



HAL
open science

Mitotic bookmarking during early *Drosophila* embryogenesis

Maelle Bellec

► **To cite this version:**

Maelle Bellec. Mitotic bookmarking during early *Drosophila* embryogenesis. Agricultural sciences. Université Montpellier, 2021. English. NNT : 2021MONTT042 . tel-03470344

HAL Id: tel-03470344

<https://theses.hal.science/tel-03470344>

Submitted on 8 Dec 2021

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 : Génétique Moléculaire

École doctorale Sciences Chimiques et Biologiques pour la Santé (CBS2)

Unité de recherche UMR5535

Mitotic bookmarking during early *Drosophila* embryogenesis

Présentée par Maëlle BELLEC
Le 19 novembre 2021

Sous la direction de Mounia LAGHA

Devant le jury composé de

Dr. David SUTER, associate professor, Ecole Polytechnique Federale de Lausanne, Switzerland.

Dr. Timothy SAUNDERS, associate professor, University of Warwick, Coventry, United Kingdom.

Dr. Pavel TOMANCAK, Max Planck Institute for Molecular Cell Biology and Genetics, Dresden, Germany.

Dr. Severine CHAMBEYRON, Directrice de Recherche, Institut de Génétique Humaine, Montpellier, France.

Dr. Yohanns BELLAÏCHE, Directeur de Recherche, Institut Curie, Paris, France.

Dr. Edouard BERTRAND, Directeur de Recherche, Institut de Génétique Humaine, Montpellier, France.

Rapporteur

Rapporteur

Examineur

Examinatrice

Examineur

Président



UNIVERSITÉ
DE MONTPELLIER

Résumé

Durant le développement de l'embryon, l'expression des gènes dans chaque cellule doit être contrôlée très précisément dans l'espace et dans le temps afin qu'elles puissent adopter leur destin. Cette régulation spatio-temporelle peut être réalisée à deux étapes clés du dogme central de la biologie moléculaire : la transcription et la traduction.

Au cours de ma thèse, mon projet principal visait à comprendre comment une cellule mère transmet son identité à ses cellules filles. Un mécanisme potentiel de support de cet héritage pourrait être la mémoire transcriptionnelle mitotique, qui permet aux cellules filles d'hériter du statut transcriptionnel de leurs mères à chaque division. Ce processus est complexe car l'information doit pouvoir persister pendant la mitose où la plupart des régulateurs transcriptionnels se séparent de leurs gènes cibles. Cependant, à la manière de marques pages, certains facteurs ont la capacité de rester associés aux chromosomes mitotiques, représentant des supports potentiels de cette mémoire. Au cours de ma thèse, nous avons voulu déterminer quels facteurs sont ces « marque pages » mitotiques et si ceux-ci pourraient favoriser la mémoire transcriptionnelle, en utilisant l'embryon précoce de drosophile comme système modèle.

Tout d'abord, nous avons identifié plusieurs facteurs de transcription, tels que le facteur associé GAGA (GAF) et dBrd4, restant sur les chromosomes mitotiques au cours de l'embryogenèse précoce. Nous avons ensuite développé un protocole d'immunoprécipitation de la chromatine mitotique suivie d'un séquençage, afin d'identifier les gènes cibles de ces facteurs au cours de la mitose à l'échelle du génome entier. Nous avons identifié des milliers de régions retenues par ce facteur en mitose, correspondant principalement à des séquences cis-régulatrices de gènes clés du développement. Ensuite, pour étudier la cinétique d'activation de ces gènes, nous nous sommes concentrés sur quelques gènes de développement, pour lesquels nous avons généré des allèles CRISPR marqués avec des boucles MS2. En utilisant l'imagerie par microscopie confocale d'embryons précoces de drosophile, nous sommes en mesure de surveiller l'état transcriptionnel des cellules et de leur descendance. Comme résultat, nous avons observé que les cellules filles issues de mères transcriptionnellement actives dans le cycle cellulaire précédent s'activent plus rapidement que les cellules filles issues de mères transcriptionnellement inactives. Ceci démontre pour la première fois sur un gène endogène du développement l'existence d'une mémoire mitotique

transcriptionnelle. De plus, nous avons montré que cette mémoire mitotique est dépendante de GAF.

Parallèlement à ce projet principal, j'étais également impliqué dans un projet qui concernait le devenir des ARNm une fois transcrits. En effet, s'il est important d'étudier la production d'ARN, il est tout aussi important de regarder où et quand la traduction d'ARN a lieu au sein des cellules de l'embryon. Mais jusqu'à présent, aucune technologie n'a été capable de mesurer la vitesse et l'efficacité de cette traduction en temps réel dans un embryon. C'est pourquoi nous avons adapté la technique SunTag pour marquer la traduction de l'ARNm *twist*, un gène essentiel pour la spécification du devenir mésodermique, dans des embryons vivants. Cette technologie nous a permis de déterminer une cinétique de traduction rapide de Twist. Nous avons également mis en évidence une traduction périnucléaire localisée permettant une importation nucléaire plus efficace. Jusqu'à présent, aucune technique ne permettait de visualiser la traduction en temps réel dans un organisme multicellulaire, et sa combinaison avec l'imagerie à haute résolution fournit désormais des informations sur la dynamique de production des protéines.

Abstract

During development, gene expression must be precisely controlled in space and time. Precision ensures the correct establishment of a variety of cell fates.

During my thesis, my main project aimed at understanding how a mother cell transmits its identity to its daughter cells. A potential mechanism supporting this inheritance could be mitotic transcriptional memory, which allows daughter cells to inherit the transcriptional status of their mothers at each division. This process is complex because information must be able to persist during mitosis, where most transcriptional regulators are evicted from their target genes. However, like bookmarks, certain factors have the capacity to remain associated with mitotic chromosomes, representing potential supports of this memory. During my thesis, we wanted to characterize new mitotic bookmarkers and determine whether bookmarking could foster transcriptional memory, using the early *Drosophila* embryo as a model system. Specifically, I focused on two pivotal and conserved regulators of transcriptional activation, GAGA associated factor (GAF) and dBrd4. After establishing that these remain associated to mitotic chromosomes during early embryogenesis, we aimed at uncovering their targets during interphase and mitosis genome-wide. For this, I developed a protocol for immunoprecipitation of mitotic chromatin followed by sequencing. We identified thousands of mitotically retained regions, bound by GAF both in interphase and mitosis, mainly corresponding to *cis*-regulatory sequences of key developmental genes. Next, to investigate the kinetics of activation of GAF-bookmarked genes, we focused on few developmental genes, for which we generated MS2 tagged CRISPR alleles. Using confocal microscopy imaging of *Drosophila* early embryos, we are able to monitor the transcriptional status of cells and their progeny. As a result, we observed that daughter cells from transcriptionally active mothers in the previous cell cycle activate faster than daughter cells from transcriptionally inactive mothers. This demonstrates for the first time on a developmental endogenous gene the existence of a transcriptional mitotic memory. In addition, we have shown that this mitotic memory is GAF-dependent.

Spatio-temporal regulation of gene expression can be achieved at two key stages of the central dogma of molecular biology: transcription and translation.

Parallel to this main project on transcriptional control, during my thesis, I also contributed to a project related to translational control. Indeed, while it is important to study RNA

production, it is just as important to look at where and when their translation takes place within the cells of the embryo. But until now, no technology has been able to measure the speed and efficiency of this translation in real time in an embryo. This is why we adapted the SunTag technique to label translation of *twist* mRNA, an essential gene for specification of the mesodermal fate, in living embryos. This technology allowed us to determine, where, when and by which kinetics *twist* mRNAs were translated. We also revealed a localized perinuclear translation allowing more efficient nuclear import.

Acknowledgements

I thank Dr. David Suter, Dr. Timothy Saunders, Dr. Pavel Tomancak, Dr. Geneviève Almouzni, Dr. Yohanns Bellaïche and Dr. Edouard Bertrand for accepting to be members of my PhD thesis committee, and for their time reading and evaluating the manuscript. I feel very honored to have you as my committee members.

I thank Dr. Giacomo Cavalli, Dr. Jonhathan Chubb and Dr. Jerome Dejardin for being part of my annual PhD committee. Thank you for all your time and guidance about my project and my future career.

I want to thank my PhD advisor Mounia Lagha for giving me the chance to work with you on these amazing projects. You were always supportive and gave me confidence over the years and I will never forget that. You gave me the perfect supervision-freedom balance I needed. I thank you for your time spent talking science, career and mentoring me and giving me such precious advices. Merci merci merci!

I thank all the members and past members of the Lagha Lab: Matthieu, Jérémy, Virginia, Antonello, Hélène, Marion, Louise, Morgane, Olivier, and Maelys. I would like to specially thank interns who worked with me: Marie Lamarque, Heloïse Faure-Gautron, Basile Poutier (La Baz) and Vihn-Hoa Le. It was such a pleasure to teach you what I know and I thank you for your great work, I hope I gave you the passion for *Drosophila* genetics ;).

I particularly thank Jérémy (Doctor Jeremy Dufourt ;) for being so supportive since the beginning of my career, pushing me to be ambitious and believing in me. Thanks for all our scientific and intellectual discussions (mostly with coffees and cigarettes ;) and for teaching me the good scientific methodology, knowing that I could always count on you was so valuable. You will always be a person I admire and I will miss you for sure!

I thank Cyril Favard for all what you taught me about physics and all our discussions. Thanks Antonello Trullo for your friendship, your cool attitude and your support. Thanks my direct labmate ;) Virginia Pimmett for your help and friendship. Thank you Matthieu Dejean for giving me self-confidence and for all your jokes, your enthusiasm and patience! You four have been ones of the most remarkable persons I met during my PhD journey.

I thank all the MRI microscopy platform for all your help, especially Virginie, Leslie, Baptiste and Sylvain. I thank the members of the J.C. Andrau lab (especially Cyril Esnault, Jean-

Christophe Andrau and Talha Magat), Dr. Francois Juge for all your advices in genetics and during our labmeetings and George Hunt for all our stimulating scientific discussions.

