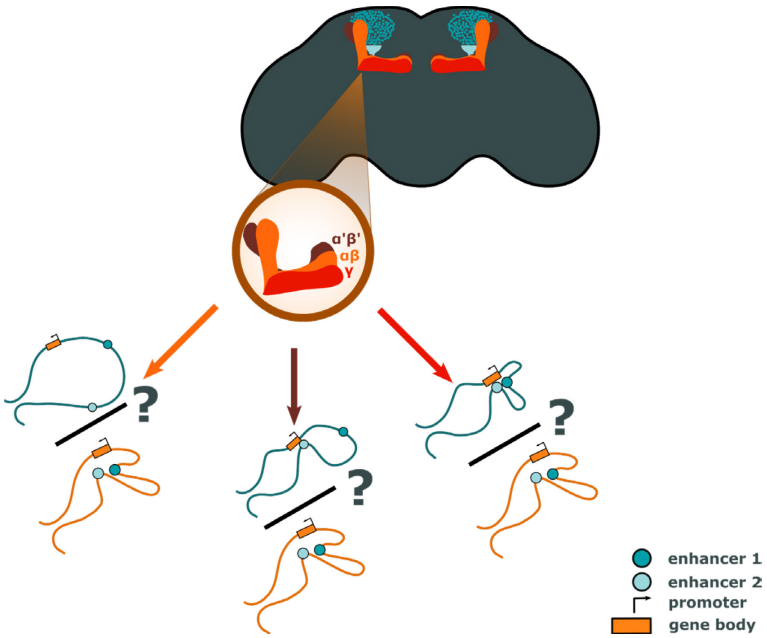


Figure 12 : Models for chromatin conformation in the *Drosophila* adult mushroom bodies - The mushroom bodies are composed of Kenyon cells, which are divided into three subtypes: γ (red), α'β' (brown), and αβ (orange). Schematic representation of chromatin conformation models: "instructive" (blue) or "permissive" (light orange).

2.2.2. Part 1 : Development of tools to simultaneously detect genome conformation and different cell types in the *Drosophila* adult brain

To determine if the three-dimensional organisation of the genome varies among different Kenyon cells (Figure 12), it is essential to simultaneously detect chromatin conformation and identify distinct cell types. I employed the Hi-M technique, developed by my host team, which enables the reconstruction of the three-dimensional structure of chromatin [Cardozzo Gizzi et al., 2019][Cardozzo Gizzi et al., 2021] (Figure 13). This technique relies on the combination of fluorescence in situ hybridization (FISH) with fluidics handling devices and super-resolution microscopy to image multiple DNA loci in single cells to reconstruct the chromatin conformation from thousands of cells in their native tissue context.

Notably, there was no existing protocol in the literature for performing DNA-FISH in the *Drosophila* adult brain at the time my thesis started. Thus, the first step of this project was to develop a protocol for Hi-M in the *Drosophila* adult brain.

2.2.2.1. Development to detect chromatin conformation with Hi-M in fly brains

Hi-M was initially developed using *Drosophila* embryos. During advanced stages of embryo development, the embryo's outer layer comprises multiple layers of nuclei, whereas in early stages (nuclear cycles 11-14), it consists of a single layer. Because of this issue, imaging late-stage embryos with widefield microscopy resulted in increased background noise, a reduced signal-to-noise ratio which made detection of single DNA spots very inefficient. To address these issues in samples with multiple layers of nuclei or cells, a former student of the group combined Hi-M with confocal microscopy [Gurgo et al., 2024].

Confocal , however, proved to have its own limitations. First, confocal based imaging is slower than widefield imaging, therefore a typical experiment with 25 cycles would require 7 days in confocal and 2 days in widefield. Second, confocal was more prone to photobleaching, therefore further reducing the efficiency of detection of DNA-FISH spots. Third, long-term imaging of whole-mount samples proved to

be challenging primarily due to sample detaching from the cover slip over extended periods of time (1 week) in the fluidics chamber. To tackle these limitations, two lab members combined cryosectioning with widefield microscopy to apply Hi-M to mouse tissues [Messina, Schaeffer et al., in preparation], similarly to the method *Mateo et al.* (2019) has implemented in the past to image chromatin in late *Drosophila* embryos [Mateo et al., 2019]. I adapted this methodology to image *Drosophila* adult brains using cryosectioning, which effectively minimised the 'out-of-focus' signal and enabled single-molecule localisation with widefield microscopy (Figure 13).

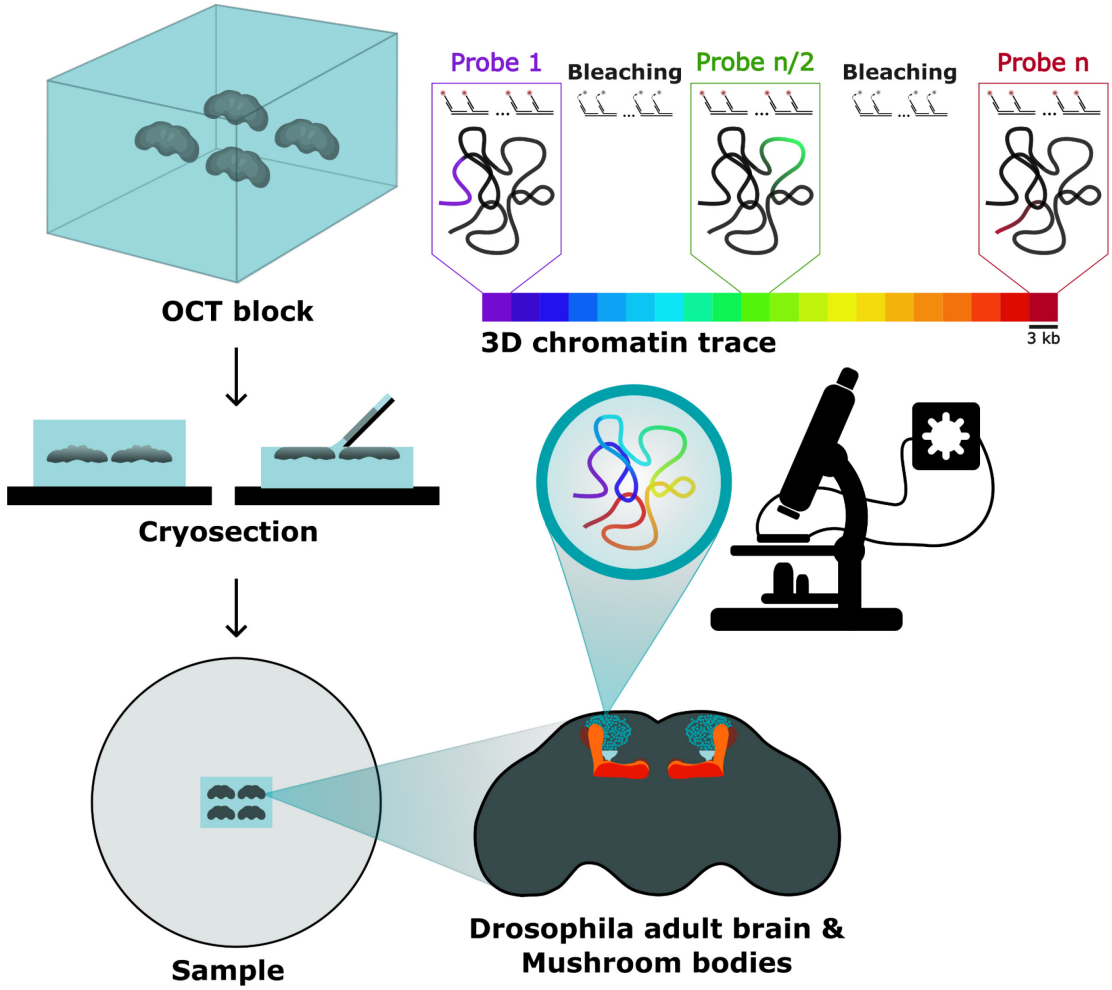


Figure 13 : Schematic representation of an Hi-M acquisition and chromatin trace reconstruction in the *Drosophila* adult brain - Schematic illustrating the imaging-based strategy used to study chromosome conformation at single cell level in cryosectioned *Drosophila* brain tissues (Hi-M). Each colour corresponds to a sub-region of 3 kb in size. All the colours together form the genomic region of interest. Each sub-region, also called probe, is imaged and then bleached to enable the acquisition of the entire locus with sequential DNA-FISH.

I implemented and tested multiple protocols (see Materials and Methods section) aimed at ensuring the preservation of both the overall brain morphology and nuclear structure during dissection, fixation and cryosectioning. Once these first steps were achieved, a coverslip functionalisation protocol was also adapted to ensure adhesion of the brain tissue to the glass slide, preventing any detachment during the numerous washing steps involved in the various staining protocols.

To evaluate the specificity of Oligopaint staining in the *Drosophila* adult brain, I initially focused on confirming the staining specificity. For this, I tested whether the fluorescence signal from a small genomic

locus appeared as a diffraction-limited spot and was confined within the volume of the nucleus (Figure 14). Additionally, various controls were implemented, including verifying the colocalization of neighbouring genomic loci. At this stage, a genomic resolution of 3 kilobase pairs (kbp) was reached, as shown in Figure 14. This resolution corresponds to the minimal size of the homology DNA region targeted by an oligopaint probe, while still being able to image the sample by DNA-FISH. This DNA-FISH resolution had already been achieved using ORCA in late embryo cryosections [Mateo et al., 2019].
Achieving this resolution was a critical objective in the project's initial phase, as it was essential for detecting physical proximity between enhancers and promoters in *Drosophila*. Indeed, in contrast to mammals, enhancers in *Drosophila* are generally within short genomic distance from the promoters they regulate.

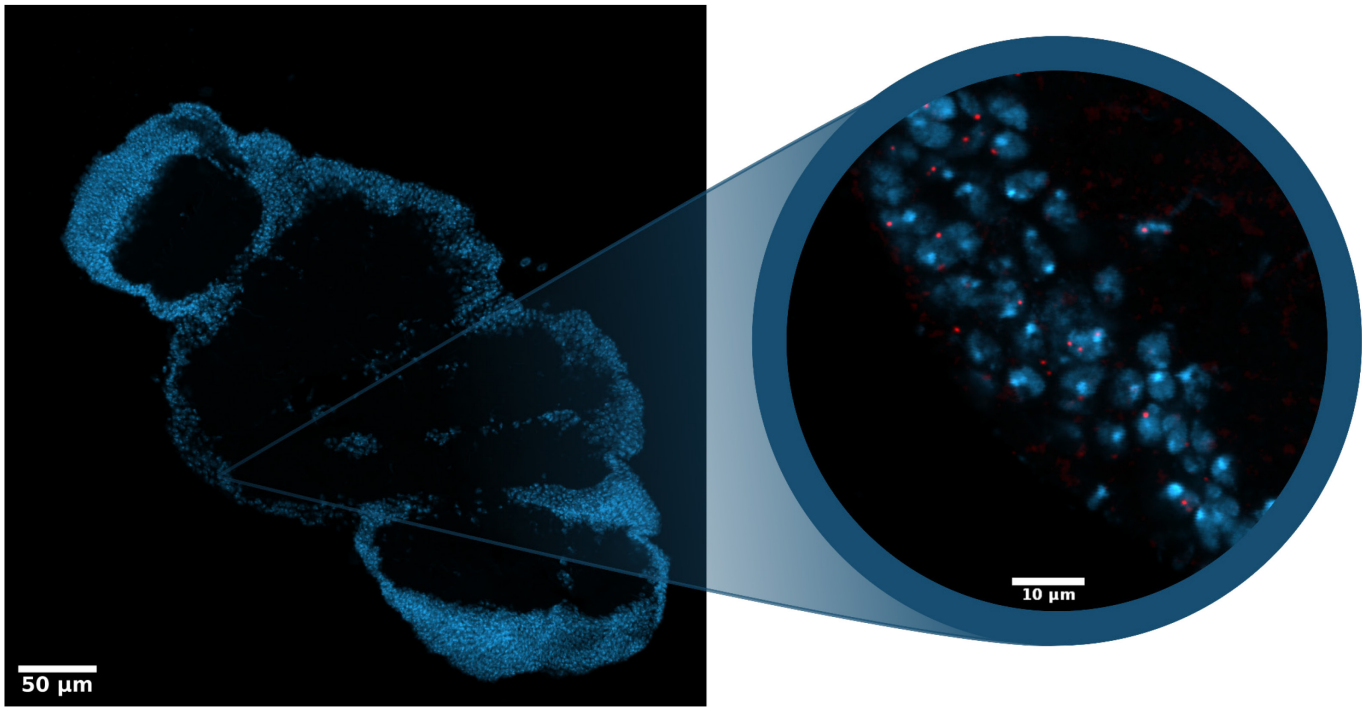


Figure 14 : Specific labelling in a *Drosophila* adult brain cryosection - Nuclei in blue (DAPI) - 3kb staining in red (Alexa Fluor 647) - 20x (left) - 60x (right)

2.2.2.2. Tools and development to detect cell types in the *Drosophila* adult brain

For cell type identification, we have selected the GAL4-UAS system (Figure 15a). In this system, GAL4 expression is controlled by a tissue-specific enhancer. Once the GAL4 protein is synthesised, it binds to the UAS region to activate the transcription of a transgene coding for the nls-GFP (nuclear localization signal - Green Fluorescent Protein) protein fusion. This allowed nuclei staining of the cell types of interest, although parts of the axons/lobes can also have a weaker staining due to the diffusion of nls-GFP through the cytoplasm (Figure 15b).

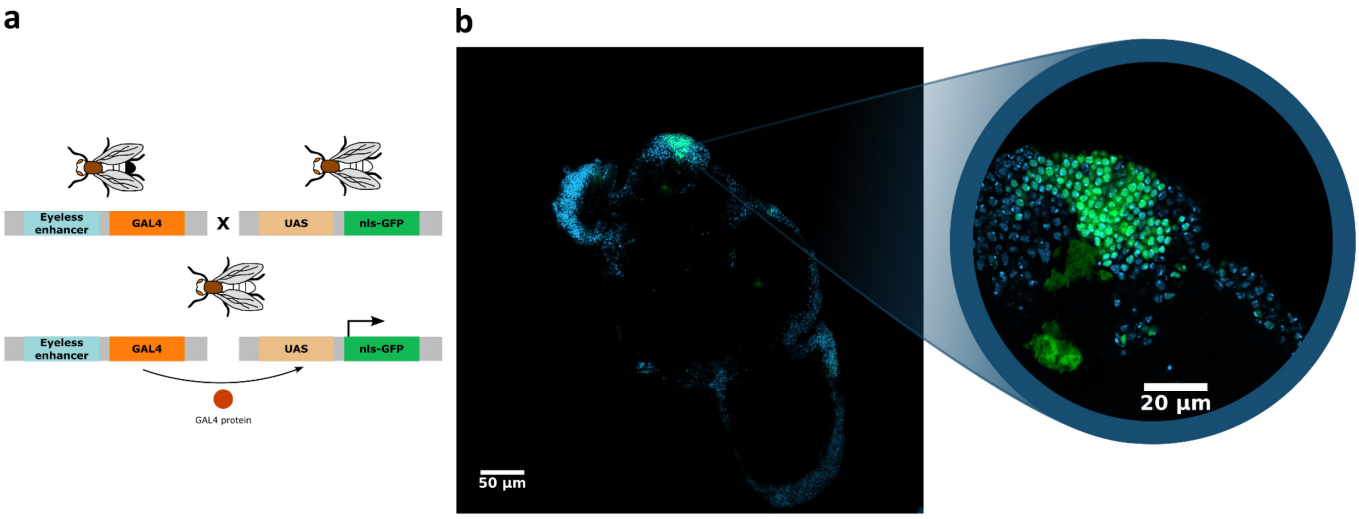


Figure 15 : GAL4-UAS schematic representation and fluorescence image of a *Drosophila* adult brain expressing GFP in the mushroom bodies - (a) Schematic representation of the OK107-GAL4 x UAS-nls-GFP cross. The OK107-GAL4 line enables the expression of the GAL4 protein exclusively in Kenyon cells due to a tissue-specific enhancer of the Eyeless gene. Following the cross with the UAS-nls-GFP line, the GAL4 protein induces nuclear GFP expression in the Kenyon cells nuclei. (b) Confocal image of a 10 µm thick cryosection of an adult *Drosophila* brain after a OK107-GAL4 x UAS-nls-GFP cross allowing specific GFP expression in the nuclear membrane of all Kenyon cells. Nuclei - DAPI (blue) ; Kenyon cells nuclei - GFP (Green).

For this project, all the GAL4 lines used in this study have been thoroughly described in *Aso et al.* (2009) and were selected to identify the following cell types of interest :

- OK107-GAL4 [Connolly et al., 1996] expressing GFP in all Kenyon cells (Figure 15b)
- H24-GAL4 [Zars et al., 2000] expressing GFP in γ Kenyon cells
- c739-GAL4 [Yang et al., 1995] expressing GFP in αβ Kenyon cells
- c305a-GAL4 [Krashes et al., 2007] expressing GFP in α'β' Kenyon cells

After performing DNA-FISH, the GFP fluorescence was lost due to conformational changes in the fluorophore structure caused by the denaturation step at 85°C. Indeed, GFP shows great stability until 70°C. Past that temperature, GFP undergoes thermal denaturation. However, even though GFP is no longer fluorescent, an anti-GFP antibody can still recognize the protein. To restore GFP localisation, I incorporated an anti-GFP immunostaining step at the end of the DNA-FISH protocol.

Each GAL4 fly line/cell type of interest had to be processed independently. Indeed, using the GAL4-UAS system alone, it is impossible to visualise two different populations of neurons in the same brain without ending up mixing their signals. Once the Hi-M libraries imaged in all the brains of interest and the anti-GFP signal from the Kenyon cells corresponding to each brain was acquired, Hi-M data were processed using the pyHi-M pipeline presented in section 1.3.4. After producing 3D chromatin traces, the anti-GFP signal (Figure 16a) was segmented (Figure 16b) using ilastik [Berg et al., 2019] (a machine learning toolkit for segmentation) and each trace was affiliated either to the Kenyon cells or to the other cells of the brain by identifying the traces that overlap with the GAL4 pattern (Figure 16c).

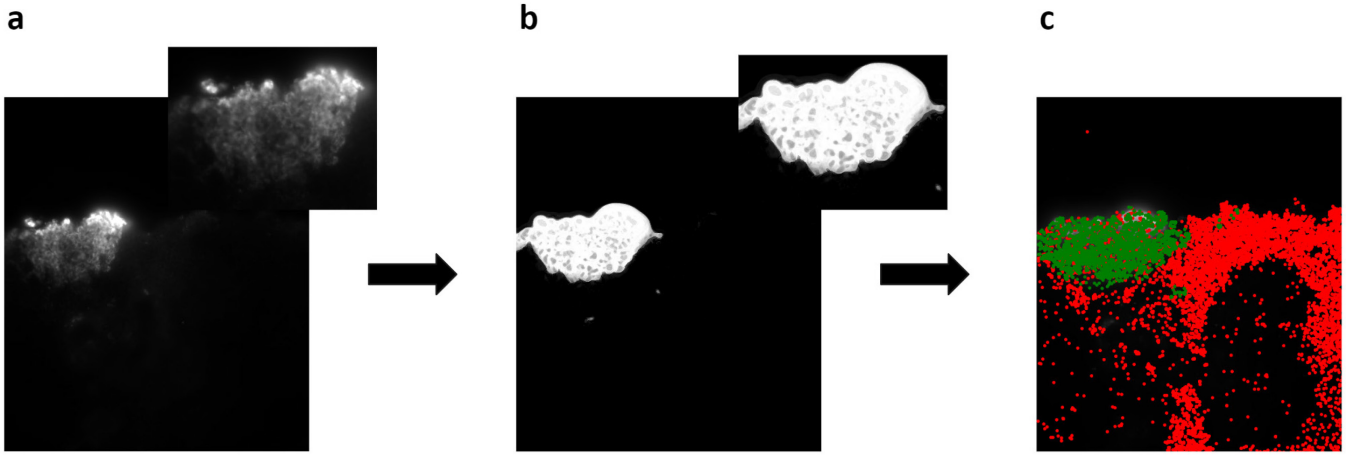


Figure 16 : Assignment of Kenyon cells and other brain cells to traces - (a) Deconvolved image of the anti-GFP staining Kenyon cells (OK107-GAL4 x UAS-nls-GFP) (b) Anti-GFP signal segmented with ilastik (c) chromatin trace assignment to Kenyon cells (green) and other cells of the brain (red).

At this stage of the project, by combining Hi-M with the GAL4-UAS system, chromatin traces were successfully matched to the cell types of interest.

2.2.2.3. Validation

To validate the multiplexed Hi-M protocol I developed, 3D chromatin traces of the doc locus in the adult *Drosophila* brain and Kenyon cells were first compared to previously obtained data from *Drosophila* embryos at nuclear division cycle 14 (nc14) [Espinola, Götz et al., 2021] (Figure 17).

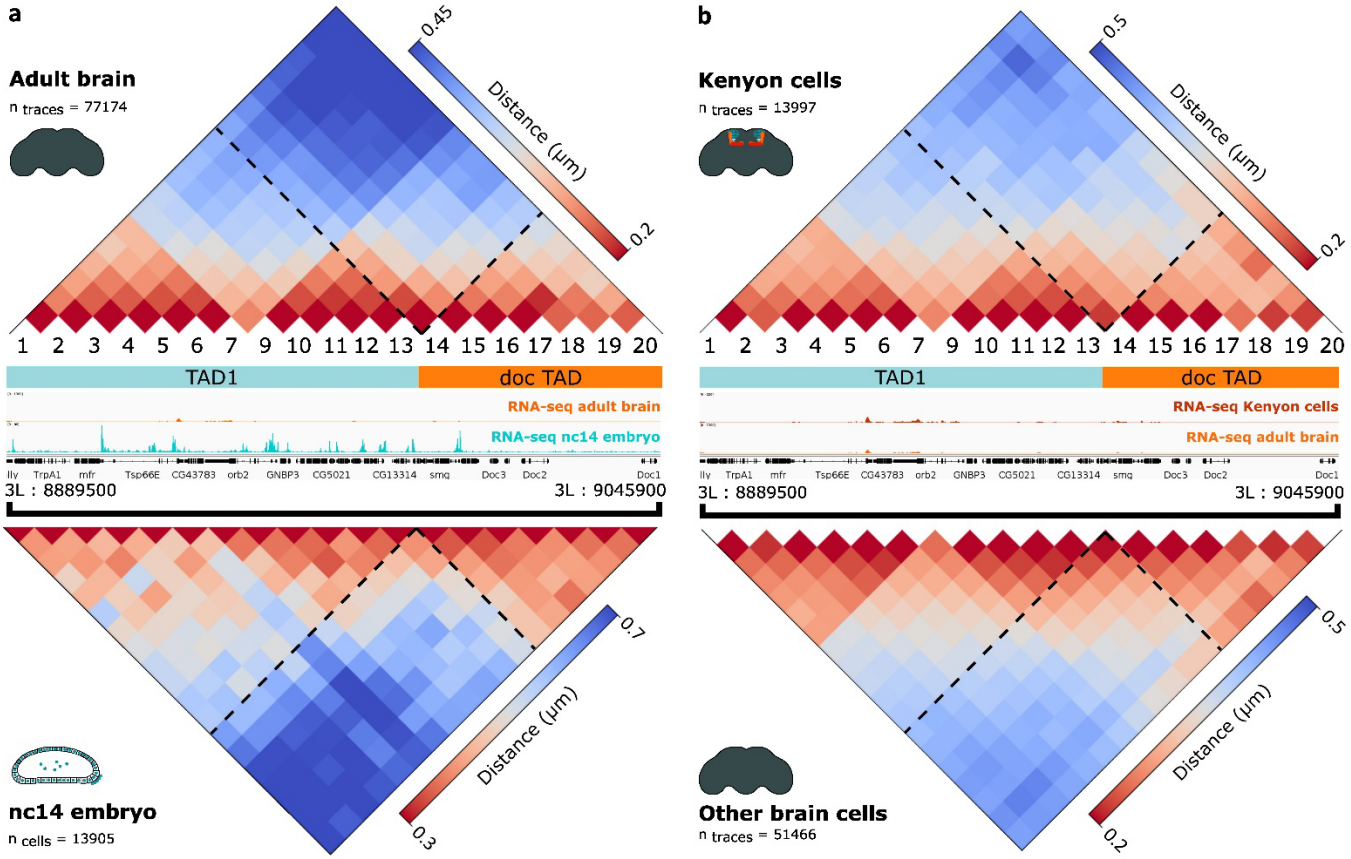


Figure 17 : Hi-M Pairwise distance single cell matrices of nuclear cycle 14 *Drosophila* embryos of the *Drosophila* adult brain and of the adult Kenyon cells - (a-b) Comparison of the three-dimensional organisation of the doc locus. Dashed lines highlight the TAD border (a) PWD median matrices at two *Drosophila* developmental stages : the adult brain (top) and the embryo at developmental cycle 14 (bottom). (b) PWD median matrices in Kenyon cells (top) and in the rest of the cells of the adult brain (bottom).

As illustrated by the pairwise distance matrices, chromatin in adult brains is more condensed compared to embryos (see scale bars in Figure 17a). This increase in condensation is likely due to the smaller nuclear size of adult brains compared to nc14 embryos. The overall TAD structure remains visible (dashed lines in Figure 17a), thereby reassuring me in my ability to reconstruct 3D chromatin structure in *Drosophila* adult brains using Hi-M (Figure 17a).

After comparing chromatin conformation between embryos and adult brains, identification of chromatin traces from all Kenyon cells (OK107 GAL4 fly lines) and the rest of the brain (non-Kenyon cells) was achieved. Subsequently, differences in pairwise distance matrices generated for the doc locus in each condition was evaluated. As depicted in Figure 17b, the two matrices exhibit a striking similarity. Importantly, this is consistent with the lack of expression of the doc genes in Kenyon cells and in other cells. Thus, from these experiments I can conclude that 3D chromatin organisation in a silent locus does not considerably change between different brain cell types.

Following these experiments, I explored chromatin organisation in loci containing genes differentially expressed in different Kenyon cells. To find these loci, I implemented and used a bioinformatics pipeline to search for highly expressed genes with differential γ, αβ, and α'β' expression, and with distant enhancers to directly test whether gene expression is coupled to differential 3D chromatin topologies in the fly brain.

2.2.3. Part 2 : Identification of genes of interest and their potential enhancers in different cell types

2.2.3.1. Bioinformatic analysis

To identify potential regions of interest, I did a bioinformatic study to select genes of interest based on publicly available ATAC-seq data, scRNA-seq, RNA-seq data, and enhancer libraries [Henry et al., 2012][Crocker et al., 2016][Croset et al., 2017][Davie et al., 2018][Janssens et al., 2022]. Subsequently, I identified highly and differentially expressed genes in the three subtypes, with enhancers specific to Kenyon cells and a distance greater than 15 kb between these enhancers and their promoter. Two genes were selected, *rutabaga* (*rut*) that is expressed in all the Kenyon cells, and *Short NeuroPeptide F* (*sNPF*) that is mainly expressed in γ and αβ neurons (Figure 18).

The *rut* gene encodes a membrane-bound Ca[2+]/calmodulin-activated adenylyl cyclase which is essential for the formation of appetitive short-term memory and long-term memory in *Drosophila* [Trannoy et al., 2011]. Furthermore, experiments involving the rescue of *rut* mutants by expressing *rut* in specific neuron types revealed that *rut* is specifically required in γ neurons for STM formation and in αβ neurons for LTM formation.

The *sNPF* gene encodes a protein that binds to the product of *sNPF-R* and activates ERK-Dilps signalling or the PKA-CREB pathway. *sNPF* is involved in the formation of appetitive olfactory memory [Knapek et al., 2013]. This study demonstrates that activation of Kenyon cells leads to a decrease in *sNPF* levels, and knockdown of *sNPF* in Kenyon cells impairs sugar-rewarded olfactory memory without affecting reflexive sugar preference or odour response. Additionally, knockdown of *sNPF* receptors outside the MB also results in deficits in appetitive memory. These findings suggest that *sNPF* acts as a functional

neuromodulator released by Kenyon cells to regulate olfactory memory in *Drosophila*. Thus, both *rut* and *sNPF* are critical for the role of Kenyon cells in memory formation.

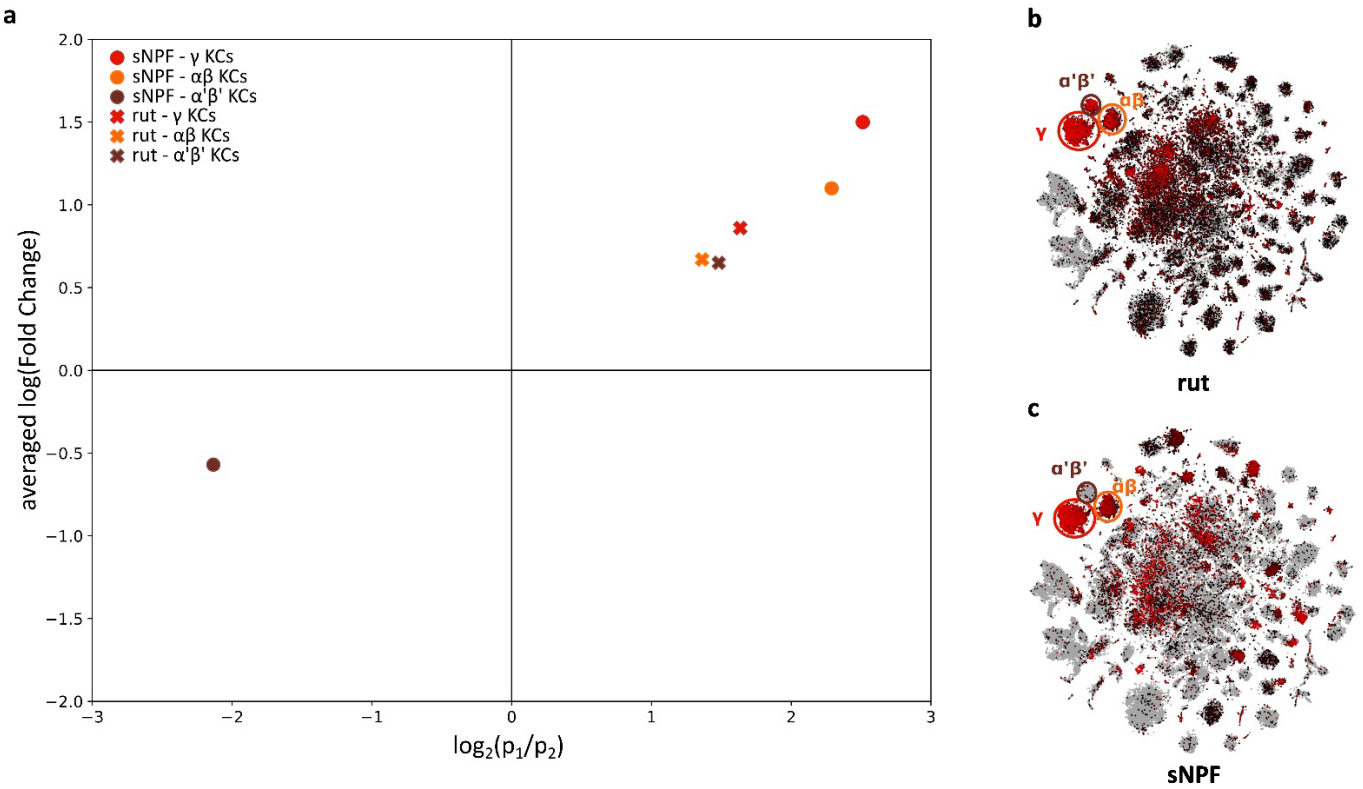


Figure 18 : Differential expression of the genes of interest in the Kenyon cells of the *Drosophila* adult brain - (a) Gene expression of *rut* (crosses) and *sNPF* (dots) in γ KCs (red), α'β' KCs (brown) and αβ KCs (orange) compared to the other cells of the brain. p_1 is the percentage of detection in the cluster of interest. p_2 is the percentage of detection in the remaining cells. averaged log(Fold Change) compares expression values between the cell type of interest and the rest of the brain cells. Values were calculated with Seurat for the given genes and cell types by [Davie et al., 2018]. (b-c) t-SNE clustering for the visualisation of scRNA-seq dataset of the *Drosophila* adult brain from [Davie et al., 2018] - γ KCs (red cluster) ; α'β' KCs (brown cluster) ; αβ KCs (orange cluster) (b) t-SNE clustering of the *rut* gene (c) t-SNE clustering of the *sNPF* gene.

After selecting these candidate genes, I set out to test the link between genome conformation and transcription. *rut* is expressed in all Kenyon cells (Figure 18b), thus one would expect similar chromatin folding patterns in different KCs. However, *rut* is not expressed at the same levels in all KCs subtypes (Figure 18a), suggesting a potential modulation in gene expression that may be directed by changes in chromatin architecture. In this case, one would expect detectable differences in chromatin architecture between MB cell subtypes due to the differential action of subtype-specific regulatory elements. In the case of *sNPF*, the differential expression in different KC sub cell types raises more complex hypotheses. *sNPF* is highly expressed in γ and αβ Kenyon cells but repressed in α'β' neurons (Figure 18a-c). Thus, if all sub cell types display similar chromatin conformations, this may suggest that this differential expression is not directly influenced by chromatin conformation but rather by cell-type-specific activators and repressors.

2.2.3.2. Library design and validation

The next step was to design oligopaint libraries for *rut* and *sNPF*. For this, I designed libraries for loci containing these genes and divided them into 25 smaller genomic regions (probes) to achieve multiplexed DNA-FISH at a resolution of around 3kb. To define the size of the library and the position of the probes, I analysed the accessibility and epigenetic profiles of the region surrounding these genes (Figure 19 and Figure 20).

***Rutabaga* (*rut*)**

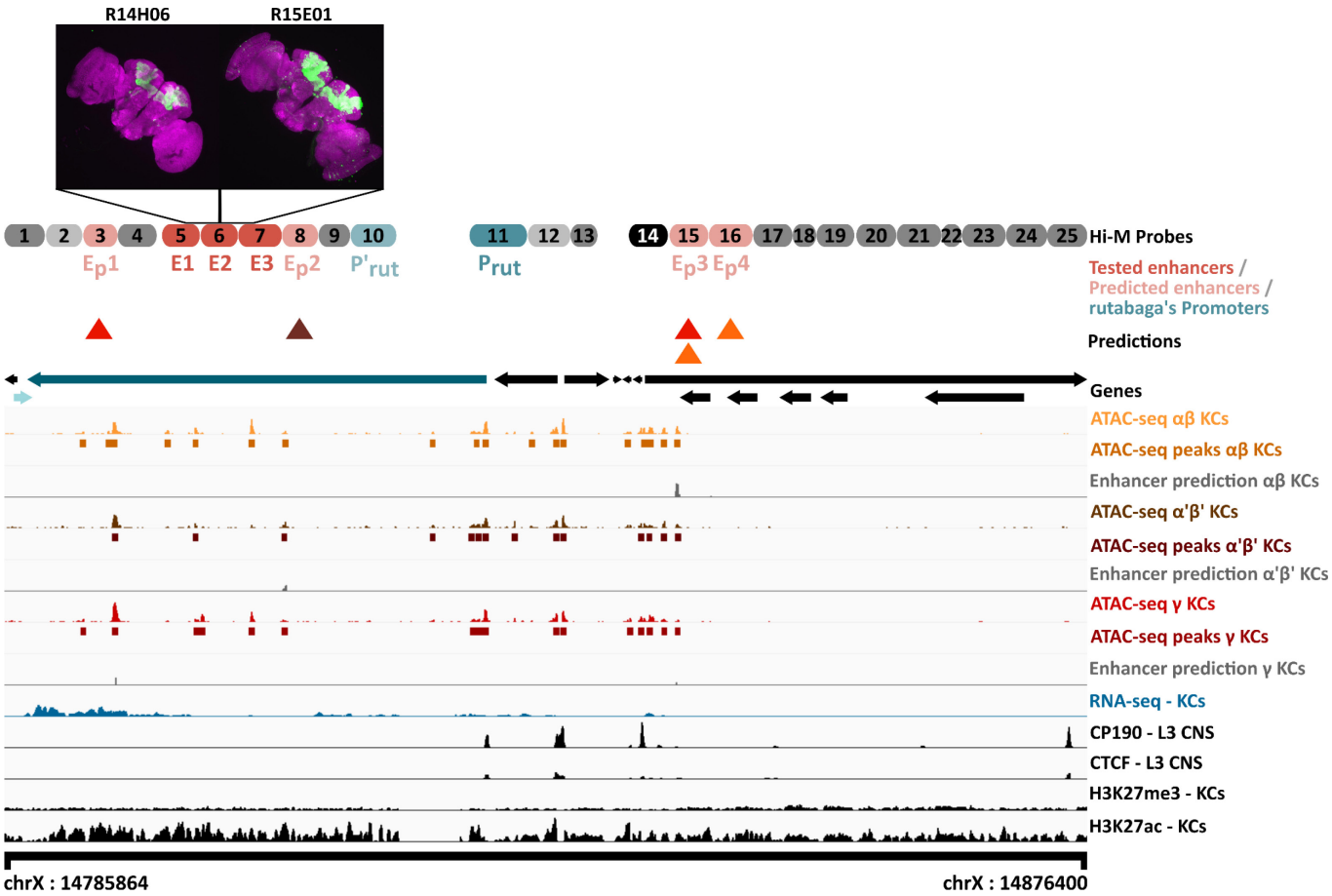


Figure 19 : Detailed view of the *rut* locus - From top to bottom : images of GAL-4 fly lines expression pattern from Janelia's Flylight project [Jenett et al., 2012] of mushroom body specific enhancers (R14H06 and R15E01). Hi-M probe repartition along the genome (*rut*'s promoters (teal) ; tested enhancers (red) ; enhancers predicted by [Janssens et al., 2022] (pink) ; accessible regions (light grey) ; inaccessible regions (dark grey)). Identification of tested and predicted enhancers. Predicted enhancers : γ specific (red triangles) ; α'β' specific (brown triangle) ; αβ specific (orange triangles) [Janssens et al., 2022]. Gene repartition (*rut*'s gene body (dark teal) ; CR45522 (light blue) ; other genes (black)). ATAC-seq data and ATAC-seq peaks from [Janssens et al., 2022] (γ KCs (orange) ; α'β' KCs (brown) ; αβ KCs (red)). Enhancer prediction tracks from [Janssens et al., 2022]. RNA-seq data in OK107 GAL-4 from [Henry et al., 2012]. CP190 and CTCF ChiP-seq data for L3 CNS from [Kaushal et al., 2021]. Histone marks in OK107 GAL-4 from [Henry et al., 2012] (H3K27me3 and H3K27ac) - Genomic coordinates of the locus (chrX:14785864-14876400).