I warmly thank Ovidiu Radulescu, Juanma Vacquerizas, Liz Ing-Simmons, Mattias Mannervik, Marcelo Nollman and Amal Makrini for our great collaborations.

I want to thank my friends for being so patient with me arriving late at dinners or guitar sessions after my experiments: thanks Anaïs, Antoine, Zoé, Manon, Maréva, Gaël, Jean-Marc, Luc-Marie, Manu, Boris, Dadao, Pascal ...

Finally, I would like thank my family, my two sisters and Maman et Papa for always being encouraging and putting your trust in me. Special thanks to Léa and Tom being always here for me, I am so proud of you guys!

"Emancipate yourselves from mental slavery. None but ourselves can free our minds."

Bob Marley (Big up for Anto and Jerem ;-))

A mes grands-mères Louissette et Jeanine

"Le problème avec ce monde est que les personnes intelligentes sont pleines de doutes tandis que les personnes stupides sont pleines de confiance."

"La chose importante est la chose évidente que personne ne dit."

Charles Bukowski

Long résumé en Français de la thèse

Ce résumé est constitué de deux chapitres. Le premier constitue le projet principal réalisé lors de ma thèse, étant volontairement plus développé, et le deuxième un projet secondaire auquel j'ai participé.

Chapitre 1 :

Introduction

Durant le développement embryonnaire, l'expression des gènes de chaque cellule doit être contrôlée de manière très précise dans l'espace et dans le temps afin que celles-ci puissent adopter leur destin. Pourtant, lors du développement, les cellules doivent se diviser des milliers de fois pour constituer un organisme sain entier. Mais alors comment une cellule mère transmet son identité à ses cellules filles ? Un des mécanismes potentiels est la mémoire transcriptionnelle mitotique, où les noyaux filles hériteraient du statut transcriptionnel de leurs mères. Certains facteurs ont la capacité de rester associé aux chromosomes mitotiques, représentant de potentiels supports de la mémoire, tels des 'marques pages'. Plusieurs facteurs ont été identifiés comme 'marque pages' surtout en culture cellulaire mESc (cellules souches embryonnaires de souris), mais très peu reste connu dans l'embryon entier.

Ces 'marques pages' peuvent être de nature différente : des facteurs de transcription, modification de queue d'histone (marques d'histone) ou des régulateurs de la chromatine permettant de reconnaître et/ou modifier ces marques d'histone. De manière non exclusive, certains loci restent 'ouvert' en mitose (Figure a). Ceci permettrait une meilleure accessibilité des facteurs de transcription après la mitose afin d'activer plus rapidement la transcription.

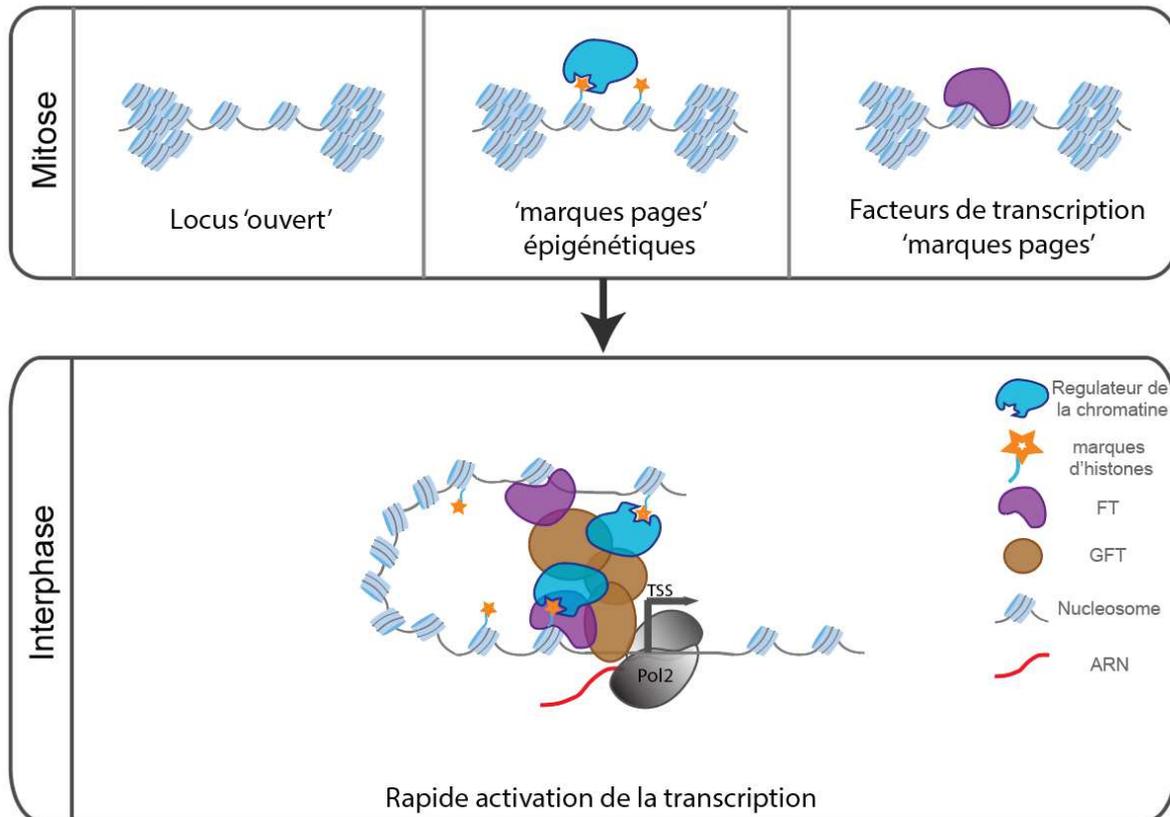


Figure a : Différents mécanismes de transmission d'information durant la mitose.

FT : Facteur de Transcription. GFT : Facteurs général de la transcription.

Chez presque tous les animaux métazoaires, la fertilisation engendre la fusion des génomes maternels et paternels et l'activation du génome zygotique débute durant les premières heures du développement et est appelée l'activation du génome zygotique (AGZ). Jusqu'à cette activation, un stock d'ARN et de protéines provenant du gamète maternel va régir l'embryogenèse précoce. L'AGZ consiste au démarrage de la transcription des gènes du génome zygotique et se doit d'être reproductible entre embryon afin d'assurer le bon développement des organismes. Durant cette première période cruciale du développement, les cellules totipotentes de l'embryon vont se différencier afin de former les trois feuillettes embryonnaires fondamentaux : ectoderme, endoderme, et mésoderme.

Chez la drosophile, l'embryogenèse commence par 14 cycles nucléaires au sein d'un syncytium. Environ deux heures après fécondation, le génome zygotique s'active et la cellularisation commence. Ce 'réveil' du génome est orchestré par des facteurs de transcription clé déposés maternellement. C'est le cas de la protéine Zelda mais aussi du

GAGA associated factor (GAF). Zelda a été montré comme absent des chromosomes lors de la mitose, alors que GAF semble rester sur la chromatine mitotique.

Afin de suivre la mémoire transcriptionnelle *in vivo*, les transcrits doivent être visualisés dans un organisme à travers plusieurs cycles cellulaires sur une échelle de temps raisonnable. De plus, un biais d'activation transcriptionnel est nécessaire pour distinguer une population de cellule « mère » active de noyaux d'une population inactive. L'embryon précoce de drosophile représente un modèle de choix pour étudier la mémoire (développement syncytial rapide avec divisions synchrones et visualisation directe de l'ARNm avec le système MS2/MCP). Grâce à cette technique, l'équipe dans laquelle j'ai effectué ma thèse a récemment pu visualiser une mémoire transcriptionnelle pour la première fois dans un organisme multicellulaire. En effet, lorsqu'un noyau mère est actif au cycle 13, ses descendants ont une probabilité quatre fois plus élevée d'activer la transcription dans le cycle suivant. Avec cet outil en main, nous avons cherché à identifier les mécanismes mis en jeu dans cette mémoire mitotique.

Pour cela, mon but était de :

- i) Identifier des facteurs 'marque pages' chez l'embryon de drosophile, déterminer quels sont leurs gènes cibles.
- ii) Mieux comprendre le rôle de leur rétention en mitose sur la dynamique transcriptionnelle lors de l'activation du génome zygotique.

Approches méthodologiques

Afin de déterminer si certains facteurs étaient présents pendant la mitose chez l'embryon précoce de drosophile, nous avons réalisé des marquages immunologiques contre la protéine GAF. Aussi, afin de visualiser GAF en temps réel, nous avons utilisé une lignée de drosophile GFP-GAF publiée par le laboratoire de Melissa Harrison. Afin de déterminer la dynamique de binding de GAF nous avons réalisé des expériences de Redistribution de Fluorescence après Photoblanchiment (FRAP) et Spectroscopie de Corrélation de Fluorescence (FCS).

Pour identifier les cibles de GAF pendant la mitose et l'interphase, nous avons développé un protocole de triage d'embryons mitotiques et d'embryons en interphase afin

de réaliser une immuno-précipitation de la chromatine (ChIP) associée à GAF suivie de séquençage haut débit (Figure b).

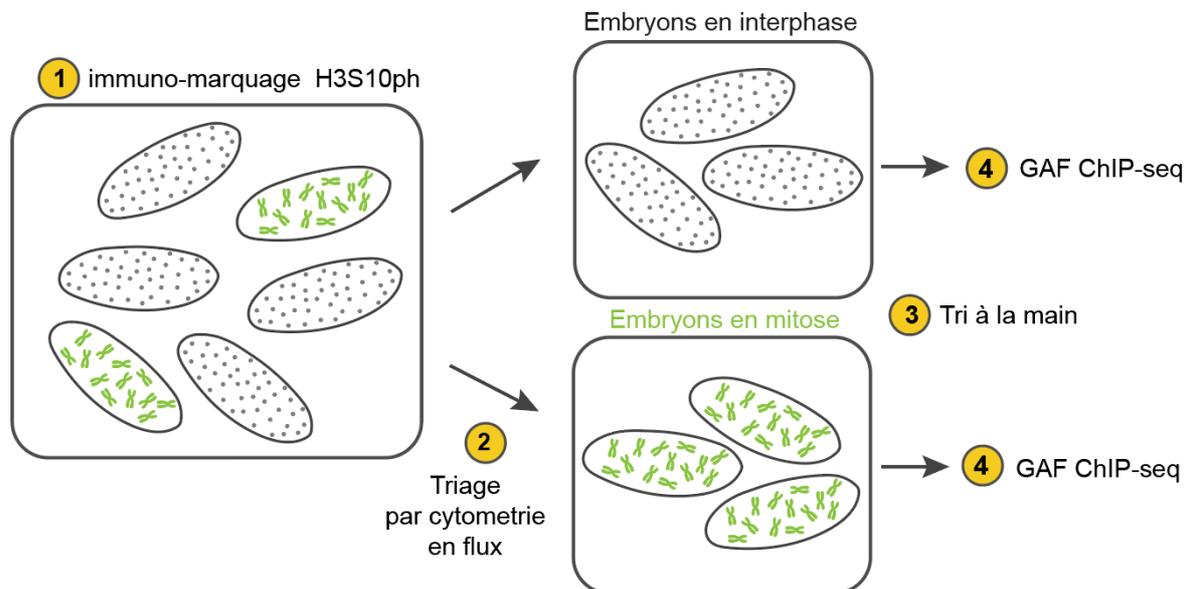


Figure b : Schéma d'expérience de ChIP mitotique.

GAF ChIP-seq: immuno-précipitation de la chromatine contre la protéine GAF suivie de séquençage haut débit.

Pour mesurer la dynamique de la transcription, j'ai sélectionné certains gènes développementaux ciblés par GAF pendant la mitose et à l'aide la technique CRISPR/Cas9, inséré des boucles MS2. Une fois transcrites, ces boucles vont être reconnues par la MS2-Coat Protein (MCP) fusionnée à une protéine fluorescente (ici Green Fluorescent Protein, GFP) ce qui créera un point vert fort au-dessus du bruit de fond. Combiné à de la microscopie à haute résolution, nous avons pu quantifier la production d'ARN en temps réel et ce, à travers plusieurs générations cellulaires.

En collaboration avec le mathématicien Dr. Ovidiu Radulescu, nous avons pu modéliser la mémoire mitotique sur le gène ciblé par GAF en mitose.

Résultats

Après avoir validé la présence de GAF sur les chromosomes mitotiques dans l'embryon précoce (Figure c), nous avons étudié sa dynamique de liaison à l'ADN. Nous avons révélé un temps résidence particulièrement de long, de l'ordre de la minute (58s). Ceci suggère que

GAF se lie de manière stable à la chromatine pouvant potentiellement expliquer sa rétention en mitose.

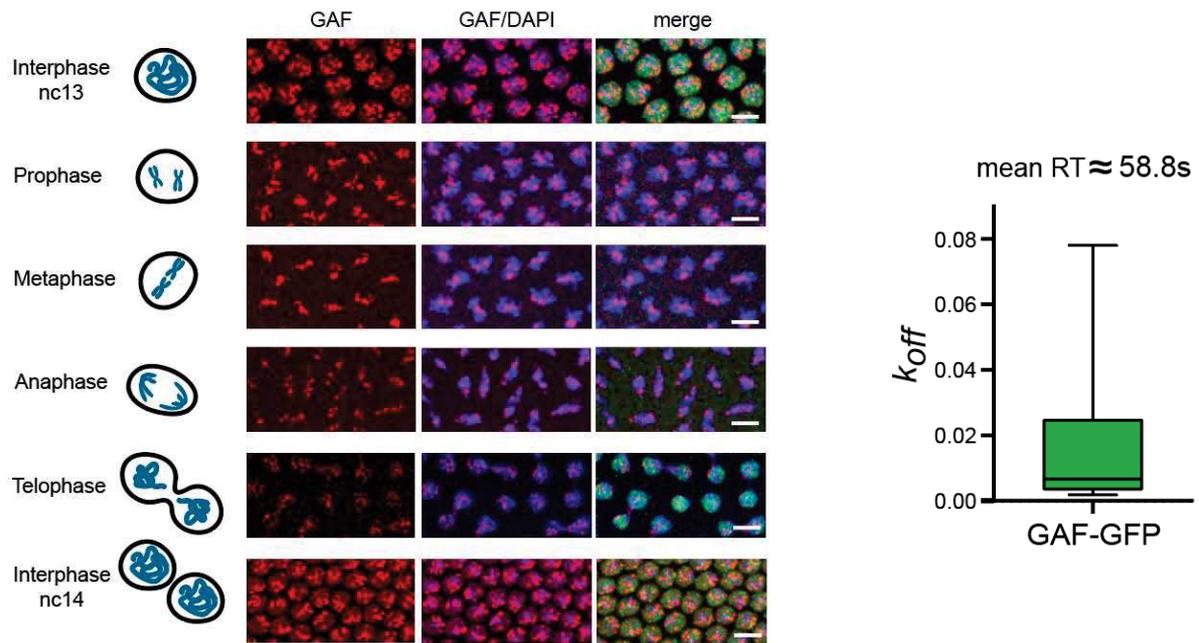


Figure c : Rétention de la protéine GAF en mitose.

(Gauche) Immuno-marquage d'embryon de type sauvage à différents stades de mitose avec anti-GAF (rouge) et DAPI (bleu). Echelle, $5\mu\text{m}$. (Droite) Boite à moustaches des k_{off} estimé par FRAP de la protéine GFP-GAF. RT : temps de résidence.

Afin de déterminer les cibles de GAF en mitose, nous avons réalisé un ChIP d'embryons en mitose et en interphase. Ceci a révélé plusieurs milliers de cibles liées par GAF en mitose, correspondant à des régions *cis*-régulatrices (promoteurs et enhanceurs) de gènes du développement (Figure d).

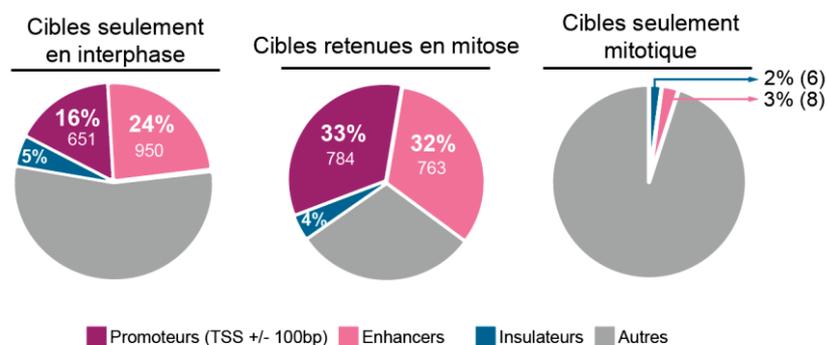


Figure d : Les cibles de GAF en mitose sont majoritairement des régions *cis*-régulatrices.

Diagrammes représentant la proportion de promoteurs, enhancers et insulateurs dans les trois catégories de cibles identifiées : seulement en interphase, retenues en mitose et seulement mitotiques.

A l'aide la technique MS2/MCP, nous avons pu suivre la transcription du gène *scylla*, retenu en mitose par GAF. La quantification du 'timing' d'activation après la mitose a permis de découvrir un biais mémoire entre les noyaux issus de noyaux actifs et les noyaux issus de noyaux inactifs. En effet, les noyaux issus de mères actives s'activent plus rapidement après la mitose que les noyaux issus de mères inactives (Figure e).

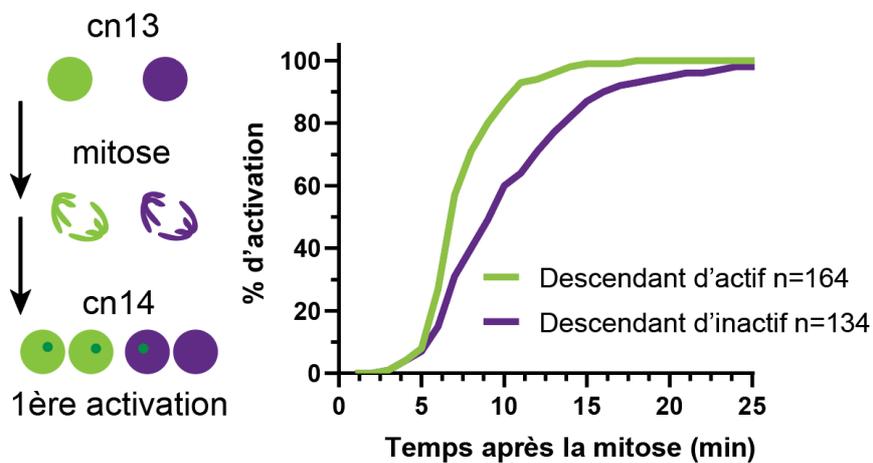


Figure e : Mémoire transcriptionnelle du gène *scylla*.

(Gauche) Schéma représentant les deux populations de noyaux : issus de noyaux actifs (vert) et issus de noyaux inactif (violet). cn : cycle nucléaire. (Droite) Courbe d'activation après la mitose des deux populations de noyaux. n : nombre de noyaux analysés, min : minutes.

Grace à l'utilisation d'une souche de drosophile exprimant l'ARN interférence contre l'ARN de GAF, nous avons pu réduire la protéine GAF dans les embryons, et mesurer le biais de mémoire transcriptionnelle du gène *scylla*. Lorsque les embryons ont une quantité réduite de GAF, le biais de mémoire est diminué. Ceci suggère que GAF jouera un rôle dans la transmission d'information entre cellule mère et cellule fille. Pour valider cet effet, nous avons modéliser les paramètres mis en jeu lors de la mémoire mitotique. Ceci a indiqué que

GAF jouera un rôle d'accélérateur d'activation de la transcription après la mitose et ce seulement sur les noyaux provenant de mères actives (Figure f).