Based on this data, I designed an oligopaint library to image a genomic region containing the gene *rut* located between chrX:14785864 and chrX:14876400 (90,5 kb). This region encompasses promoters of thirteen different genes, with *rut*'s promoter located at probe 11 (P_{rut}) and an alternative *rut* promoter at probe 10 (P'_{rut}). Notably, a long non-coding RNA (lncRNA) is located within probe 1. Although the lncRNA CR45522 has been identified, its function remains unknown.

Within the *rut* locus, two mushroom body-specific enhancers have been validated using an enhancer assay conducted by Janelia's Flylight project [Jenett et al., 2012]. These mushroom body-specific enhancers overlap with probes 5, 6 and 7 which are designated as E1, E2 and E3, respectively (see expression pattern - Figure 19, top) .

DeepFlyBrain, a deep learning model trained on enhancer sequences from Stein Aerts' lab, predicted thousands of additional fly brain enhancers [Janssens et al., 2022]. In the *rut* locus, I found five accessible regions that were predicted by DeepFlyBrain as KC enhancers with more than a 50% probability. These predicted enhancers are predicted to have different KC specificities, and are located at probes 3, 8, 15 and 16. I identified them as Ep1 for Enhancer prediction 1 (γ specific), Ep2 (α'β' specific), Ep3 (γ and αβ specific) and Ep4 (αβ specific), respectively.

Although cell-type specific transcription in the adult *Drosophila* brain and in the mushroom body has been extensively studied [Crocker et al., 2016][Croset et al., 2017][Davie et al., 2018], information on binding patterns of transcription factors, insulators or epigenetic marks in the fly brain is unfortunately scarce. Based on the only accessible dataset, the locus is globally active and does not exhibit polycomb marks in adult Kenyon cells [Henry et al., 2012] (Figure 19).

To assess chromatin insulation in this region, I used published data on the central nervous system (CNS) from the third larval stage (L3) [Kaushal et al., 2021]. ChIP-seq data from this study identifies probes with known insulator protein binding sites for CP190 and CTCF (probes 11, 12, 14 and 25). Probe 14 correlates with the position of a TAD border, consistent with the colocalization of both CTCF and CP190, although *Drosophila* TAD borders are generally less enriched in CTCF proteins than those in mammals [Sexton, 2012][Hou et al., 2012].

As previously mentioned in the Introduction, insulator proteins often bind at TAD boundaries, but they can also facilitate long-range interactions between CREs [Özdemir and Gambetta, 2019]. With *rut*'s promoter located next to an insulator binding site (detected in the CNS at L3), CP190 might help regulate its interactions with surrounding enhancers. These observations are currently hypotheses, and a comprehensive analysis of the *rutabaga* locus with Hi-M will help determine their validity.

short NeuroPeptide F (*sNPF*)

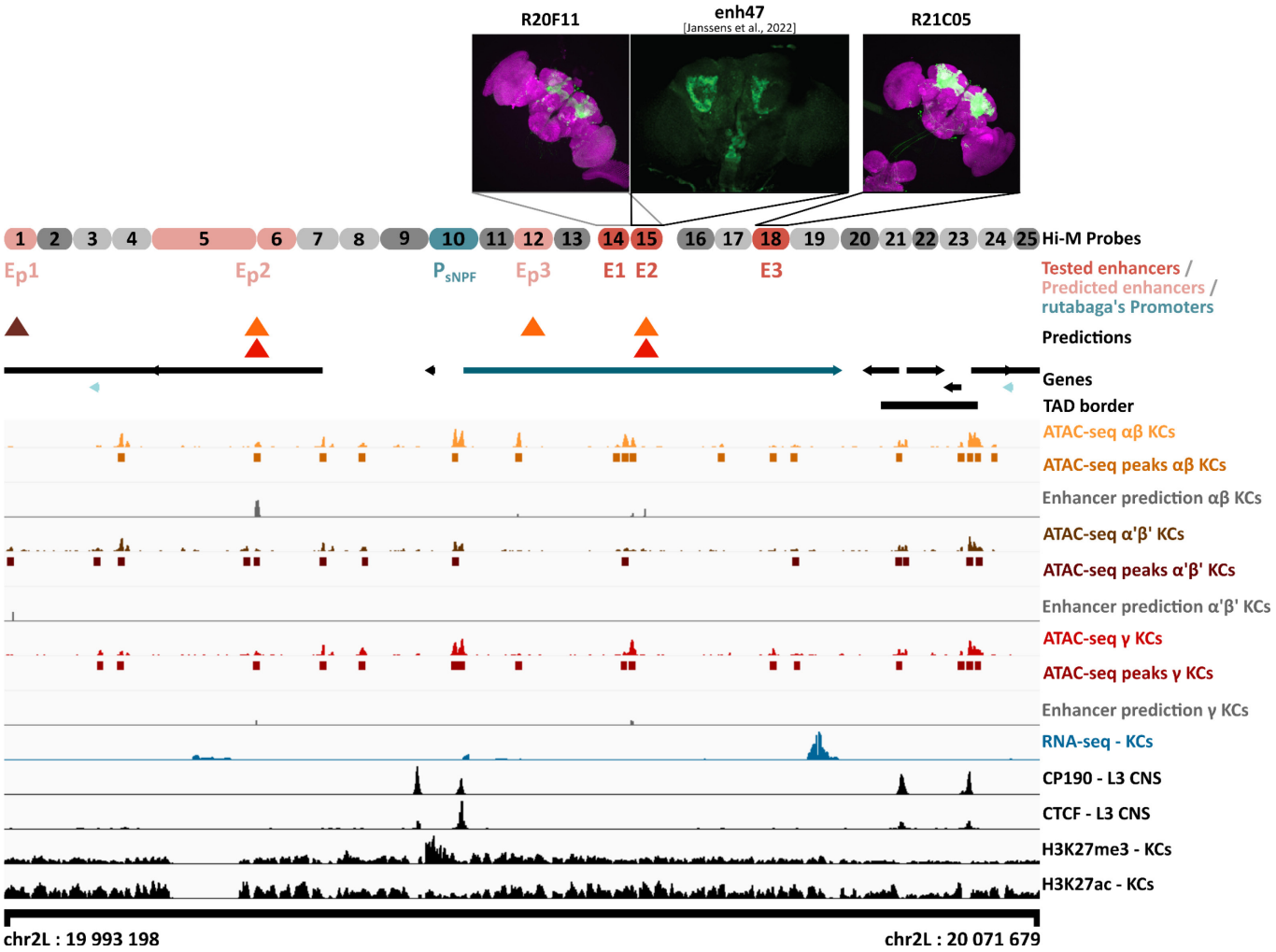


Figure 20 : Detailed view of the *sNPF* locus - From top to bottom : images of GAL-4 fly lines expression pattern from Janelia's Flylight project [Jenett et al., 2012] (R20F11 and R21C05) and from Stein Aerts' lab [Janssens et al., 2022] (enh47) of mushroom body specific enhancers. Hi-M probe repartition along the genome (*sNPF*'s promoter (teal) ; tested enhancers (red) ; enhancers predicted by [Janssens et al., 2022] (pink) ; accessible regions (light grey) ; inaccessible regions (dark grey)). Identification of tested and predicted enhancers. Predicted enhancers (γ specific (red triangles) ; α'β' specific (brown triangle) ; αβ specific (orange triangles) [Janssens et al., 2022]. Gene repartition (*sNPF*'s gene body (dark teal) ; lncRNAs (light blue) ; other genes (black)). TAD border defined by [Hug et al., 2017] (black rectangle). ATAC-seq data and ATAC-seq peaks from [Janssens et al., 2022] (γ KCs (orange) ; α'β' KCs (brown) ; αβ KCs (red)). Enhancer prediction tracks from [Janssens et al., 2022]. RNA-seq data in OK107 GAL-4 from [Henry et al., 2012]. CP190 and CTCF ChiP-seq data for L3 CNS from [Kaushal et al., 2021]. Histone marks in OK107 GAL-4 from [Henry et al., 2012] (H3K27me3 and H3K27ac). Genomic coordinates of the locus (chr2L:19993198-20071679).

Based on similar analyses, I designed an oligopaint library for the *sNPF* locus that spans the following genomic coordinates: chr2L:19993198-20071679 (78,5kb). Within this locus, there are promoters of seven different genes, two of which are lncRNAs (light blue genes). One of these is an antisense RNA (asRNA) called CR44907. Antisense RNAs are known to be involved in gene regulation and RNA degradation [Pelechano and Steinmetz, 2013].

The promoter of *sNPF* is located at probe 10 (P_{sNPF}). *sNPF* does not present any alternative promoter. In this region, three mushroom body-specific enhancers were validated through enhancer assays conducted by Janelia's Flylight project [Jenett et al., 2012] and Stein Aerts' lab [Janssens et al., 2022]. These enhancers overlap with probes 14 (R20F11), 15 (R20F11 and enh47) and 18 (R21C05) (see expression pattern on Figure 20) and are annotated E1, E2 and E3, respectively.

Four additional accessible regions were predicted by DeepFlyBrain [Janssens, 2022] to have more than a 50% probability of being KC-specific enhancers. These predicted enhancers are located at probes 1 (Ep1- α'β' specific), 5 (Ep2 - αβ & γ specific), 6 (Ep2 - αβ & γ specific) and 12 (Ep3 - αβ specific).
Furthermore, the locus contains four CTCF and CP190 binding sites [Kaushal et al., 2021]. All four are either situated near promoters or at the TAD border (probe 23) [Hug et al., 2017]. A slight increase in H3K27me3 is located between probes 9 and 10, before *sNPF*'s promoter.

We note that these loci are particularly complex, as they include many enhancers and promoters. In addition, we do not know the exact position of IBPs in the adult brain and our knowledge on cell type specific enhancers are, in some instances, based on predictions. Nonetheless, finding how chromatin is organised in different MB cell types will be important to understand what the role of 3D chromatin structure in gene regulation in the Drosophila adult brain may be.

Validation of Hi-M oligopaint libraries

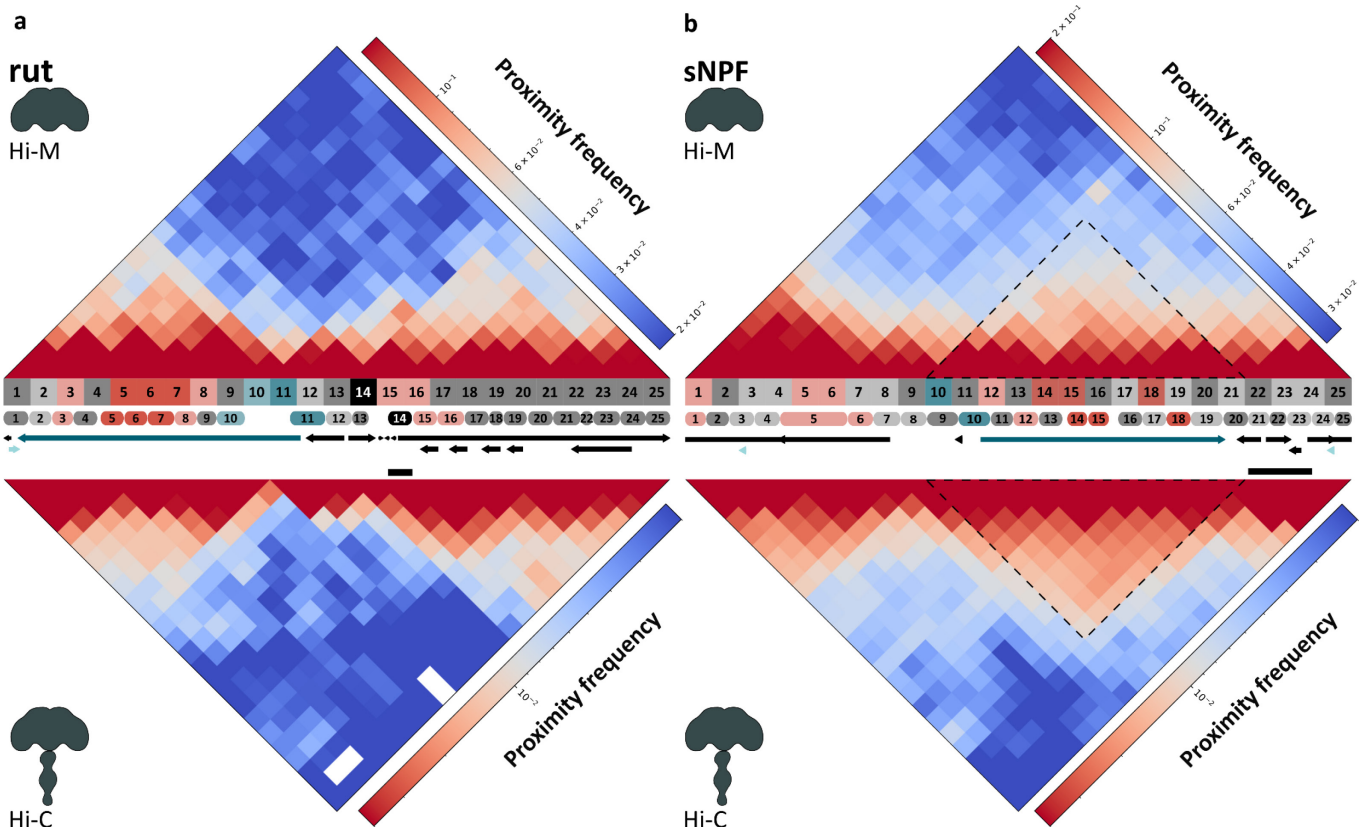


Figure 21 : Validation of the Hi-M libraries design of the *rut* and *sNPF* loci - (a-b) Comparison of the three-dimensional organisation between Hi-M and Hi-C methods. Hi-M probes are coloured accordingly : promoter (teal) ; tested enhancers (red) ; enhancers predicted by [Janssens et al., 2022] (pink) ; accessible regions (light grey) ; inaccessible regions (dark grey)). Genes are identified as follows : gene of interest (dark teal), lncRNAs (light blue) and other genes (black). Black rectangle corresponds to the TAD border. Dashed triangles represent a loss of interactions between Hi-M and Hi-C matrices. (a) Chromatin conformation of the *rut* locus : Hi-M matrix in the *Drosophila* adult

brain (top) and Hi-C matrix in the *Drosophila* adult CNS [Mohana et al., 2023] (bottom) in log scale. (b) Chromatin conformation of the *sNPF* locus : Hi-M matrix in the *Drosophila* adult brain (top) and Hi-C matrix in the *Drosophila* adult CNS [Mohana et al., 2023] (bottom) in log scale.

Once the libraries were bioinformatically designed and biochemically amplified, I conducted a series of validation experiments to further assess Hi-M in fly brains by comparing pairwise distance maps with the recently published Hi-C dataset in the adult CNS by [Mohana et al., 2023]. For this, I interpolated the Hi-C matrix using the coordinates of the probes used for Hi-M and compared the resulting interpolated Hi-C map to my data. The *rut* Hi-M matrix shows high similarities to the *rut* Hi-C map (Pearson correlation coefficient of 0.85, Figure 21a). For the *sNPF* locus, Hi-M and Hi-C maps are highly correlated (Pearson correlation coefficient of 0.9, Figure 21b), but we note that the Hi-M matrix exhibits slightly fewer interactions in the region located between probe 10 and probe 21 compared to the Hi-C matrix (dashed triangle, Figure 21b). Nevertheless, these results validate the ability of Hi-M to reconstruct chromatin conformation in the adult *Drosophila* brain.

Once the tools were developed and validated, I moved on to studying variations in interactions between gene regulatory regions in Kenyon cells to establish a link between chromatin conformation and transcription.

2.2.4. Part 3 : Investigation of the variations in interactions between gene regulatory regions in Kenyon cells to establish a link between chromatin conformation and transcription

2.2.4.1. 3D chromatin conformation of a gene similarly expressed in different cell types of the *Drosophila* adult brain in the *rut* locus

I first focused on the *rutabaga* gene due to its expression across all three subtypes of Kenyon cells and aimed to investigate the relationship between chromatin folding and gene regulation.

2.2.4.1.1. Chromatin conformation differences between the Kenyon cells and other brain cells in the *rut* locus

The chromatin conformation of the *rut* locus in Kenyon cells was compared to other brain cells in the brain (Figure 22, Supp. Figure 1 and 2).

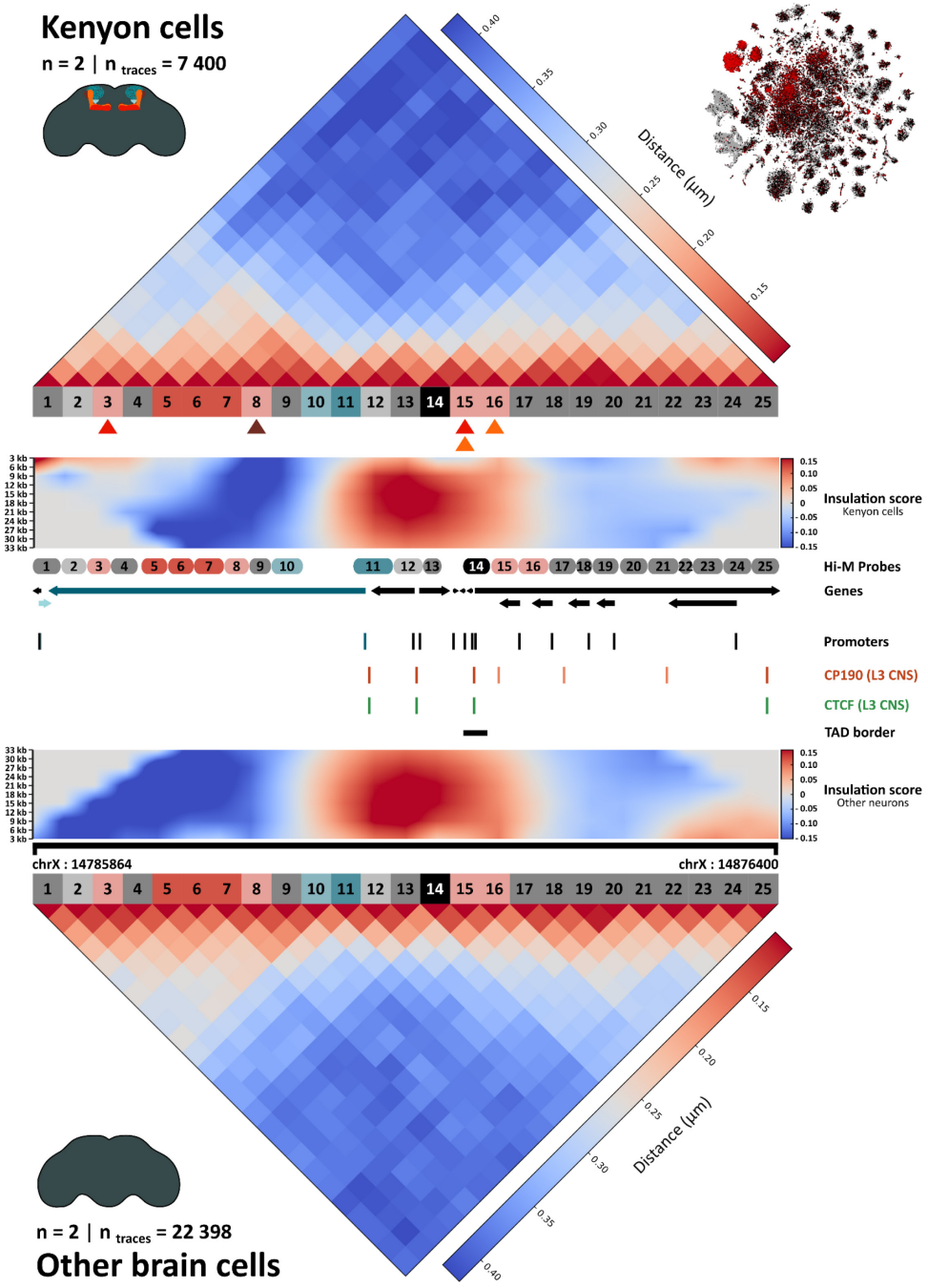


Figure 22 : Comparison of PWDs matrices between Kenyon cells and other brain cells for the *rut* locus - From top to bottom : PWDs median matrix of the *rut* locus in the Kenyon cells - t-SNE clustering for the visualisation of scRNA-seq dataset of the *Drosophila* adult brain from [Davie et al., 2018] of the *rut* gene - Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey) - Predicted enhancers : γ specific (red triangles), α'β' specific (brown triangle) and αβ specific (orange triangles) [Janssens et al., 2022] - Insulation score for different genomic bin sizes (from 3kb to 33kb) in the Kenyon cells - Repartition of the Hi-M probes along the genome - Genes are identified as follows : *rut*'s gene body (dark teal), lncRNAs (light blue) and other genes (black) - Promoters are identified as vertical lines : *rut* (dark teal), lncRNAs (light blue) and other genes (black) - CP190 binding sites in L3 CNS (orange) [Kaushal et al., 2021] - CTCF binding sites in L3 CNS (green) [Kaushal et al., 2021] - Black rectangle corresponds to the TAD border - Insulation score for different genomic bin sizes (from 3kb to 33kb) in the non-Kenyon cells/other cells of the brain - Genomic coordinates of the locus (chr2L:19993198-20071679) - PWDs median matrix of the *rut* locus in the other cells of the *Drosophila* adult brain

While the pairwise distance matrices (PWDs) for both conditions appear similar (Figure 22), the difference matrix reveals shorter distances between enhancers and promoters (black arrows), among promoters (green arrow), and among enhancers specifically in Kenyon cells (red arrow, Figure 23). Notably, the *rut* promoter is in close proximity to the three mushroom-body-specific enhancers (E1, E2, and E3) in KCs where *rut* is expressed, consistent with a classical contact model of enhancer action (Figure 23).

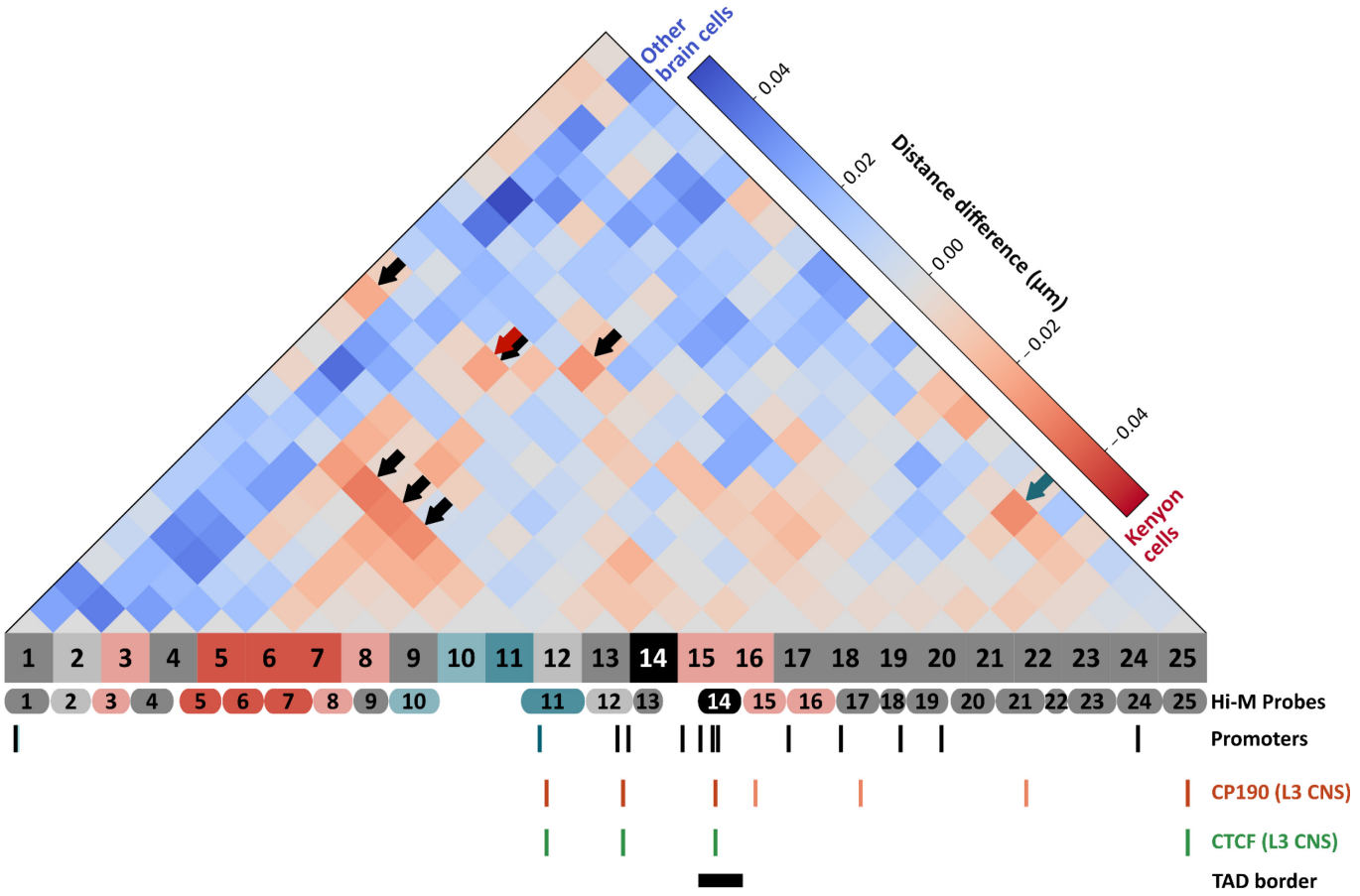


Figure 23 : Differences between Kenyon cells and the other brain cells of the *Drosophila* adult brain for the rut locus - From top to bottom : PWDs difference matrix of Kenyon cells minus other brain cells, interactions specific to Kenyon cells are in red and interactions specific to the rest of the brain are in blue. E-P interactions (black arrows), E-E interactions (red arrow) and P-P interactions (teal arrow) - Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey) - Repartition of the Hi-M probes along the genome - Promoters are identified as vertical lines : *rut* (dark teal), lncRNAs (light blue) and other genes (black) - CP190 binding sites in L3 CNS (orange) [Kaushal et al., 2021] - CTCF binding sites in L3 CNS (green) [Kaushal et al., 2021] - Black rectangle corresponds to the TAD border.

To provide a more quantitative analysis of chromatin organisation in this locus, insulation scores were extracted from the PWDs matrices [Crane et al., 2015][Hug et al., 2017][Götz et al., 2022][Messina et al., 2022] (Figure 12b-c-d). With our Hi-M data, insulation scores are calculated based on PWD matrices, strong insulation is characterised by positive values [Götz et al., 2022], [Messina et al., 2022], while in Hi-C data it corresponds to negative values [Crane et al., 2015], [Hug et al., 2017].

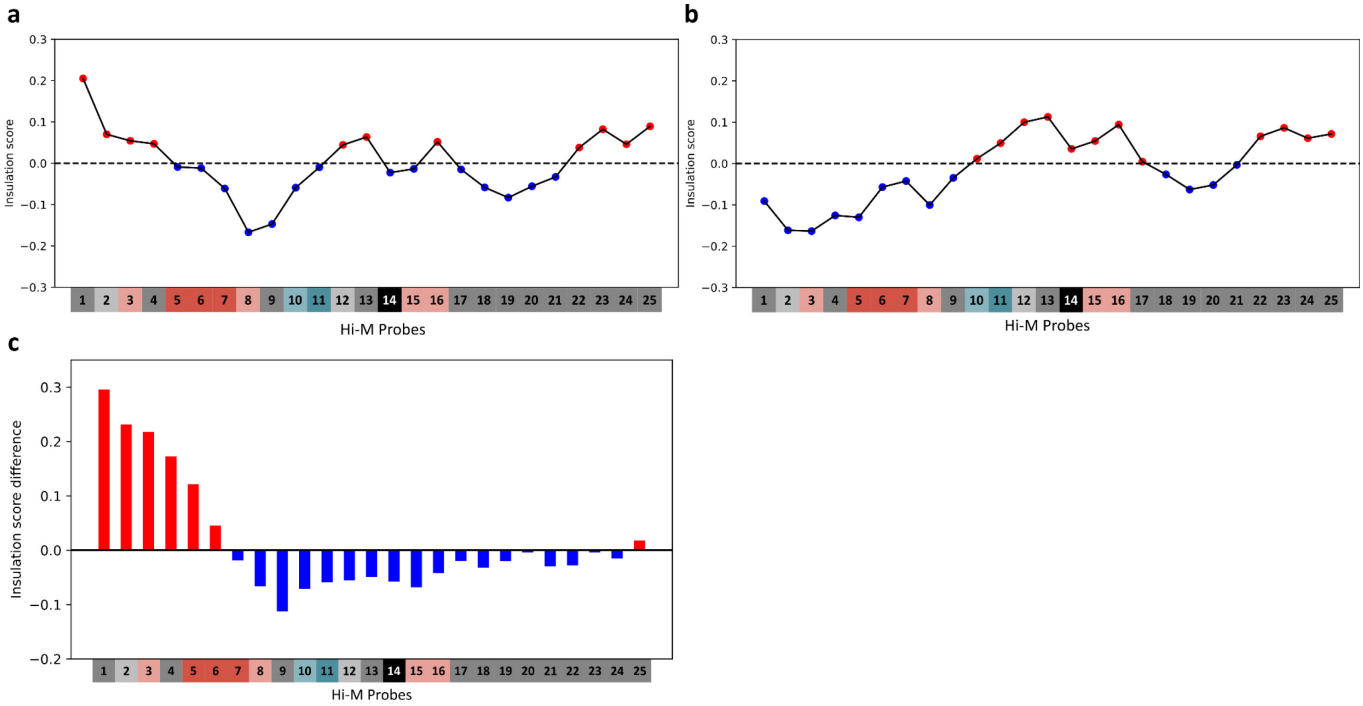


Figure 24 : Insulation in Kenyon cells and the other cells of the *Drosophila* adult brain for the *rut* locus - (a-b-c) Insulation score of the *rut* locus calculated for one bin of a 3kb size. Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). (a-b) Low insulation scores are in blue. High insulation scores are in red. (a) Insulation score in the Kenyon cells (b) Insulation score in the other cells of the brain (c) Insulation score difference between the Kenyon cells and the other cells of the brain calculated for one bin of a 3kb size. Red highlights that the values are more Kenyon cells specific while blue values are more other cells specific.

Notable variations in insulation were observed between Kenyon cells and other cells (Figure 24c). While insulation in other cells steadily increases toward the topologically associated domain (TAD) border (between probe 3 and 13) (Figure 24b), insulation in Kenyon cells decreases until probe 8 before following a similar trend in other brain cells (Figure 24a). This distinct insulation pattern around mushroom body-specific enhancers and the *rut* promoter suggests the existence of a smaller regulatory domain potentially involved in *rutabaga* transcriptional regulation within Kenyon cells. The shorter distances between these enhancers and the *rut* promoter in the PWDs difference matrix support this hypothesis (Figure 23).

Kenyon cell enhancers more frequently co-localise with the *rut* promoter in KCs with respect to other brain cells (Figure 23). We also note additional changes in 3D chromatin organisation between KC and brain cells that do not involve the *rut* promoter or its enhancers. These differences may involve the action of insulators or interactions between the many other regulatory elements in this locus.

Next, I examined whether differences in E-P interactions or overall chromatin conformation exist among the different subtypes of Kenyon cells, despite *rutabaga*'s similar gene expression levels in αβ, α'β', and γ Kenyon cells.

2.2.4.1.2. Relationship between chromatin folding and gene expression of the *rut* locus within the Kenyon cells subtypes

2.2.4.1.2.1. Global chromatin structure and insulation of the *rut* locus

Following the comparison of the Kenyon cells against the other cells of the brain, the chromatin conformation of the *rut* locus in the three main KC subtypes were compared (Figure 25, Supp. Figure 3, 4, 5 and 6). The Hi-M maps show that KC subtypes share several common features, such as a strong TAD border or a sub-TAD inside the first TAD (located between probe 4 and 11). However, they also present differences in their global chromatin conformation: For instance, only the α'β' cell type exhibits, in the second TAD, two substructures and a strong interaction around probes 16 (Ep3) and 23 (Figure 25a), which does not fit with an E-P interaction. In contrast, αβ cells show reduced interactions within the second TAD and a sharper border around probe 11 (Figure 25b). To further explore these observations, the insulation score was examined for the three conditions.

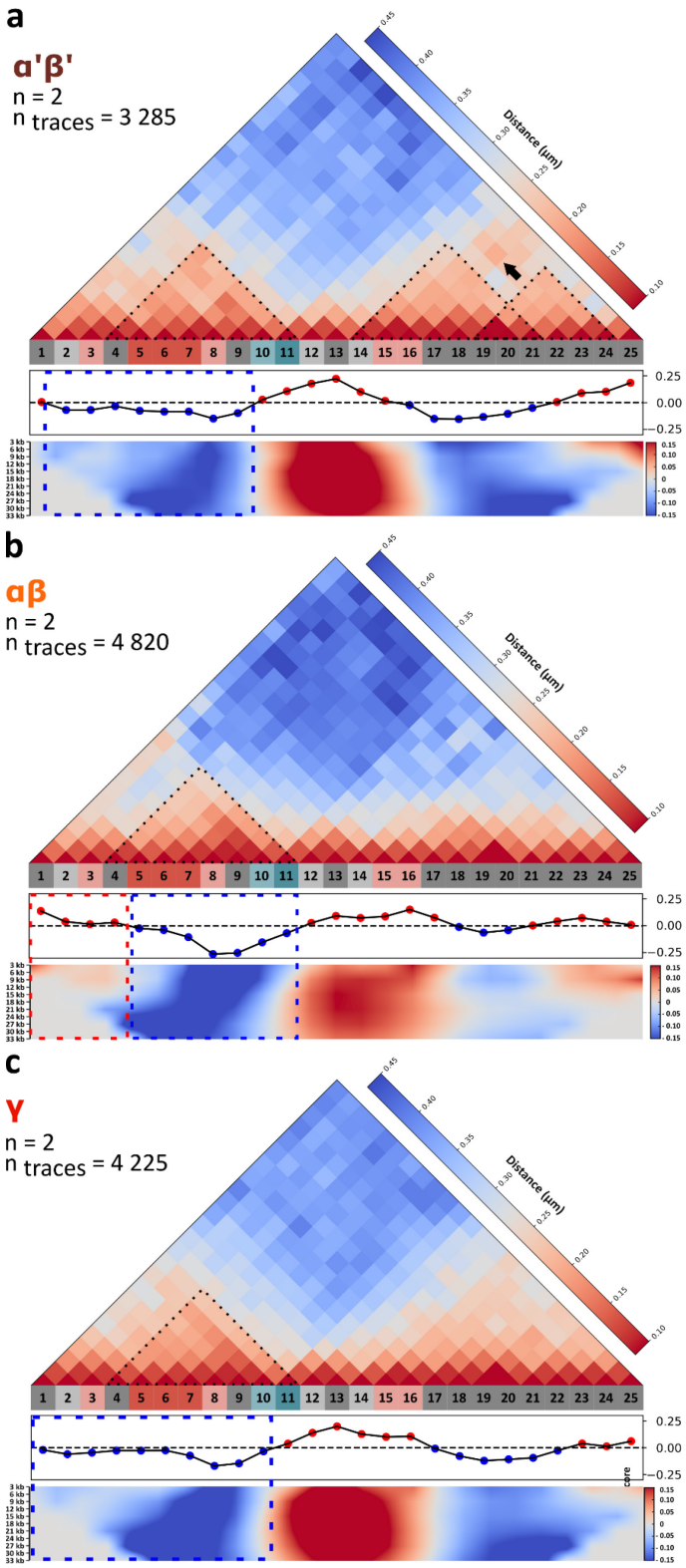


Figure 25 : Comparison of PWDs matrices and insulation scores in Kenyon cells subtypes for the *rut* locus

- (a-b-c) From top to bottom : PWDs median matrix of the *rut* locus in the different Kenyon cells subtypes (a) α'β' Kenyon cells (b) αβ Kenyon cells and (c) γ Kenyon cells. Dashed black lines highlight the different substructures of the region. Dashed red and blue lines highlight high or low insulation scores, respectively, in the first TAD. Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). - Insulation score of the *rut* locus calculated for one bin of a 3kb size. Low insulation scores are in blue. High insulation scores are in red. - Insulation score for different genomic bin sizes (from 3kb to 33kb).

All three subtypes exhibit strong insulation in the middle of the *rut* locus, consistent with the two TADs being insulated in all cell types.

The first TAD presents insulation differences in αβ cells compared to the other two subtypes. While a sub-TAD was visible on the PWDs matrices of all three cell types, only the αβ Kenyon cells show a strong insulation transition on the insulation score diagram and the domainogram, giving rise to what seems like two sub-TADs (dashed red and blue lines, Figure 25b). The first one of the two sub-TADs is highly insulated between probes 1 and 4 (dashed red lines, Figure 25b), while the second one, located between probes 5 and 11 (dashed blue lines, Figure 25b), presents two-fold less insulation in αβ cells than the other KCs subtypes. The PWD matrix of αβ cells presents fewer interactions in the second TAD compared to the other two conditions (Figure 25b). The insulation score shows similar results : the αβ condition presents a weaker TAD border and a globally more insulated second TAD than the other two conditions (Figure 25b). These findings suggest that even though these three sub cell types belong to the same neuronal type, they have different genome folding patterns.

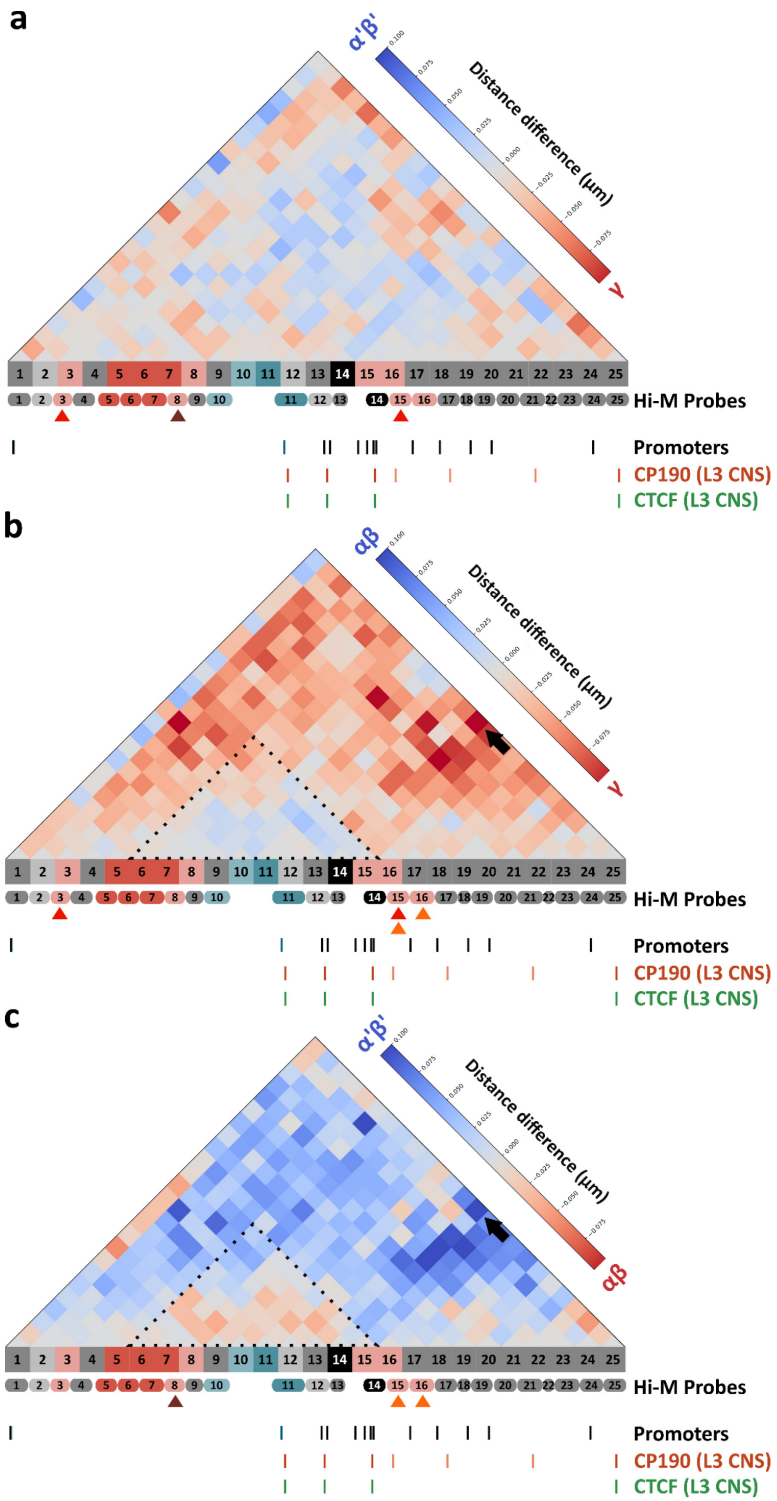


Figure 26 : PWDs difference matrices of Kenyon cells subtypes for the *rut* locus - (a-b-c) From top to bottom : PWDs difference matrices for (a) γ minus α'β' (b) γ minus αβ and (c) αβ minus α'β' Kenyon cells. Dashed black lines highlight cell type specific interactions within the locus. Black arrows indicate interactions between CTCF and CP190 binding sites. Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). Then, repartition of the Hi-M probes along the genome. Predicted enhancers are identified in the following manner : γ specific (red triangles), α'β' specific (brown triangles) and αβ specific (orange triangles) [Janssens et al., 2022]. Promoters are identified as vertical lines : *rut* (dark teal), lncRNAs (light blue) and other genes (black). CP190 binding sites in L3 CNS are identified as orange vertical lines [Kaushal et al., 2021] - CTCF binding sites in L3 CNS are identified as green vertical lines [Kaushal et al., 2021].

PWDs difference matrices were generated (by subtracting the different KCs subtypes matrices) to extract further differences between the sub cell types. In αβ cells, the *rut* locus is generally less condensed than in the other two conditions (i.e. most bins are red in 26b and blue in 26c), except between probe 6 (enhancer) and probe 15 (promoter of CG12539 and predicted enhancer) (dashed lines, Figure 26b-c). This region includes *rut*'s promoter (including the alternative one), tested enhancers, predicted enhancers, and the TAD border. These lower distances between regulatory elements in αβ cells suggest a potential difference in transcription regulation of *rut* in αβ cells compared to the other two sub cell types despite their similar expression levels.

In contrast, larger distances for αβ cells compared to γ and α'β' cells were observed around CP190 binding sites (measured in L3 CNS [Kaushal, 2021] between probes 14 and 25 (black arrows) (Figure 26b-c). Probe 14 is not only a binding site for CTCF and CP190, but it also corresponds to the TAD border and to the position of three promoters (CG14411, CG14407, and Flo2). This enrichment of interactions for the region bordered by probes 14 and 25 could highlight the role of these regions in interactions mediation for γ and α'β' Kenyon cells. Although all KC subtypes express *rut* similarly but at slightly different levels, they exhibit differences in their global chromatin folding that involve *rut*'s regulatory and non-regulatory regions.

In this section, we did not explore the interactions between P_{rut} and most of the enhancers of the locus since their differential distances are fairly small and within the error range of the values (see Supp. Figures 4b-d, Figure 6b-d). In most cases, trying to correlate E-P cell-type specific interactions with *rut*'s expression in the Kenyon cells does not seem statistically relevant.

Nevertheless, by comparing the chromatin conformations of αβ, γ and α'β' Kenyon cells, I was able to identify cell type-specific interactions, structures and insulation patterns. The αβ cells present a region with increased interactions compared to the other cell types, including P_{rut}, E2, E3, Ep2 (α'β' specific) and Ep3 (αβ and γ specific). To study in more detail the interactions between *rut*'s promoter and the surrounding genomic region, I turned to the double Gaussian fitting method presented in the following section.

2.2.4.1.2.2. Introduction to one of the analysis tools used to extract complementary information from Hi-M data

Hi-M, like many other sequential DNA-FISH techniques [Wang et al., 2016][Mateo et al., 2019][Cardozzo Gizzi et al., 2019][Liu et al., 2020][Su et al., 2020][Takei et al., 2021], relies on measuring distances between different genomic regions to reconstruct the 3D chromatin conformation of the locus of interest. Pairwise distances between each pair of probes within the locus are recorded and represented in a histogram map (Figure 27a).

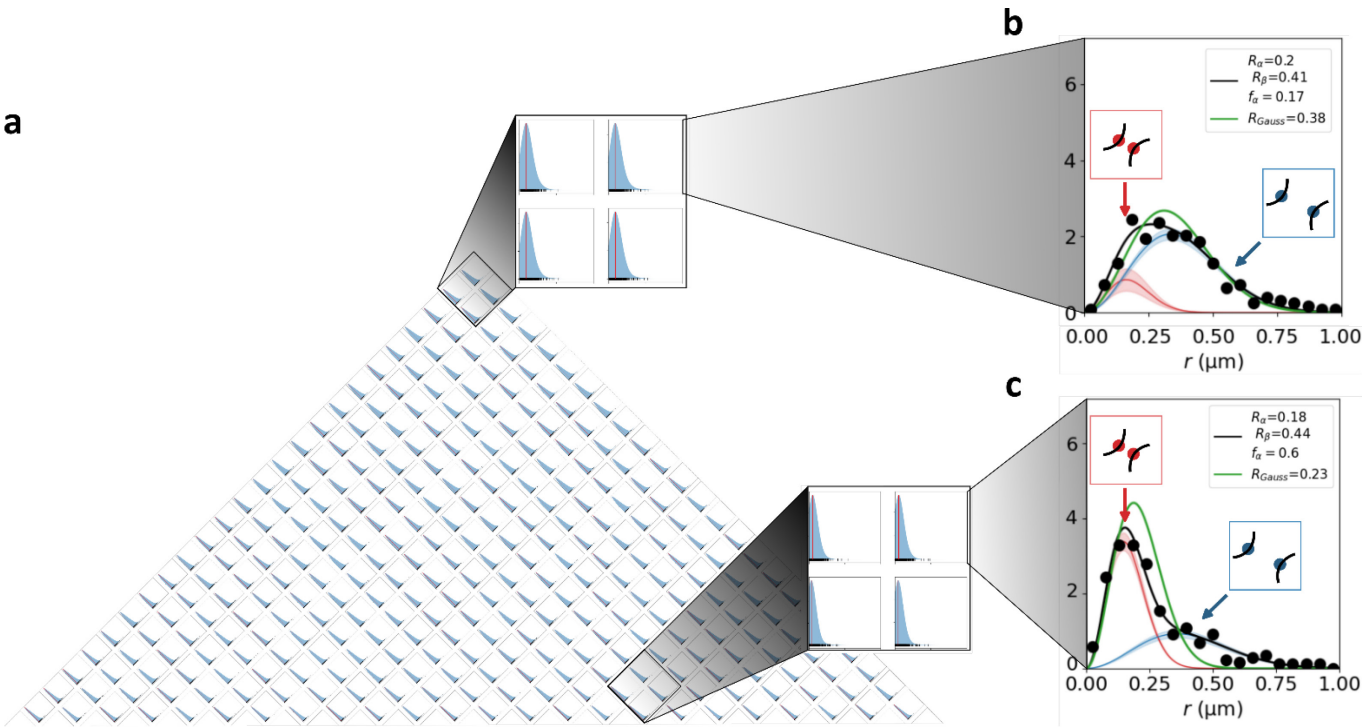


Figure 27 : PWD histograms and double gaussian fit model - (a) Distances between each pair of probes within the locus represented in a histogram (b) Double gaussian fitting model for genomically distant probes - experimental data (black curve and black dots) - one Gaussian fit (green) - two Gaussian fit : alpha phase (red) and beta phase (blue) (c) Double gaussian fitting model for genomically close probes - experimental data (black curve and black dots) - one Gaussian fit (green) - two Gaussian fit : alpha phase (red) and beta phase (blue)

These histograms are commonly fitted by a Kernel density estimator that considers the data as arising from a Chi distribution which assumes the distribution in distances for each dimension is represented by a single Gaussian. Recently, a paper from a collaborator [Remini et al., 2024] has suggested that these distances might encapsulate two polymer regimes, thus advocating for a fitting of the PWD distributions assuming that the distances in each dimension could be represented by two Gaussians rather than one (Figure 27a). The green curve shows the "one gaussian model" that fails to fit the experimental data (Figure 27b). This novel approach allows for the detection of more than one chromatin conformation.

The analysis by *Remini et al.* (2024) was originally conducted on multiplexed DNA-FISH data from *Bintu et al.* (2018) in human cells and was adapted, with the assistance of Loucif Remini and Andrea Parmeggiani to analyse our Hi-M data from *Drosophila* adult brains.

The two Gaussian distributions can be extracted for each histogram of probe pairs. The first distribution, termed “alpha phase”, depicts short distances, typically under 250 nm that could describe a bound state or a close conformation (Figure 27b-c). The second population, the “beta phase”, accounts for the histogram tail, indicating more distant conformations or unbound state (Figure 27b-c). From these distributions, three parameters can be derived :

- The mean values of the alpha (R_α) and beta (R_β) Gaussians
- The fraction of cells in the alpha phase (f_α), which indicates the prevalence of probes to be in the close conformation. This value ranges from 0 to 1, with 1 indicating that only the alpha phase is represented in all the measured cells.

These parameters offer a new perspective on our data, providing insights into probe-probe interactions through f_α.

2.2.4.1.2.3. Analysis of *rut*'s promoter interactions with the surrounding region using the fraction of the alpha phase

Here, we will delve into the analysis of the fraction of the alpha phase (f_α). Here, the f_α is studied using *rut*'s promoter as an anchor. Thus, we will be looking at how often *rut*'s promoter is in the alpha phase (short distances) with the other probes of the locus. This analysis summarises a common work made with Loucif Remini and Andrea Parmeggiani and should lead to a common publication in the near future.

The goodness of fit (ΔR^2) describes the quality of fit of the theoretical model compared to experimental results (Figure 28). Although the two gaussians model always work, the fit of histograms of distant probes showed high uncertainties in the determination of the fit parameters. This means that many sets of parameters would have worked, even the ones in which f_α might have near extreme values (0 or 1). Thus, we established a criterion ΔR^2 that estimates the goodness of both fits. Based on this criterion we always privilege the minimal model (namely the gaussian one). Values within the green window are better represented by a single gaussian and therefore one single typical distance (short or large, i.e., alpha or beta depending on the threshold used), in any case, a single phase and not two possible phases. Therefore, in the following analysis, we have ignored probes located within this green window (Figure 28). For example, a switch in the fit can be observed and correlated with the TAD border (Figure 28). Interestingly, the TAD border limits close proximity between the region on the left and right of the border.

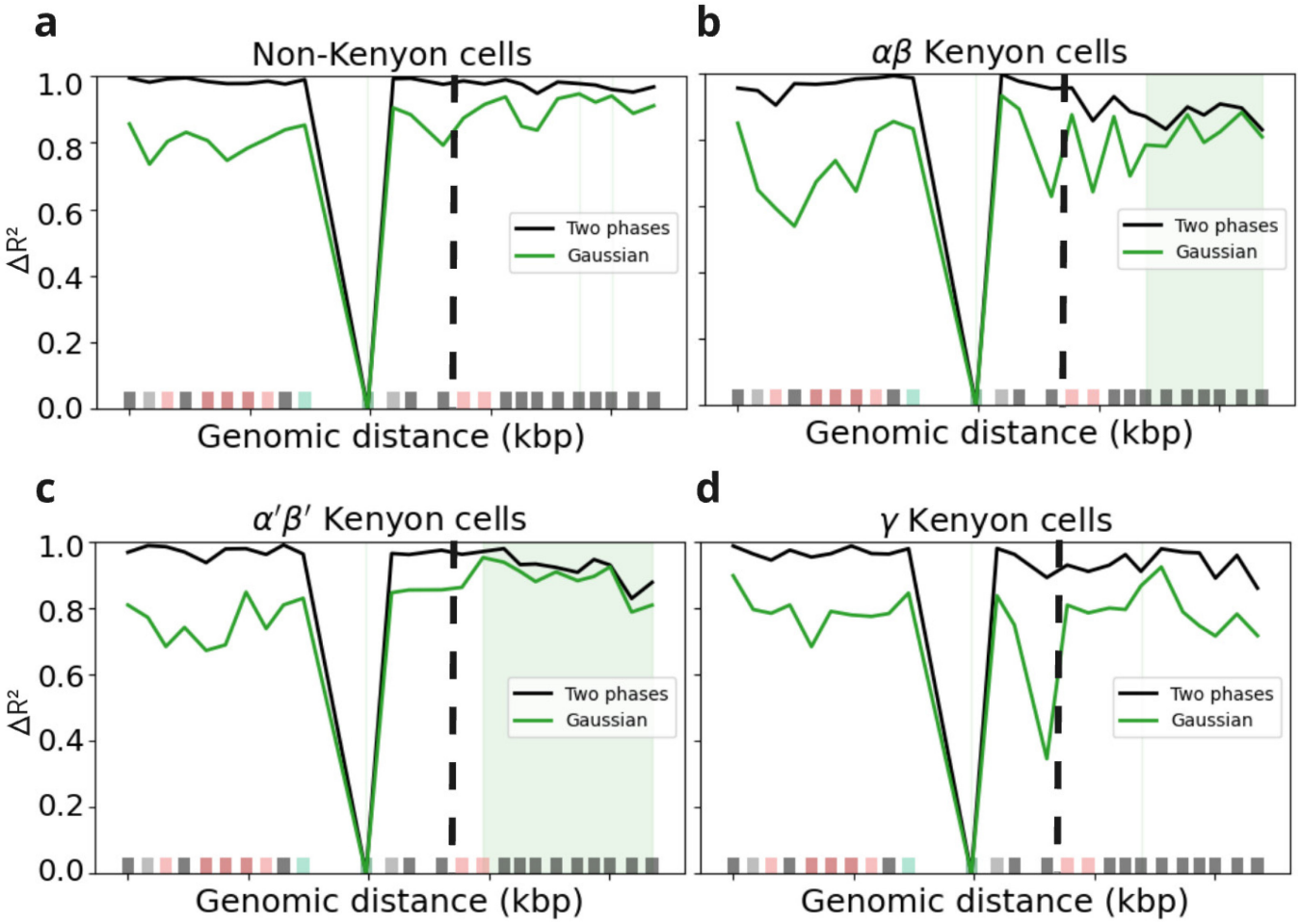


Figure 28 : Goodness of fit of the theoretical model compared to experimental results - (a-b-c-d) Goodness of fit as a function of the genomic distance with *rut*'s promoter (teal) as anchor of the four conditions : (a) non-Kenyon cells/other cells of the brain (b) αβ Kenyon cells (c) α'β' Kenyon cells and (d) γ Kenyon cells. Two Gaussians fit (black curve). One Gaussian fit (green curve). The genomic location for which the one Gaussian fits best the experimental data, thus for our analysis these data have been ignored as it does not present the two dynamic regimes (light green vertical lines). TAD border (dashed vertical black lines). Repartition of the Hi-M probes along the genome and coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey).

Common enhancers

In all three KC sub cell types, the distances between the promoter and enhancers E2 and E3 are predominantly shorter (red arrows) compared to the rest of the locus (Figures 29b-c-d). On average, the f_α of P-E2 and P-E3 are above the averaged f_α ($f_{\alpha\ avg}$) (for P-E2 : $f_{\alpha\ avg}$(αβ) = 44%, $f_{\alpha\ avg}$ (α'β') = 43% and $f_{\alpha\ avg}$ (γ) = 36% ; for P-E3 : $f_{\alpha\ avg}$ (αβ) = 41%, $f_{\alpha\ avg}$ (α'β') = 42% and f_α (γ) = 35%) of the *rut* locus, although the number of cells in alpha phase varies between cell types (for P-E2 : f_α(αβ) = 51% ± 5%, f_α (α'β') = 50% ± 3% and f_α (γ) = 69% ± 4% ; for P-E3 : f_α (αβ) = 51% ± 3%, f_α (α'β') = 61% ± 5% and f_α (γ) = 57% ± 3%). In other brain cells, these two enhancers do not show significantly shorter distances with the promoter compared to the rest of the locus (Figure 29a). This result is consistent with our previous observations of relative pairwise distance changes between Kenyon cells and other cells of the brain. As these two enhancers are active in the mushroom body (Figure 8), this enrichment in the proximal state is consistent with E2 and E3 regulating *rut* expression in all three sub cell types by a contact mode of action.

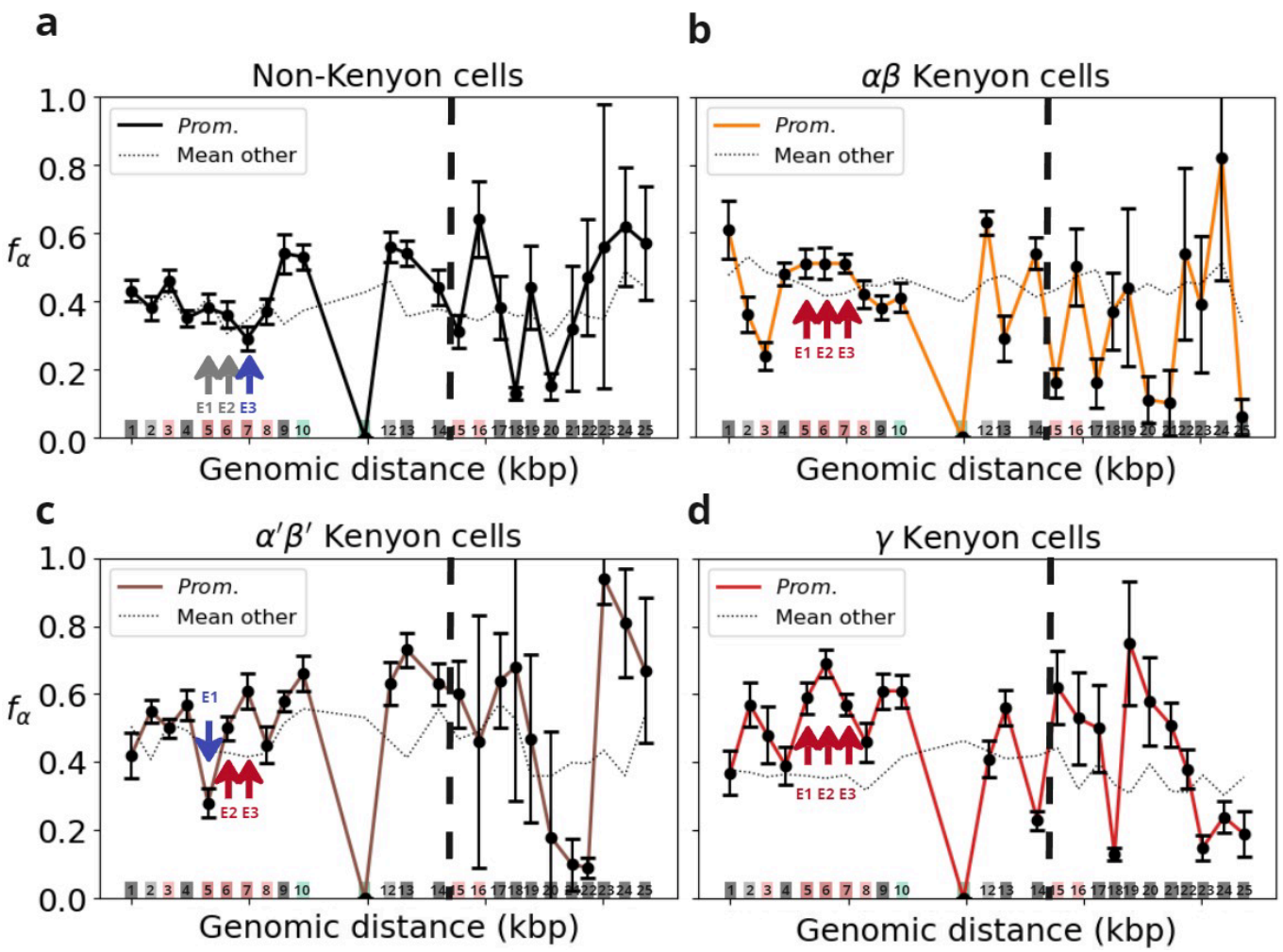


Figure 29 : The fraction of the alpha phase highlights cell type specific interactions between *rut*'s promoter and the surrounding region - (a-b-c-d) Fraction of the alpha phase f_α as a function of the genomic distance with *rut*'s promoter (teal) as anchor of four conditions (a) non-Kenyon cells/other cells (black curve) (b) αβ Kenyon cells (orange curve) (c) α'β' Kenyon cells (brown curve)and (d) γ Kenyon cells (red curve). Error bars represent the standard deviation of f_α. Averaged f_α of the *rut* entire locus is represented with a horizontal black dashed line. Red arrows are for increased f_α between *rut*'s promoter and the enhancer compared to the averaged f_α of the *rut* entire locus. Blue arrows are for decreased f_α between *rut*'s promoter and the enhancer compared to the averaged f_α of the *rut* entire locus. Grey arrow is for unchanged f_α between *rut*'s promoter and the enhancer compared to the averaged f_α of the *rut* entire locus. The TAD border is a dashed vertical black line. Hi-M probes are positioned according to their genomic location and are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey).

Subtype-specific enhancers and predictions

Interestingly, interactions between *rut*'s promoter and enhancer E1 differ from the other two mushroom body-specific enhancers (E2 and E3). P_{rut} and E1 are in proximity in only 28% ± 4% of α'β' Kenyon cells (Figure 18c) compared to 59% ± 4% (Figure 29d) and 51% ± 4% (Figure 29a) in γ and αβ cell types, respectively. As a reminder, E1 and E2 Hi-M probes are located within the same fragment of DNA tested by Janelia's fly light project (Figure 19). Based on these results and the E-P contact model, E1 is most likely active in γ and αβ Kenyon cells and inactive in the α'β' cell type.

From a prediction perspective, data from *Janssens et al* (2022), do not predict E1, E2 and E3 as enhancers. However, E1 and E3 are accessible in all KCs subtypes while E2 presents an accessibility peak only in γ KCs. The accessibility of E3 is consistent with the activity of this enhancer in mushroom bodies (Figure 18). The enhancer E1 is accessible in all three subtypes although the accessibility is lower in α'β' Kenyon cells (Figure 19), which seems to be consistent with our results. The lack of accessibility of E2 in γ cell type, while still presenting E-P interactions in all three PWDs matrices is surprising, this result possibly highlights that accessibility may not be the only factor for E-P interactions. Thus, understanding cell type specificity of enhancers in the *rut* locus from prediction and accessibility is complex and awaits further experimental validation.

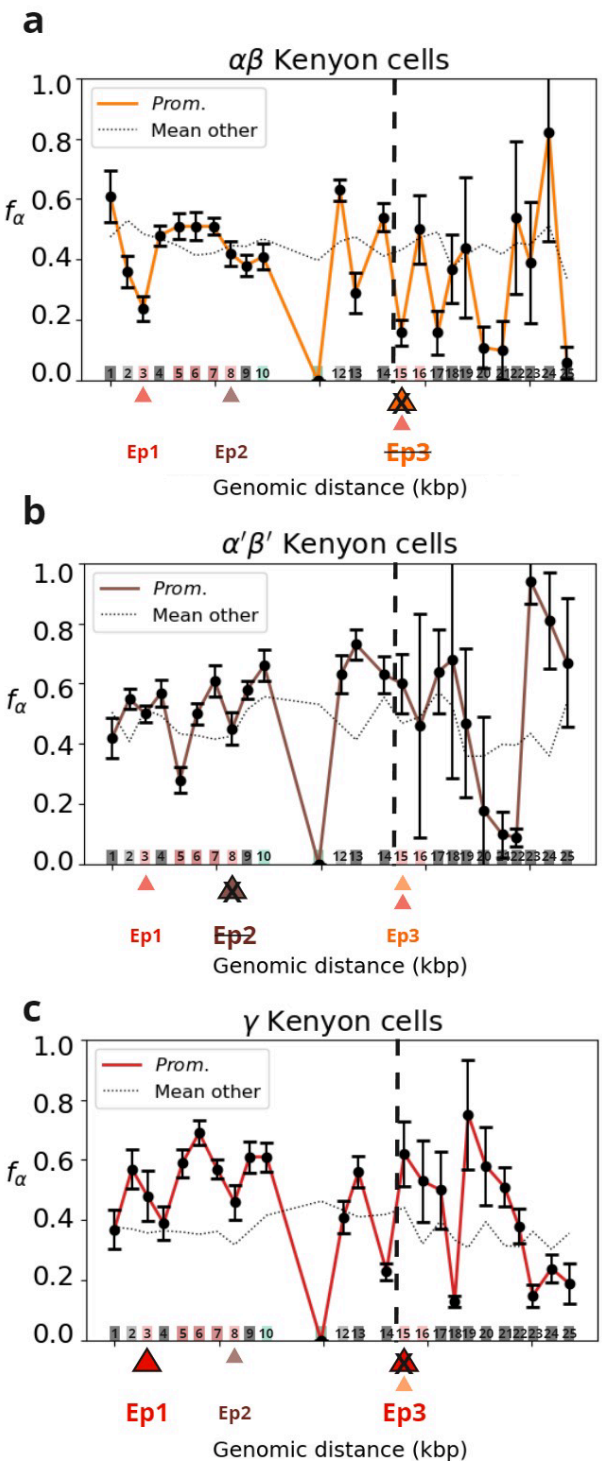


Figure 30 : The fraction of the alpha phase highlights differences between Hi-M data and enhancer predictions for the *rut* locus - (a-b-c-d) Fraction of the alpha phase f_α as a function of the genomic distance with *rut*'s promoter (teal) as anchor of four conditions (a) αβ Kenyon cells (orange curve) (b) α'β' Kenyon cells (brown curve) and (c) γ Kenyon cells (red curve). Error bars represent the standard deviation of f_α. Averaged f_α of the *rut* entire locus is represented with a horizontal black dashed line. The TAD border is a dashed vertical black line. Hi-M probes are positioned according to their genomic location and are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). Outlined and enlarged triangles highlight the position of cell type specific predicted enhancers : αβ specific (orange triangles), α'β' specific (brown triangles) and γ specific (red triangles) [Janssens et al., 2022]. Crosses indicate that the prediction does not fit with the results.

Interactions between *rut*'s promoter and the predicted enhancers also present interesting behaviours (Figure 30). Earlier in the manuscript, Ep1 (probe 3) was identified as γ specific, Ep2 (probe 8) was predicted to be α'β' specific and Ep3 (probe 15) to be αβ and γ specific [Janssens et al., 2022]. From our analysis of f_α, the prediction of Ep1 seems to be consistent with more frequent interactions with *rut*'s promoter only in γ Kenyon cells (Figure 30c). Then, Ep2 does not present shorter distances in α'β' Kenyon cells as we could have expected. Although Ep3 is predicted in both γ and αβ Kenyon cells, only γ and α'β' cells present shorter distances with P_{rut}. The regulation of *rut*'s promoter by Ep2 and Ep3 might be different from the one of Ep1, thus explaining the differences between predictions and observed interactions. Instead of following the contact mode of action, these two enhancers might fit with the "action-at-a-distance" model for E-P interactions. Thus, enhancer predictions do not consistently account for enhancer-promoter interactions observed with the double gaussian model. To validate whether our observations account for enhancer activity, enhancer activity assays will be needed.

By using the fraction of f_α, differences that were not visible in the PWDs matrices were extracted and the goodness factor ΔR^2 helped exclude interactions in which a two dynamics regime was not present. This analysis highlights cell type-specific interactions between a promoter and surrounding cell-type specific enhancers even in a gene similarly expressed in all three sub cell types, such as *rut*.

2.2.4.1.3. Key points of the Hi-M analysis of the *rut* locus

Variations in chromatin folding patterns

α'β' cells

- Exhibit unique substructures in the second TAD

αβ cells

- Exhibit a unique insulation signature in the first TAD
- Show reduced interactions within the second TAD
- Present sharper TAD borders

Insulation Score

- Significant differences in insulation scores were noted among the Kenyon cell subtypes

Enhancer-promoter proximities

Common enhancers (E2 and E3)

- The distances between *rut*'s promoter and enhancers E2 and E3 were predominantly shorter in Kenyon cells compared to other cells, indicating a role in gene regulation

γ Kenyon cells

- Exhibit higher interaction frequencies between the *rut* promoter and surrounding enhancers, correlating with slightly stronger expression of *rut* in these cells

Cell type-specific enhancer (E1)

- The interaction between *rut*'s promoter and enhancer E1 varied significantly among Kenyon cell subtypes, suggesting cell type-specific regulation

Enhancer prediction

- Enhancers predicted by [Janssens et al., 2022] do not consistently account for enhancer-promoter contact

2.2.4.2. 3D chromatin conformation of a gene differentially expressed in different cell types of the *Drosophila* adult brain for the *sNPF* locus

Following the results of the *rutabaga* gene, I investigated further the relationship between chromatin folding and gene regulation in a locus containing the differentially expressed gene *sNPF*.

2.2.4.2.1. Chromatin conformation differences between the Kenyon cells and the other cells for the *sNPF* locus

The chromatin conformation of the *sNPF* locus in Kenyon cells was first compared to other cells in the brain (Figure 31, Supp. Figure 7 and 8). Similarly to the *rut* locus, the PWDs matrices of Kenyon cells and other cells in the *sNPF* locus are almost identical (Pearson correlation coefficient of ~1.0) (Supp. Figure 7c). However, distinctions become evident when examining the difference matrix and insulation score (Figure 31).

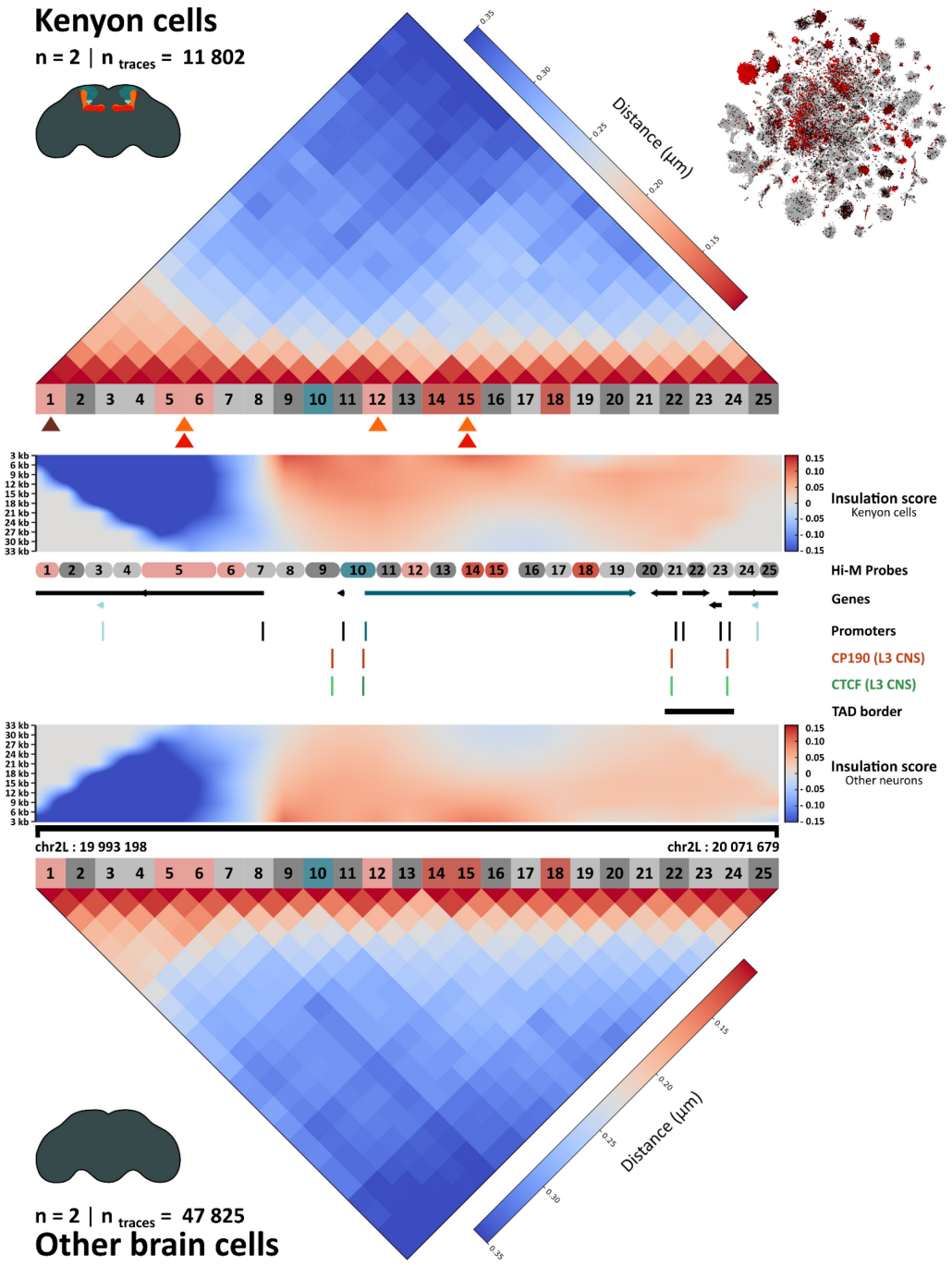


Figure 31 : Comparison of PWDs matrices between Kenyon cells and other brain cells for the *sNPF* locus - From top to bottom : PWDs median matrix of the *sNPF* locus in the Kenyon cells - t-SNE clustering for the visualisation of scRNA-seq dataset of the *Drosophila* adult brain from [Davie et al., 2018] of the *sNPF* gene - Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey) - Predicted enhancers : γ specific (red triangles), α'β' specific (brown triangle) and αβ specific (orange triangles) [Janssens et al., 2022] - Insulation score for different genomic bin sizes (from 3kb to 33kb) in the Kenyon cells - Repartition of the Hi-M probes along the genome - Genes are identified as follows : *sNPF*'s gene body (dark teal), lncRNAs (light blue) and other genes (black) - Promoters are identified as vertical lines : *sNPF* (dark teal), lncRNAs (light blue) and other genes (black) - CP190 binding sites in L3 CNS (orange) [Kaushal et al., 2021] - CTCF binding sites in L3 CNS (green) [Kaushal et al., 2021] - Black rectangle corresponds to the TAD border - Insulation score for different genomic bin sizes (from 3kb to 33kb) in the non-Kenyon cells/other cells of the brain - Genomic coordinates of the locus (chr2L:19993198-20071679) - PWDs median matrix of the *sNPF* locus in the other brain cells of the adult *Drosophila*.

In contrast to the *rut* difference matrix, the *sNPF* matrix shows subtler (maximum pairwise difference between probes is 30 nm for *sNPF* vs. 45 nm for *rut*) and more generalised changes between regions that are not necessarily accessible (dark grey probes Figure 32). Nevertheless, some differences involving E-E interactions (red arrows) and E-P interactions (black arrows) specific to Kenyon cells are observed (Figure 32).

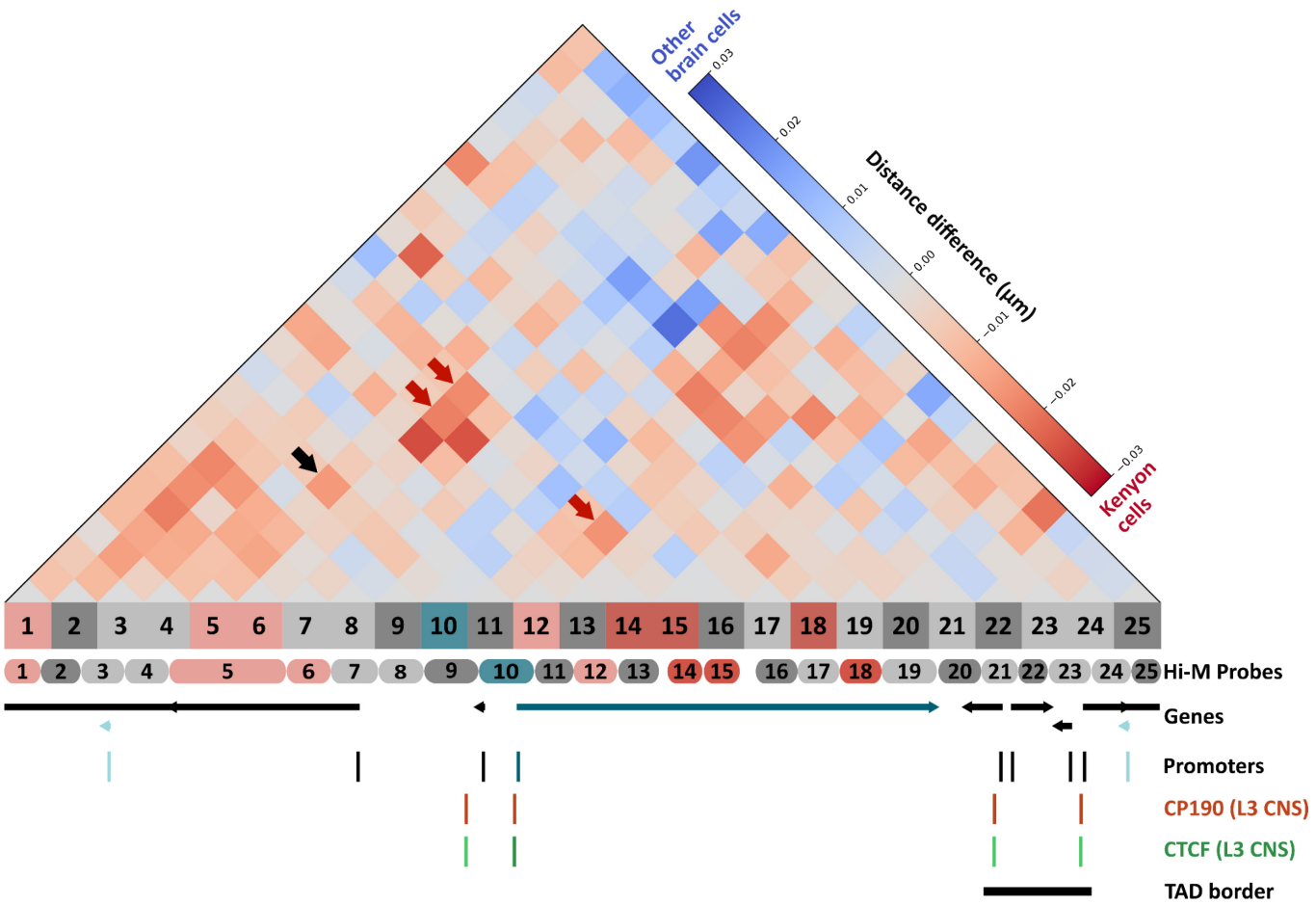


Figure 32 : Differences between Kenyon cells and the other brain cells of the *Drosophila* adult brain for the *sNPF* locus - PWDs difference matrix of Kenyon cells minus other cells of the brain, interactions specific to Kenyon cells are in red and interactions specific to the rest of the brain are in blue. E-P interactions (black arrows) and E-E interactions (red arrow)- Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). Repartition of the Hi-M probes along the genome. Genes are identified as follows : *sNPF*'s gene body (dark teal), lncRNAs (light blue) and other genes (black). Promoters are identified as vertical lines : *sNPF* (dark teal), lncRNAs (light blue) and other genes (black). CP190 binding sites in L3 CNS (orange) [Kaushal et al., 2021]. CTCF binding sites in L3 CNS (green) [Kaushal et al., 2021]. Black rectangle corresponds to the TAD border.