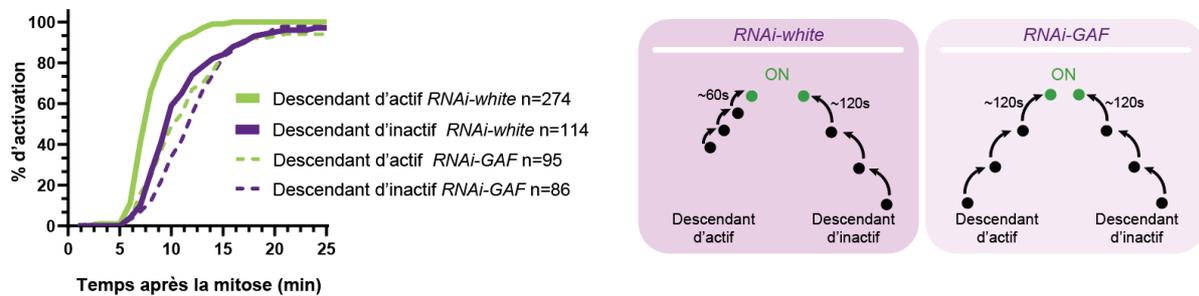


Figure f : La mémoire transcriptionnelle du gène *scylla* dépend de la présence de GAF.

(Gauche) Courbe d'activation après la mitose des deux populations de noyaux dans des embryons de type sauvage (*RNAi-white*) ou déplété en GAF (*RNAi-GAF*). *n* : nombre de noyaux analysés. *min* : minutes. (Droite) Schéma du modèle de mémoire mitotique amenée par GAF.

Conclusion

Durant ma thèse, nous avons cherché à déterminer comment la régulation des gènes par un facteur de transcription pourrait se propager durant la mitose dans un embryon en développement. En utilisant une combinaison d'imagerie quantitative en temps réel et de génomique, nous fournissons la preuve que le facteur GAF agit comme un 'marque-page' mitotique stable lors de l'activation du génome zygotique dans les embryons de drosophile. A ma connaissance la plupart des facteurs de transcription dans l'embryon, et dont leur temps de résidence sur l'ADN a été mesuré, se lie de manière transitoire et rapide (de l'ordre de la seconde). Dans le cas de GAF, cette liaison est longue rendant ce facteur assez unique.

Pour la première fois, nous avons révélé l'existence d'un biais d'activation de transcription en fonction du statut transcriptionnel de la cellule mère sur un gène endogène. La mémoire mitotique est essentielle lors de la spécification des tissus normaux dans un embryon en développement et des dérégulations pourraient conduire à de nombreuses pathologies telles que des anomalies congénitales et des cancers. Ainsi, il est crucial d'identifier de tels marque-pages ainsi que leur gènes cibles pour potentiellement contrôler cette mémoire. Il serait maintenant intéressant de regarder le rôle exact de cette mémoire dans le cas du gène *scylla*

dans l'embryon mais qui restent difficiles à analyser avec les techniques disponibles pour désactiver GAF seulement pendant la mitose.

Chapitre 2 :

Introduction

Le système MS2/MCP a été une véritable avancée pour l'approfondissement de nos connaissances concernant la dynamique de transcription. Pourtant, nous savons que le dogme central de la biologie moléculaire consiste en deux étapes clés : la transcription et la traduction. Cependant, très peu reste connu sur la dynamique de la traduction, particulièrement lors du développement embryonnaire. Cela est d'autant plus important de l'étudier lorsque l'on sait que le niveau d'un ARNm produit et la quantité de protéine qu'il code ne sont pas directement corrélés. Ceci peut être en partie expliqué par la localisation de certains ARNm dans des compartiments subcellulaires spéciaux, favorables ou non à une meilleure traduction.

De récentes études ont développé des techniques d'imagerie en temps réel de la production de protéine notamment avec le système SunTag (Figure g) dans des cellules en cultures. C'est pourquoi nous avons déployé cette technique pour la première fois dans un embryon vivant.

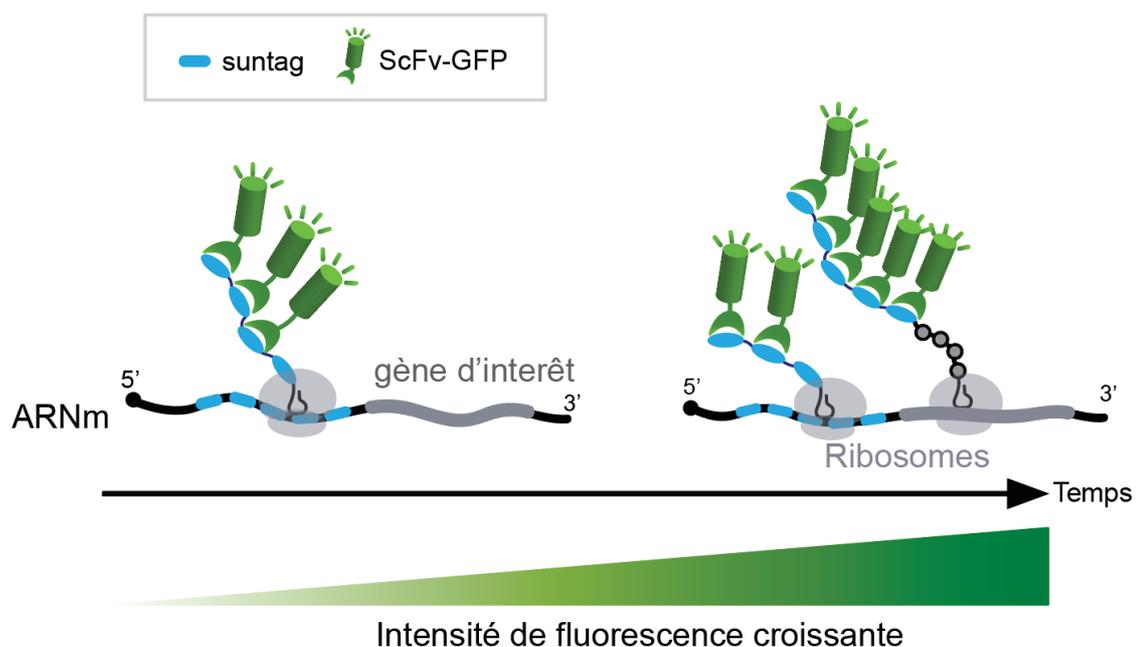


Figure g : Principe de la technique SunTag.

Quand l'ARN messenger (ARNm) est traduit, les répétitions suntag vont être reconnues par des anticorps à chaîne unique fusionnés avec une protéine fluorescente (scFv-GFP), ici la Green Fluorescent Protein (GFP). Ceci va créer un point de fluorescence augmentant au fur et à mesure que la protéine est produite.

Dans ce projet, initié par le Dr. Jérémie Dufourt dans le laboratoire de Mounia Lagha, nous avons voulu implémenter la technique de visualisation de la traduction en temps réel dans l'embryon précoce de drosophile, afin de pouvoir répondre à des questions fondamentales en biologie : toutes les molécules d'ARN sont-elles engagées dans la traduction ? Existe-t-il une régulation spatiale et/ou temporelle ? Quelles sont les principales caractéristiques quantitatives que nous pouvons extraire de la traduction des gènes du développement ?

Approches méthodologiques

Nous avons introduit des séquences suntag dans un des gènes majeurs pour le déroulement du premier événement morphologique, la gastrulation, qui est le gène *twist*. Nous avons inséré en position 3' de ce gène, 32 répétitions de suntag de manière endogène à l'aide de la technique CRISPR/Cas9. La création de lignée de drosophile transgénique exprimant l'anticorps scFv-GFP nous permet de visualiser la production naissante de la protéine (Figure g).

L'imagerie à haute résolution en microscopie confocale d'embryons fixés et hybridé avec des sondes contre les ARN suntag, nous a permis de quantifier les molécules uniques d'ARNm en traduction ou non. L'imagerie en temps réel nous a permis de suivre les polysomes en mouvement dans plusieurs compartiments subcellulaires afin de mesurer leur vitesse de déplacement par suivi de particule unique (SPT). Il a aussi été possible de mesurer la vitesse de traduction par corrélation de la fluctuation d'intensité dans le temps de point unique de scFv-GFP.

Résultats

Nous avons tout d'abord pu observer la traduction naissante de Twist sous forme de points verts fluorescents dans un embryon vivant (Figure h). Nous avons validé que ces points correspondent à la traduction par injection d'inhibiteur de traduction.

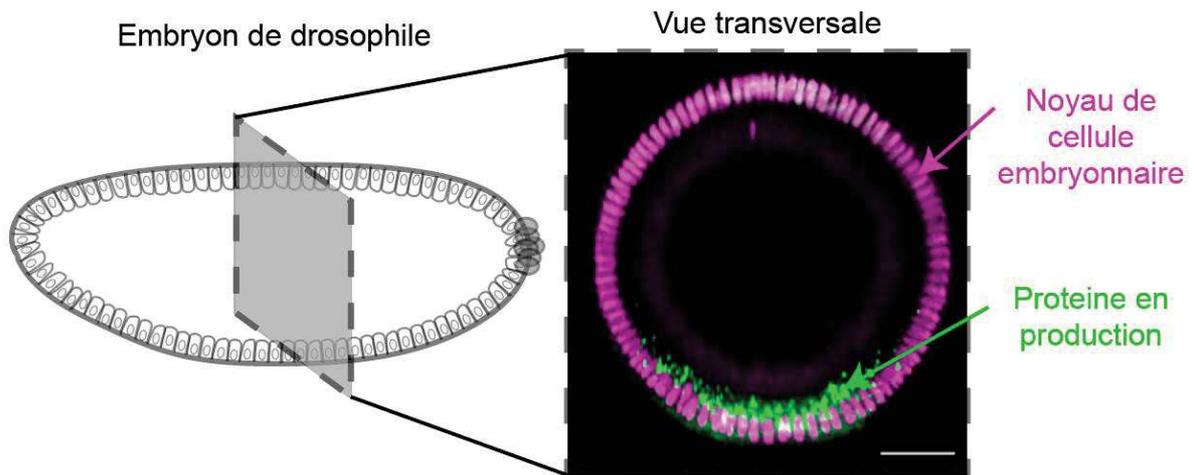


Figure h : Traduction de Twist dans l'embryon de drosophile.

(Gauche) Schéma d'un embryon de drosophile. (Droite) Vue transversal issue d'un film d'embryon *twist_suntag_CRISPR* et comportant la *scFv-GFP*.

Ensuite nous avons pu mesurer la vitesse à laquelle Twist est traduite révélant une traduction rapide de l'ordre de 35 acides aminés par seconde. Avec l'observation que Twist est traduit localement dans la partie basale de la cellule, nous proposons que cette traduction rapide et localisée permet une meilleure réimportation de la protéine dans le noyau afin qu'elle exerce sa fonction de facteur de transcription (Figure i). En effet Twist est une protéine majeure pour l'activation de gènes qui vont permettre le bon déroulement de la gastrulation.

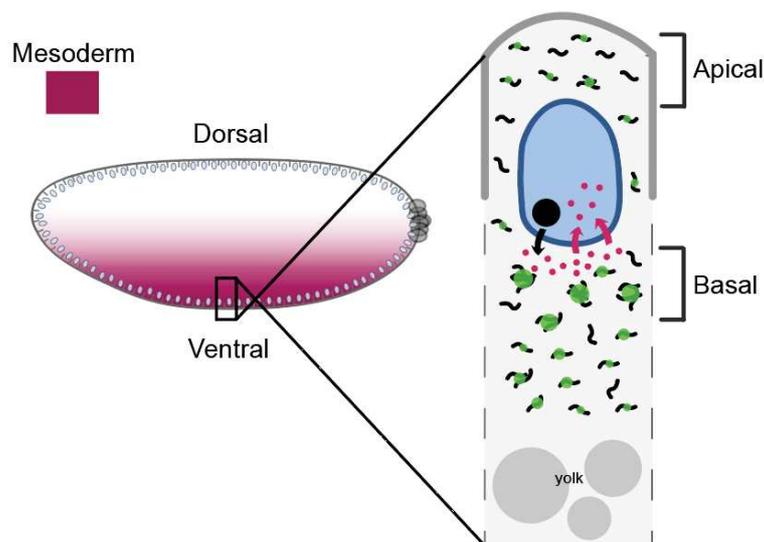


Figure i : Modèle de travail sur la traduction de la protéine Twist dans l'embryon de drosophile.

Lorsque le gène twist est transcrit, il est exporté dans le cytoplasme avec une accumulation dans la partie basale de la cellule. Il est ensuite traduit de manière efficace et rapide localement pour que la protéine puisse être importée dans le noyau et effectuer ses fonctions.

Conclusion

Grâce à l'implémentation de la technique SunTag dans l'embryon et sa combinaison à de l'imagerie à haute résolution il est maintenant possible d'étudier la dynamique de production des protéines dans un embryon vivant. Ceci a révélé la mise en place d'une régulation de la traduction durant le développement embryonnaire. Beaucoup reste à explorer dans ce domaine notamment lorsque l'on sait qu'environ 70% des ARN de l'embryon de drosophile sont localisés. Cependant rien n'est connu concernant leur traduisibilité.

Ces travaux ouvrent la voie vers l'étude de la dynamique spatio-temporelle de traduction d'autres protéines, étape clef pour le bon déroulement des fonctions cellulaires.

Table of contents

Résumé	1
Abstract	3
Acknowledgements	5
Long résumé en Français de la thèse	8
Table of contents	19
Chapter 1. Introduction: Inheritance of cell fate	25
A. How do cells know what to become?	26
A.1. Cell fate during embryogenesis	26
A.2. The importance of regulation in space and time of transcriptional programs	29
A.2.1. Gene expression regulation in space	29
A.2.2. Gene expression regulation in time.....	31
A.2.3. Robustness of gene expression	32
B. Regulation of transcription	32
B.1. Transcription factors.....	33
B.1.1. Pioneer or pioneer-like TFs.....	34
B.1.2. Dynamic aspect of TF binding	36
B.1.3. Nuclear microenvironment	42
B.2. TFs in embryonic development	44
C. Maintenance and inheritance of transcriptional programs.....	45
C.1. Transcriptional memory	45
C.2. Memory and mitosis	46
C.2.1. Repressed chromatin states.....	48
C.2.2. Active chromatin states	49
C.3. Mitotic bookmarking	49
C.3.1. Open loci during mitosis.....	50
C.3.2. Transcription factors in mitosis	51
C.3.3. Chromatin associated factors.....	53
C.3.4. Histone marks.....	54
C.4. Mitotic memory: towards a faster transcription re-activation after mitosis?.....	55

C.4.1. Monitoring transcription activation after mitosis	55
C.4.2. Transcription is a multi-step process.....	56
C.4.3. Bookmarking as a facilitator for post mitotic re-activation of transcription.....	58
C.5. Mitotic memory: towards a stronger transcription re-activation after mitosis?	59
D. Early <i>Drosophila</i> embryogenesis.....	60
D.1. Synchronous and fast first divisions.....	60
D.2. Maternal support of early embryogenesis	62
D.3. The awakening of the zygotic genome	63
D.3.1. Major factors at play during ZGA	65
D.3.2. Genome organization during ZGA.....	66
E. Mitotic bookmarking during <i>Drosophila</i> embryogenesis	67
E.1. <i>Female sterile (1) homeotic (Fs(1)h)</i>	67
E.2. Absent, Small, or Homeotic discs 1 protein (Ash1)	69
E.3. GAGA associated factor (GAF).....	70
Aims.....	71
Chapter 2. Results: Mitotic bookmarking by the pioneer-like factor GAF.....	72
Material and Methods.....	97
Chapter 3. Results: Trithorax Group proteins and mitotic bookmarking	132
Chapter 4. Discussion.....	148
Chapter 5. Introduction: Translation dynamics during embryogenesis	155
A. Mechanisms of translation.....	156
A.1. Initiation phase of translation	157
A.2. Elongation phase of translation.....	158
A.3. Termination phase of translation.....	159
B. Regulation of translation in space.....	160
B.1. mRNA localization	160
B.2. Local translation.....	161
C. Regulation of translation in time.....	162
D. Translation in the early <i>Drosophila</i> embryo.....	164
Aims.....	166
Chapter 6. Results: Imaging translation dynamics of <i>twist</i>	167
Chapter 7. Discussion.....	168

Chapter 8. General conclusion	172
References	175
Annexes	206
1. Transcription activation of <i>doc</i> transgene	206
2. Transcriptional memory of the <i>Tom</i> gene	208
3. Publication: Cis-regulatory chromatin loops during early <i>Drosophila</i> development.....	210
4. Publication: Remembering the past, mitotic bookmarking in a developing embryo	211
5. Publication: Zelda le maestro du réveil du génome zygotique	212

Table of illustrations

Chapter 1 :

Figure 1. 1: Autonomous specification in the Patella (mollusk).....	27
Figure 1. 2: Conditional specification in the frog embryo.	28
Figure 1. 3: Syncytial specification in <i>Drosophila melanogaster</i> embryo.....	29
Figure 1. 4 : Ectopic expression of bicoid mRNA in <i>Drosophila</i> embryo.	30
Figure 1. 5 : Example of a mis-expression phenotype.	31
Figure 1. 6 : Different classes of regulatory elements.	33
Figure 1. 7 : Enhancer-promoter gene regulation.	34
Figure 1. 8 : Mechanism of action of pioneer factors.	35
Figure 1. 9 : DNA structure with chemical 'signatures'.	36
Figure 1. 10 : Scheme of Fluorescence Recovery After Photobleaching (FRAP) experiment.	38
Figure 1. 11 : Scheme of Single Molecule Tracking (SMT) experiment.	38
Figure 1. 12 : Transcription factor binding kinetics.	39
Figure 1. 13 : Scheme of transcriptional factor 'hub' inside nucleus.	42
Figure 1. 14 : Scheme of nuclear 'hub' with cis-regulatory elements.....	44
Figure 1. 15 : Representation of the memory time window involved in transcriptional memory.	46
Figure 1. 16 : Transcription memory between mother and daughter cells in <i>Dictyostellium</i>	47
Figure 1. 17 : Transcriptional memory in the <i>Drosophila</i> embryo.	48
Figure 1. 18 : Potential mechanisms of mitotic inheritance.	50
Figure 1. 19 : Mechanism of mitotic inheritance via open chromatin states.	51
Figure 1. 20 : Mechanism of mitotic inheritance via mitotic retention of TFs.	52
Figure 1. 21 : Mid-scale imaging screen in mESc reveals a high fraction of TF remaining on mitotic chromosome.	52
Figure 1. 22 : Mechanism of mitotic inheritance via mitotic retention of chromatin associated factors.	54
Figure 1. 23 : Mechanism of mitotic inheritance via mitotic retention of histone tail modifications.	55
Figure 1. 24 : MS2/MCP system to visualize transcription in live.	56
Figure 1. 25 : Steps of transcription.	57
Figure 1. 26 : Hypothesis of bookmarking role on post-mitotic transcription activation steps.	58
Figure 1. 27 : Scheme of transcriptional memory.	59
Figure 1. 28 : First hours of <i>Drosophila</i> development.	61
Figure 1. 29 : Synchronized nuclear cycles of a <i>Drosophila</i> embryo.	61
Figure 1. 30 : The process of cellularisation during the n.c.14 of <i>Drosophila</i> embryogenesis.	62
Figure 1. 31 : Scheme of the <i>Drosophila</i> Maternal to Zygotic transition (MZT).	63
Figure 1. 32 : Cascade of gene activation for antero-posterior patterning in the early <i>Drosophila</i> embryo.	64
Figure 1. 33 : The formation of germ layers from dorso-ventral patterning in the <i>Drosophila</i> embryo.	65
Figure 1. 34 : Genome organization during <i>Drosophila</i> ZGA.	66