Regarding the insulation scores, differences are less pronounced than for the *rut* locus. The *sNPF* locus in Kenyon cells exhibits similar insulation tendencies (Figure 33a-b-c), reflecting the small and distributed differences in pairwise distances.

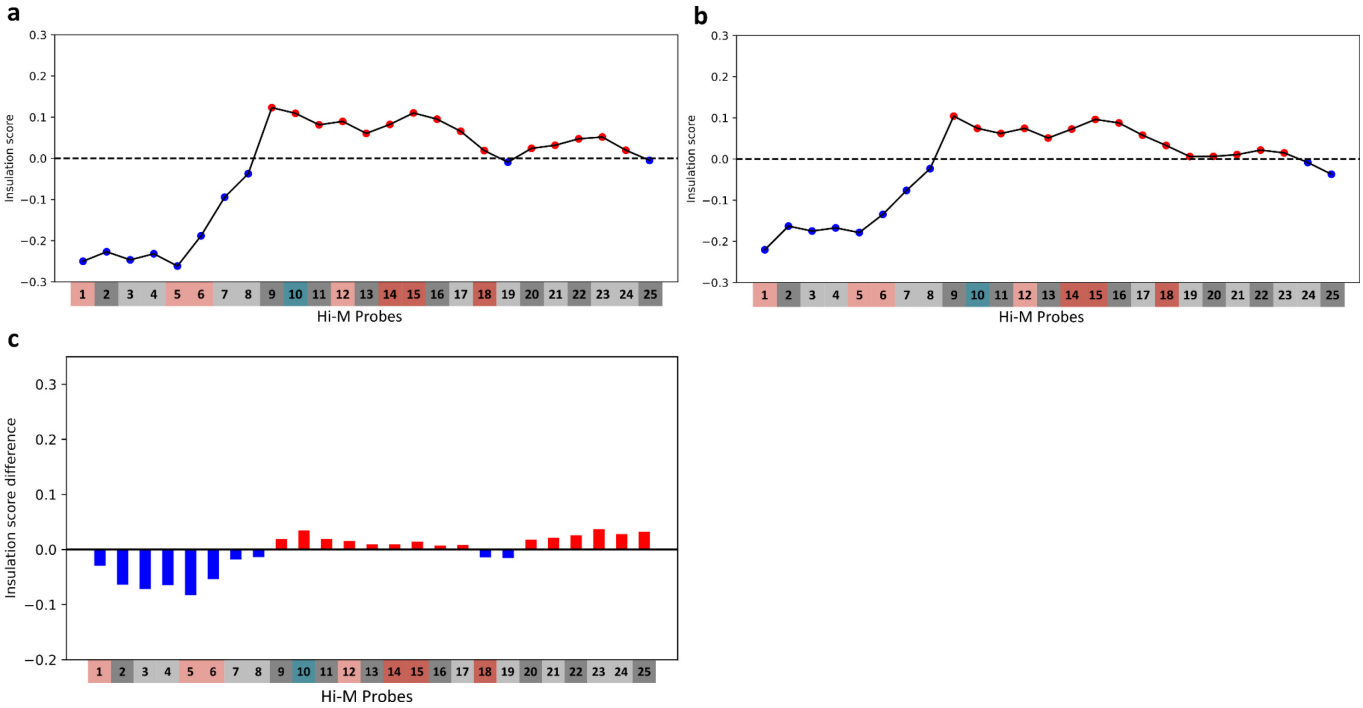


Figure 33 : Insulation differences between Kenyon cells and the other brain cells of the *Drosophila* adult brain for the *sNPF* locus - (a-b) Insulation score of the *sNPF* locus calculated for one bin of a 3kb size. Low insulation scores are in blue. High insulation scores are in red. (a) Insulation score in the Kenyon cells (b) Insulation score in the other cells of the brain (c) Insulation score difference between the Kenyon cells and the other cells of the brain calculated for one bin of a 3kb size. Red highlights that the values are more Kenyon cells specific while blue values are more other cells specific.

The lack of strong differences in 3D chromatin organisation between KC and non-KC cells for this locus might stem from the mixed population of Kenyon cells, some of which express *sNPF* (αβ and γ) and some that do not (α'β') (Figure 18 a-c). This mix could smooth potential differences. To further investigate this hypothesis, I analysed differences in chromatin folding between Kenyon subtypes.

2.2.4.2.2. Relationship between chromatin folding and gene expression of the *sNPF* locus within the Kenyon cells subtypes

2.2.4.2.2.1. Global chromatin structure and insulation of the *sNPF* locus

Following the comparison of the Kenyon cells against the other cells of the brain, chromatin conformations of the *sNPF* locus from the three main subtypes of the Kenyon cells were compared (Figure 34, Supp. Figures 9, 10, 11, 12, 13 and 14). The γ Kenyon cells exhibit shorter distances than the other conditions (Figure 34). All cell types share a sub-TAD structure between probe 1 and probe 10 (Figure 34). To further decipher changes in global conformation, the insulation score was extracted from the Hi-M matrices.

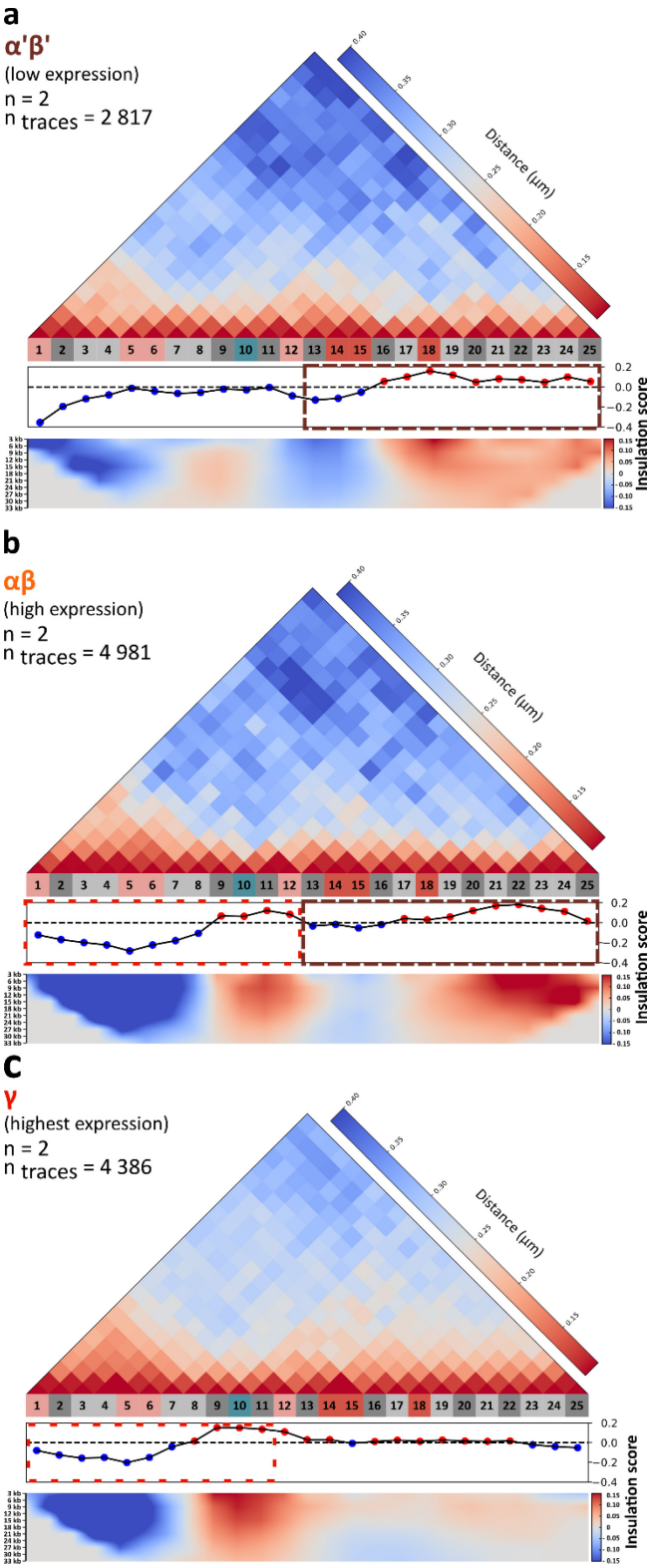


Figure 34 : Comparison of PWDs matrices and insulation among Kenyon cells subtypes for the *sNPF* locus

- (a-b-c) From top to bottom : PWDs median matrix of the *sNPF* locus for (a) α'β' Kenyon cells (b) αβ Kenyon cells (c) γ Kenyon cells. Hi-M probes are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). Insulation score of the *sNPF* locus calculated for one bin of a 3kb size. Low insulation scores are in blue. High insulation scores are in red. Insulation score for different genomic bin sizes (from 3kb to 33kb). Similarities between α'β' KCs and αβ KCs are highlighted with a dashed brown rectangle. Similarities between αβ KCs and γ KCs are highlighted with a dashed red rectangle.

The insulation score reveals notable differences among the three conditions (Figure 34). Close to the *sNPF* promoter region (probes 9-11), γ Kenyon cells display the strongest insulation. In αβ KCs this insulation is decreased, and in α'β' KCs it reaches negative values. This correlates with the expression of *sNPF*, which is highest in γ-KCs, diminishes in αβ KCs, and is null in α'β' KCs (Figure 18). The insulators CTCF and CP190 (located in probe 9 and 10) may facilitate gene regulation by encouraging interactions between *sNPF*'s promoter (located at probe 10) and surrounding enhancers (positioned within probes 12, 14, 15 and 18). Alternatively, this increased insulation may be the consequence of transcription. In any case, we observe a correlation between local 3D chromatin conformation around the promoter region and transcriptional levels.

Interestingly, the PWDs matrix from αβ KCs (high expression) shares similarities with the maps of the other KC subtypes (Figures 34a-b-c). Its insulation pattern aligns with γ KCs (highest expression) until probe 12. Both conditions display a negative insulation score from probe 1 until probes 8/9, after which the insulation of both cell types becomes positive (Figures 34b-c). Between probes 16-25, the insulation of αβ and α'β' Kenyon cells (no expression) share similar trends for the rest of the *sNPF* locus (Figure 34a-b).

2.2.4.2.2.2. Promoter interactions of the *sNPF* gene with the surrounding region

This section of the manuscript focuses on the fraction of the alpha phase, and we will explore the interactions of *sNPF*'s promoter with the other probes of the locus (Figure 35 and Supp. Figure 15).

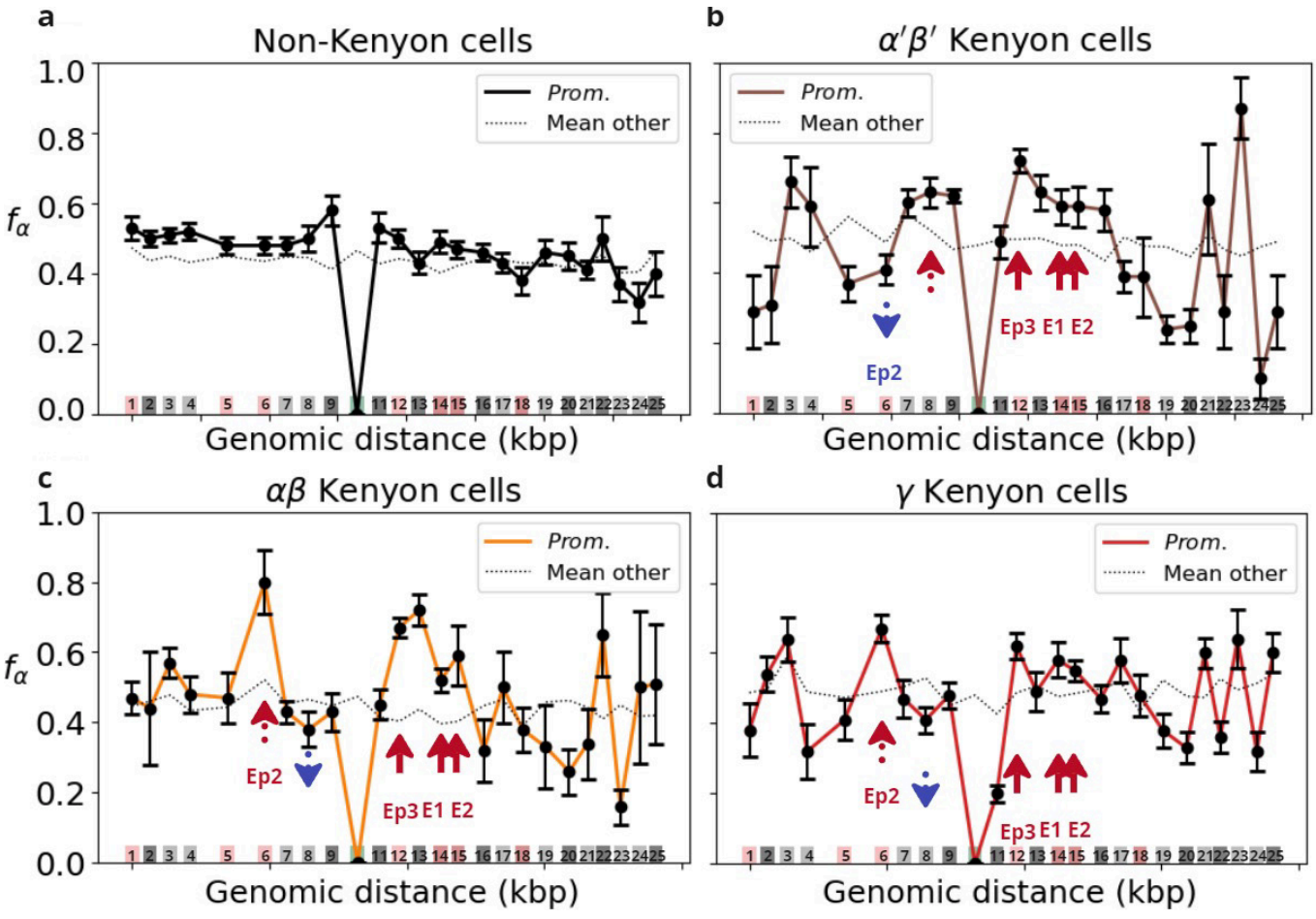


Figure 35 : Proximity between *sNPF*'s promoter and surrounding regions - (a-b-c-d) Fraction of the alpha phase f_α as a function of the genomic distance with *sNPF*'s promoter (teal) as anchor of four conditions (a) non-Kenyon cells/other cells of the brain (black curve) (b) α'β' Kenyon cells (brown curve) (c) αβ Kenyon cells (orange curve) and (d) γ Kenyon cells (red curve). Error bars represent the standard deviation of f_α. Averaged f_α of the entire *sNPF* locus is represented with a horizontal black dashed line. Red arrows are for increased f_α between *sNPF*'s promoter and the enhancer compared to the averaged f_α of the entire *sNPF* locus. Blue arrows are for decreased f_α between *sNPF*'s promoter and the enhancer compared to the averaged f_α of the entire *sNPF* locus. Hi-M probes are positioned according to their genomic location and are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). Red vertical lines highlight differential cell type specific interactions for tested enhancers. Pink vertical lines highlight differential cell type specific interactions for predicted enhancers.

Commonly shared enhancers

Enhancers Ep3, E1 and E2 present f_α values above the average f_α for all three cell types (Figure 35b-c-d and Table1). This is consistent with higher frequencies of short E-P distances observed in KCs with respect to other cells of the brain (Figure 35a). The main hypothesis that arises from this result is that all three enhancers are in proximity to *sNPF*'s promoter but are bound by different transcription factors inducing the differential transcription activity of the gene within the subtypes, as previously observed during early *Drosophila* development [Espinola, Götz et al., 2021].

	$f_\alpha >$ average f_α	$f_\alpha <$ average f_α
Common to all cell types	12 (Ep3) ; 14 (E1) ; 15 (E2)	19 ; 20
Common to αβ and γ KCs	6 (Ep2)	8
Common to αβ and α'β' KCs	3 ; 13	
Common to α'β' and γ KCs		1 ; 5 ; 22 ; 24

Table 1 : Commonly and differentially shared regions of the *sNPF* locus - f_α values above the average f_α and below the average f_α for the different KCs subtypes. Probes are coloured accordingly : tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey).

Enhancer predictions

Based on the predictions from *Janssens et al.* (2022), out of the three tested enhancers only E2 is predicted to be Kenyon cells specific. However, as previously highlighted, E1 and E2 present similar behaviours in all three subtypes although only E2 is supposed to be accessible (based on ATAC-seq data – Figure 20). Thus, similarly to our previous observations in the *rut* locus, accessibility may not necessarily reflect of enhancer activity or encourage E-P proximity.

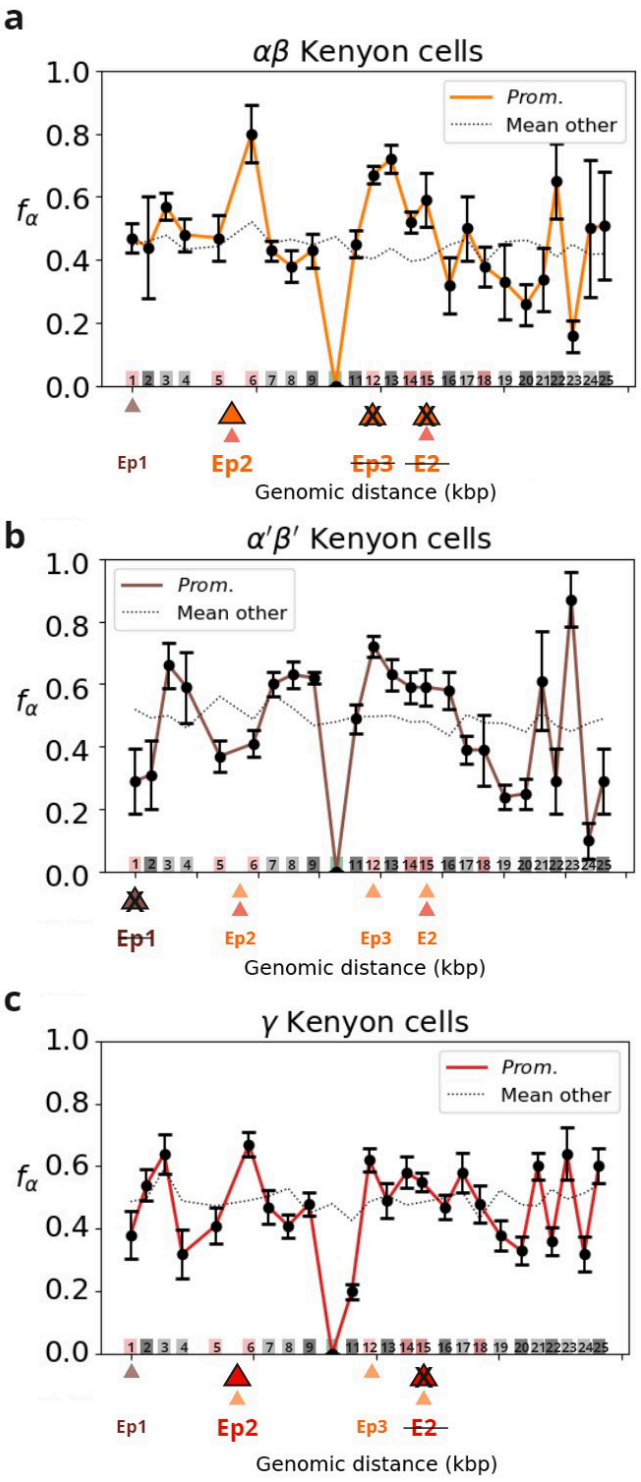


Figure 36 : The fraction of the alpha phase highlights differences between Hi-M data and enhancer predictions for the *sNPF* locus - (a-b-c-d) Fraction of the alpha phase f_α as a function of the genomic distance with *sNPF*'s promoter (teal) as anchor of four conditions (a) αβ Kenyon cells (orange curve) (b) α'β' Kenyon cells (brown curve) and (c) γ Kenyon cells (red curve). Error bars represent the standard deviation of f_α. Averaged f_α of the *rut* entire locus is represented with a horizontal black dashed line. Hi-M probes are positioned according to their genomic location and are coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey). Outlined and enlarged triangles highlight the position of cell type specific predicted enhancers : αβ specific (orange triangles), α'β' specific (brown triangles) and γ specific (red triangles) [Janssens et al., 2022]. Crosses indicate that the prediction does not fit with the results.

Interactions between *sNPF*'s promoter and the predicted enhancers also present interesting behaviours (Figure 36). Earlier in the manuscript, Ep1 (probe 1) was identified as α'β' specific, Ep2 (probe 5/6) was predicted to be γ and αβ specific and Ep3 (probe 12) to be αβ specific (Figure 20). From our analysis of f_{α}, the prediction of Ep1 is inconsistent with our results. This predicted enhancer does not present increased contacts with *sNPF*'s promoter in the α'β' Kenyon cells, it even seems to be quite the opposite. The low f_{α} overserved in the α'β' subtype (f_{α} = 29% ± 10%) could reflect of a more distal type of action of Ep1 to regulate sNPF's expression in this cell type (Figure 36b). Although, Ep3 present an f_{α} above the $f_{\alpha\ avg}$ in αβ KCs which might be consistent with the predictions and a contact mode of action of the enhancer, α'β' and γ subtypes present similar behaviours. On the other hand, Ep2 seem to follow exactly the predictions as γ and α'β' KCs present shorter distances with P_{sNPF}. Thus, enhancer predictions do not consistently account for enhancer-promoter proximity which might highlight different types of E-P interactions than the contact mode.

Potential origin of *sNPF*'s differential expression

As previously mentioned, the enhancer Ep2 is predicted to be αβ and γ specific (Figure 20). In both subtypes, f_{α} is larger than its average. Although both conditions share an increased alpha phase fraction, the value of the f_{α} varies between conditions (f_{α} = 80% ± 9% for αβ and f_{α} = 67% ± 4% for γ) (Figure 35c-d). In contrast, in the α'β' KC subtype, P_{sNPF} and Ep2 are in proximity in a even lower proportion than in non-KC brain cells (f_{α} = 41% ± 4% for αβ, Figure 35b). Overall, these analyses indicate that Ep2 is more frequently in a close distance to P_{sNPF} in KC subtypes where *sNPF* is expressed, similarly to our previous observations for the *rut* locus.

Previously, I focused on similarities between αβ and γ KCs, which both highly express *sNPF*, at regions with potential enhancers. Now, I will explore similarities between αβ and α'β' to understand the lower *sNPF* expression in αβ compared to γ KCs. Based on the f_{α} average threshold, αβ and α'β' cell types share two positive values (probes 3 and 13) (Table 1). Probe 3 (accessible chromatin) contains the promoter of lncRNA CR44908, while probe 13 (not accessible) does not appear to contain any enhancer. Similar results had been observed in *Pollex et al.* (2024), where lncRNAs presented differential interactions specific to the nervous system that changed during embryonic development.

By exploring distances between the promoter of *sNPF* and surrounding regions, we have identified differential and common enhancer-promoter interactions in Kenyon cell subtypes.

2.2.4.2.3. Key points of the Hi-M analysis of the *sNPF* locus

Variations in chromatin folding patterns

Insulation Score

- Kenyon cells and other cells have a highly similar insulation throughout their locus
- Insulation may play a role in regulating *sNPF* expression in the three Kenyon cells subtypes

Cell type specific domain

- αβ and γ Kenyon cells share a sub-TAD at the beginning of the locus
    - although the insulation of the rest of the locus is most similar between αβ and α'β' KCs

Enhancer-promoter and potential CREs proximities

Shared enhancers

- E1, E2, and Ep3 are commonly close to the promoter of *sNPF* in all three sub types

Cell type-specific enhancer (Ep2)

- Differential regulation by Ep2 across sub-cell types might encourage *sNPF*'s expression in αβ and γ Kenyon cells

Enhancer prediction

- Similarly to the *rut* locus, enhancers predicted by [Janssens et al., 2022] do not consistently account for enhancer-promoter proximity

2.2.5. Conclusion & perspectives

During my PhD, I have developed tools to simultaneously detect genome conformation for different cell types in the *Drosophila* adult brain. In this work, I have identified two genes of interest and their potential enhancers in different mushroom body cell types, using data from previously published studies [Henry et al., 2012][Crocker et al., 2016][Croset et al., 2017][Davie et al., 2018][Janssens et al., 2022].

In the Kenyon cells, the two genomic regions examined displayed global chromatin conformation changes involving multiple shared or cell-type-specific enhancers as well as other regulatory (promoters, lnRNAs) or non-regulatory sequences. Specifically, enhancers previously tested by Janelia's Flylight project [Jenett et al., 2012] or predicted to be cell-type-specific by DeepFlyBrain [Janssens et al., 2022] showed significantly higher interaction frequencies in mushroom body cells with respect to other cells of the. These results are consistent with the specific expression of *rut* and *sNPF* in the mushroom bodies.

In Kenyon cells subtypes, the *rut* locus has presented both very different global chromatin organisation including unique cell-type-specific insulation signatures or TAD substructures. Different E-P interactions were also observed even though the gene is similarly expressed in all three sub cell types. These results highlight the differential action of subtype-specific regulatory elements in gene expression.

For the *sNPF* locus, the insulation score analysis has highlighted significant differences across the three cell types, particularly near the *sNPF* promoter region. The insulation pattern aligned with *sNPF*'s expression levels, which was highest in γ KCs (highest expression), lower in αβ KCs (high expression), and absent in α'β' KCs (low expression). In addition, our analysis of the fraction of the alpha phase has indicated that a predicted enhancer (Ep2) was more frequently in a close proximity to the promoter of *sNPF* in KC subtypes where *sNPF* is expressed. Thus, in this manuscript we were able to highlight a clear correlation between global chromatin conformation, E-P interactions and gene expression in the Kenyon cells subtypes.

Most of our results aligned with the E-P contact model, showing that cell-type-specific enhancers are often closer to promoters (under 250 nm). This observation suggests that chromatin can adopt different conformations depending on the cell type, although not all E-P interactions seem to be instructive. To further explore these results, future studies could delve deeper into E-E, P-P, and E-P interactions beyond the *rut* and *sNPF* promoters. Notably, research by [Espinola et al., 2021] has shown enhancer and promoter clustering during embryogenesis, which could be similarly explored in the *Drosophila* adult brain using our data.

Interestingly, enhancers predicted by [Janssens et al., 2022] did not consistently account for E-P interactions, despite some of them displaying differential interactions in Kenyon cell subtypes. This may indicate different E-P mode of actions or cell-type-specific transcription factors modulating gene activity. To explore this, future work could study the impact of transcription factors on E-P interactions using fly lines with mutated transcription factor binding sites within the *sNPF* enhancer E2 [Janssens et al., 2022]. Following the mutation of two TFs binding sites (*ey* and *mef2*), they showed a loss of activity for enhancer E2. Similar mutation in the endogenous locus could be highly informative, and if combined with Hi-M, these lines could provide insights into the relationship between enhancer activity and E-P interactions.

Chapter 3 : Discussion

Both studies presented in the results chapter have explored the role of E-P interactions in gene expression in *Drosophila* pupal leg discs and adult brains. Based on these data, we aimed to determine which model best described genome folding in these different tissues.

In *Denaud et al. (2024)*, we examined the chromatin conformation of various tarsal segments of the *Drosophila* pupal leg disc. In wild-type (WT) conditions, only the first segment expresses the *dac* gene, while the cells of other segments do not. Using Hi-M, we did not observe any significant increase in proximity between the cell-type-specific enhancer and the promoter of the gene of interest in the first tarsal segment. Similarly to previous studies in early *Drosophila* embryogenesis [Espinola, Götz et al., 2021][Ing-Simmons et al., 2021], our observations suggest that modulation of physical proximity between an enhancer and a promoter is not necessarily a major determinant of gene activation as no strong interaction was observed between the promoter of *dac* and the ring enhancer across the different segments. These findings align with the "permissive" model, for the regulation of the target gene by its enhancer [De Laat et al., 2013].

Consistently with recently published results by *Pollex et al*. (2024), I have observed numerous cell-type-specific enhancer-promoter proximities in the *Drosophila* adult brain. These results support the instructive model hypothesised for E-P interaction during differentiation [De Laat et al., 2013]. However, common E-P interactions were also identified in all Kenyon cell subtypes within both loci, which could indicate that instructive E-P topologies are formed on pre-existing chromatin landscapes [Ghavi-Helm et al., 2014][Espinola, Götz et al., 2021][Ing-Simmons et al., 2021][Pollex et al., 2024].

By using the double Gaussian fit model, specifically the fraction of the alpha phase, we extracted cell-type-specific interactions between promoters and surrounding enhancers that were not visible in PWD matrices. Our results show that the method developed by [Remini et al., 2024] is also adapted to decipher enhancer-promoter interactions in complex tissues. The two detected phases of the distance distribution are consistent with previous observations made in [Chen et al., 2018] for the *even-skipped* locus. The beta phase could correspond to the first state they have detected, where enhancers and promoters are not in proximity before transcription activation. The alpha phase could correspond to the other two states where the promoter and the enhancers are either "within range" or in close proximity to induce gene transcription.

To better understand the role of E-P interactions in gene regulation, both studies from the results chapter could have benefited from simultaneous detection of gene expression and chromatin folding. Further exploration of the dynamic interplay between enhancer-promoter topology and gene activity [Bartman et al., 2016][Chen et al., 2018][Levo et al., 2022] could be achieved through live imaging of enhancer-promoter localisation and transcription in the *Drosophila* adult brain. Techniques similar to those used by *Chen et al. (2018)* or *Levo et al. (2022)* could be applied to a relevant locus in the *Drosophila* adult brain without causing significant phenotypic consequences. These methods, combined with long-term brain imaging [Huang et al., 2018] or ex-vivo long-term adult brain culture [Ayaz et al., 2008], could enable multicolour live imaging to study E-P dynamics and gene expression simultaneously at the single-cell level within *Drosophila* adult brains.

Moreover, using the method from *Huang et al. (2018)*, living adult *Drosophila* could be mounted onto a microscope to study E-P dynamics and gene expression. The flies could then be unmounted, trained with a specific memory paradigm, and remounted to observe how E-P interactions and gene expression evolve with memory formation. Although multicolour live imaging limits the number of genomic regions studied, the same fly could undergo brain dissection and a Hi-M protocol to reconstruct chromatin folding at a larger

scale. Despite the complexity and low throughput of this experiment, it could provide unprecedented insights into E-P dynamics, gene expression, global locus conformation, and memory formation within the same *Drosophila* adult brain.

To explore the hypothesis of a pre-existing E-P landscape established during early embryogenesis to regulate gene expression throughout the fly's nervous system development, lineage tracing techniques could be beneficial [Yu et al., 2009][Chow et al., 2021]. Imaging chromatin folding in the neural ectoderm, neuroblasts, and fully differentiated mushroom bodies could help identify whether these common E-P interactions in Kenyon cell subtypes are instructive E-P loops maintained throughout differentiation, serving as a scaffold for gene regulation. Combining intMEMOIR [Chow et al., 2021] with Hi-M could exploit the strength of Hi-M in maintaining spatial information of tissues, a feature not fully utilised in the *Drosophila* adult brain project.

Finally, we hope that developing a protocol for sequential DNA-FISH imaging in the *Drosophila* adult brain will provide the community studying the *Drosophila* adult brain a powerful tool to study chromatin structure in cells. In the next few months, we aim to achieve the first simultaneous detection of genome folding and gene expression at the single-cell level in the *Drosophila* adult brains. Hopefully, this advancement will bring us closer to understanding how and why chromatin folds the way it does in the *Drosophila* adult brain, potentially leading to new insights into the fundamental mechanisms of gene regulation.

Chapter 4 : Materials and methods of the *Drosophila* brain project

Fly lines

The transgene used was UAS-nls-GFP (from Bloomington #BL4770). All the GAL4 lines used in this study have been thoroughly described elsewhere [Aso et al., 2009]. The GAL4 lines used are the following : OK107-GAL4 expressed in all Kenyon cells, c739-GAL4 expressed in αβ Kenyon cells, c305a-GAL4 expressed in α'β' Kenyon cells, and H24-GAL4 expressed in γ Kenyon cells (from J.M. Dura). All flies and progeny for GAL4-UAS crosses were raised on standard cornmeal yeast extract medium at 25°C.

***Drosophila* brain dissection and freezing**

Adult male and female flies were 2 to 5 days old when collected (from previously cited fly lines) after a CO_2 anaesthesia. Flies were briefly dewaxed by dipping them for 2-5s in 70% ethanol using forceps. Then, flies were washed twice with 2x saline sodium citrate (SSC). Male and female flies were transferred to their own pre labelled (male or female) 2mL Eppendorf tube with a maximum of 15-20 flies per tube. Flies were then fixed and permeabilized with 4% paraformaldehyde (PFA) / 0.5% PBT (1x phosphate-buffered saline (PBS) + 0.5% triton X-100) on a rotating wheel for 3h at room temperature (RT). After fixation, tubes were washed for 15 min, 4 times with 0.5% PBT followed by two PBS washes. During dissection, flies were stored in PBS on ice. *Drosophila* adult brains were dissected on a Silgar polymer plate in ice cold PBS. Dissection was timed to ensure that the first brain dissected did not spend more than 20 min on the Silgar plate to limit potential brain damage. Brains were post-fixed in 1.8 mL 4% PFA/0.5% PBT for 20 min followed by three 0.5% PBT washes for 20 min. Previously described steps were made on a rotating wheel at RT and solutions were removed extremely slowly to avoid brain resuspension. After the last wash, 0.5% PBT was slowly removed, leaving the brains in 200 µL of 0.5% PBT.

For cryoprotection, brains were incubated for 48h in a 30% filtered sucrose solution at 4°C on a rotating wheel. 1 mL of solution was removed from the 2 mL Eppendorf, brains were softly resuspended and transferred onto a square cryomold (TED PELLA, cat. no. 27147-2). For *sNPF* samples, male and female brains were aligned to enable sex identification during the acquisition phase. Once brains were aligned, the Optimal Cutting Temperature (OCT) embedding matrix was poured into the cryomold. Cryomold was then placed in a -80°C freezer and samples were stored for up to several months until cryosectioning.

Glass slides functionalization and brains cryosection

Prior to cryosectioning, 40 mm round coverslips (Bioptechs, cat. no. 40-1313-0319) were cleaned with 70% ethanol and activated with air plasma for 30s. Coverslips were placed onto glass petri dishes and covered with 100 µL of 3-aminopropyltrimethoxysilane for 5 min at RT. Slides were then covered for 10 min with ddH2O. Petri dishes were moved onto an orbital shaker and washed two other times with ddH_2O for 10 min. Coverslips were incubated in a 0,5% glutaraldehyde/PBS solution for 30 min and then washed twice with ddH2O. Once air dried, slides were covered with 20% poly-L-lysine (Sigma-Aldrich, cat. no.

P8920) /PBS for 1h and incubated overnight (O/N) in sterile ddH2O. The next day, slides were air dried for 1-2h at room temperature.

Then, embedded *Drosophila* adult brains were cutted into 10 µm thick sections using a cryostat and cryosections were transferred onto silane-coated coverslips. Slides with tissue sections were then air dried for 1-2h at room temperature. For long term storage, samples were transferred onto ice for 15 min and then stored at -20°C. For RNA-FISH labelling, samples were dried for 30 min at RT and labelling protocol was started right after.

Sequential RNA-FISH library design

Sequential RNA-FISH libraries were designed with Oligostan, a R script that identifies optimal probe sequences for single molecule RNA-FISH [Tsanov et al., 2016]. Then, probes were filtered with a homemade script which removes potential off target probes after verifying gene specificity with fly blast (http://flybase.org/blast/). RNA-FISH probes were designed with a maximum of 38 probes per target gene with each probe having a mRNA homology sequence size between 26 and 32 nt. The shortest distance between probes was 5 nt. The probes have the following construction :

1. two 20 nt readout sequences separated by an A
2. one 26 to 32 nt mRNA homology region
3. two 20 nt readout sequences separated by an A

Probes were designed with either an exonic or an intronic FISH design. The exonic design was used for Kenyon cells identification. RNA FISH probes were designed for *eyeless* (*ey*) and *portabella* (*prt*) due to their high specificity to the neuron type. To identify Kenyon cells subtypes, probes were designed for CG13055 (γ), bru2 (αβ and γ) and GstD11 (α'β' and γ). Intron FISH probes were constructed for *rut* and *sNPF* to identify the cells that were actively transcribing the gene of interest. RNA FISH probes were synthesised and purchased as Opools (DNA pooled oligos) from Integrated DNA Technologies (IDT).