Figure 1. 35 : Conservation of the Bromodomain protein family	68
Figure 1. 36 : Scheme of the different isoforms for dBrd4 in Drosophila.	69
Figure 1. 37 : Scheme of the different isoforms for GAF.	70

Chapter 3 :

Figure 3. 1 : Maternally deposited dBrd4 and Ash1 are retained during mitosis in the Drosophila embryo.	134
Figure 3. 2 : Maternally deposited dBrd4 is retained during all phases of mitosis.	134
Figure 3. 3 : dBrd4 forms nuclear clusters.....	135
Figure 3. 4 : dBrd4 puncta do not exclusively co-localize with Zelda.	136
Figure 3. 5 : dBrd4 depletion generate chromosome conformation rearrangements.....	138
Figure 3. 6 : dBrd4 is retained on specific genes during mitosis.....	139
Figure 3. 7 : Experimental set-up of transcription activation quantification.	140
Figure 3. 8 : Maternal Ash1 reduction delay post-mitotic transcription activation of the sna-transgene.....	141
Figure 3. 9 : Depleting maternal Ash1 do not alter the transcriptional memory bias.	142
Figure 3. 10 : Reducing dBrd4 level does not affect post-mitotic transcriptional timing of activation.	143
Figure 3. 11 : Reducing maternal dBrd4 affects the transcriptional memory.	144
Figure 3. 12 : Scheme of the dBrd4 optogenetic tool.	146
Figure 3. 13 : Optogenetic experiment in S2 cells.....	146
Figure 3. 14 : mCherry-LEXY optogenetic system in the Drosophila embryo.	147

Chapter 5 :

Figure 5. 1 : Three main steps of mRNA translation.	157
Figure 5. 2 : Scheme of the subunits which forms a eukaryotic ribosome.	158
Figure 5. 3 : Scheme of the elongation phase of translation.....	159
Figure 5. 4 : Examples of localized mRNAs.	161
Figure 5. 5 : Different ribosomal composition in a mouse embryo.	162
Figure 5. 6 : Scheme of the SunTag system principle.	163
Figure 5. 7 : In-situ hybridization of diverse mRNAs in Drosophila embryos.....	165

Tables :

Table 1. 1 : Non-exhaustive list of TFs residence time.	41
Table 1. 2 : List of TF binding kinetics which have been measured in interphase and mitosis.....	53

This thesis manuscript is separated in eight chapters. I will first present the main project of my PhD on the mitotic bookmarking during *Drosophila* embryogenesis. This part is developed in four chapters (Chapter 1, 2, 3 and 4) containing introduction, results and discussion. In the course of my PhD, I also worked on a second project related to imaging translation dynamics in the early *Drosophila* embryo, which will be described less extensively than my main project. This part is summarized in Chapters 5, 6 and 7 containing introduction, results and discussion. This manuscript is then concluded with a general conclusion. I included additional results in the Annexe section as well as three publications in which I participated.

Chapter 1. Introduction: Inheritance of cell fate

How an entire organism is progressively built from one unique single cell, is a fascinating question. This question is particularly interesting knowing that cells will divide thousands of times, undergo a series of morphological movements and yet have to adopt a precise fate. Almost all metazoan experience the same steps of development during early embryonic development known as embryogenesis. During this process, cells will be organized in three germ layers, endoderm, mesoderm and ectoderm followed by the first morphogenetic event, gastrulation (Solnica-Krezel and Sepich 2012). The set of genes regulating these steps were first discovered in *Drosophila*. Indeed, in the 1980s Eric Wieschaus, Christiane Nüsslein-Volhard, Edward B. Lewis and their labs discovered and classified several genes responsible of directing undifferentiated and identical cells to form a new fly organism (Wieschaus and Nüsslein-Volhard 2016). Remarkably, many years later, this set of genes was described to be conserved in most metazoan, playing similar functions (Schneider and Amemiya 2016). Therefore, all animals development is governed by the same group of genes, referred to as the developmental-genetic toolkit (Cañestro et al. 2007). Considerable discoveries since the 1980s revealed the clear conservation of the gene toolkit. A classic example is the *Hox* genes family, which controls anterior–posterior patterning in both fly and mammals (McGinnis et al. 1984).

A. How do cells know what to become?

A.1. Cell fate during embryogenesis

In all animals, embryogenesis starts with a pool of initially identical totipotent¹ cells. During the course of development, these cells acquire specific identities to form distinct tissue types. There are different types of specification: autonomous, conditional or syncytial (Davidson 1990). A cell is capable of autonomous specification if it can differentiate without receiving an external signal (Gilbert 2000a). In this case, if a group of cells is removed from an embryo early during its development, those isolated cells will produce the same tissue type than that they would have adopted if they were still part of the embryo (Figure 1.1). This takes place

¹ Undifferentiated cells that can, each of them, form an entire organism.

1during embryogenesis, when cell fate specification arises from the cleavage of cells with asymmetrically expressed maternal cytoplasmic determinants (proteins, mRNAs or small RNAs (Houston 2013)).

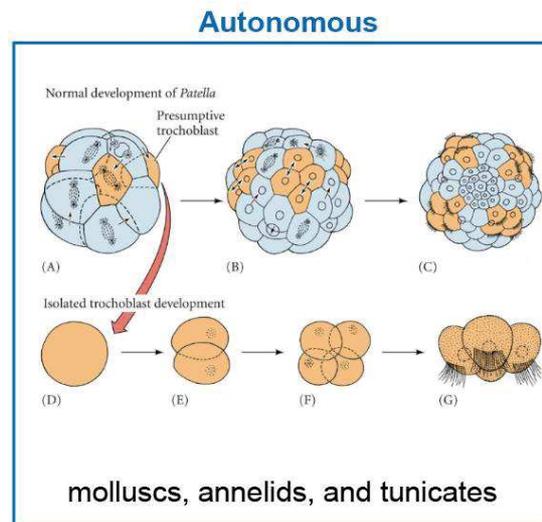


Figure 1. 1: Autonomous specification in the Patella (mollusk).

This experiment was performed by Chabry in 1887. If you transplant cells to another location, they will still keep their original identity. Image from Gilbert 2000a.

Conditional specification consists of cells that will be differentiated depending on the surrounding environment. In an embryonic context, what a cell becomes will depend on its position within the embryo (Figure 1.2). Its fate will be determined by interactions with neighboring cells. Therefore if a group of cells is isolated and misplaced in the embryo, these transplanted cells will give rise to alternative cell types (Gilbert 2000a).

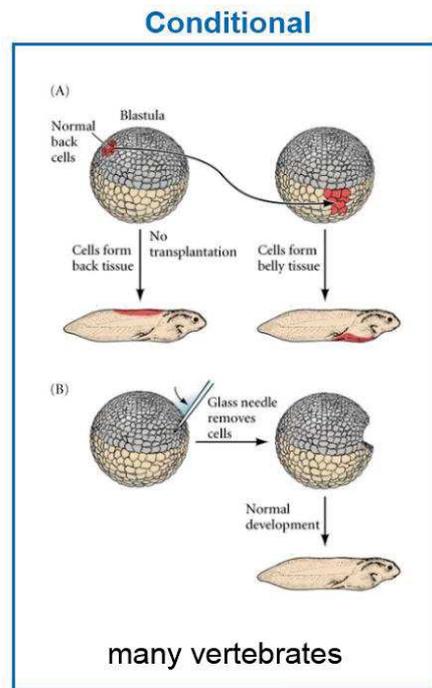


Figure 1. 2: Conditional specification in the frog embryo.

The cell fate is according to its environment. If cells are transplanted to another location they will take the identity of the surrounding cells. Image from Gilbert 2000a.

Syncytial specification involves both of autonomous and conditional. Here, specification does not occur between cells but within parts of the embryo, as nuclei share the same cytoplasm (Figure 1.3). It is the case for insects such as *Drosophila* (Gilbert 2000a). This makes this organism a good model to study both of these phenomena.

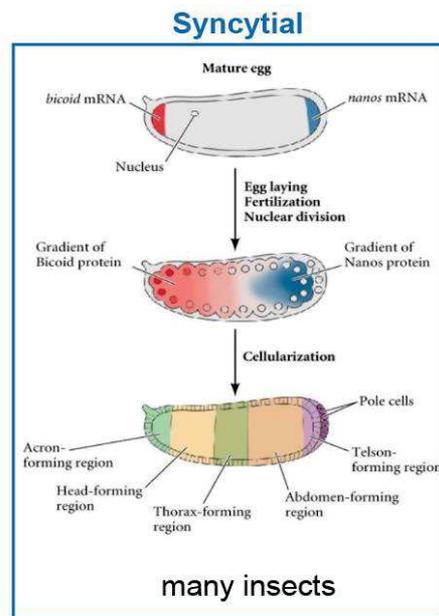


Figure 1. 3: Syncytial specification in *Drosophila melanogaster* embryo.

Anterior-posterior axis specification originates from gradients within the egg cell. Image from Gilbert 2000a.

Then, it is the different combination of transcription factors that will elicit transcription programs that are cell-type specific and will give an identity to each cell. For a harmonious development, these transcription programs must be maintained in specific group of cells and during a certain amount of time throughout embryonic development.