Hi-M library design and amplification

For this project, two libraries were designed for *rut* and *sNPF* loci. Each primary oligopaint library used 42nt sequences with genome homology obtained from the oligopaint public database (http://genetics.med.harvard.edu/oligopaints). Synthesis was performed by CustomArray. The *rut* library covers 90,5 kb and the *sNPF* library covers 78,5kb. Each probe contains 45 primary oligonucleotides covering in average 3,2 kb (9,3 - 26,5 probes/kb) for *rut* and 3 kb (5,6 - 22,5 probes/kb) for *sNPF*. Primary oligopaint oligonucleotides are composed of 5 different regions :

1. one 21 nt forward priming region for library amplification
2. two 20 nt probe-specific readout sequence
3. one 42nt genome homology region
4. two 20 nt probe-specific readout sequence
5. one 21 nt reverse universal priming region

Forward and reverse priming sequences are the following :

- For *rut*
    - Forward - BB287 : CGCTCGGTCTCCGTTCGTCTC

- Reverse - BB288 : GCTGAACCCTGTACCTAGCCC

- For *sNPF*
  - Forward - BB297 : GACTGGTACTCGCGTGACTTG
  - Reverse - BB299 : CCAGTCCAGAGGTGTCCCTAC

The procedure for oligopaint library amplification was previously described [Cardozo Gizzi et al., 2019], [Cardozo Gizzi et al., 2020], [Espinola, Götz et al., 2021]. The protocol consists of seven steps : (1) an emulsion PCR (emPCR) to extract each library from the oligonucleotide pool using primers targeting 21 nt regions ; (2) a limited-cycle PCR performed on the emPCR to identify the most efficient amplification cycle ; (3) a large scale PCR with a T7 promoter on the reverse primer ; (4) an in vitro T7 transcription ; (5) a reverse transcription to transform RNAs into single-stranded DNA (ssDNA) ; (6) an alkaline hydrolysis for the removal of the intermediate RNA and (7) a ssDNA purification and concentration.

Each readout sequence is unique and specific to a specific genomic region. The imaging oligonucleotide (complementary to the readout sequence) is labelled with an Alexa Fluor 647 fluorophore. The fluorescently labelled part of the imaging probe is attached via a disulfide bridge that can be cleaved (chemical bleaching) during the sequential imaging of FISH probes. The fiducial library, used for drift correction during the pyHi-M analysis, includes sequences specifically designed to target satellites in *Drosophila Melanogaster* [Li et al., 2017][Jagannathan et al., 2017] with a complementary secondary probe bound to a non-cleavable Atto-488 fluorophore. For locus segmentation, we use an adapter oligo complementary to the reverse primer of the library with the imaging oligonucleotide labelled with an Alexa Fluor 647 fluorophore. Adapters and fluorescently labelled imaging oligos were synthesised and purchased from Integrated DNA Technologies (IDT).

RNA-FISH labelling

Adult brain cryosections were post-fixed for 15 min with 4% PFA/PBS followed by three PBS washes at RT. Samples were then permeabilized in 70% EtOH at 4°C for 60h. After permeabilization, brains were cleared at RT for 20 min with a 8% SDS/PBS solution. Slides were washed once with cold 70% EtOH and once with 30% formamide for 5 min at RT. Samples were incubated in 30% formamide for 3h at 37°C in a humid incubator. After the pre-hybridization step, slides were placed upside down in a glass petri dish and sealed onto the RNA-FISH primary probe hybridization solution. The hybridization lasted for 36h at 37°C in a humid incubator. Hybridization solution comprised probes diluted in hybridization buffer (HB) (30% formamide, 10% dextran sulphate, 5% tRNA stock (20 mg/mL), 1% RVC stock (200 mM), 0.4% of nuclease-free BSA stock (50 mg/mL), 2x SSC) at a concentration of 20 nM. After hybridization, samples were washed with 2x SSC three times, with 30% formamide twice for 30 min at 37°C in a humid incubator and then again three times with 2x SSC. Until imaging, samples were stored at 4°C in 2x SSC.

Hi-M labelling

Drosophila adult brain cryosections were removed from the -20°C freezer and allowed to equilibrate at RT for 1h. Then, slides were permeabilized with 10 mM Sodium Citrate, first for 5 min at RT, then for 25 min at 80°C in a water bath. Samples were then left to cool down for 1h. Slides were washed for 5 min with 2x SSC and incubated for 2h at RT in 50% formamide/2x SSC (wash buffer). After incubation, slides were placed upside-down in a glass petri dish and sealed onto the DNA-FISH primary probe hybridization solution. The previously mentioned solution comprises 160-190 pmols of the primary oligopaint library

diluted in 21 µL of FISH Hybridization Buffer (FHB) (50% formamide, 10% dextran sulphate, Salmon Sperm DNA 0.5 mg/mL, 2x SSC) and 1µL of 100 µM satellite solution. Prior to a heat shock of 5 min at 85°C, samples were incubated for 3h at 45°C in a water bath. Post heat shock, glass petri dishes were placed overnight at 37°C in a humid incubator. The next day, cryosections were washed 2 times with 50% wash buffer (WB) for 1h and sequentially washed with four different concentrations of WB : 40%, 30%, 20%, 10% formamide during 40 min for 40% and 20 min for the other concentrations. These wash steps were performed onto an orbital shaker set at 50 rpm. Finally, slides were washed for 20 min with 2x SSC, post-fixed with 4% PFA/PBS for 10 min and washed again three times with 2x SSC. Samples were stored at 4°C in 2x SSC.

Immunolabeling

DNA-FISH labelling induces a loss of GFP signal. To detect Kenyon cells by fluorescence, an immunostaining against GFP was used. For the first step of this protocol, slides were incubated for 2h in 0.1% bovine serum albumin (BSA) in 0.1% PBT (1x PBS + 0,1% triton X-100). Chicken GFP polyclonal antibody (anti-GFP) was diluted 1:200 (Invitrogen, catalogue no. A10262) in 0,1% BSA-PBT and was incubated overnight at 4°C in a humid chamber. The next day, brain cryosections were washed three times for 20 min in PBT. Samples were then incubated at RT in 1:500 goat anti-chicken secondary antibody (Invitrogen, catalogue no. A21449) diluted in PBT for 3h. This final step of antibody hybridisation was followed by three PBT washing steps of 10 min each. To remove traces of triton, samples were washed rapidly two times with PBS. Slides were post-fixed for 5 min with 4% PFA diluted in PBS. Finally, samples were washed three times in PBS and kept in 2x SSC at 4°C until imaging.

Hi-M imaging system

Hi-M experiments were performed with a homemade widefield and epifluorescence microscope. This setup includes a Rapid Automated Modular Microscope (RAMM) (Applied Scientific Instrumentation, Oregon, US) coupled with a microfluidic device previously described in [Barho et al., 2022]. The microscope and fluidics system were controlled using Qudi-HiM (our homemade hardware control package)[Barho et al., 2022]. The fluidics system permitted the automated and sequential hybridizations of probes. The solutions were delivered to the sample by a combination of three eight-way valves (HVXM 8-5, Hamilton), a negative pressure pump (MFCS-EZ, Fluigent) and a FCS2 flow chamber (Bioptechs, cat. no. 03060319-2-NH). The fluorescence excitation was performed by three different lasers : 405 nm (Obis 405, 100 mW, Coherent), 488 nm (Sapphire 561 LP, 150 mW, Coherent) and 642 nm (VFL-0-1000-642-OEM1, 1W, MPB communications Inc.). The fluorescence was collected through a Nikon APO x60 1.2 NA water immersion objective lens mounted on a closed-loop piezoelectric stage (Nano-F100, Mad City Labs Inc.). Images were acquired using a sCMOS camera (ORCA Flash 4.0 V3, Hamamatsu) with an effective optical pixel size of 106 nm. To correct axial drift in real time, we used a homemade autofocus system composed of a 785 nm laser (OBIS 785, 100 mW, Coherent) and an infrared sensitive camera (CMOS - DCC1545M, Thorlabs).

Acquisition of Hi-M datasets

The slide was mounted onto the flow chamber. Adult *Drosophila* brain cryosections were first incubated with the imager probe complementary to the fiducial library (25 nM of Atto-488, 2x SSC, 40% v:v formamide) for 20 min and then washed with a washing buffer solution (2x SSC, 40% v:v formamide)

for 10 min. After a 5 min wash with 2x SSC, we proceeded with nuclei staining with 0.5-1 µg/mL of DAPI in PBS for 20 min. After another 5 min wash with 2x SSC, the imaging buffer (1× PBS, 5% Glucose, 0.5 mg/ml glucose oxidase and 0.05 mg/ml catalase) was injected to limit fiducial photobleaching during the acquisition. An image stack was acquired for each of the 10-20 regions of interest (ROIs) of 200 µm x 200 µm. The DAPI, the fiducial and the immunostaining against the GFP were imaged sequentially (using 405 nm, 488 nm and 647 nm lasers) with a z-step size of 250 nm for a total range of 17.5 µm.
Next, readout probes were sequentially hybridised, acquired and photobleached to image the whole oligopaint library. The following steps were performed for each of the 25 probes : (1) readout probes (40 nM secondary probe, 2x SSC, 40% v:v formamide) injection and incubation for 10 min; (2) 10 min wash with washing buffer solution; (3) 5 min wash with 2x SSC; (4) imaging buffer injection and sequential acquisition of fiducial and probe with 488 and 641 nm lasers; (5) Chemical bleaching (2× SCC, 50 mM tris(2-carboxyethyl)phosphine (TCEP, Sigma-Aldrich, cat. no. 646547)) of the imaging probe; (6) 10 min wash with 2x SSC before a new cycle of hybridization. At the end of the experiment, the entire locus was imaged by staining the reverse priming sequence with a complementary oligo bound to a AF647 fluorescent probe. The images produced at the end of the acquisition step were used for loci mask segmentation.

Image processing and Hi-M analysis

Raw TIFF images were deconvolved using Huygens Professional 21.04 (Scientific Volume Imaging, https://svi.nl). Hi-M analysis was performed using pyHi-M, a homemade analysis pipeline [Devos, Fiche et al., 2024]. First, images were z-projected by applying either sum for DAPI channels or maximum intensity for probes, fiducial and loci mask. For each cycle of hybridization, fiducial images were used to register the corresponding probe image using global and local registration methods. Probes and locus images were segmented in 3D using a neural network, followed by 3D localisation of the centre of each probe mask [Devos, Fiche et al., 2024]. Therefore, we built chromatin traces by combining the DNA-FISH spots colocalizing within single locus masks. Pairwise distances (PWD) matrices were calculated for each single chromatin trace. Each chromatin trace contained at least 12% of the probes and no duplicated probes. Traces not fitting this first criterion were removed. Proximity frequencies for each probe combination were calculated as the number of chromatin traces in which pairwise distances were < 100 nm, normalised by the total number of chromatin traces containing both probes.

After final outputs of pyHi-M were generated, the anti-GFP signal was segmented using ilastik [Berg et al., 2019] (a machine learning toolkit for segmentation) and a homemade software (developed by Jean-Bernard Fiche) to assign each trace to either the cell type of interest or the other cells of the brains. The number of chromatin traces from two replicates per condition are the following :

genes	**Kenyon cells / OK107**	**Other cells of the brain / non-OK107**	**αβ Kenyon cells / c739**	**α'β' Kenyon cells / c305a**	**γ Kenyon cells / H24**
rut	7 400 traces	22 398 traces	4 820 traces	3 285 traces	4 225 traces
sNPF	11 802 traces	47 825 traces	4 981 traces	2 817 traces	4 386 traces

To evaluate reproducibility between the replicates, Wilcoxon two-sided rank tests between the pairwise distance distributions were performed. This tested the hypothesis that two independent samples

followed the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). We estimated the distance error between probes by performing bootstrapping analysis. For this, we performed 1000 bootstrapping cycles drawn from the experimental distribution of pairwise distances to estimate the standard deviation in the determination of the median distance (see Supplementary figures).

Insulation score extracted from Hi-M dataset

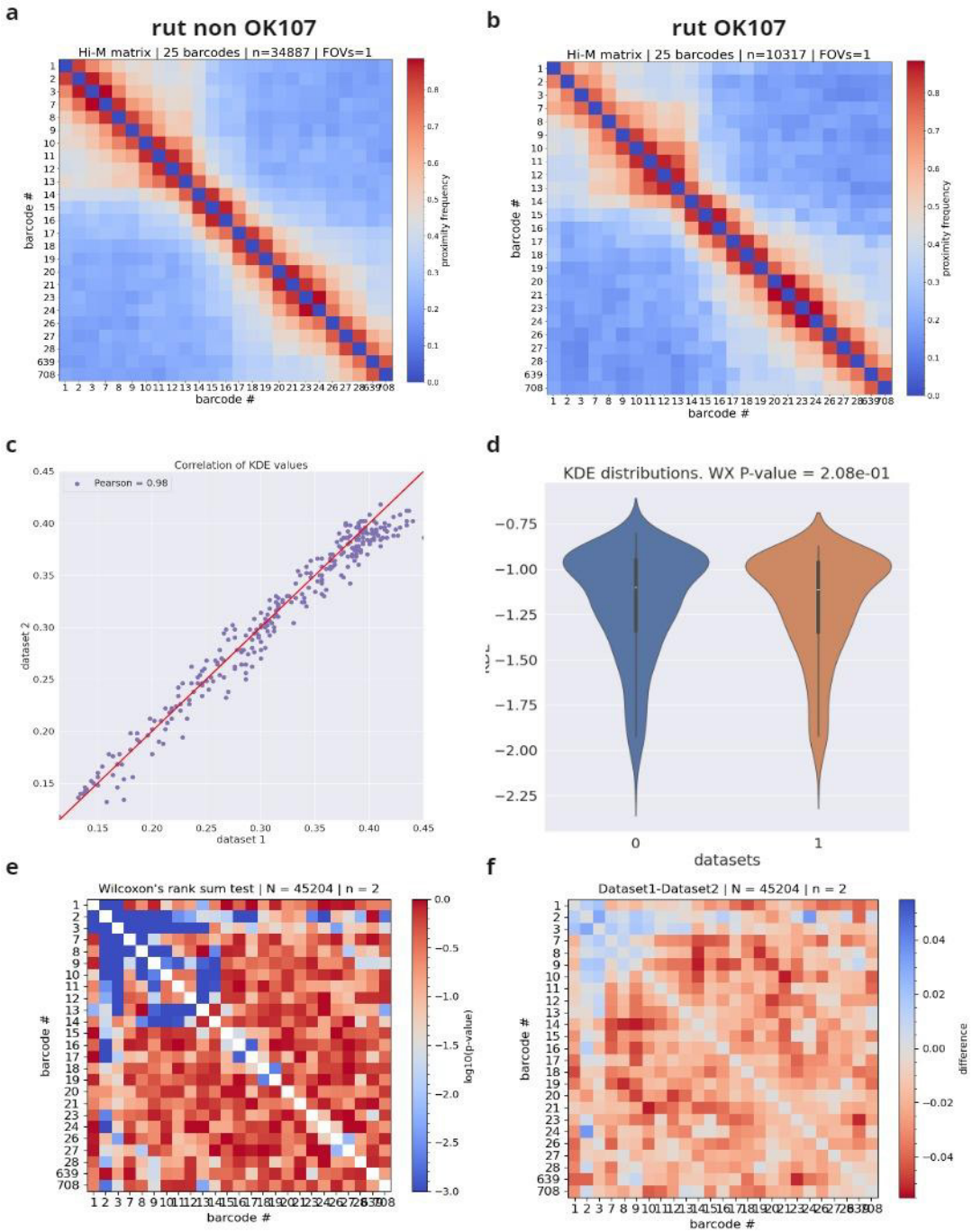
Insulation scores were extracted from Hi-M datasets by sliding an n-by-n square window along the diagonal of the median PWD matrix and by summing the distances within this square. Domainograms were computed by smoothing a matrix obtained by calculating the insulation score with an increased window size (from 1-by-1 to 6-by-6) over the complete locus [Messina et al., 2023].

Analysis of chromatin traces with the double gaussian model fitting

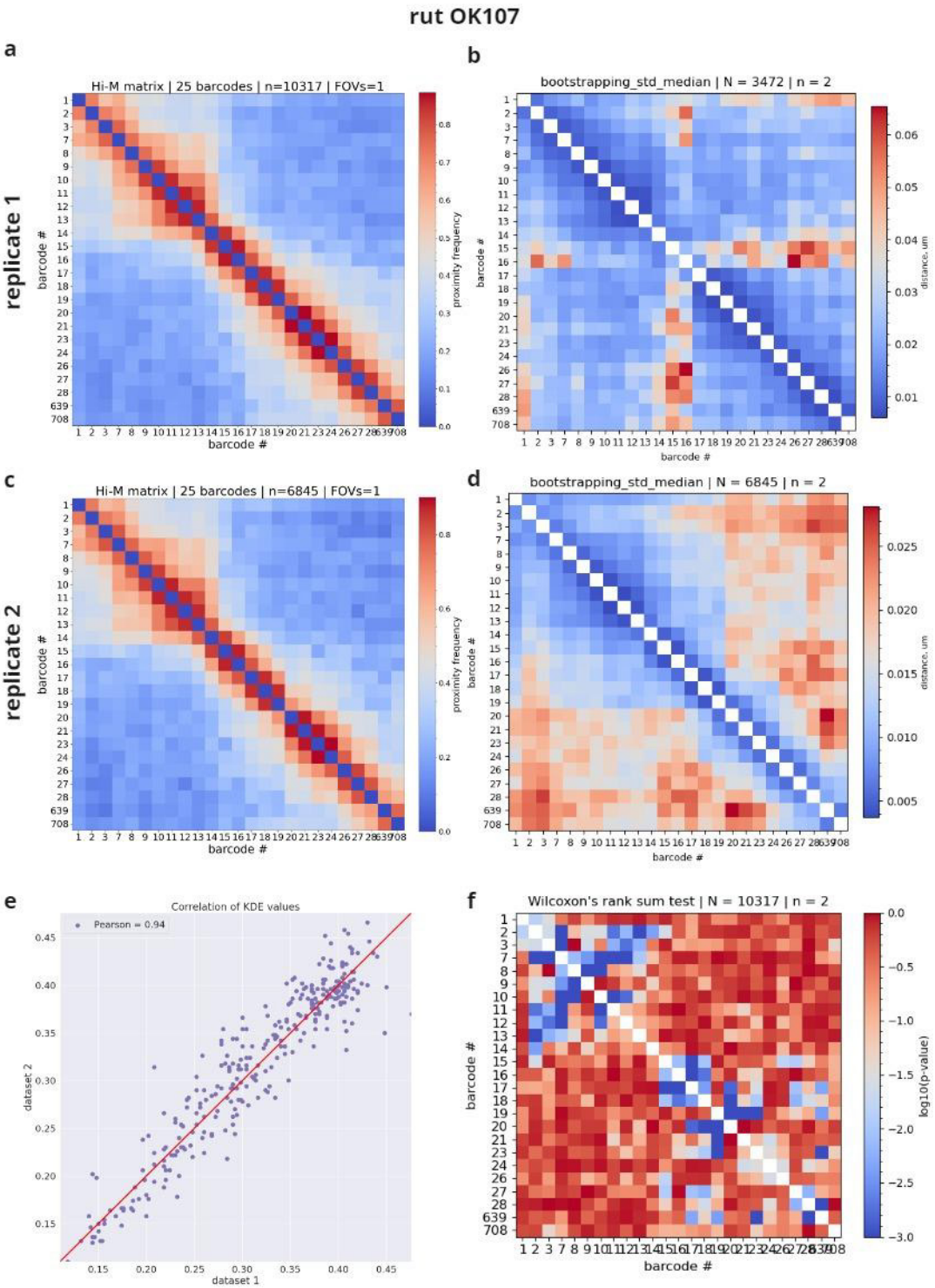
Models, fits and results were produced by Loucif Remini using the analysis pipeline he developed in [Remini et al., 2024]. The mean values of the alpha (R_α) and beta (R_β) Gaussians, the fraction of cells in the alpha phase (f_α) and scaling behaviours of the polymer (ν_α, ν_β and ν_t) were calculated on non-filtered traces (all traces kept, no minimal number of probes per traces) and for the following conditions : all brain cells , Kenyon cells, non-Kenyon cells (or other cells of the brain), αβ Kenyon cells, α'β' Kenyon cells and γ Kenyon cells.

Appendices

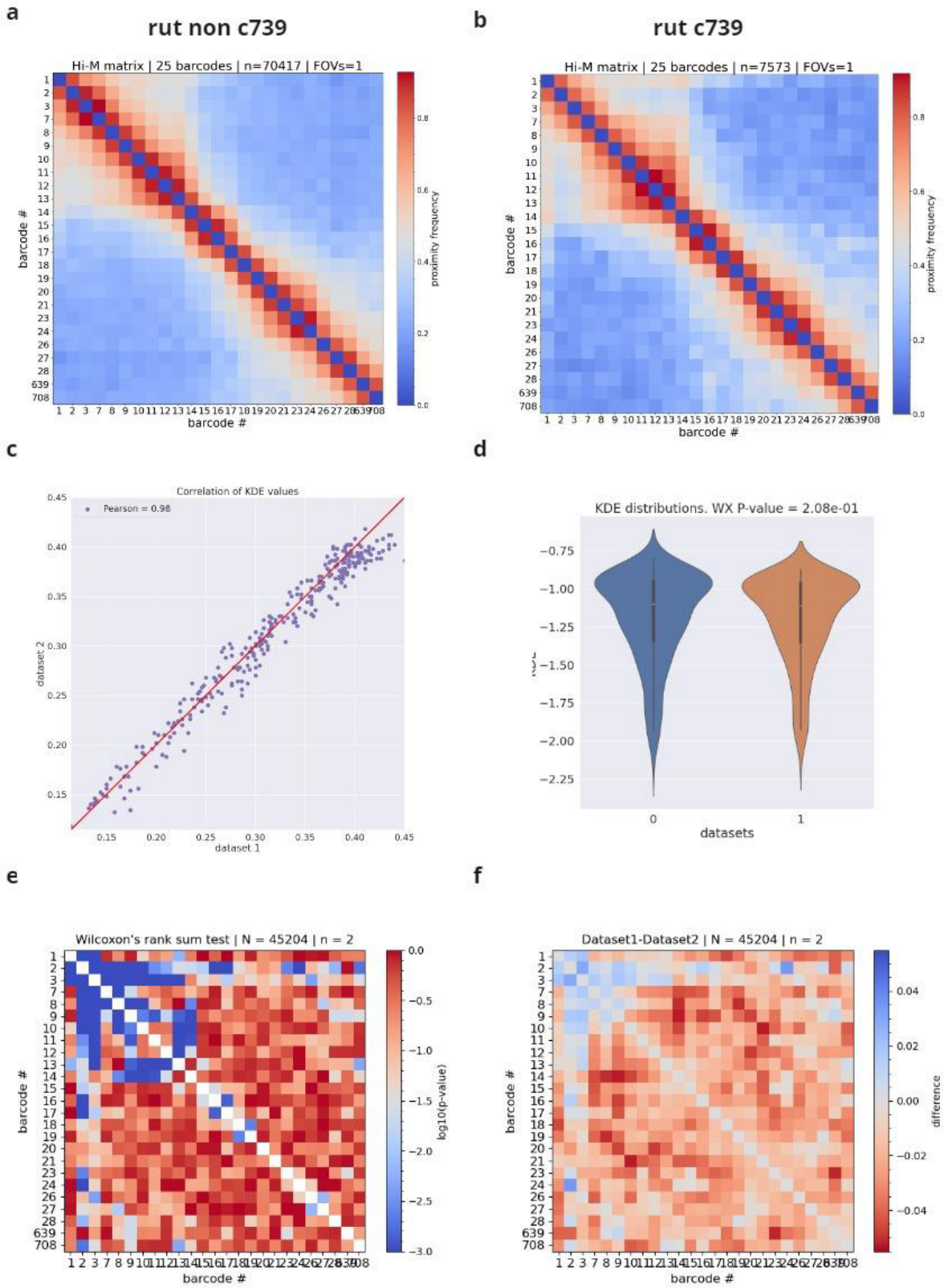
1. Supplementary figures for section 2.2.



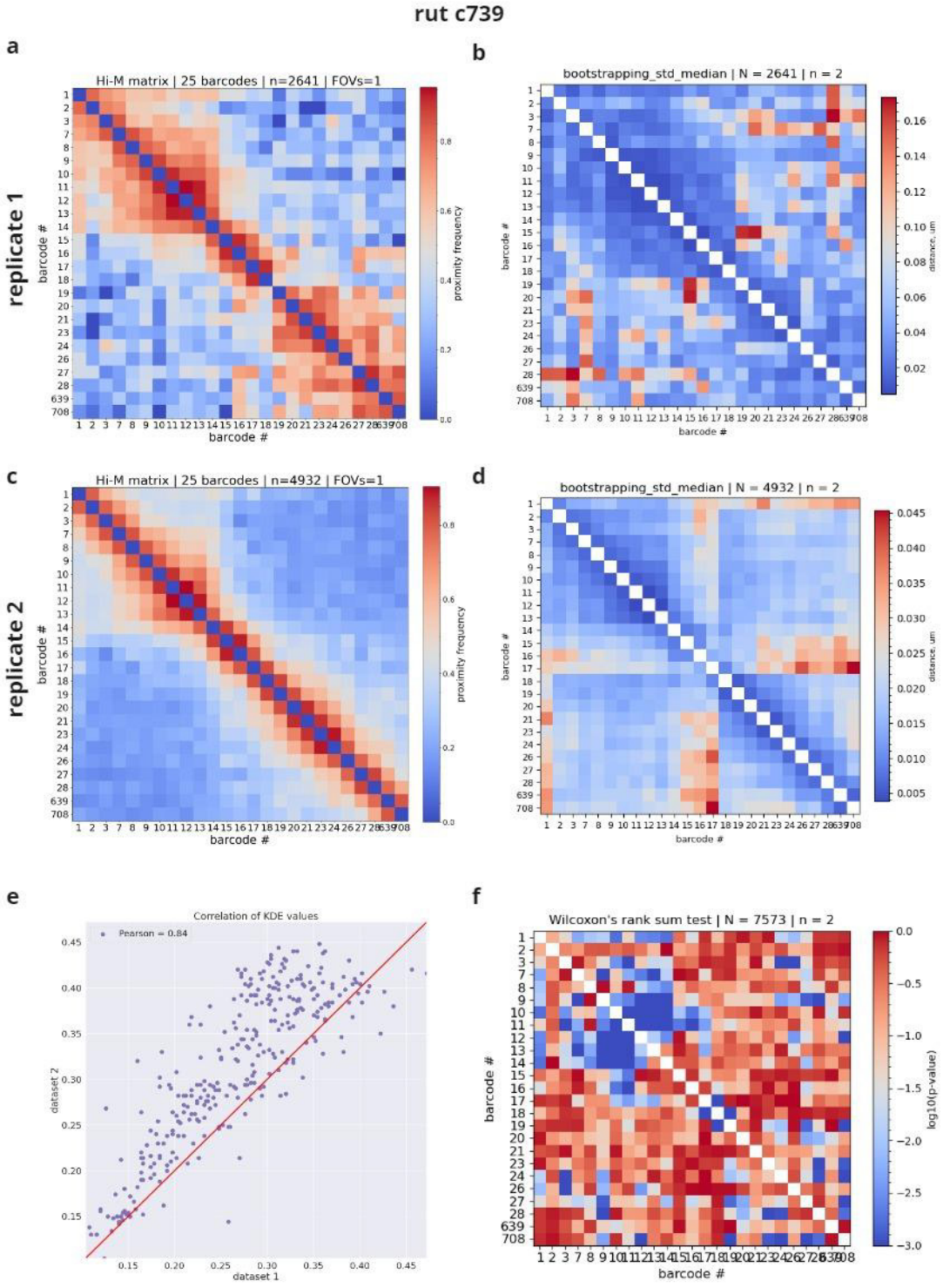
Supplementary figure 1 : Comparison between Kenyon cells and other brain cells in the *rut* locus - (a) other brain cells Hi-M matrix (non OK107–dataset2) (b) Kenyon cells Hi-M matrix (OK107–dataset1) (c) Pearson correlation between the matrices of the OK107 and non OK107 (d) Distributions of distances inside OK107 and non OK107 datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of OK107 and non OK107 conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of OK107 vs non OK107



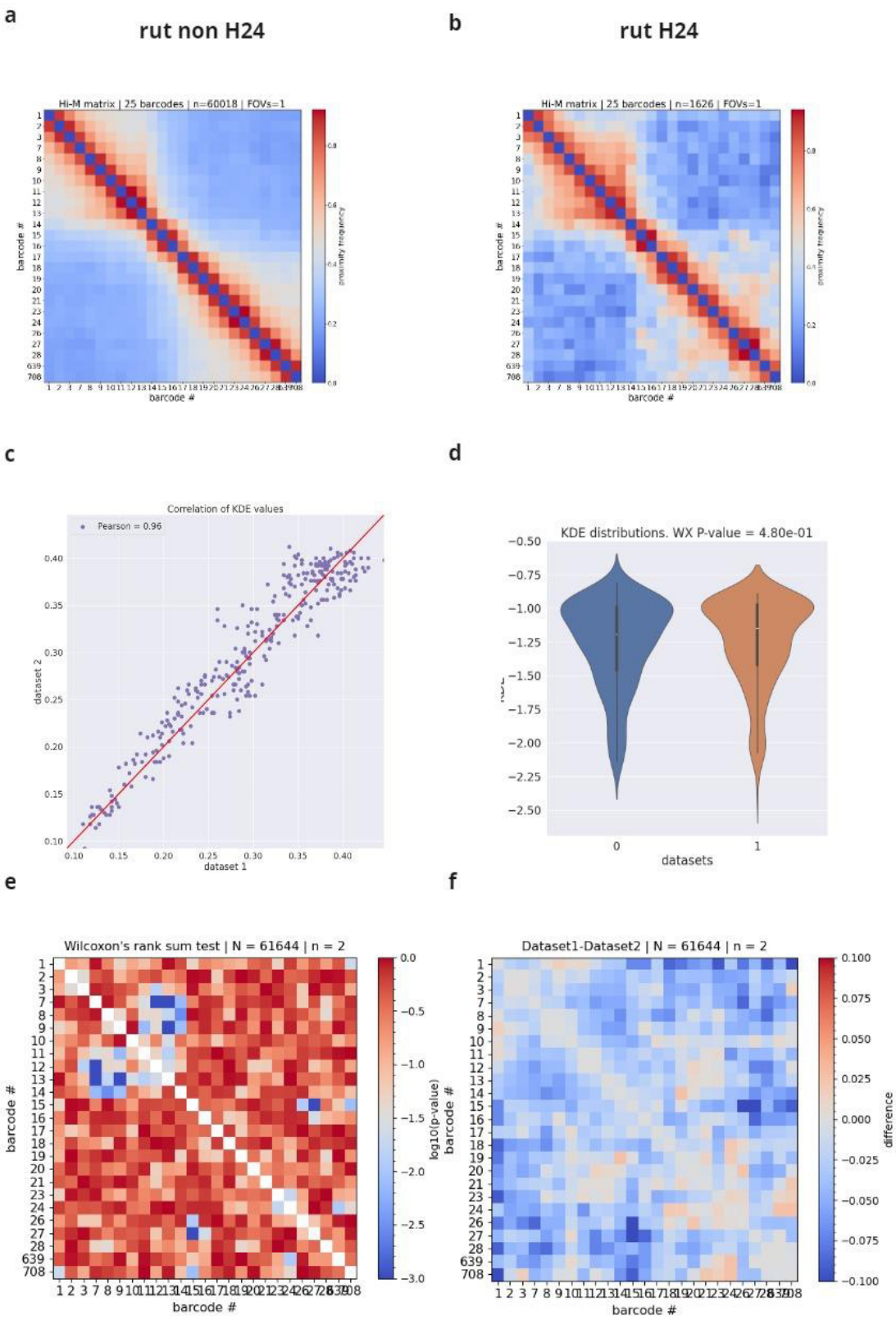
Supplementary figure 2 : Comparison between two replicates for the Kenyon cells in the *rut* locus - (a) Kenyon cells/OK107 Hi-M matrix of the first replicate (b) Distance error between probes matrix. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (c) Kenyon cells/OK107 Hi-M matrix of the second replicate (d)Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two OK107 replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two OK107 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



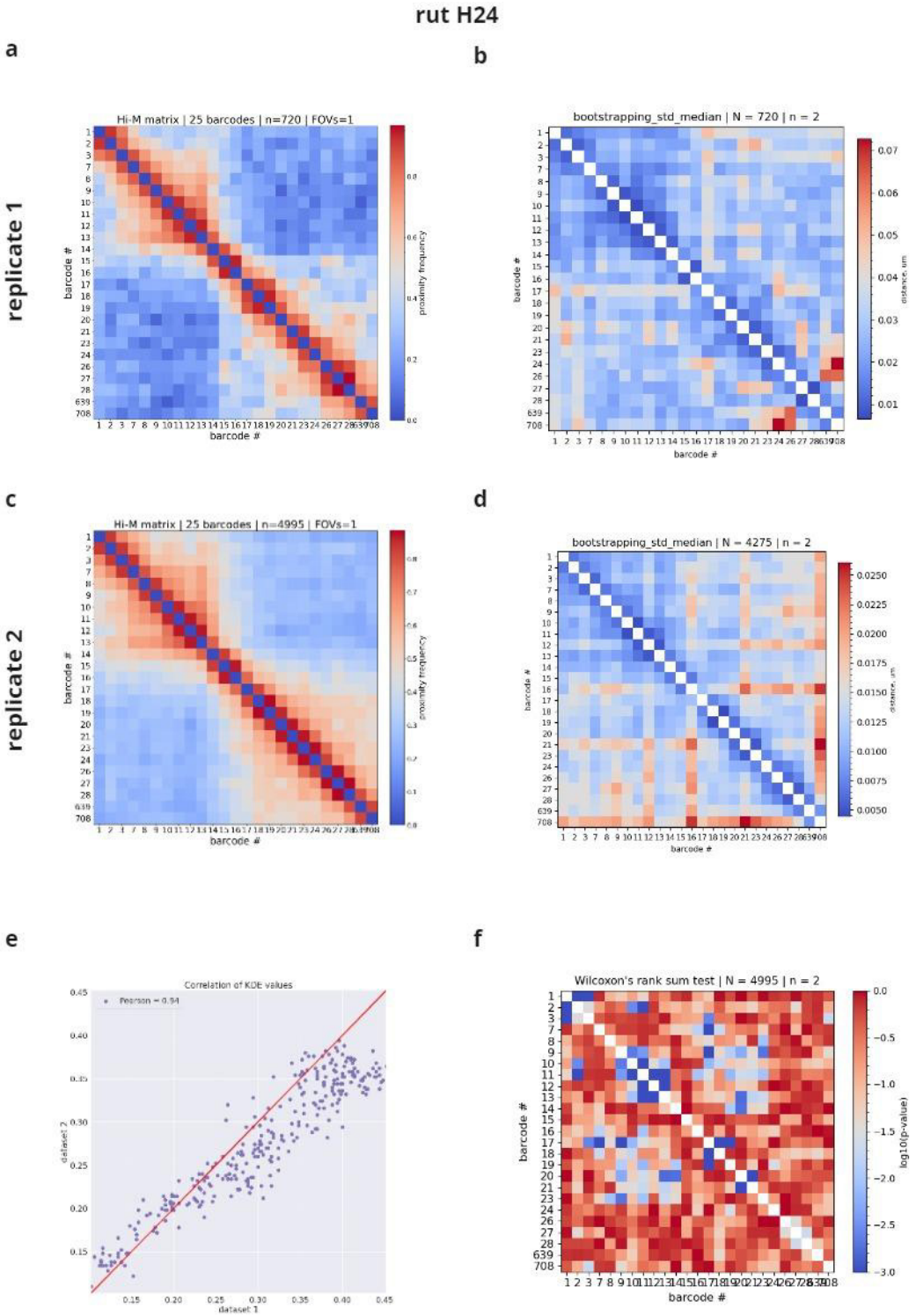
Supplementary figure 3 : Comparison between αβ Kenyon cells and other brain cells in the *rut* locus - (a) other brain cells Hi-M matrix (non c739–dataset2)) (b) αβ Kenyon cells Hi-M matrix (c739-dataset1) (c) Pearson correlation between the matrices of the c739 and non c739 (d) Distributions of distances inside c739 and non c739 datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of c739 and non c739 conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of c739 vs non c739



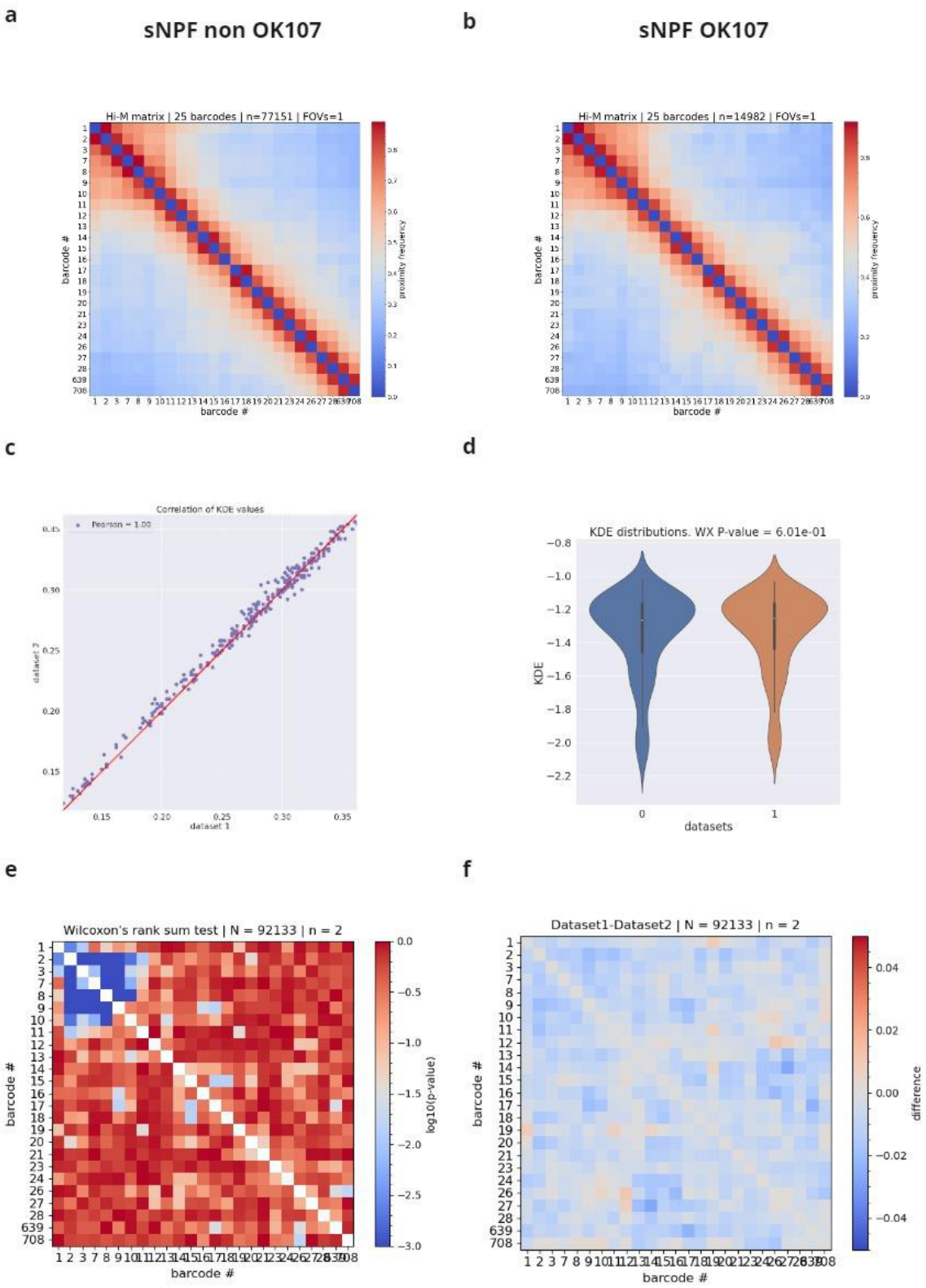
Supplementary figure 4 : Comparison between two replicates for the αβ Kenyon cells in the *rut* locus - (a) αβ Kenyon cells/c739 Hi-M matrix of the first replicate (b) Distance error between probes matrix. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (c) αβ Kenyon cells/c739 Hi-M matrix of the second replicate (d)Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two c739 replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two c739 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



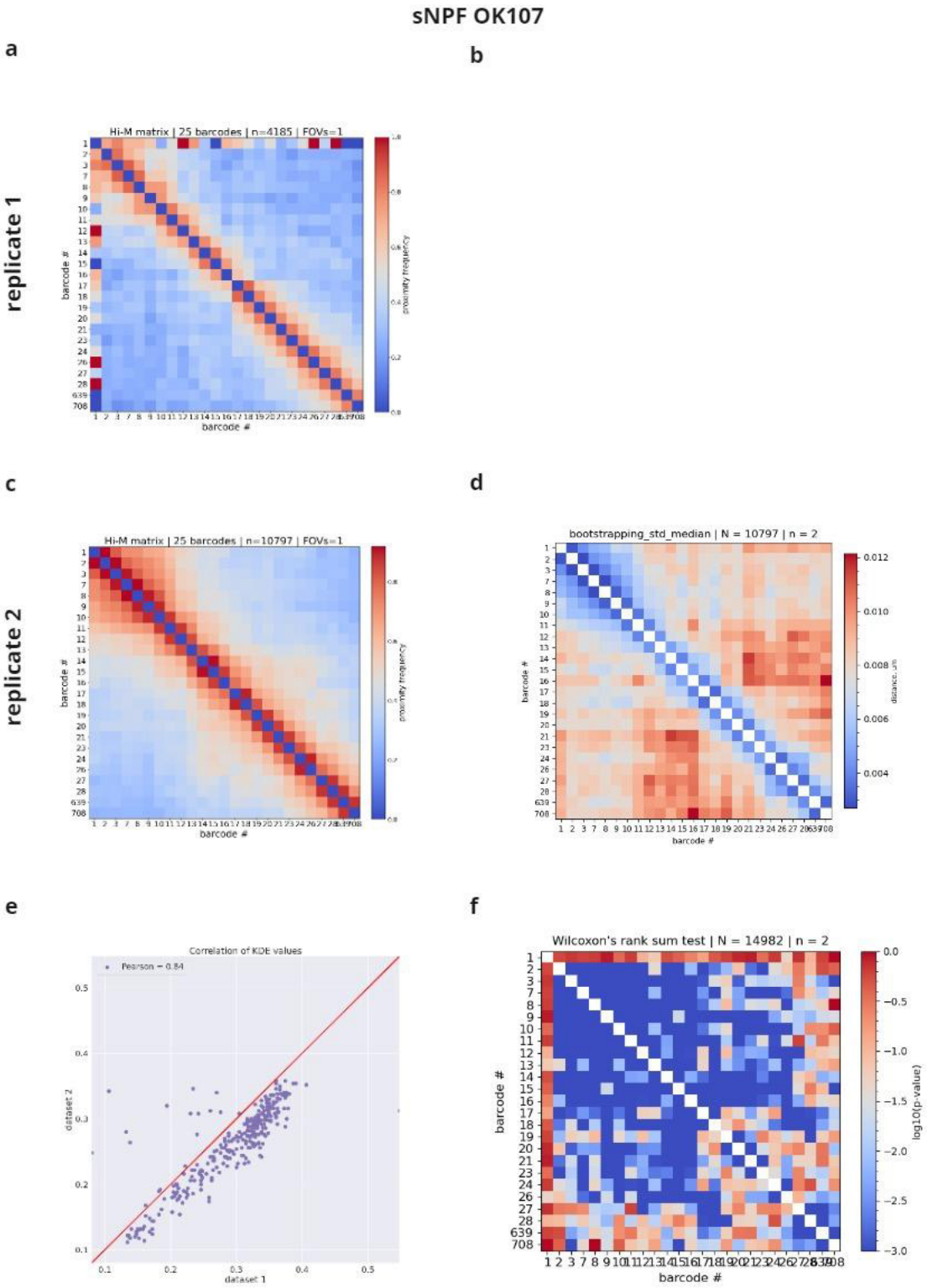
Supplementary figure 5 : Comparison between γ Kenyon cells and other brain cells in the *rut* locus - (a) other brain cells Hi-M matrix (non H24–dataset2)) (b) γ Kenyon cells Hi-M matrix (H24-dataset1) (c) Pearson correlation between the matrices of the H24 and non H24 (d) Distributions of distances insideH24 and non H24 datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of H24 and non H24 conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of H24 vs non H24



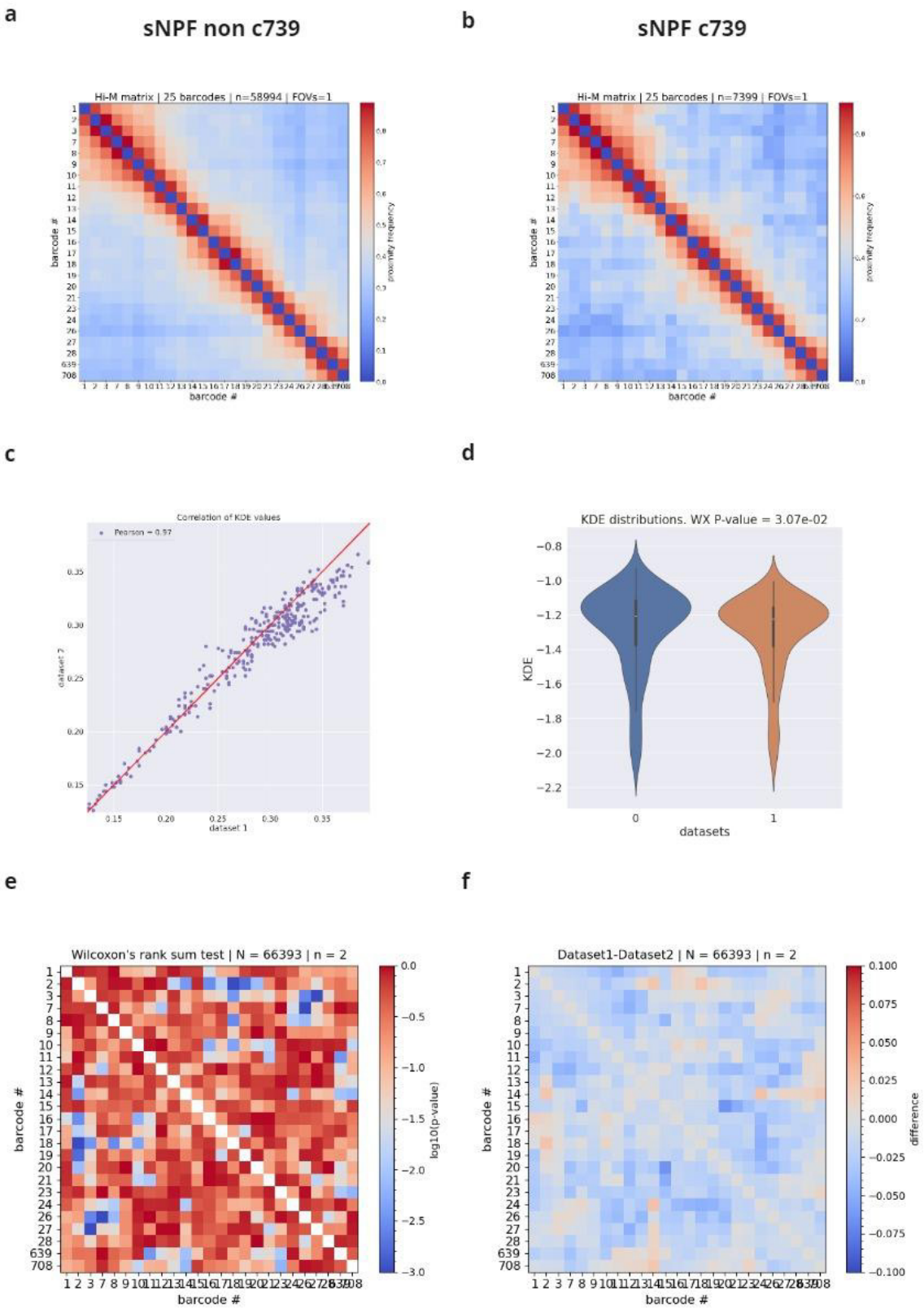
Supplementary figure 6 : Comparison between two replicates for the γ Kenyon cells in the *rut* locus - (a) γ Kenyon cells/H24 Hi-M matrix of the first replicate (b) Distance error between probes matrix of the first replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (c) γ Kenyon cells/H24 Hi-M matrix of the second replicate (d) Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two H24 replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two c739 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



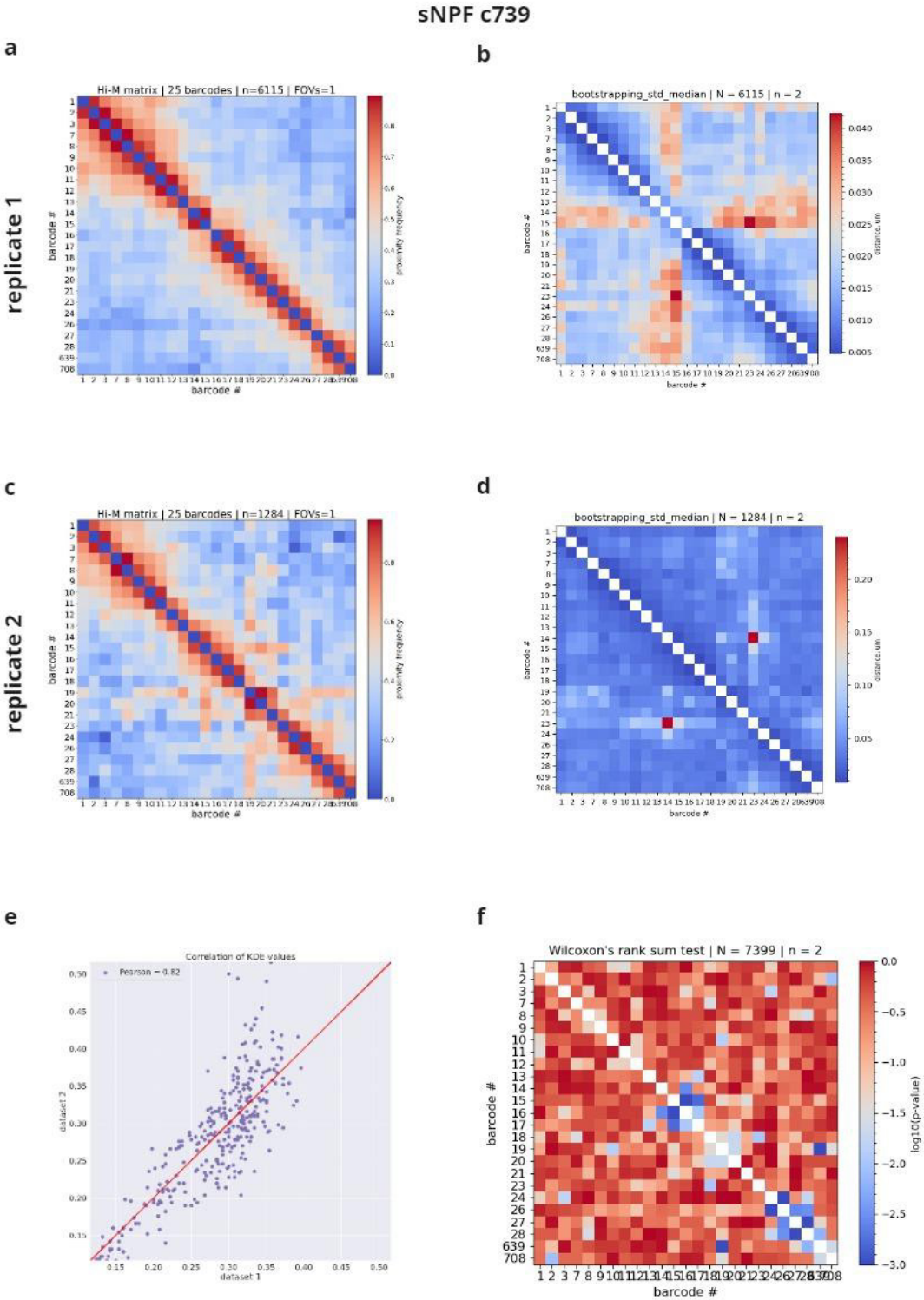
Supplementary figure 7 : Comparison between Kenyon cells and other brain cells in the *sNPF* locus - (a) other brain cells Hi-M matrix (non OK107–dataset2)) (b) Kenyon cells Hi-M matrix (OK107–dataset1) (c) Pearson correlation between the matrices of the OK107 and non OK107 (d) Distributions of distances inside OK107 and non OK107 datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of OK107 and non OK107 conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of OK107 vs non OK107



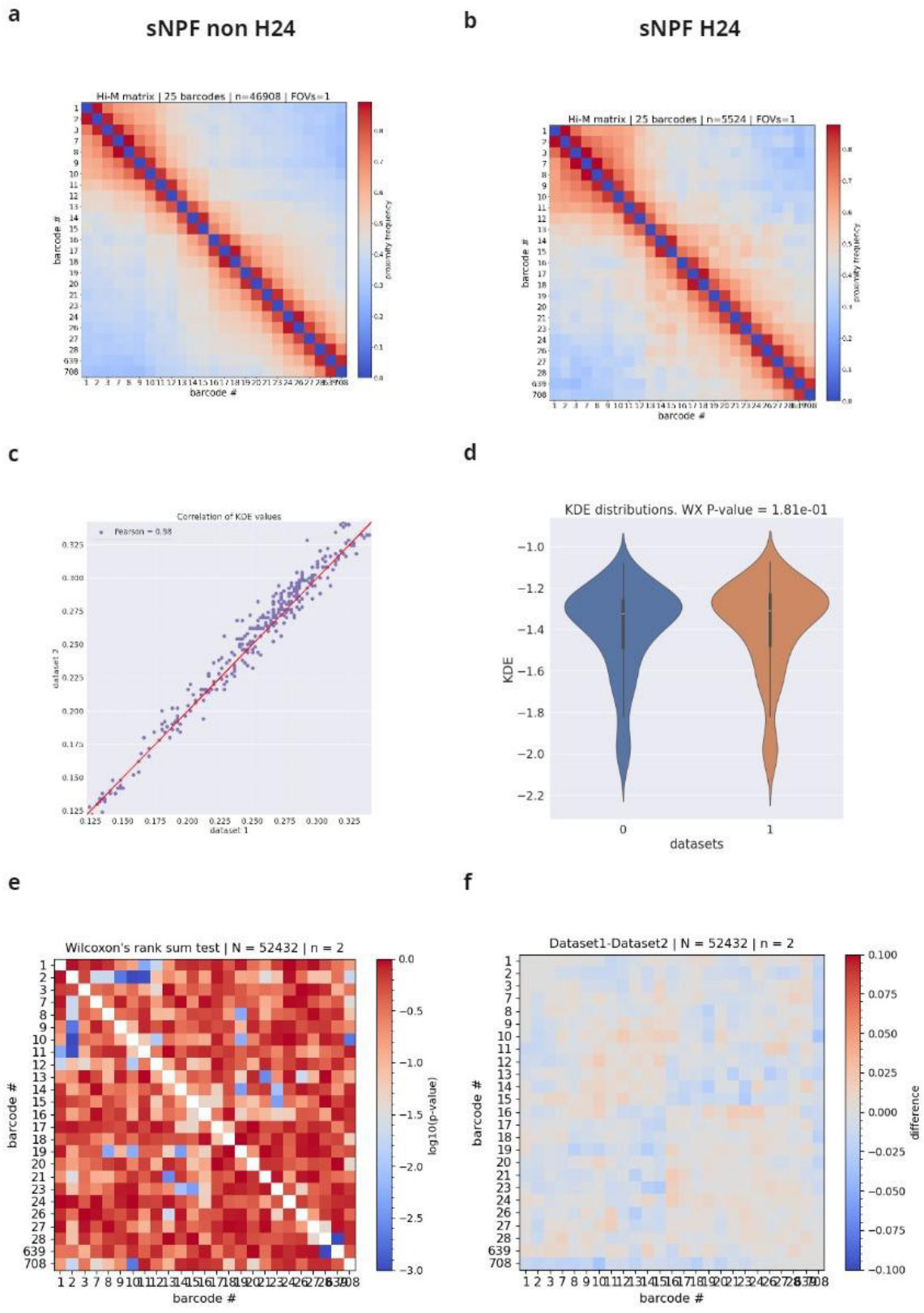
Supplementary figure 8 : Comparison between two replicates for the Kenyon cells in the *sNPF* locus - (a) Kenyon cells/OK107 Hi-M matrix of the first replicate (c) Kenyon cells/OK107 Hi-M matrix of the second replicate (d)Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two OK107 replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two OK107 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



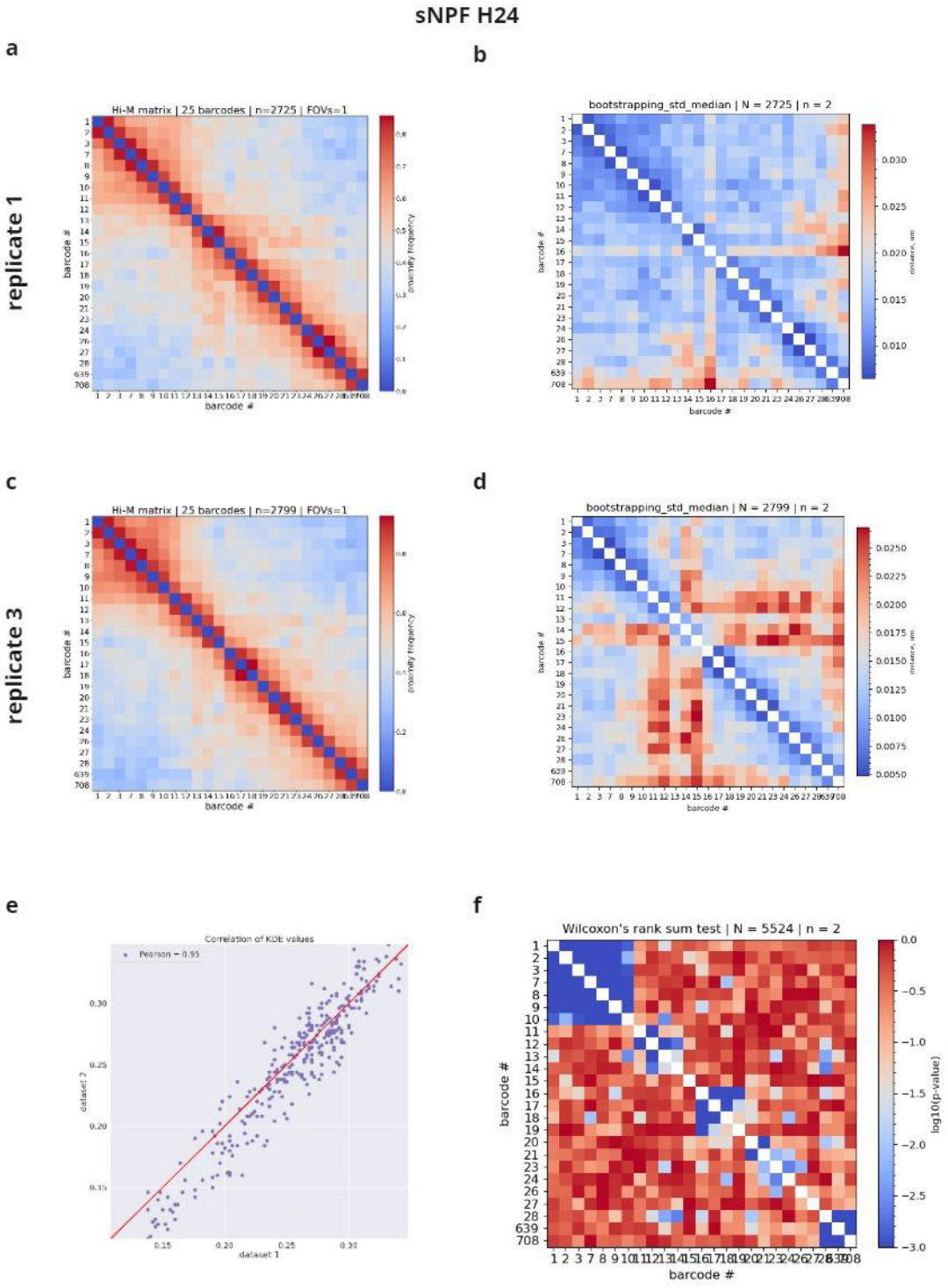
Supplementary figure 9 : Comparison between αβ Kenyon cells and other brain cells in the *sNPF* locus - (a) other brain cells Hi-M matrix (non c739–dataset2) (b) αβ Kenyon cells Hi-M matrix (c739–dataset1) (c) Pearson correlation between the matrices of the c739 and non c739 (d) Distributions of distances inside c739 and non c739 datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of c739 and non c739 conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of c739 vs non c739



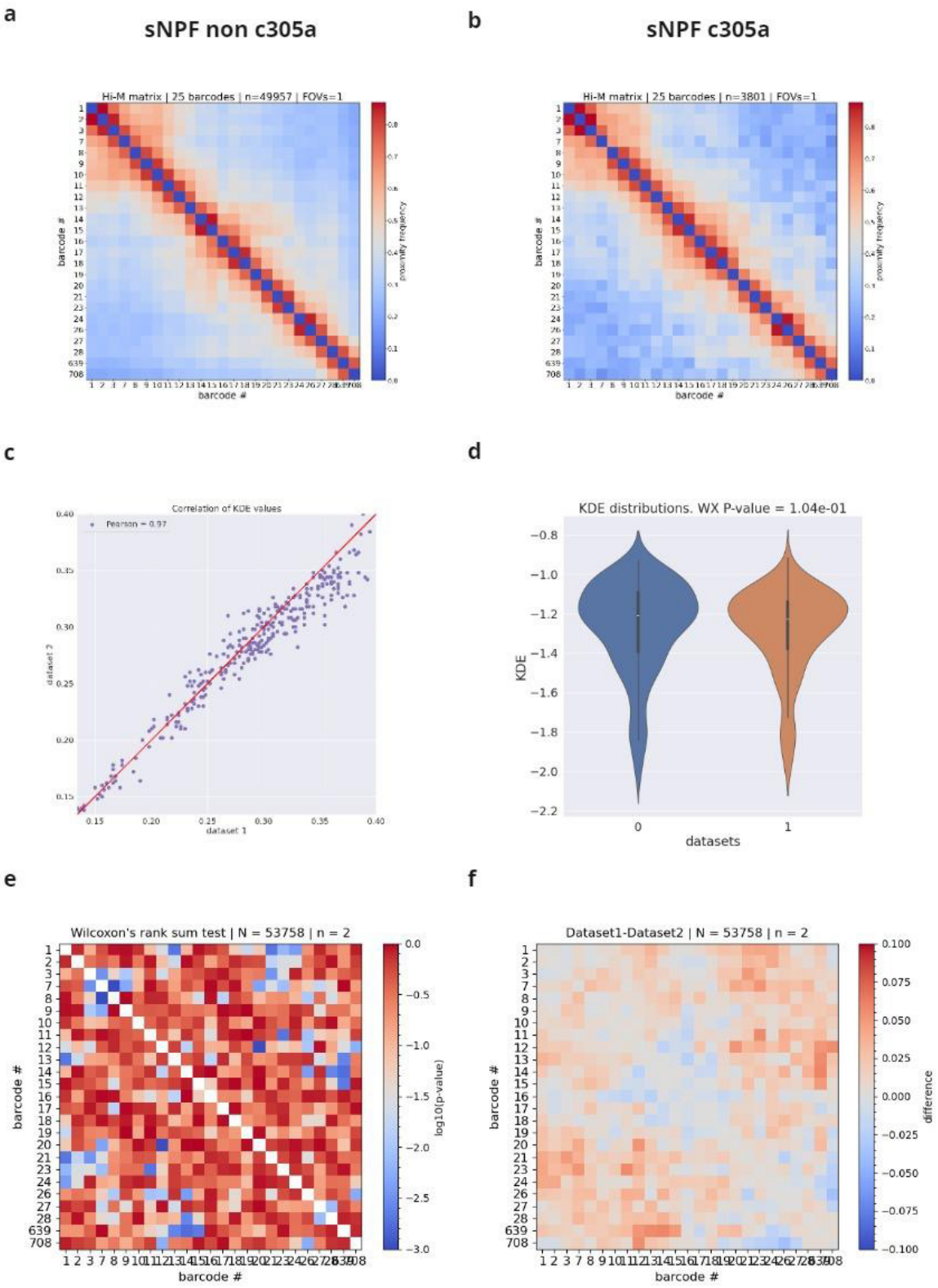
Supplementary figure 10 : Comparison between two replicates for the αβ Kenyon cells in the *sNPF* locus - (a) αβ Kenyon cells/c739 Hi-M matrix of the first replicate (b) Distance error between probes matrix. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (c) αβ Kenyon cells/c739 Hi-M matrix of the second replicate (d)Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two c739 replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two c739 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



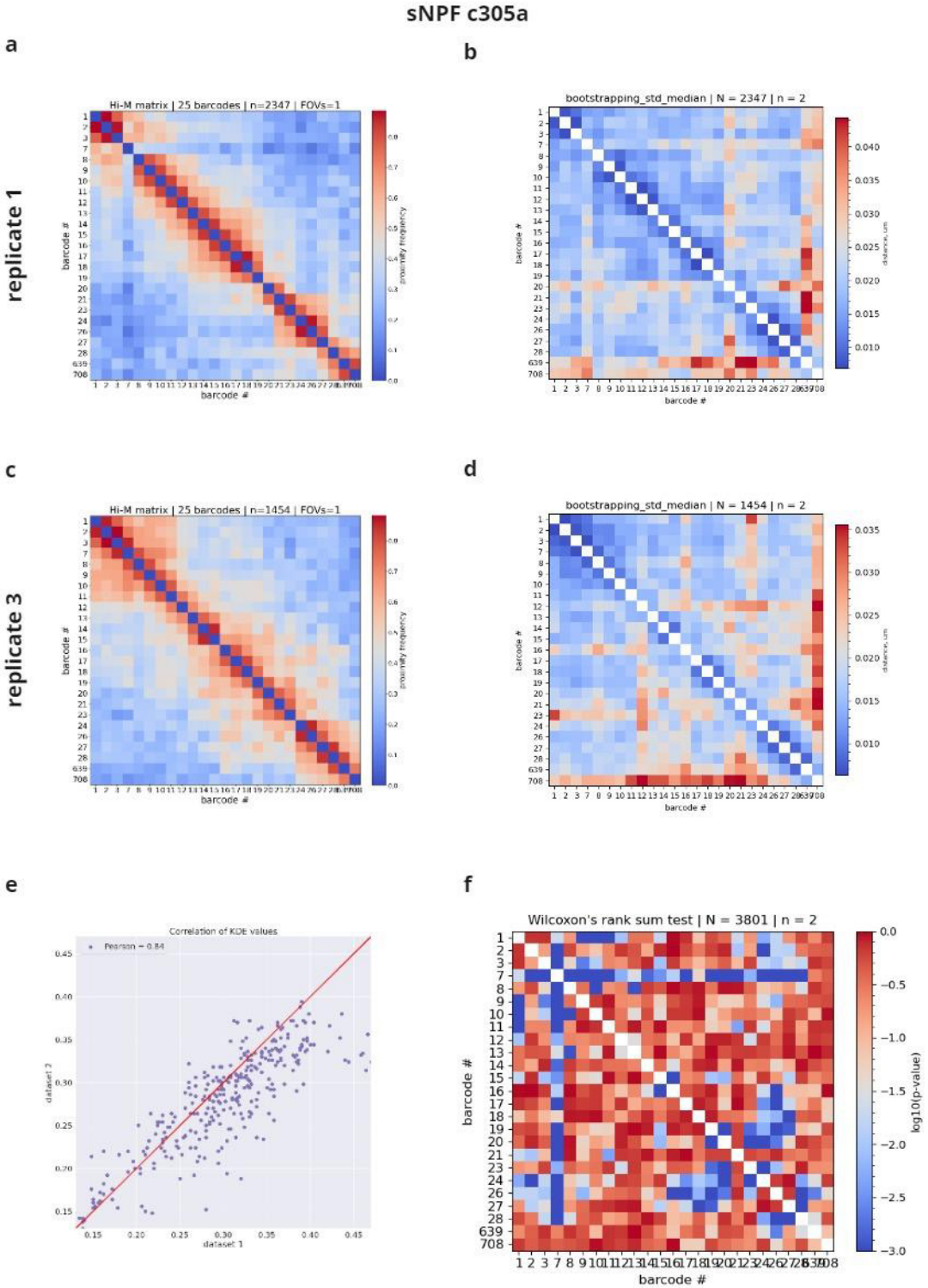
Supplementary figure 11 : Comparison between γ Kenyon cells and other brain cells in the *sNPF* locus - (a) other brain cells Hi-M matrix (non H24–dataset2)) (b) γ Kenyon cells Hi-M matrix (H24–dataset1) (c) Pearson correlation between the matrices of the H24 and non H24 (d) Distributions of distances insideH24 and non H24 datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of H24 and non H24 conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of H24 vs non H24



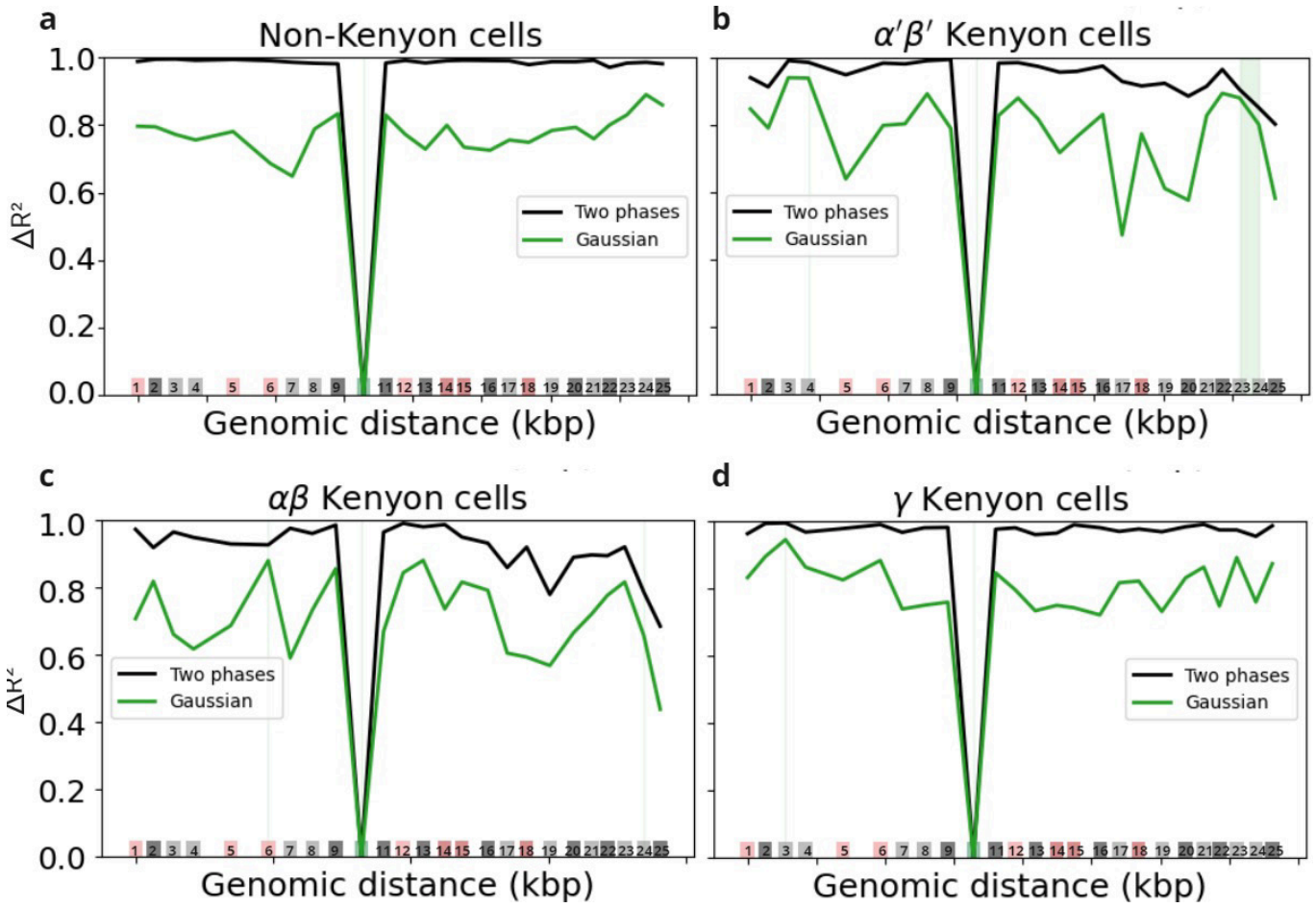
Supplementary figure 12 : Comparison between two replicates for the γ Kenyon cells in the *sNPF* locus - (a) γ Kenyon cells/H24 Hi-M matrix of the first replicate (b) Distance error between probes matrix of the first replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (c) γ Kenyon cells/H24 Hi-M matrix of the second replicate (d) Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two H24 replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two c739 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



Supplementary figure 13 : Comparison between α'β' Kenyon cells and other brain cells in the *sNPF* locus - (a) other brain cells Hi-M matrix (non c305a–dataset2) (b) α'β' Kenyon cells Hi-M matrix (c305a–dataset1) (c) Pearson correlation between the matrices of the c305a and non c305a (d) Distributions of distances inside c305a and non c305a datasets (e) Wilcoxon two-sided rank tests between the pairwise distance distributions of c305a and non c305a conditions. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level). (f) Distance difference matrix of c305a vs non c305a



Supplementary figure 14 : Comparison between two replicates for the α'β' Kenyon cells in the *sNPF* locus - (a) α'β' Kenyon cells/c305a Hi-M matrix of the first replicate (b) Distance error between probes matrix. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (c) α'β' Kenyon cells/c305a Hi-M matrix of the second replicate (d) Distance error between probes matrix of the second replicate. Performed with bootstrapping analysis (1000 bootstrapping cycles) to estimate the standard deviation in the determination of the median distance (e) Pearson correlation between the matrices of the two c305a replicates (f) Wilcoxon two-sided rank tests between the pairwise distance distributions of the two c739 replicates. Hypothesis : two independent samples follow the same distribution. P-values < 0.05 were considered significant to reject the hypothesis (i.e. 5% significance level).



Supplementary figure 15 : Goodness of fit of the theoretical model compared to experimental results for the sNPF locus - (a-b-c-d) Goodness of fit as a function of the genomic distance with *sNPF*'s promoter (teal) as anchor of the four conditions : (a) non-Kenyon cells/other cells of the brain (b) αβ Kenyon cells (c) α'β' Kenyon cells and (d) γ Kenyon cells. Two Gaussians fit (black curve). One Gaussian fit (green curve). The genomic location for which the one Gaussian fits best the experimental data, thus for our analysis these data have been ignored as it does not present the two dynamic regimes (light green vertical lines). Repartition of the Hi-M probes along the genome and coloured accordingly : promoter (teal), tested enhancers (red), enhancers predicted by [Janssens et al., 2022] (pink), accessible regions (light grey) and inaccessible regions (dark grey).

2. Additional publications

Qudi-HiM: an open-source acquisition software package for highly multiplexed sequential and combinatorial optical imaging

This work published in Open Research Europe in 2022 presents Qudi-HiM, a software package written in Python 3. Qudi-HiM was mainly developed by Franziska Barho and Jean-Bernard Fiche, two members or former members of my host lab. This acquisition software for multiplexed and combinatorial optical microscopy is highly flexible as it offers the possibility to control various independent hardware components. Qudi-HiM automates acquisition of multicolour microscopy images, controls microfluidic injections, and the remote monitoring of ongoing acquisitions. In this research project, I actively participated in the validation of the acquisition software for months by identifying and reporting bugs. Qudi-HiM is now used as a routine software in the lab for Hi-M acquisition and is available in open access for the microscopy community.

Title: Qudi-HiM: an open-source acquisition software package for highly multiplexed sequential and combinatorial optical imaging
Authors: Franziska Barho, Jean-Bernard Fiche, **Marion Bardou**, Olivier Messina, Alexandre Martiniere, Christophe Houbron, Marcelo Nollmann
Journal: Open Research Europe
Status: Published
Year: 2022
URL: https://pubmed.ncbi.nlm.nih.gov/37645324/190

pyHiM, a new open-source, multi-platform software package for spatial genomics based on multiplexed DNA-FISH imaging

This work has been published in Genome Biology in 2024 and presents pyHiM a software package that allows the analysis of multiplexed DNA-FISH data. In this research project, I actively participated in benchmarking and debugging activities. pyHiM is now used as a routine software in the lab for Hi-M data analysis.

Title: pyHiM, a new open-source, multi-platform software package for spatial genomics based on multiplexed DNA-FISH imaging
Authors: Devos Xavier, Fiche Jean-Bernard, **Bardou Marion**, Messina Olivier, Houbron Christophe, Gurgo Julian, Schaeffer Marie, Götz Markus, Walter Thomas, Mueller Florian, Nollmann Marcelo
Journal: Genome Biology
Status: Published
Year: 2024
URL: https://doi.org/10.1186/s13059-024-03178-x

Bibliography

Adel, Mohamed and Leslie C. Griffith (June 2021). "The Role of Dopamine in Associative Learning in Drosophila: An Updated Unified Model". In: *Neuroscience Bulletin* 37 (6), pp. 831–852. ISSN: 19958218. DOI: 10.1007/s12264-021-00665-0.

Ali, Tamer, Rainer Renkawitz, and Marek Bartkuhn (Apr. 2016). *Insulators and domains of gene expression*. DOI: 10.1016/j.gde.2015.11.009.

Allen, Benjamin L. and Dylan J. Taatjes (Mar. 2015). *The Mediator complex: A central integrator of transcription*. DOI: 10.1038/nrm3951.