A.2. The importance of regulation in space and time of transcriptional programs

A.2.1. Gene expression regulation in space

In all *Metazoan*, the early steps of development are crucial for the cells to get their identity in terms of transcriptional program². Indeed, in a period spanning from three hours to 48 hours in human (Vastenhouw et al. 2019), maternally deposited mRNAs and proteins generate localized patterns of gene activity, such as polarity axis and segmentation stripes of gene expression (Gilbert 2000b; Blair 2008). These maternal determinants will generate a 'cascade' of gene regulation, highly controlled in space and time. Expression of a gene in a wrong

² I define transcriptional program as a specific expressed set of genes in a cell, giving its cellular identity. It also encompasses sets of transcription factors which cooperate to achieve this co-regulation of genes.

location has been shown to lead to defects in development of *Drosophila* embryo, such as pole plasm localized *nanos* mRNA (Gavis and Lehmann 1992) or ectopic injection of *bicoid* mRNA leading to the formation of anterior structures (Driever et al. 1990). In this last study, they injected *bicoid* mRNA which is normally localized to the anterior pole of the embryo, at the posterior pole (Figure 1.4). This leads to the formation of an embryo with two heads (but of course they do not fully develop) (Driever et al. 1990). To me this beautiful experiment revealed the importance of master maternally deposited factors, and it is crazy to think that only one factor can generate complex structures such as the head in the embryo. Another famous example is the mis-expression of the *antennapedia* homeotic gene (also conserved in mammals) expressed in head precursor cells of the embryo inducing the formation of legs instead of antenna (Schneuwly et al. 1987) (Figure 1.5). The spatial regulation of gene expression is therefore very important to generate a complete well organized organism.

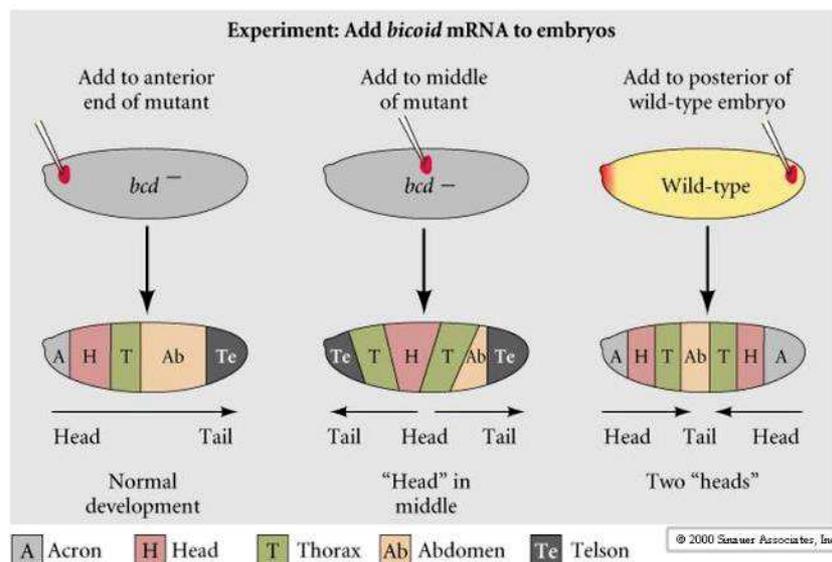


Figure 1. 4 : Ectopic expression of *bicoid* mRNA in *Drosophila* embryo.

In this experiment (Driever et al. 1990), *bicoid* mRNA was injected in different parts of the embryo either in *bicoid* mutant embryos (left and middle schematics) or at the posterior part of a wild type embryo (right schematic). Each time this leads to the formation of head structures at the place where the mRNA was injected, suggesting the importance of mRNA localization in morphological structures formation.

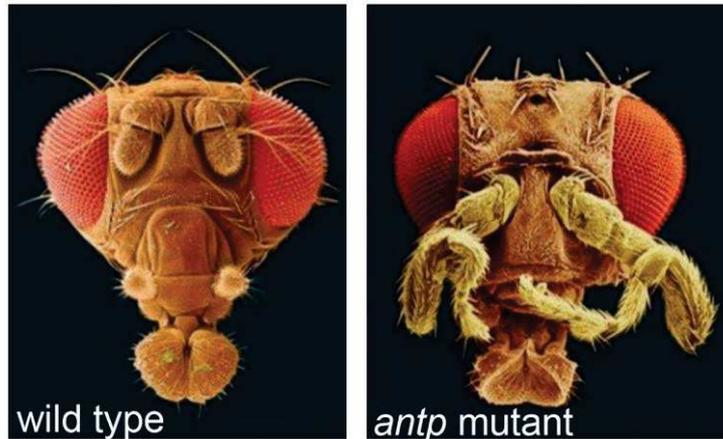


Figure 1. 5 : Example of a mis-expression phenotype.

(left) Wild type *Drosophila* head. (right) *Drosophila* head in *antennapedia (antp)* dominant gain-of-function mutant³. Note the legs formation instead of antenna.

A.2.2. Gene expression regulation in time

Gene expression regulation in time is still not well understood, mainly due to technical limitations, but few examples can be given (Balaskas et al. 2012; Lagha et al. 2013; McDaniel et al. 2019; Kogler et al. 2021; Huang et al. 2017). In the study of Lagha et al., the authors have shown that changing transcription synchrony between cells by changing the promoter of the *snail* gene in *Drosophila* embryos, leads to gastrulation defects. More recently, optogenetics tools allowed the study of protein perturbation in time. In 2017, Huang et al. were able to inactivate the transcription factor Bicoid in a time-controlled and reversible manner in the *Drosophila* embryo. They showed that persistent Bicoid-dependent transcription activity from early in development is required for robust embryonic patterning, and ultimately embryonic viability. In the study of McDaniel et al., they used the same approach to temporally inactivate Zelda, a master regulator of gene activation in the *Drosophila* embryo. They showed that the continuous expression of Zelda is required for the proper development. In Kogler et al., they inactivated the transcription factor Twist at different developmental stages of *Drosophila* embryogenesis and determined a precise time window, during which Twist expression is required for proper development.

³ A type of mutation in which the altered gene product possesses a new molecular function or a new pattern of gene expression. Gain-of-function mutations are almost always dominant.

Temporal control of gene expression patterning has also been highlighted during the nervous system development of *Drosophila* but also vertebrates (Holguera and Desplan 2018; Kohwi and Doe 2013). The sequential production of different cell types at the same location has been proposed to be one of the sources of generation of cell type diversity among neurons. Thus, in this case cell type specification does not only rely on spatial cues (Sagner et al. 2020). In this study, the authors provide the evidence of a general temporal program of neuroblast subtypes identity. By birth-dating neurons with EdU labeling of mouse embryos, they were able to identify distinct cohorts of transcription factors that are induced at different developmental stages in neurons in the spinal cord. They were also able to identify factors (TGF β signaling) responsible for the speed of neuronal progenitor maturation.

A.2.3. Robustness of gene expression

As development seems sensitive to changing in gene expression in space and time, this must be regulated in a robust manner. In 1942, Conrad Hal Waddington brought the concept of the existence of several “redundant” mechanisms which lead a group of individuals to have the same phenotypes, regardless of some genetics or environmental variations. This process was called canalization (Waddington, 1942). It is believed that gene expression is regulated by several mechanisms that leads to precision to ensure a proper embryonic development (Lagha et al. 2012; Bentovim et al. 2017). For example, the presence of shadow enhancers, groups of enhancers that control the same target gene and drive similar expression patterns (Hong et al. 2008), is one of these mechanisms to faithfully express critical patterning genes (Perry et al. 2010; Frankel et al. 2010). As I discuss in Chapter 4, I believe that one process of gene expression robustness would be the memory of transcriptional programs during development, as it would ensure the propagation of the information of gene expression (Ferraro et al. 2016; Chubb 2016).

B. Regulation of transcription

At the center of gene expression regulation, stand the Transcription Factors (TFs). This class of factors are defined by their ability to directly bind DNA (Latchman 1993) and can have a positive or negative action on RNA polymerase II (Pol II) transcription (Roberts 2000).

Chromatin regulators are cofactors which interact with TFs and can also act positively or negatively regulate transcription (Wolffe et al. 1997). Here I will focus on TFs and chromatin regulators which act as activators of transcription. I will also mainly present in this section TFs and chromatin regulators which play major roles in embryonic development.

Transcription is regulated in space and time by TFs which bind to *cis*-regulatory sequences. *Cis*-regulatory sequences are promoters, enhancers and insulators (Noonan and McCallion 2010). These elements are non-coding sequences where factors can bind to activate or repress transcription (Figure 1.6).

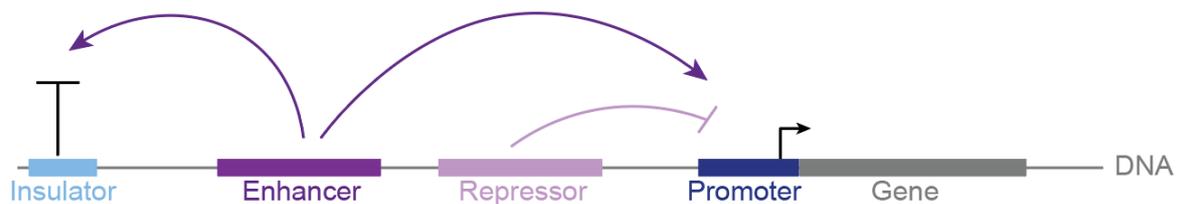


Figure 1. 6 : Different classes of regulatory elements.

Enhancers can activate transcription (purple arrow) and repressors can repress (pink arrow) transcription of a given gene. Promoters are necessary for transcription activation (black arrow). Insulators inhibit the action of enhancers and repressors to neighboring genes.

B.1. Transcription factors

Transcription factors are proteins which interact directly with DNA through a DNA binding domain (Garvie and Wolberger 2001). These TFs have multiple different partners to regulate a specific target gene. Each regulated gene involves a network of TFs, co-activators, histone tails modifiers. The combinatorial action of these regulators creates a particular local chromatin environment, with various consequences for transcription. Such environments are present on *cis*-regulatory regions (promoters and enhancers) and could represent large complexes (Figure 1.7).