Almeida, Bernardo P. de, Christoph Schaub, Michaela Pagani, Stefano Secchia, Eileen E.M. Furlong, and Alexander Stark (Feb. 2024). "Targeted design of synthetic enhancers for selected tissues in the Drosophila embryo". In: *Nature* 626 (7997), pp. 207–211. ISSN: 14764687. DOI: 10.1038/s41586-023-06905-9.

Amrein, Hubert and Natasha Thorne (Sept. 2005). "Gustatory perception and behavior in Drosophila melanogaster". In: *Current Biology* 15 (17). ISSN: 09609822. DOI: 10.1016/j.cub.2005.08.021.

Anreiter, Ina, Stephanie D. Biergans, and Marla B. Sokolowski (2019). "Epigenetic regulation of behavior in Drosophila melanogaster". In: *Current Opinion in Behavioral Sciences* 25 (Figure 1), pp. 44–50. ISSN: 23521546. DOI: 10.1016/j.cobeha.2018.06.010. URL: https://doi.org/10.1016/j.cobeha.2018.06.010.

Arendt, Detlev, Jacob M. Musser, Clare V.H. Baker, Aviv Bergman, Connie Cepko, Douglas H. Erwin, Mihaela Pavlicev, Gerhard Schlosser, Stefanie Widder, Manfred D. Laubichler, and Gunter P. Wagner (Dec. 2016). *The origin and evolution of cell types*. DOI: 10.1038/nrg.2016.127.

Arnold, Cosmas D., Christoph Stelzer Daniel Gerlach, Łukasz M. Boryn, Martina Rath, and Alexander Stark (Mar. 2013). "Genome-Wide Quantitative EnhancerActivity Maps Identified by STARR-seq". In: *Science* 339 (6123), pp. 1074–1077. ISSN: 10959203. DOI: 10.1126/science.1232518.

Aso, Yoshinori, Kornelia Grübel, Sebastian Busch, Anja B. Friedrich, Igor Siwanowicz, and Hiromu Tanimoto (2009). "The mushroom body of adult Drosophila characterized by GAL4 drivers". In: *Journal of Neurogenetics* 23 (1-2), pp. 156–172. ISSN: 15635260. DOI: 10.1080/01677060802471718.

Aso, Yoshinori, Daisuke Hattori, Yang Yu, Rebecca M. Johnston, Nirmala A. Iyer, Teri T.B. Ngo, Heather Dionne, L. F. Abbott, Richard Axel, Hiromu Tanimoto, and Gerald M. Rubin (2014). "The neuronal architecture of the mushroom body provides a logic for associative learning". In: *eLife* 3, e04577. ISSN: 2050084X. DOI: 10.7554/eLife.04577.

Awasaki, Takeshi, Mai Saito, Masaki Sone, Emiko Suzuki, Ryoko Sakai, Kei Ito, and Chihiro Hama (2000). "The Drosophila Trio Plays an Essential Role in Patterning of Axons by Regulating Their Directional Extension Recent studies using cultured cells have shown that the members of the Rho family of small GTPases, which include Rho, Rac, and Cdc42, function in the receptor-mediated signaling cascades that control the motility of growth cones (Hall, 1998). For example, activation of". In: *Neuron* 26, pp. 119–131.

Ayaz, Derya, Maarten Leyssen, Marta Koch, Jiekun Yan, Mohammed Srahna, Vasu Sheeba, Keri J. Fogle, Todd C. Holmes, and Bassem A. Hassan (June 2008). "Axonal injury and regeneration in the adult brain of Drosophila". In: *Journal of Neuroscience* 28 (23), pp. 6010–6021. ISSN: 02706474. DOI: 10.1523/JNEUROSCI.0101-08.2008.

Balasubramanian, Deevitha, Pedro Borges Pinto, Alexia Grasso, Severine Vincent, Helene Tarayre, Damien Lajoignie, and Yad Ghavi-Helm (Feb. 2024). "Enhancer–promoter interactions can form independently of genomic distance and be functional across TAD boundaries". In: *Nucleic Acids Research* 52 (4), pp. 1702–1719. ISSN: 13624962. DOI: 10.1093/nar/gkad1183.

Banerji, Julian, Sandro Rusconi, and Walter Schaffner (1981). "Expression of a Beta Globin Gene Is Enhanced by Remote SV40 DNA Sequences". In: *Cell* 27, pp. 299–308.

Barho, Franziska, Jean-Bernard Fiche, Marion Bardou, Olivier Messina, Alexandre Martiniere, Christophe Houbron, and Marcelo NOLLMANN (Apr. 2022). "Qudi-HiM: an open-source acquisition software package for highly multiplexed sequential and combinatorial optical imaging". In: *Open Research Europe* 2, p. 46. DOI: 10.12688/openreseurope.14641.1.

Bartman, Caroline R., Sarah C. Hsu, Chris C.S. Hsiung, Arjun Raj, and Gerd A. Blobel (Apr. 2016). "Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping". In: *Molecular Cell* 62 (2), pp. 237–247. ISSN: 10974164. DOI: 10.1016/j.molcel.2016.03.007.

Bauman, JG, J Wiegant, P Borst, and P van Duijn (1980). "A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labelled RNA". In: *Academic Press* 128, pp. 485–490.

Beliveau, Brian J., Alistair N. Boettiger, Maier S. Avendaño, Ralf Jungmann, Ruth B. McCole, Eric F. Joyce, Caroline Kim-Kiselak, Frédéric Bantignies, Chamith Y. Fonseka, Jelena Erceg, Mohammed A. Hannan, Hien G. Hoang, David Colognori, Jeannie T. Lee, William M. Shih, Peng Yin, Xiaowei Zhuang, and Chao Ting Wu (May 2015). "Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes". In: *Nature Communications* 6. ISSN: 20411723. DOI: 10.1038/ncomms8147.

Beliveau, Brian J., Eric F. Joyce, Nicholas Apostolopoulos, Feyza Yilmaz, Chamith Y. Fonseka, Ruth B. McCole, Yiming Chang, Jin Billy Li, Tharanga Niroshini Senaratne, Benjamin R. Williams, Jean Marie Rouillard, and Chao Ting Wu (Dec. 2012). "Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes". In: *Proceedings of the National Academy of Sciences of the United States of America* 109 (52), pp. 21301–21306. ISSN: 00278424. DOI: 10.1073/pnas.1213818110.

Beliveau, Brian J., Jocelyn Y. Kishi, Guy Nir, Hiroshi M. Sasaki, Sinem K. Saka, Son C. Nguyen, Chao ting Wu, and Peng Yin (Mar. 2018). "OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes". In: *Proceedings of the National Academy of Sciences of the United States of America* 115 (10), E2183–E2192. ISSN: 10916490. DOI: 10.1073/pnas.1714530115.

Belle, J Steven de and Martin Heisenberg (1994). "Associative Odor Learning in Drosophila Abolished by Chemical Ablation of MushroomBodies". In: *Science* 263, pp. 692–695.

Berg, Stuart, Dominik Kutra, Thorben Kroeger, Christoph N Straehle, Bernhard X Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, Kemal Eren, Jaime I Cervantes, Buote Xu, Fynn Beuttenmueller, Adrian Wolny, Chong Zhang, Ullrich Koethe, Fred A Hamprecht, and Anna Kreshuk (2019). "ilastik: interactive machine learning for (bio)image analysis". In: *Nature Methods* 16, pp. 1226–1232. URL: https://forum.image.sc/tags/ilastik..

Betzig, Eric, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess (Sept. 2006). "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". In: *Science* 313, pp. 1642–1645. ISSN: 00368075. DOI: 10.1126/science.1130258.

Bhattacharya, Mallika, Savanna F. Lyda, and Elissa P. Lei (Aug. 2024). *Chromatin insulator mechanisms ensure accurate gene expression by controlling overall 3D genome organization*. DOI: 10.1016/j.gde.2024.102208.

Bickmore, Wendy A. and Bas Van Steensel (Mar. 2013). "Genome architecture: Domain organization of interphase chromosomes". In: *Cell* 152 (6), pp. 1270–1284. ISSN: 10974172. DOI: 10.1016/j.cell.2013.02.001.

Bintu, Bogdan, Leslie J. Mateo, Jun Han Su, Nicholas A. Sinnott-Armstrong, Mirae Parker, Seon Kinrot, Kei Yamaya, Alistair N. Boettiger, and Xiaowei Zhuang (2018). "Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells". In: *Science* 362 (6413). ISSN: 10959203. DOI: 10.1126/science.aau1783.

Blair, Seth S. (2008). "Segmentation in animals". In: *Current Biology* 18 (21), PR991–R995. URL: http://www.archaeologiemuseum.it/index_ice.html.

Blum, Allison L., Wanhe Li, Mike Cressy, and Josh Dubnau (Aug. 2009). "Short- and Long-Term Memory in Drosophila Require cAMP Signaling in Distinct Neuron Types". In: *Current Biology* 19 (16), pp. 1341–1350. ISSN: 09609822. DOI: 10.1016/j.cub.2009.07.016.

Bonev, Boyan, Netta Mendelson Cohen, Quentin Szabo, Lauriane Fritsch, Giorgio L. Papadopoulos, Yaniv Lubling, Xiaole Xu, Xiaodan Lv, Jean Philippe Hugnot, Amos Tanay, and Giacomo Cavalli (2017). "Multiscale 3D Genome Rewiring during Mouse Neural Development". In: *Cell* 171 (3), 557–572.e24. ISSN: 10974172. DOI: 10.1016/j.cell.2017.09.043. URL: https://doi.org/10.1016/j.cell.2017.09.043.

Borggrefe, Tilman and Xiaojing Yue (2011). *Interactions between subunits of the Mediator complex with gene-specific transcription factors*. DOI: 10.1016/j.semcdb.2011.07.022.

Bowman, Sarah K., Aimee M. Deaton, Heber Domingues, Peggy I. Wang, Ruslan I. Sadreyev, Robert E. Kingston, and Welcome Bender (July 2014). "H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex". In: *eLife* 3 (July2014), pp. 1–13. ISSN: 2050084X. DOI: 10.7554/eLife.02833.

Boyle, Shelagh, Susan Gilchrist, Joanna M Bridger, Nicola L Mahy, Juliet A Ellis, and Wendy A Bickmore (2001). "The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells". In: *Human Molecular Genetics* 10 (3).

Brivanlou, Ali H and James E Darnell (2002). "Signal Transduction and the Control of Gene Expression". In: *Science* 295, pp. 813–818. URL: www.sciencemag.org.

Bushey, Ashley M., Elizabeth R. Dorman, and Victor G. Corces (Oct. 2008). *Chromatin Insulators: Regulatory Mechanisms and Epigenetic Inheritance*. DOI: 10.1016/j.molcel.2008.08.017.

Bushey, Ashley M., Edward Ramos, and Victor G. Corces (June 2009). "Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions". In: *Genes and Development* 23 (11), pp. 1338–1350. ISSN: 08909369. DOI: 10.1101/gad.1798209.

Calderon, Diego, Ronnie Blecher-Gonen, Xingfan Huang, Stefano Secchia, James Kentro, Riza M. Daza, Beth Martin, Alessandro Dulja, Christoph Schaub, Cole Trapnell, Erica Larschan, Kate M. O'Connor-Giles, Eileen E.M. Furlong, and Jay Shendure (Aug. 2022). "The continuum of Drosophila embryonic development at single-cell resolution". In: *Science* 377 (6606). ISSN: 10959203. DOI: 10.1126/science.abn5800.

Calo, Eliezer and Joanna Wysocka (Mar. 2013). *Modification of Enhancer Chromatin: What, How, and Why?* DOI: 10.1016/j.molcel.2013.01.038.

Cardozo-Gizzi, Andrés M., Diego I. Cattoni, Jean Bernard Fiche, Sergio M. Espinola, Julian Gurgo, Olivier Messina, Christophe Houbron, Yuki Ogiyama, Giorgio L. Papadopoulos, Giacomo Cavalli, Mounia Lagha, and Marcelo Nollmann (2019). "Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms". In: *Molecular Cell* 74 (1), 212–222.e5. ISSN: 10974164. DOI: 10.1016/j.molcel.2019.01.011.

Cardozo-Gizzi, Andrés M., Sergio M. Espinola, Julian Gurgo, Christophe Houbron, Jean Bernard Fiche, Diego I. Cattoni, and Marcelo Nollmann (2020). "Direct and simultaneous observation of transcription and chromosome architecture in single cells with Hi-M". In: *Nature Protocols* 15 (3), pp. 840–876. ISSN: 17502799. DOI: 10.1038/s41596-019-0269-9. URL: http://dx.doi.org/10.1038/s41596-019-0269-9.

Cavalheiro, Gabriel R, Charles Girardot, Rebecca R Viales, Tim Pollex, T B Ngoc Cao, Perrine Lacour, Songjie Feng, Adam Rabinowitz, and Eileen E M Furlong (2023). "CTCF, BEAF-32, and CP190 are not required for the establishment of TADs in early Drosophila embryos but have locus-specific roles". In: *Science Advances* 9 (eade1085). URL: https://www.science.org.

Chen, Hongtao, Michal Levo, Lev Barinov, Miki Fujioka, James B. Jaynes, and Thomas Gregor (Sept. 2018). "Dynamic interplay between enhancer–promoter topology and gene activity". In: *Nature Genetics* 50 (9), pp. 1296–1303. ISSN: 15461718. DOI: 10.1038/s41588-018-0175-z.

Chen, Zhuoxin, Valentina Snetkova, Grace Bower, Sandra Jacinto, Benjamin Clock, Atrin Dizehchi, Iros Barozzi, Brandon J. Mannion, Ana Alcaina-Caro, Javier Lopez-Rios, Diane E. Dickel, Axel Visel, Len A. Pennacchio, and Evgeny Z. Kvon (Apr. 2024). "Increased enhancer–promoter interactions during developmental enhancer activation in mammals". In: *Nature Genetics* 56 (4), pp. 675–685. ISSN: 15461718. DOI: 10.1038/s41588-024-01681-2.

Choi, Yoon Ha and Jong Kyoung Kim (2019). *Dissecting cellular heterogeneity using single-cell RNA sequencing*. DOI: 10.14348/molcells.2019.2446.

Chow, Ke-Huan K., Mark W. Budde, Maria Cabrera Alejandro A. Granados, Shinae Yoon, Soomin Cho, Ting-hao Huang, Noushin Koulena, Kirsten L. Frieda, Long Cai, Carlos Lois, and Michael B. Elowitz (2021). "Imaging cell lineage with a synthetic digital recording system". In: *Science* 372. URL: https://doi.org/10.1126/science.abb3099.

Comet, Itys, Ekaterina Savitskaya, Bernd Schuettengruber, Nicolas Negre, Sergey Lavrov, Aleksander Parshikov, François Juge, Elena Gracheva, Pavel Georgiev, and Giacomo Cavalli (July 2006). "PRE-Mediated Bypass of Two Su(Hw) Insulators Targets PcG Proteins to a Downstream Promoter". In: *Developmental Cell* 11 (1), pp. 117–124. ISSN: 15345807. DOI: 10.1016/j.devcel.2006.05.009.

Connolly, John B, Ian J H Roberts, J Douglas Armstrong, Kim Kaiser, Michael Forte, Tim Tully, and Cahir J O'kane (1996). "Associative Learning Disrupted by Impaired G s Signaling in Drosophila Mushroom Bodies". In: *Science* 274, pp. 2104–2107. URL: www.sciencemag.org.

Cooper, Sara J., Nathan D. Trinklein, Elizabeth D. Anton, Loan Nguyen, and Richard M. Myers (Jan. 2006). "Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome". In: *Genome Research* 16 (1), pp. 1–10. ISSN: 10889051. DOI: 10.1101/gr.4222606.

Costa, Marta, James D. Manton, Aaron D. Ostrovsky, Steffen Prohaska, and Gregory S.X.E. Jefferis (July 2016). "NBLAST: Rapid, Sensitive Comparison of Neuronal Structure and Construction of Neuron Family Databases". In: *Neuron* 91 (2), pp. 293–311. ISSN: 10974199. DOI: 10.1016/j.neuron.2016.06.012.

Cramer, Patrick (Sept. 2019). *Organization and regulation of gene transcription*. DOI: 10.1038/s41586-019-1517-4.

Crane, Emily, Qian Bian, Rachel Patton McCord, Bryan R. Lajoie, Bayly S. Wheeler, Edward J. Ralston, Satoru Uzawa, Job Dekker, and Barbara J. Meyer (July 2015). "Condensin-driven remodelling of X chromosome topology during dosage compensation". In: *Nature* 523 (7559), pp. 240–244. ISSN: 14764687. DOI: 10.1038/nature14450.

Cremer, T and C Cremer (2001). "CHROMOSOME TERRITORIES, NUCLEAR ARCHITECTURE AND GENE REGULATION IN MAMMALIAN CELLS". In: *Nature Reviews* 2, pp. 292–301. URL: www.nature.com/reviews/genetics.

Cremer, Thomas and Marion Cremer (2010). "Chromosome territories". In: *Cold Spring Harbor perspectives in biology* 2 (3). ISSN: 19430264. DOI: 10.1101/cshperspect.a003889.

Creyghton, Menno P., Albert W. Cheng, G. Grant Welstead, Tristan Kooistra, Bryce W. Carey, Eveline J. Steine, Jacob Hanna, Michael A. Lodato, Garrett M. Frampton, Phillip A. Sharp, Laurie A. Boyer, Richard A. Young, and Rudolf Jaenisch (Dec. 2010). "Histone H3K27ac separates active from poised enhancers and predicts developmental state". In: *Proceedings of the National Academy of Sciences of the United States of America* 107 (50), pp. 21931–21936. ISSN: 00278424. DOI: 10.1073/pnas.1016071107.

Crittenden, Jill R, Efthimios M C Skoulakis, Kyung-An Han, Daniel Kalderon, and Ronald L Davis (1998). *Tripartite Mushroom Body Architecture Revealed by Antigenic Markers*.

Crocker, Amanda, Xiao Juan Guan, Coleen T. Murphy, and Mala Murthy (2016). "Cell-Type-Specific Transcriptome Analysis in the Drosophila Mushroom Body Reveals Memory-Related Changes in Gene Expression". In: *Cell Reports* 15 (7), pp. 1580–1596. ISSN: 22111247. DOI: 10.1016/j.celrep.2016.04.046. URL: http://dx.doi.org/10.1016/j.celrep.2016.04.046.

Croset, Vincent, Christoph D. Treiber, and Scott Waddell (2017). "Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics". In: *bioRxiv*, pp. 1–31. DOI: 10.1101/237818.

Davie, Kristofer, Jasper Janssens, Duygu Koldere, Maxime De Waegeneer, Uli Pech, Łukasz Kreft, Sara Aibar, Samira Makhzami, Valerie Christiaens, Carmen Bravo González-Blas, Suresh Poovathingal, Gert Hulselmans, Katina I. Spanier, Thomas Moerman, Bram Vanspauwen, Sarah Geurs, Thierry Voet, Jeroen Lammertyn, Bernard Thienpont, Sha Liu, Nikos Konstantinides, Mark Fiers, Patrik Verstreken, and Stein Aerts (2018). "A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain". In: *Cell* 174 (4), 982–998.e20. ISSN: 10974172. DOI: 10.1016/j.cell.2018.05.057.

Davis, Ronald L. (Aug. 2023). "Learning and memory using Drosophila melanogaster: A focus on advances made in the fifth decade of research". In: *Genetics* 224 (4). ISSN: 19432631. DOI: 10.1093/genetics/iyad085.

Denaud, Sandrine, Marion Bardou, Giorgio-Lucio Papadopoulos, Stefan Grob, Marco Di, Gonzalo Sabarís, Marcelo Nollmann, Bernd Schuettengruber, and Giacomo Cavalli (n.d.). *A PRE loop at the dac locus acts as a topological chromatin structure that restricts and specifies enhancer promoter communication*.

Deng, Hao, Gaochen Jin, and Bomyi Lim (2022). "Unveiling dynamic enhancer-promoter interactions in Drosophila melanogaster". In: *Biochemical Society Transactions* 50 (6), pp. 1633–1642. ISSN: 14708752. DOI: 10.1042/BST20220325.

Deng, Wulan, Jongjoo Lee, Hongxin Wang, Jeff Miller, Andreas Reik, Philip D. Gregory, Ann Dean, and Gerd A. Blobel (June 2012). "Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor". In: *Cell* 149 (6), pp. 1233–1244. ISSN: 10974172. DOI: 10.1016/j.cell.2012.03.051.

Devos, Xavier, Jean Bernard Fiche, Marion Bardou, Olivier Messina, Christophe Houbron, Julian Gurgo, Marie Schaeffer, Markus Götz, Thomas Walter, Florian Mueller, and Marcelo Nollmann (Dec. 2024). "pyHiM: a new open-source, multi-platform software package for spatial genomics based on multiplexed DNA-FISH imaging". In: *Genome Biology* 25 (1). ISSN: 1474760X. DOI: 10.1186/s13059-024-03178-x.

Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren (May 2012). "Topological domains in mammalian genomes identified by analysis of chromatin interactions". In: *Nature* 485 (7398), pp. 376–380. ISSN: 00280836. DOI: 10.1038/nature11082.

Dong, Pengfei, Xiaoyu Tu, Po Yu Chu, Peitao Lu, Ning Zhu, Donald Grierson, Baijuan Du, Pinghua Li, and Silin Zhong (Dec. 2017). "3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments". In: *Molecular Plant* 10 (12), pp. 1497–1509. ISSN: 17529867. DOI: 10.1016/j.molp.2017.11.005.

Dorkenwald, Sven, Claire E. McKellar, Thomas Macrina, Nico Kemnitz, Kisuk Lee, Ran Lu, Jingpeng Wu, Sergiy Popovych, Eric Mitchell, Barak Nehoran, Zhen Jia, J. Alexander Bae, Shang Mu, Dodam Ih, Manuel Castro, Oluwaseun Ogedengbe, Akhilesh Halageri, Kai Kuehner, Amy R. Sterling, Zoe Ashwood, Jonathan Zung, Derrick Brittain, Forrest Collman, Casey Schneider-Mizell, Chris Jordan, William Silversmith, Christa Baker, David Deutsch, Lucas Encarnacion-Rivera, Sandeep Kumar, Austin Burke, Doug Bland, Jay Gager, James Hebditch, Selden Koolman, Merlin Moore, Sarah Morejohn, Ben Silverman, Kyle Willie, Ryan Willie, Szi chieh Yu, Mala Murthy, and H. Sebastian Seung (Jan. 2022). "FlyWire: online community for whole-brain connectomics". In: *Nature Methods* 19 (1), pp. 119–128. ISSN: 15487105. DOI: 10.1038/s41592-021-01330-0.

Dubnau, Josh and Tim Tully (1998). "GENE DISCOVERY IN DROSOPHILA: New Insights for Learning and Memory". In: *Annu. Rev. Neurosci* 21, pp. 407–451.

Dubnau.J, Grady.L, Kitamoto.T, and Tully.T (2001). "Disruption of neurotrans mission in Drosophila mushroom body blocks retrieval but not acquisition of memory". In: *Nature* 411, pp. 476–480.

Ernst, Jason and Manolis Kellis (Aug. 2010). "Discovery and characterization of chromatin states for systematic annotation of the human genome". In: *Nature Biotechnology* 28 (8), pp. 817–825. ISSN: 10870156. DOI: 10.1038/nbt.1662.

Espinola, Sergio Martin, Markus Götz, Maelle Bellec, Olivier Messina, Jean Bernard Fiche, Christophe Houbron, Matthieu Dejean, Ingolf Reim, Andrés M. Cardozo Gizzi, Mounia Lagha, and Marcelo Nollmann (2021). "Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development". In: *Nature Genetics* 53 (4), pp. 477–486. ISSN: 15461718. DOI: 10.1038/s41588-021-00816-z. URL: http://dx.doi.org/10.1038/s41588-021-00816-z.

Fahrbach, Susan E. (2006). "Structure of the mushroom bodies of the insect brain". In: *Annual Review of Entomology* 51 (56), pp. 209–232. ISSN: 00664170. DOI: 10.1146/annurev.ento.51.110104.150954.

Filion, Guillaume J., Joke G. van Bemmel, Ulrich Braunschweig, Wendy Talhout, Jop Kind, Lucas D. Ward, Wim Brugman, Ines J. de Castro, Ron M. Kerkhoven, Harmen J. Bussemaker, and Bas van Steensel (Oct. 2010). "Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells". In: *Cell* 143 (2), pp. 212–224. ISSN: 00928674. DOI: 10.1016/j.cell.2010.09.009.

Francis, Nicole J, Robert E Kingston, and Christopher L Woodcock (2004). "Chromatin Compaction by a Polycomb Group Protein Complex". In: *Science* 306, pp. 1574–1577. URL: https://www.science.org.

Freire-Pritchett, Paula, Stefan Schoenfelder, Csilla Vá rnai, Steven W Wingett, Jonathan Cairns, Amanda J Collier, Raquel García-Vílchez, Mayra Furlan-Magaril, Cameron S Osborne, Peter Fraser, Peter J Rugg-Gunn, and Mikhail Spivakov (2017). "Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells". In: *eLIFE* 6 (e21926). DOI: 10.7554/eLife.21926.001. URL: http://osf.io/sdbg4.

Fuda, Nicholas J., M. Behfar Ardehali, and John T. Lis (Sept. 2009). *Defining mechanisms that regulate RNA polymerase II transcription in vivo*. DOI: 10.1038/nature08449.

Fudenberg, Geoffrey, Maxim Imakaev, Carolyn Lu, Anton Goloborodko, and Leonid A Mirny (2016). "Formation of Chromosomal Domains by Loop Extrusion". In: 15 (9), pp. 2038–2049. DOI: 10.1016/j.celrep.2016.04.085.Formation.

Furlong, Eileen E.M. and Michael Levine (Sept. 2018). "Developmental enhancers and chromosome topology". In: *Science* 361 (6409), pp. 1341–1345. ISSN: 10959203. DOI: 10.1126/science.aau0320.

Gambetta, Maria Cristina and Eileen E.M. Furlong (Sept. 2018). "The insulator protein CTCF is required for correct hox gene expression, but not for embryonic development in Drosophila". In: *Genetics* 210 (1), pp. 129–136. ISSN: 19432631. DOI: 10.1534/genetics.118.301350.

Gaszner, Miklos and Gary Felsenfeld (Sept. 2006). *Insulators: Exploiting transcriptional and epigenetic mechanisms*. DOI: 10.1038/nrg1925.

George, Orphanides and Reinberg Danny (2002). "A Unified Theory of Gene Expression". In: *Cell* 108, pp. 439–451.

Gerasimova, Tatiana I, David A Gdula, Denis V Gerasimov, Olga Simonova, and Victor 0 Domes (1995). *A Drosophila Protein That Imparts Directionality on a Chromatin Insulator Is an Enhancer of Position-Effect Variegation*, pp. 587–597.

Ghavi-Helm, Yad, Felix A. Klein, Tibor Pakozdi, Lucia Ciglar, Daan Noordermeer, Wolfgang Huber, and Eileen E.M. Furlong (2014). "Enhancer loops appear stable during development and are associated with paused polymerase". In: *Nature* 512 (1), pp. 96–100. ISSN: 14764687. DOI: 10.1038/nature13417.

Gilchrist, Daniel A., Gilberto Dos Santos, David C. Fargo, Bin Xie, Yuan Gao, Leping Li, and Karen Adelman (Nov. 2010). "Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation". In: *Cell* 143 (4), pp. 540–551. ISSN: 00928674. DOI: 10.1016/j.cell.2010.10.004.

Gisselbrecht, Stephen S., Luis A. Barrera, Martin Porsch, Anton Aboukhalil, Preston W. Estep, Anastasia Vedenko, Alexandre Palagi, Yongsok Kim, Xianmin Zhu, Brian W. Busser, Caitlin E. Gamble, Antonina Iagovitina, Aditi Singhania, Alan M. Michelson, and Martha L. Bulyk (Aug. 2013). "Highly parallel assays of tissue-specific enhancers in whole Drosophila embryos". In: *Nature Methods* 10 (8), pp. 774–780. ISSN: 15487091. DOI: 10.1038/nmeth.2558.

Gisselbrecht, Stephen S., Alexandre Palagi, Jesse V. Kurland, Julia M. Rogers, Hakan Ozadam, Ye Zhan, Job Dekker, and Martha L. Bulyk (Jan. 2020). "Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts". In: *Molecular Cell* 77 (2), 324–337.e8. ISSN: 10974164. DOI: 10.1016/j.molcel.2019.10.004.

Gotz, Markus, Olivier Messina, Sergio Espinola, Jean Bernard Fiche, and Marcelo Nollmann (2022). "Multiple parameters shape the 3D chromatin structure of single nuclei at the doc locus in Drosophila". In: *Nature communications* 13 (1), p. 5375. ISSN: 20411723. DOI: 10.1038/s41467-022-32973-y.

Gurgo, Julian, Jean Charles Walter, Jean Bernard Fiche, Christophe Houbron, Marie Schaeffer, Giacomo Cavalli, Frédéric Bantignies, and Marcelo Nollmann (May 2024). "Multiplexed chromatin imaging reveals predominantly pairwise long-range coordination between Drosophila Polycomb genes". In: *Cell Reports* 43 (5). ISSN: 22111247. DOI: 10.1016/j.celrep.2024.114167.

Gyurkovics, H, J Gausz, J Kummer', and F Karch' (1990). *A new homeotic mutation in the Drosophila bithorax complex removes a boundary separating two domains of regulation*, pp. 2579–2585.

H., Cai and M. Levine (1995). "Modulation of enhancer-promoter interactions by insulators in the Drosophila embryo". In: *Nature* 376, pp. 533–536.

Haberle, Vanja and Alexander Stark (Oct. 2018). *Eukaryotic core promoters and the functional basis of transcription initiation*. DOI: 10.1038/s41580-018-0028-8.

Halfon, Marc S., Steven M. Gallo, and Casey M. Bergman (Jan. 2008). "REDfly 2.0: An integrated database of cis -regulatory modules and transcription factor binding sites in Drosophila". In: *Nucleic Acids Research* 36 (SUPPL. 1). ISSN: 03051048. DOI: 10.1093/nar/gkm876.

Hediger, Florence and Susan M. Gasser (Apr. 2006). *Heterochromatin protein 1: Don't judge the book by its cover!* DOI: 10.1016/j.gde.2006.02.013.

Heilemann, Mike, Sebastian Van De Linde, Mark Schuttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer (Aug. 2008). "Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes". In: *Angewandte Chemie - International Edition* 47 (33), pp. 6172–6176. ISSN: 14337851. DOI: 10.1002/anie.200802376.

Heintzman, Nathaniel D., Gary C. Hon, R. David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F. Harp, Zhen Ye, Leonard K. Lee, Rhona K. Stuart, Christina W. Ching, Keith A. Ching, Jessica E. Antosiewicz-Bourget, Hui Liu, Xinmin Zhang, Roland D. Green, Victor V. Lobanenkov, Ron Stewart, James A. Thomson, Gregory E. Crawford, Manolis Kellis, and Bing Ren (May 2009). "Histone modifications at human enhancers reflect global cell-type-specific gene expression". In: *Nature* 459 (7243), pp. 108–112. ISSN: 00280836. DOI: 10.1038/nature07829.

Heinz, Sven, Casey E. Romanoski, Christopher Benner, and Christopher K. Glass (Mar. 2015). *The selection and function of cell type-specific enhancers*. DOI: 10.1038/nrm3949.

Henry, Gilbert L., Fred P. Davis, Serge Picard, and Sean R. Eddy (2012). "Cell type-specific genomics of Drosophila neurons". In: *Nucleic Acids Research* 40 (19), pp. 9691–9704. ISSN: 03051048. DOI: 10.1093/nar/gks671.

Hess, Samuel T., Thanu P.K. Girirajan, and Michael D. Mason (2006). "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy". In: *Biophysical Journal* 91 (11), pp. 4258–4272. ISSN: 00063495. DOI: 10.1529/biophysj.106.091116.

Hou, Chunhui, Li Li, Zhaohui S. Qin, and Victor G. Corces (Nov. 2012). "Gene Density, Transcription, and Insulators Contribute to the Partition of the Drosophila Genome into Physical Domains". In: *Molecular Cell* 48 (3), pp. 471–484. ISSN: 10972765. DOI: 10.1016/j.molcel.2012.08.031.

Huang, Cheng, Jessica R. Maxey, Supriyo Sinha, Joan Savall, Yiyang Gong, and Mark J. Schnitzer (Dec. 2018). "Long-term optical brain imaging in live adult fruit flies". In: *Nature Communications* 9 (1). ISSN: 20411723. DOI: 10.1038/s41467-018-02873-1.

Hug, Clemens B., Alexis G. Grimaldi, Kai Kruse, and Juan M. Vaquerizas (2017). "Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription". In: *Cell* 169 (2), 216–228.e19. ISSN: 10974172. DOI: 10.1016/j.cell.2017.03.024. URL: http://dx.doi.org/10.1016/j.cell.2017.03.024.

Ing-Simmons, Elizabeth, Roshan Vaid, Xin Yang Bing, Michael Levine, Mattias Mannervik, and Juan M. Vaquerizas (2021). "Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning". In: *Nature Genetics* 53 (4), pp. 487–499. ISSN: 15461718. DOI: 10.1038/s41588-021-00799-x. URL: http://dx.doi.org/10.1038/s41588-021-00799-x.

Ito, Kei, Wakae Awano, Kazumi Suzuki, Yasushi Hiromi, and Daisuke Yamamoto (1997). "The Drosophila mushroom body is a quadruple structure of clonal units eachof which contains a virtually identical set of neurones and glial cells". In: *Development* 124, pp. 761–771.

Ito, Satomi, Adriana Magalska, Manuel Alcaraz-Iborra, Jose P. Lopez-Atalaya, Victor Rovira, Bruno Contreras-Moreira, Michal Lipinski, Roman Olivares, Jose Martinez-Hernandez, Blazej Ruszczycki, Rafael Lujan, Emilio Geijo-Barrientos, Grzegorz M. Wilczynski, and Angel Barco (2014). "Loss of neuronal 3d chromatin organization causes transcriptional and behavioural deficits related to serotonergic dysfunction". In: *Nature Communications* 5. ISSN: 20411723. DOI: 10.1038/ncomms5450.

Jagannathan, Madhav, Natalie Warsinger-Pepe, George J. Watase, and Yukiko M. Yamashita (2017). "Comparative analysis of satellite DNA in the drosophila melanogaster species complex". In: *G3: Genes, Genomes, Genetics* 7 (2), pp. 693–704. ISSN: 21601836. DOI: 10.1534/g3.116.035352.

Janssens, Jasper, Sara Aibar, Ibrahim Ihsan Taskiran, Joy N. Ismail, Alicia Estacio Gomez, Gabriel Aughey, Katina I. Spanier, Florian V. De Rop, Carmen Bravo González-Blas, Marc Dionne, Krista Grimes, Xiao Jiang Quan, Dafni Papasokrati, Gert Hulselmans, Samira Makhzami, Maxime De Waegeneer, Valerie Christiaens, Tony Southall, and Stein Aerts (2022). "Decoding gene regulation in the fly brain". In: *Nature* 601 (7894), pp. 630–636. ISSN: 14764687. DOI: 10.1038/s41586-021-04262-z. URL: http://dx.doi.org/10.1038/s41586-021-04262-z.

Jenett, Arnim, Gerald M. Rubin, Teri T.B. Ngo, David Shepherd, Christine Murphy, Heather Dionne, Barret D. Pfeiffer, Amanda Cavallaro, Donald Hall, Jennifer Jeter, Nirmala Iyer, Dona Fetter, Joanna H. Hausenfluck, Hanchuan Peng, Eric T. Trautman, Robert R. Svirskas, Eugene W. Myers, Zbigniew R. Iwinski, Yoshinori Aso, Gina M. DePasquale, Adrianne Enos, Phuson Hulamm, Shing Chun Benny Lam, Hsing Hsi Li, Todd R. Laverty, Fuhui Long, Lei Qu, Sean D. Murphy, Konrad Rokicki, Todd Safford, Kshiti Shaw, Julie H. Simpson, Allison Sowell, Susana Tae, Yang Yu, and Christopher T. Zugates (2012). "A GAL4-Driver Line Resource for Drosophila Neurobiology". In: *Cell Reports* 2 (4), pp. 991–1001. ISSN: 22111247. DOI: 10.1016/j.celrep.2012.09.011. URL: http://dx.doi.org/10.1016/j.celrep.2012.09.011.

Johard, Helena A.D., Lina E. Enell, Elisabeth Gustafsson, Pierre Trifilieff, Jan A. Veenstra, and Dick R. Nässel (Apr. 2008). "Intrinsic neurons of Drosophila mushroom bodies express short neuropeptide F: Relations to extrinsic neurons expressing different neurotransmitters". In: *Journal of Comparative Neurology* 507 (4), pp. 1479–1496. ISSN: 10969861. DOI: 10.1002/cne.21636.

Joshi, Chintan J., Wenfan Ke, Anna Drangowska-Way, Eyleen J. O'Rourke, and Nathan E. Lewis (July 2022). "What are housekeeping genes?" In: *PLoS Computational Biology* 18 (7 July). ISSN: 15537358. DOI: 10.1371/journal.pcbi.1010295.

Kaaij, Lucas J.T., Robin H. van der Weide, Rene F. Ketting, and Elzo de Wit (July 2018). "Systemic Loss and Gain of Chromatin Architecture throughout Zebrafish Development". In: *Cell Reports* 24 (1), 1–10.e4. ISSN: 22111247. DOI: 10.1016/j.celrep.2018.06.003.

Kamath, M. B., I. B. Houston, A. J. Janovski, X. Zhu, S. Gowrisankar, A. G. Jegga, and R. P. DeKoter (2008). "Dose-dependent repression of T-cell and natural killer cell genes by PU.1 enforces

myeloid and B-cell identity". In: *Leukemia* 22 (6), pp. 1214–1225. ISSN: 14765551. DOI: 10.1038/leu.2008.67.

Kaushal, Anjali, Giriram Mohana, Julien Dorier, Isa Özdemir, Arina Omer, Pascal Cousin, Anastasiia Semenova, Michael Taschner, Oleksandr Dergai, Flavia Marzetta, Christian Iseli, Yossi Eliaz, David Weisz, Muhammad Saad Shamim, Nicolas Guex, Erez Lieberman Aiden, and Maria Cristina Gambetta (2021). "CTCF loss has limited effects on global genome architecture in Drosophila despite critical regulatory functions". In: *Nature Communications* 12 (1). ISSN: 20411723. DOI: 10.1038/s41467-021-21366-2. URL: http://dx.doi.org/10.1038/s41467-021-21366-2.

Kawasaki, Koji and Takashi Fukaya (2024). "Regulatory landscape of enhancer-mediated transcriptional activation". In: *Trends in Cell Biology*, pp. 1–12. ISSN: 18793088. DOI: 10.1016/j.tcb.2024.01.008. URL: https://doi.org/10.1016/j.tcb.2024.01.008.

Kellum, Rebecca and Paul Schedl (1991). "A Position-Effect Assay for Boundaries of Higher Order Chromosomal Domains". In: *Cell* 64, pp. 941–950.

Kharchenko, Peter V., Artyom A. Alekseyenko, Yuri B. Schwartz, Aki Minoda, Nicole C. Riddle, Jason Ernst, Peter J. Sabo, Erica Larschan, Andrey A. Gorchakov, Tingting Gu, Daniela Linder-Basso, Annette Plachetka, Gregory Shanower, Michael Y. Tolstorukov, Lovelace J. Luquette, Ruibin Xi, Youngsook L. Jung, Richard W. Park, Eric P. Bishop, Theresa K. Canfield, Richard Sandstrom, Robert E. Thurman, David M. MacAlpine, John A. Stamatoyannopoulos, Manolis Kellis, Sarah C.R. Elgin, Mitzi I. Kuroda, Vincenzo Pirrotta, Gary H. Karpen, and Peter J. Park (Mar. 2011). "Comprehensive analysis of the chromatin landscape in Drosophila melanogaster". In: *Nature* 471 (7339), pp. 480–486. ISSN: 00280836. DOI: 10.1038/nature09725.

Khater, Ismail M., Ivan Robert Nabi, and Ghassan Hamarneh (June 2020). "A Review of Super-Resolution Single-Molecule Localization Microscopy Cluster Analysis and Quantification Methods". In: *Patterns* 1 (3). ISSN: 26663899. DOI: 10.1016/j.patter.2020.100038.

Knapek, Stephan, Lily Kahsai, Åsa M.E. Winther, Hiromu Tanimoto, and Dick R. Nässel (2013). "Short neuropeptide F acts as a functional neuromodulator for olfactory memory in kenyon cells of Drosophila mushroom bodies". In: *Annals of Internal Medicine* 158 (6), pp. 5340–5345. ISSN: 00034819. DOI: 10.1523/JNEUROSCI.2287-12.2013.

Kong, Suming, Daniel Bohl, Chunhua Li, and Dorothy Tuan (1997). "Transcription of the HS2 Enhancer toward a cis-Linked Gene Is Independent of the Orientation, Position, and Distance of the Enhancer Relative to the Gene". In: *MOLECULAR AND CELLULAR BIOLOGY* 17 (7), pp. 3955–3965.

Konstantinides, Nikolaos, Katarina Kapuralin, Chaimaa Fadil, Luendreo Barboza, Rahul Satija, and Claude Desplan (July 2018). "Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features". In: *Cell* 174 (3), 622–635.e13. ISSN: 10974172. DOI: 10.1016/j.cell.2018.05.021.

Krashes, Michael J., Alex C. Keene, Benjamin Leung, J. Douglas Armstrong, and Scott Waddell (Jan. 2007). "Sequential Use of Mushroom Body Neuron Subsets during Drosophila Odor Memory Processing". In: *Neuron* 53 (1), pp. 103–115. ISSN: 08966273. DOI: 10.1016/j.neuron.2006.11.021.

Laat, Wouter De and Denis Duboule (2013). *Topology of mammalian developmental enhancers and their regulatory landscapes*. DOI: 10.1038/nature12753.

Lasko, Paul (Nov. 2020). *Patterning the Drosophila embryo: A paradigm for RNA-based developmental genetic regulation*. DOI: 10.1002/wrna.1610.

Latchman, David S (1997). "Transcription Factors: An Overview". In: *Int. J. Biochem. Crll Biol* 29, pp. 1305–1312.

Lelek, Mickael, Melina T. Gyparaki, Gerti Beliu, Florian Schueder, Juliette Griffie, Suliana Manley, Ralf Jungmann, Markus Sauer, Melike Lakadamyali, and Christophe Zimmer (Dec. 2021). *Single-molecule localization microscopy*. DOI: 10.1038/s43586-021-00038-x.

Levine, Michael, Claudia Cattoglio, and Robert Tjian (2014). "Looping back to leap forward: Transcription enters a new era". In: *Cell* 157 (1), pp. 13–25. ISSN: 10974172. DOI: 10.1016/j.cell.2014.02.009. URL: http://dx.doi.org/10.1016/j.cell.2014.02.009.

Levo, Michal, João Raimundo, Xin Yang Bing, Zachary Sisco, Philippe J. Batut, Sergey Ryabichko, Thomas Gregor, and Michael S. Levine (May 2022). "Transcriptional coupling of distant regulatory genes in living embryos". In: *Nature* 605 (7911), pp. 754–760. ISSN: 14764687. DOI: 10.1038/s41586-022-04680-7.

Lewis, E B (1978). "A gene complex controlling segmentation in Drosophila". In: *Nature* 276, pp. 565–570.

Li, Feng, Jack Lindsey, Elizabeth C. Marin, Nils Otto, Marisa Dreher, Georgia Dempsey, Ildiko Stark, Alexander Shakeel Bates, Markus William Pleijzier, Philipp Schlegel, Aljoscha Nern, Shinya Takemura, Nils Eckstein, Tansy Yang, Audrey Francis, Amalia Braun, Ruchi Parekh, Marta Costa, Louis Scheffer, Yoshinori Aso, Gregory S.X.E. Jefferis, L. F. Abbott, Ashok Litwin-Kumar, Scott Waddell, and Gerald M. Rubin (2020). "The connectome of the adult drosophila mushroom body provides insights into function". In: *eLife* 9, pp. 1–217. ISSN: 2050084X. DOI: 10.7554/eLife.62576.

Li, Hongjie, Felix Horns, Qijing Xie, Qijing Xie, Tongchao Li, David J. Luginbuhl, Liqun Luo, and Stephen R. Quake (Nov. 2017). "Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing". In: *Cell* 171 (5), 1206.e22–1220.e22. ISSN: 10974172. DOI: 10.1016/j.cell.2017.10.019.

Lichter, P, T Cremer, J Borden, L Manuelidis, and D C Ward (1988). "Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries". In: *Hum Genetics* 80, pp. 224–234.

Lieberman-Aiden, Erez, Nynke L. Van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R. Lajoie, Peter J. Sabo, Michael O. Dorschner, Richard Sandstrom, Bradley Bernstein, M. A. Bender, Mark Groudine, Andreas Gnirke, John Stamatoyannopoulos, Leonid A. Mirny, Eric S. Lander, and Job Dekker (Oct. 2009). "Comprehensive mapping of long-range interactions reveals folding principles of the human genome". In: *Science* 326 (5950), pp. 289–293. ISSN: 00368075. DOI: 10.1126/science.1181369.

Luger, Karolin, Armin W Mader, Robin K Richmond, David F Sargent, and Timothy J Richmond (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution". In: *Nature* 389, pp. 251–260.

Margulies, Carla, Tim Tully, and Josh Dubnau (2005). "Deconstructing memory in Drosophila". In: *Current Biology* 15 (17), pp. 700–713. ISSN: 09609822. DOI: 10.1016/j.cub.2005.08.024.

Markaki, Yolanda, Daniel Smeets, Susanne Fiedler, Volker J. Schmid, Lothar Schermelleh, Thomas Cremer, and Marion Cremer (May 2012). "The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for studies of nuclear architecture". In: *BioEssays* 34 (5), pp. 412–426. ISSN: 02659247. DOI: 10.1002/bies.201100176.

Mateo, Leslie J., Sedona E. Murphy, Antonina Hafner, Isaac S. Cinquini, Carly A. Walker, and Alistair N. Boettiger (2019). "Visualizing DNA folding and RNA in embryos at single-cell resolution". In: *Nature* 568 (7750), pp. 49–54. ISSN: 14764687. DOI: 10.1038/s41586-019-1035-4. URL: http://dx.doi.org/10.1038/s41586-019-1035-4.

Matthews, Nicholas E. and Rob White (Sept. 2019). "Chromatin Architecture in the Fly: Living without CTCF/Cohesin Loop Extrusion?: Alternating Chromatin States Provide a Basis for Domain Architecture in Drosophila". In: *BioEssays* 41 (9). ISSN: 15211878. DOI: 10.1002/bies.201900048.

McBride, Sean M. J., Giuliani Giovanna, Choi Catherine, Krause Paul, Correale Dana, Watson Karli, Baker Glenn, and Siwicki Kathleen K. (1999). "Mushroom Body Ablation Impairs Short-Term Memory and Long-Term Memory of Courtship Conditioning in Drosophila melanogaster". In: *Neuron* 24, pp. 967–977.

Melnikov, Alexandre, Anand Murugan, Xiaolan Zhang, Tiberiu Tesileanu, Li Wang, Peter Rogov, Soheil Feizi, Andreas Gnirke, Curtis G. Callan, Justin B. Kinney, Manolis Kellis, Eric S. Lander, and Tarjei S. Mikkelsen (2012). "Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay". In: *Nature Biotechnology* 30 (3), pp. 271–277. ISSN: 15461696. DOI: 10.1038/nbt.2137.

Messina, Olivier, Flavien Raynal, Julian Gurgo, Jean-bernard Fiche, Vera Pancaldi, and Marcelo Nollmann (2022). "3D chromatin interactions involving Drosophila insulators are infrequent but preferential and arise before TADs and transcription". In: *Nature Communications*, p. 2022.12.12.520036 DOI: 10.1038/s41467-023-42485-y. URL: http://biorxiv.org/content/early/2022/12/14/2022.12.12.520036.abstract.

Millan-Zambrano, Gonzalo, Adam Burton, Andrew J. Bannister, and Robert Schneider (Sept. 2022). *Histone post-translational modifications — cause and consequence of genome function*. DOI: 10.1038/s41576-022-00468-7.

Mogno, Ilaria, Jamie C. Kwasnieski, and Barak A. Cohen (Nov. 2013). "Massively parallel synthetic promoter assays reveal the in vivo effects of binding site variants". In: *Genome Research* 23 (11), pp. 1908–1915. ISSN: 10889051. DOI: 10.1101/gr.157891.113.

Mohana, Giriram, Julien Dorier, Xiao Li, Aleksander Jankowski, and Michael S Levine (2023). "Chromosome-level organization of the regulatory genome in the Drosophila nervous system". In: pp. 1–19. DOI: 10.1016/j.cell.2023.07.008. URL: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE214704.

Mollá-Albaladejo, Rubén and Juan A. Sánchez-Alcañiz (Nov. 2021). *Behavior Individuality: A Focus on Drosophila melanogaster*. DOI: 10.3389/fphys.2021.719038.

Moon, Hanlim, Galina Filippova, Dmitry Loukinov, Elena Pugacheva, Qi Chen, Sheryl T. Smith, Adam Munhall, Britta Grewe, Marek Bartkuhn, Rüdiger Arnold, Les J. Burke, Renate Renkawitz-Pohl, Rolf Ohlsson, Jumin Zhou, Rainer Renkawitz, and Victor Lobanenkov (2005). "CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator". In: *EMBO Reports* 6 (2), pp. 165–170. ISSN: 1469221X. DOI: 10.1038/sj.embor.7400334.

Morcillo, Patrick, Christina Rosen, Mary K Baylies, and Dale Dorsett (1997). "Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in Drosophila". In: *GENES DEVELOPMENT* 11, pp. 2729–2740.

Moreau, P, R Hen, B Wasylyk, R Everett, M P Gaub, and P Chambon (1981). "The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants". In: *Nucleic Acids Research* 9 (22), pp. 6047–6068. URL: https://academic.oup.com/nar/article/9/22/6047/2380057.

Natoli, Gioacchino and Jean Christophe Andrau (2012). *Noncoding transcription at enhancers: General principles and functional models*. DOI: 10.1146/annurev-genet-110711-155459.

Negre, Nicolas, Christopher D. Brown, Lijia Ma, Christopher Aaron Bristow, Steven W. Miller, Ulrich Wagner, Pouya Kheradpour, Matthew L. Eaton, Paul Loriaux, Rachel Sealfon, Zirong Li,

Haruhiko Ishii, Rebecca F. Spokony, Jia Chen, and Kevin P. White (2011). "A Cis-Regulatory Map of the Drosophila Genome". In: *Nature* 23 (1), pp. 1–7. ISSN: 14726963. DOI: 10.1038/nature09990.A. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf%0Ahttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3835697/pdf/nihms521772.pdf.

Nolis, Ilias K, Daniel J Mckay, Eva Mantouvalou, Stavros Lomvardas, Menie Merika, and Dimitris Thanos (2009). "Transcription factors mediate long-range enhancer-promoter interactions". In: *PNAS* 106 (48), pp. 20222–20227. URL: www.pnas.org/cgi/content/full/.

Nollmann, Marcelo, Isma Bennabi, Markus Götz, and Thomas Gregor (2021). "The Impact of Space and Time on the Functional Output of the Genome". In: *Cold Spring Harbor Perspectives in Biology*, a040378. DOI: 10.1101/cshperspect.a040378.

Noonan, James P. and Andrew S. McCallion (2010). "Genomics of long-range regulatory elements". In: *Annual Review of Genomics and Human Genetics* 11, pp. 1–23. ISSN: 15278204. DOI: 10.1146/annurev-genom-082509-141651.

Nora, Elphege P., Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, Nynke L. Van Berkum, Johannes Meisig, John Sedat, Joost Gribnau, Emmanuel Barillot, Nils Blüthgen, Job Dekker, and Edith Heard (May 2012). "Spatial partitioning of the regulatory landscape of the X-inactivation centre". In: *Nature* 485 (7398), pp. 381–385. ISSN: 00280836. DOI: 10.1038/nature11049.

Nusslein-Volhard, Christlane, Hans Georg Frohnhofer, and Ruth Lehmann (1987). "Determination of Anteroposterior Polarity in Drosophila". In: *Science* 238, pp. 1675–1681.

O'connor, Leigh, Jane Gilmour, and Constanze Bonifer (2016). "The Role of the Ubiquitously Expressed Transcription Factor Sp1 in Tissue-specific Transcriptional Regulation and in Disease". In: *Yale Journal of Biology and Medecine* 89, pp. 513–525.

Ogienko, Anna A., Evgeniya S. Omelina, Oleg V. Bylino, Mikhail A. Batin, Pavel G. Georgiev, and Alexey V. Pindyurin (Oct. 2022). *Drosophila as a Model Organism to Study Basic Mechanisms of Longevity*. DOI: 10.3390/ijms231911244.

Ogiyama, Yuki, Bernd Schuettengruber, Giorgio L. Papadopoulos, Jia Ming Chang, and Giacomo Cavalli (2018). "Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during Drosophila Development". In: *Molecular Cell* 71 (1), 73–88.e5. ISSN: 10974164. DOI: 10.1016/j.molcel.2018.05.032. URL: https://doi.org/10.1016/j.molcel.2018.05.032.

Oudelaar, A. Marieke and Douglas R. Higgs (2021). "The relationship between genome structure and function". In: *Nature Reviews Genetics* 22 (3), pp. 154–168. ISSN: 14710064. DOI: 10.1038/s41576-020-00303-x.

Ozdemir, Isa and Maria Cristina Gambetta (2019). "The role of insulation in patterning gene expression". In: *Genes* 10 (10). ISSN: 20734425. DOI: 10.3390/genes10100767.

Pai, Chi Yun, Elissa P. Lei, Dolanchanpa Ghosh, and Victor G. Corces (Dec. 2004). "The centrosomal protein CP190 is a component of the gypsy chromatin insulator". In: *Molecular Cell* 16 (5), pp. 737–748. ISSN: 10972765. DOI: 10.1016/j.molcel.2004.11.004.

Panigrahi, Anil and Bert W. O'Malley (Dec. 2021). *Mechanisms of enhancer action: the known and the unknown*. DOI: 10.1186/s13059-021-02322-1.

Panigrahi, Anil K., Charles E. Foulds, Rainer B. Lanz, Ross A. Hamilton, Ping Yi, David M. Lonard, Ming Jer Tsai, Sophia Y. Tsai, and Bert W. O'Malley (May 2018). "SRC-3 Coactivator Governs Dynamic Estrogen-Induced Chromatin Looping Interactions during Transcription". In: *Molecular Cell* 70 (4), 679–694.e7. ISSN: 10974164. DOI: 10.1016/j.molcel.2018.04.014.

Pankratz, Michael J. and Herbet Jäckle (1990). "Making stripes in the Drosophila embryo". In: *Trends in Genetics* 6, pp. 287–292.

Pardue, Lou and Joseph G Gall (1969). "MOLECULAR HYBRIDIZATION OF RADIOACTIVE DNA TO THE DNA OF CYTOLOGICAL PREPARATIONS". In: *National Academy of Sciences* 64 (2), pp. 600–604. URL: https://www.pnas.org.

Parkhurst, Susan M. and Victor G. Corces (Jan. 1986). "Interactions among the Gypsy Transposable Element and the Yellow and the Suppressor of Hairy-Wing Loci in Drosophila melanogaster". In: *Molecular and Cellular Biology* 6 (1), pp. 47–53. DOI: 10.1128/mcb.6.1.47-53.1986.

Patwardhan, Rupali P., Joseph B. Hiatt, Daniela M. Witten, Mee J. Kim, Robin P. Smith, Dalit May, Choli Lee, Jennifer M. Andrie, Su In Lee, Gregory M. Cooper, Nadav Ahituv, Len A. Pennacchio, and Jay Shendure (Mar. 2012). "Massively parallel functional dissection of mammalian enhancers in vivo". In: *Nature Biotechnology* 30 (3), pp. 265–270. ISSN: 10870156. DOI: 10.1038/nbt.2136.

Pelechano, Vicent and Lars M. Steinmetz (Dec. 2013). "Gene regulation by antisense transcription". In: *Nature Reviews Genetics* 14 (12), pp. 880–893. ISSN: 14710056. DOI: 10.1038/nrg3594.

Pfeiffer, Barret D., Teri T.B. Ngo, Karen L. Hibbard, Christine Murphy, Arnim Jenett, James W. Truman, and Gerald M. Rubin (Oct. 2010). "Refinement of tools for targeted gene expression in Drosophila". In: *Genetics* 186 (2), pp. 735–755. ISSN: 00166731. DOI: 10.1534/genetics.110.119917.

Phillips-Cremins, Jennifer E., Michael E.G. Sauria, Amartya Sanyal, Tatiana I. Gerasimova, Bryan R. Lajoie, Joshua S.K. Bell, Chin Tong Ong, Tracy A. Hookway, Changying Guo, Yuhua Sun, Michael J. Bland, William Wagstaff, Stephen Dalton, Todd C. McDevitt, Ranjan Sen, Job Dekker, James Taylor, and Victor G. Corces (June 2013). "Architectural protein subclasses shape 3D organization of genomes during lineage commitment". In: *Cell* 153 (6), pp. 1281–1295. ISSN: 10974172. DOI: 10.1016/j.cell.2013.04.053.

Pickersgill, Helen, Bernike Kalverda, Elzo De Wit, Wendy Talhout, Maarten Fornerod, and Bas Van Steensel (Sept. 2006). "Characterization of the Drosophila melanogaster genome at the nuclear lamina". In: *Nature Genetics* 38 (9), pp. 1005–1014. ISSN: 10614036. DOI: 10.1038/ng1852.

Pinkel, D, J Landegentt, C Collins, J Fuscoet, R Segraves, J Lucas, and J Gray (1988). "Fluorescence in situ hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocations of chromosome 4". In: *Proc. Nati. Acad. Sci. USA* 85, pp. 9138–9142.

Pollex, Tim, Adam Rabinowitz, Maria Cristina Gambetta, Raquel Marco-ferreres, Rebecca R Viales, Aleksander Jankowski, Christoph Schaub, and Eileen E M Furlong (2024). "Enhancer – promoter interactions become more instructive in the transition from cell-fate specification to tissue differentiation". In: DOI: 10.1038/s41588-024-01678-x.

Quinn, William G, William A Harris, and Seymour Benzer (1974). "Conditioned Behavior in Drosophila melanogaster (learnin/memory/odor discrimination/color vision)". In: *Proc.Nat.Acad.Sci.USA* 71 (3), pp. 708–712. URL: https://www.pnas.org.

Raj, Arjun and Alexander van Oudenaarden (Oct. 2008). *Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences*. DOI: 10.1016/j.cell.2008.09.050.

Ramírez, Fidel, Vivek Bhardwaj, Laura Arrigoni, Kin Chung Lam, Björn A. Grüning, José Villaveces, Bianca Habermann, Asifa Akhtar, and Thomas Manke (Dec. 2018). "High-resolution TADs reveal DNA sequences underlying genome organization in flies". In: *Nature Communications* 9 (1). ISSN: 20411723. DOI: 10.1038/s41467-017-02525-w.

Rao, Suhas S.P., Su Chen Huang, Brian Glenn St Hilaire, Jesse M. Engreitz, Elizabeth M. Perez, Kyong Rim Kieffer-Kwon, Adrian L. Sanborn, Sarah E. Johnstone, Gavin D. Bascom, Ivan D. Bochkov, Xingfan Huang, Muhammad S. Shamim, Jaeweon Shin, Douglass Turner, Ziyi Ye, Arina D. Omer, James T. Robinson, Tamar Schlick, Bradley E. Bernstein, Rafael Casellas,

Eric S. Lander, and Erez Lieberman Aiden (Oct. 2017). "Cohesin Loss Eliminates All Loop Domains". In: *Cell* 171 (2), 305–320.e24. ISSN: 10974172. DOI: 10.1016/j.cell.2017.09.026.

Rao, Suhas S.P., Miriam H. Huntley, Neva C. Durand, Elena K. Stamenova, Ivan D. Bochkov, James T. Robinson, Adrian L. Sanborn, Ido Machol, Arina D. Omer, Eric S. Lander, and Erez Lieberman Aiden (Dec. 2014). "A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping". In: *Cell* 159 (7), pp. 1665–1680. ISSN: 10974172. DOI: 10.1016/j.cell.2014.11.021.

Remini, Loucif, Midas Segers, John Palmeri, Jean Charles Walter, Andrea Parmeggiani, and Enrico Carlon (2024). "Chromatin structure from high resolution microscopy: Scaling laws and microphase separation". In: *Physical Review E* 109 (2). ISSN: 24700053. DOI: 10.1103/PhysRevE.109.024408.

Reuveni, Eli, Dmitry Getselter, Oded Oron, and Evan Elliott (Dec. 2018). "Differential contribution of cis and trans gene transcription regulatory mechanisms in amygdala and prefrontal cortex and modulation by social stress". In: *Scientific Reports* 8 (1). ISSN: 20452322. DOI: 10.1038/s41598-018-24544-3.

Richmond, Timothy J and Curt A Davey (2003). "The structure of DNA in the nucleosome core". In: *Nature* 423, pp. 145–150. URL: www.nature.com/nature.

Roberts, S G E (2000). "Mechanisms of action of transcription activation and repression domains". In: *CMLS, Cell. Mol. Life Sci* 57, pp. 1149–1160.

Robson, Michael I., Alessa R. Ringel, and Stefan Mundlos (June 2019). *Regulatory Landscaping: How Enhancer-Promoter Communication Is Sculpted in 3D*. DOI: 10.1016/j.molcel.2019.05.032.

Rubin, Adam J., Brook C. Barajas, Mayra Furlan-Magaril, Vanessa Lopez-Pajares, Maxwell R. Mumbach, Imani Howard, Daniel S. Kim, Lisa D. Boxer, Jonathan Cairns, Mikhail Spivakov, Steven W. Wingett, Minyi Shi, Zhixin Zhao, William J. Greenleaf, Anshul Kundaje, Michael Snyder, Howard Y. Chang, Peter Fraser, and Paul A. Khavari (Oct. 2017). "Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation". In: *Nature Genetics* 49 (10), pp. 1522–1528. ISSN: 15461718. DOI: 10.1038/ng.3935.

Rudan, Matteo Vietri, Christopher Barrington, Stephen Henderson, Christina Ernst, Duncan T. Odom, Amos Tanay, and Suzana Hadjur (Mar. 2015). "Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture". In: *Cell Reports* 10 (8), pp. 1297–1309. ISSN: 22111247. DOI: 10.1016/j.celrep.2015.02.004.

RUDKIN, GEORGE T. and B. D. STOLLAR (1977). "High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence". In: *Nature* 265, pp. 472–474.

Rust, Michael J., Mark Bates, and Xiaowei Zhuang (Oct. 2006). "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". In: *Nature Methods* 3 (10), pp. 793–795. ISSN: 15487091. DOI: 10.1038/nmeth929.

Schauer, Tamás, Yad Ghavi-Helm, Tom Sexton, Christian Albig, Catherine Regnard, Giacomo Cavalli, Eileen EM Furlong, and Peter B Becker (Oct. 2017). "Chromosome topology guides the Drosophila Dosage Compensation Complex for target gene activation". In: *EMBO reports* 18 (10), pp. 1854–1868. ISSN: 1469-221X. DOI: 10.15252/embr.201744292.

Scheffer, Louis K. et al. (Sept. 2020). "A connectome and analysis of the adult drosophila central brain". In: *eLife* 9, pp. 1–74. ISSN: 2050084X. DOI: 10.7554/ELIFE.57443.

Schones, Dustin E., Kairong Cui, Suresh Cuddapah, Tae Young Roh, Artem Barski, Zhibin Wang, Gang Wei, and Keji Zhao (Mar. 2008). "Dynamic Regulation of Nucleosome Positioning in the Human Genome". In: *Cell* 132 (5), pp. 887–898. ISSN: 00928674. DOI: 10.1016/j.cell.2008.02.022.

Schuettengruber, Bernd and Giacomo Cavalli (Nov. 2009). *Recruitment of Polycomb group complexes and their role in the dynamic regulation of cell fate choice*. DOI: 10.1242/dev.033902.

Schwartz, Yuri B., Tatyana G. Kahn, Per Stenberg, Katsuhito Ohno, Richard Bourgon, and Vincenzo Pirrotta (Jan. 2010). "Alternative epigenetic chromatin states of polycomb target genes". In: *PLoS Genetics* 6 (1). ISSN: 15537390. DOI: 10.1371/journal.pgen.1000805.

Scott, Kristin C, Aaron D Taubman, and Pamela K Geyer (1999). "Enhancer Blocking by the Drosophila gypsy Insulator Depends Upon Insulator Anatomy and Enhancer Strength". In: *Genetics* 153, pp. 787–798. URL: https://academic.oup.com/genetics/article/153/2/787/6048000.

Sexton, Tom, Eitan Yaffe, Ephraim Kenigsberg, Frederic Bantignies, Benjamin Leblanc, Michael Hoichman, Hugues Parrinello, Amos Tanay, and Giacomo Cavalli (Feb. 2012). "Three-dimensional folding and functional organization principles of the Drosophila genome". In: *Cell* 148 (3), pp. 458–472. ISSN: 10974172. DOI: 10.1016/j.cell.2012.01.010.

Siegel, Richard W, Jeffrey C Halo, and P D Boyer (1979). "Conditioned responses in courtship behavior of normal and mutant Drosophila". In: *Genetics* 76 (7), pp. 3430–3434. URL: https://www.pnas.org.

Siyuan, Wang, Jun-Han Su, Brian J. Beliveau, Bintu Bogdan, Moffitt Jeffrey R., Wu Chao-ting, and Zhuang Xiaowei (Aug. 2016). "The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design". In: *Science* 353 (6299), pp. 594–598. ISSN: 10959203. DOI: 10.1126/science.aaf8993.

Smale, Stephen T. and James T. Kadonaga (2003). "The RNA polymerase II core promoter". In: *Annual Review of Biochemistry* 72, pp. 449–479. ISSN: 00664154. DOI: 10.1146/annurev.biochem.72.121801.161520.

Sokolowski, Marla B. (2001). "Drosophila: Genetics meets behaviour". In: *Nature Reviews Genetics* 2 (11), pp. 879–890.

Solovei, Irina, Antonio Cavallo, Lothar Schermelleh, Françoise Jaunin, Catia Scasselati, Dusan Cmarko, Christoph Cremer, Stanislav Fakan, and Thomas Cremer (2002). "Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH)". In: *Experimental Cell Research* 276 (1), pp. 10–23. ISSN: 00144827. DOI: 10.1006/excr.2002.5513.

Sparmann, Anke and Maarten Van Lohuizen (Nov. 2006). *Polycomb silencers control cell fate, development and cancer*. DOI: 10.1038/nrc1991.

Stadler, Michael R, Jenna E Haines, and Michael B Eisen (2017). "Convergence of topological domain boundaries, insulators, and polytene interbands revealed by high-resolution mapping of chromatin contacts in the early Drosophila melanogaster embryo". In: *eLIFE* 6 (e29550). DOI: 10.7554/eLife.29550.001. URL: https://doi.org/10.7554/eLife.29550.001.

Steensel, Bas van and Andrew S. Belmont (May 2017). *Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression*. DOI: 10.1016/j.cell.2017.04.022.

Steiner, Florian A., Paul B. Talbert, Sivakanthan Kasinathan, Roger B. Deal, and Steven Henikoff (Apr. 2012). "Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling". In: *Genome Research* 22 (4), pp. 766–777. ISSN: 10889051. DOI: 10.1101/gr.131748.111.

Strausfeld, Nicholas J, Lars Hansen, Yongsheng Li, Robert S Gomez, and Kei Ito (1998). *Evolution, Discovery, and Interpretations of Arthropod Mushroom Bodies*.

Struhl, Kevin and Zarmik Moqtaderi (1998). "The TAFs in the HAT". In: *Cell* 94, pp. 1–4.

Su, Jun Han, Pu Zheng, Seon S. Kinrot, Bogdan Bintu, and Xiaowei Zhuang (2020). "Genome-Scale Imaging of the 3D Organization and Transcriptional Activity of Chromatin". In: *Cell*

182 (6), 1641–1659.e26. ISSN: 10974172. DOI: 10.1016/j.cell.2020.07.032. URL: https://doi.org/10.1016/j.cell.2020.07.032.

Sun, Fang-Lin and Sarah C R Elgin (1999). "Putting Boundaries on Silence". In: *Cell* 99, pp. 459–462.

Symmons, Orsolya, Veli Vural Uslu, Taro Tsujimura, Sandra Ruf, Sonya Nassari, Wibke Schwarzer, Laurence Ettwiller, and François Spitz (2014). "Functional and topological characteristics of mammalian regulatory domains". In: *Genome Research* 24 (3), pp. 390–400. ISSN: 15495469. DOI: 10.1101/gr.163519.113.

Takahashi, Hiroshi, Kaoru Takahashi, and Fu-Chin Liu (2009). "FOXP Genes, Neural Development, Speech and Language Disorders". In: *Advances in Experimental Medicine and Biology* 665, pp. 117–118.

Takei, Yodai, Shiwei Zheng, Jina Yun, Sheel Shah, Nico Pierson, Jonathan White, Simone Schindler, Carsten H Tischbirek, Guo-Cheng Yuan, and Long Cai (2021). "Single-cell nuclear architecture across cell types in the mouse brain". In: *Science* 374, pp. 586–594. URL: https://www.science.org.

Talbert, Paul B. and Steven Henikoff (Feb. 2017). "Histone variants on the move: Substrates for chromatin dynamics". In: *Nature Reviews Molecular Cell Biology* 18 (2), pp. 115–126. ISSN: 14710080. DOI: 10.1038/nrm.2016.148.

Taskiran, Ibrahim I., Katina I. Spanier, Hannah Dickmänken, Niklas Kempynck, Alexandra Pančíková, Eren Can Ekşi, Gert Hulselmans, Joy N. Ismail, Koen Theunis, Roel Vandepoel, Valerie Christiaens, David Mauduit, and Stein Aerts (2024). "Cell-type-directed design of synthetic enhancers". In: *Nature* 626 (7997), pp. 212–220. ISSN: 14764687. DOI: 10.1038/s41586-023-06936-2.

Tempel, Bruce L, Nancy Boninit, Douglas R Dawsont, and William G Quinn (1983). "Reward learning in normal and mutant Drosophila". In: *Proc. Nati Acad. Sci. USA* 80, pp. 1482–1486. URL: https://www.pnas.org.

Tollervey, James R. and Victoria V. Lunyak (2012). *Judge, jury and executioner of stem cell fate*. DOI: 10.4161/epi.21141.

Tolwinski, Nicholas S. (Sept. 2017). "Introduction: Drosophila-A model system for developmental biology". In: *Journal of Developmental Biology* 5 (3). ISSN: 22213759. DOI: 10.3390/jdb5030009.

Tomchik, Seth M. and Ronald L. Davis (Nov. 2009). "Dynamics of Learning-Related cAMP Signaling and Stimulus Integration in the Drosophila Olfactory Pathway". In: *Neuron* 64 (4), pp. 510–521. ISSN: 08966273. DOI: 10.1016/j.neuron.2009.09.029.

Trannoy, Séverine, Christelle Redt-Clouet, Jean Maurice Dura, and Thomas Preat (Oct. 2011). "Parallel processing of appetitive short- and long-term memories in Drosophila". In: *Current Biology* 21 (19), pp. 1647–1653. ISSN: 09609822. DOI: 10.1016/j.cub.2011.08.032.

Tsanov, Nikolay, Aubin Samacoits, Racha Chouaib, Abdel Meneem Traboulsi, Thierry Gostan, Christian Weber, Christophe Zimmer, Kazem Zibara, Thomas Walter, Marion Peter, Edouard Bertrand, and Florian Mueller (Dec. 2016). "SmiFISH and FISH-quant - A flexible single RNA detection approach with super-resolution capability". In: *Nucleic Acids Research* 44 (22). ISSN: 13624962. DOI: 10.1093/nar/gkw784.

Udvardy, Andor, Eleanor Maine, and Paul Schedl (1985). "The 87A7 Chromomere Identification of Novel Chromatin Structures Flanking the Heat Shock Locus that may Define the Boundaries of Higher Order Domains". In: *J. Mol. Biol* 185, pp. 341–358.

Ulianov, Sergey V., Ekaterina E. Khrameeva, Alexey A. Gavrilov, Ilya M. Flyamer, Pavel Kos, Elena A. Mikhaleva, Aleksey A. Penin, Maria D. Logacheva, Maxim V. Imakaev, Alexander Chertovich,

Mikhail S. Gelfand, Yuri Y. Shevelyov, and Sergey V. Razin (Jan. 2016). "Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains". In: *Genome Research* 26 (1), pp. 70–84. ISSN: 15495469. DOI: 10.1101/gr.196006.115.

Wiegant, Joop, Thomas Ried, Petra M Nederlof, Mels Van Der Ploeg, Hans J Tanke, and Anton K Raap (1994). "In situ hybridization with fluoresceinated DNA". In: *Nucleic Acids Research* 19 (12), pp. 3237–3241.

Winding, Michael, Benjamin D. Pedigo, Christopher L. Barnes, Heather G. Patsolic, Youngser Park, Tom Kazimiers, Akira Fushiki, Ingrid V. Andrade, Avinash Khandelwal, Javier Valdes-Aleman, Feng Li, Nadine Randel, Elizabeth Barsotti, Ana Correia, Richard D. Fetter, Volker Hartenstein, Carey E. Priebe, Joshua T. Vogelstein, Albert Cardona, and Marta Zlatic (Mar. 2023). "The connectome of an insect brain". In: *Science* 379 (6636). ISSN: 10959203. DOI: 10.1126/science.add9330.

Winick-Ng, Warren, Alexander Kukalev, Izabela Harabula, Luna Zea-Redondo, Dominik Szabó, Mandy Meijer, Leonid Serebreni, Yingnan Zhang, Simona Bianco, Andrea M. Chiariello, Ibai Irastorza-Azcarate, Christoph J. Thieme, Thomas M. Sparks, Sílvia Carvalho, Luca Fiorillo, Francesco Musella, Ehsan Irani, Elena Torlai Triglia, Aleksandra A. Kolodziejczyk, Andreas Abentung, Galina Apostolova, Eleanor J. Paul, Vedran Franke, Rieke Kempfer, Altuna Akalin, Sarah A. Teichmann, Georg Dechant, Mark A. Ungless, Mario Nicodemi, Lonnie Welch, Gonçalo Castelo-Branco, and Ana Pombo (2021). "Cell-type specialization is encoded by specific chromatin topologies". In: *Nature* 599 (7886), pp. 684–691. ISSN: 14764687. DOI: 10.1038/s41586-021-04081-2.

Yamada, Tomoko, Yue Yang, Pamela Valnegri, Ivan Juric, Armen Abnousi, Kelly H. Markwalter, Arden N. Guthrie, Abigail Godec, Anna Oldenborg, Ming Hu, Timothy E. Holy, and Azad Bonni (2019). "Sensory experience remodels genome architecture in neural circuit to drive motor learning". In: *Nature* 569 (7758), pp. 708–713. ISSN: 14764687. DOI: 10.1038/s41586-019-1190-7. URL: http://dx.doi.org/10.1038/s41586-019-1190-7.

Yang, Jin H. and Anders S. Hansen (July 2024). "Enhancer selectivity in space and time: from enhancer–promoter interactions to promoter activation". In: *Nature Reviews Molecular Cell Biology*. ISSN: 14710080. DOI: 10.1038/s41580-024-00710-6.

Yang, Ming Yao, '* Llya Vilinsky S J Douglas Armstrong, Nicholas J Strausfeld, and Kim Kaiser (1995). *Subdivision of the Drosophila Mushroom Bodies by Enhancer-Trap Expression Patterns*, pp. 45–54.

Yu, Dinghui, David Benjamin G. Akalal, and Ronald L. Davis (Dec. 2006). "Drosophila α/β Mushroom Body Neurons Form a Branch-Specific, Long-Term Cellular Memory Trace after Spaced Olfactory Conditioning". In: *Neuron* 52 (5), pp. 845–855. ISSN: 08966273. DOI: 10.1016/j.neuron.2006.10.030.

Yu, Hung Hsiang, Chun Hong Chen, Lei Shi, Yaling Huang, and Tzumin Lee (July 2009). "Twin-spot MARCM to reveal the developmental origin and identity of neurons". In: *Nature Neuroscience* 12 (7), pp. 947–953. ISSN: 10976256. DOI: 10.1038/nn.2345.

Zars, T, M Fischer, R Schulz, and M Heisenberg (2000). "Localization of a Short-Term Memory in Drosophila". In: *Science* 288, pp. 672–675. URL: www.sciencemag.org.

Zhao, Keji, Craig M Hart, and Ulrich K Laemmli (1995). *Visualization of Chromosomal Domains with Boundary Element-Associated Factor BEAF-32*, pp. 879–889.

Zheng, Hui and Wei Xie (2019). "The role of 3D genome organization in development and cell differentiation". In: *Nature Reviews Molecular Cell Biology* 20 (9), pp. 535–550. ISSN: 14710080. DOI: 10.1038/s41580-019-0132-4. URL: http://dx.doi.org/10.1038/s41580-019-0132-4.

Zlatanova, Jordanka, Sanford H. Leuba, and Kensal Van Holde (1998). “Chromatin fiber structure: Morphology, molecular determinants, structural transitions”. In: *Biophysical Journal* 74 (5), pp. 2554–2566. ISSN: 00063495. DOI: 10.1016/S0006-3495(98)77963-9.