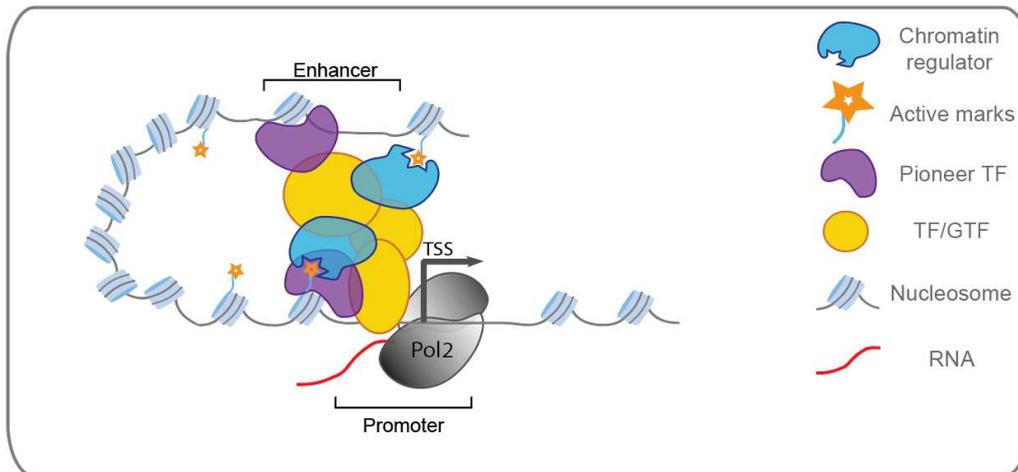


Figure 1.7 : Enhancer-promoter gene regulation.

Scheme of enhancer-promoter loop with regulation factors to activate transcription.

B.1.1. Pioneer or pioneer-like TFs

Pioneer factors were discovered when researchers wanted to understand what are the first factors able to induce regulatory complexes on a given gene in mouse embryo (Gualdi et al. 1996). It is now well assumed that TFs can be considered as pioneer if they fulfill specific features (Zaret 2020a) (Figure 1.8). First, they are able to directly target DNA on nucleosome. This central feature of pioneer factors is usually tested *in vitro* by looking at the binding of the purified TF incubated with mononucleosomes (Cirillo et al. 1998; McDaniel et al. 2019). Second, when bound to DNA pioneer factors are able to displace nucleosome, thereby creating a local 'open' chromatin state. This opening can be assessed by Micrococcal Nuclease (MNase) digestion (Cirillo et al. 2002; Iwafuchi-Doi and Zaret 2016) or Assay of Transposase Accessible Chromatin sequencing (ATAC-seq)⁴ (Buenrostro et al. 2013) and can be made without the recruitment of chromatin remodelers (Cirillo et al. 2002). In fact, this property elicited the coining of the term "pioneer" (Cirillo et al. 2002). In this study but also in a previous paper (Cirillo and Zaret 1999), FoxA1 protein was shown to engage its target sites on compacted nucleosome arrays with a subsequent opening of chromatin, independently of ATP-dependent nucleosome remodelers (Clapier et al. 2017). Interestingly,

⁴ One distinction between MNase and ATAC-seq is that MNase resolved the accessibility at the single nucleosome level, whereas ATAC-seq do not.

the recombinant FoxA₁, but not other transcription factors, bound to its sites on mono- and dinucleosomes (Cirillo and Zaret 1999; Shim et al. 1998) and helped GATA₄ to bind to adjacent sequence (Cirillo and Zaret 1999). This shows that pioneer factor binding is then usually followed by recruitments of other TFs. Another example is the pioneer factor Zelda that has been shown to locally deplete nucleosome, being necessary for another TF, Dorsal, to bind to its targets (Sun et al. 2015) in *Drosophila* embryos. Additionally, this newly created chromatin environment is then associated with covalent modifications of histones tail such as histone H₃ methylation⁵ on lysine 4 (H₃K₄me) and H₃K₉ as well as H₃K₂₇ acetylation⁶ (ENCODE Project Consortium 2011; Kharchenko et al. 2011). To date, the pioneer factor category represents only small number of transcription factors and even a smaller subset have been studied in depth.

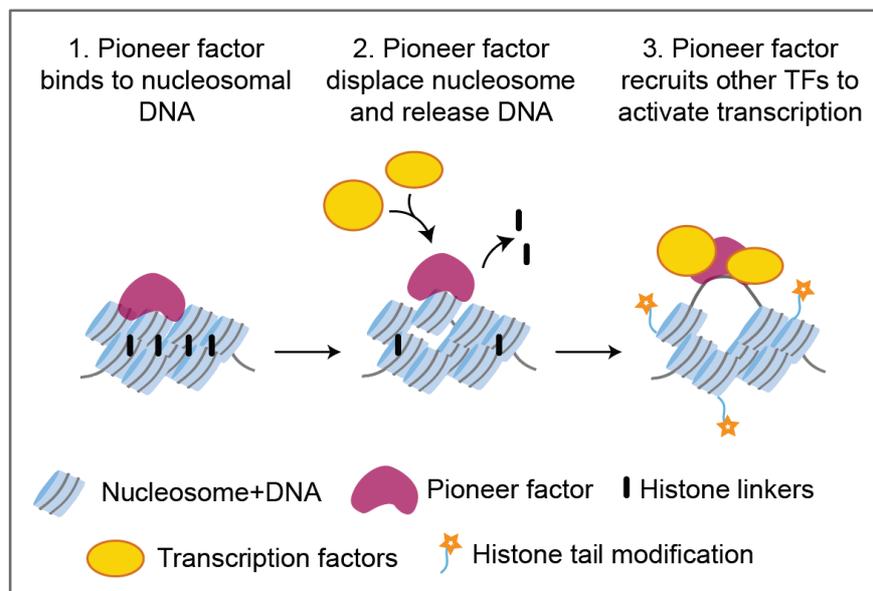


Figure 1. 8 : Mechanism of action of pioneer factors.

Pioneer factors interact with closed chromatin (1.). This leads to nucleosome displacement (like through histone linker release) and DNA unmasking (2.). Then TFs are recruited in order to activate transcription (3.).

⁵ Post translational chemical modification, here on a lysine, which corresponds to an addition of a CH₃ group.

⁶ Post translational chemical modification, here on a lysine, which corresponds to an addition of an acetyl C₂H₃O group.

B.1.2. Dynamic aspect of TF binding

B.1.2.1. TFs binding configuration

The double helix of DNA is a polymer of negatively charged sugar-phosphate backbone. One of the four base (Adenine, Thymine, Cytosine, Guanine) is attached to each sugar, and have a part exposed in the major and minor grooves (Garvie and Wolberger 2001). Each of the four DNA base pair has a characteristic chemical composition and this gives a specific chemical 'signature' exposed in the minor or major groove of the double helix (Figure 1.9). A particular DNA motif (called consensus binding sequence) will create a specific 'signature', recognized and directly bound by a corresponding TF.

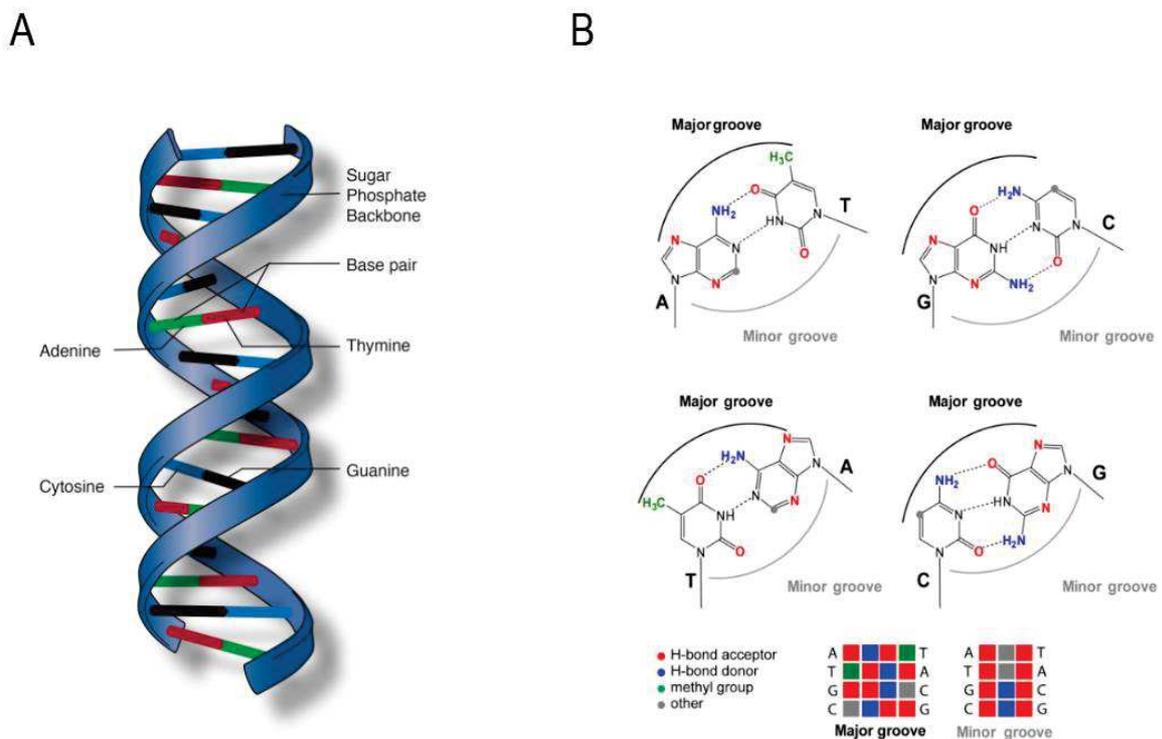


Figure 1.9 : DNA structure with chemical 'signatures'.

(A) Scheme representing the structure of the DNA double helix with the sugar-phosphate backbone (blue) and the four base pairs (green, blue, black and red). Illustration is from www.genome.gov/genetics-glossary/Double-Helix.

(B) Chemical structure of the four bases showing the different groups present in the major (black half-circles) or minor (grey half circles) groove. Each combination of bases gives a chemical 'signature', represented with the colored squares. Figure from (Harteis and Schneider 2015).

In addition, DNA is wrapped around nucleosomes which cover most of the genome (McGinty and Tan 2015) and are thought to be displaced by transcription factors in regions that direct gene expression. However, the modes of interaction between transcription factors and nucleosomal DNA remain largely unknown. Recently, nucleosome DNA binding was tested for 220 TFs from different structural families (Zhu et al. 2018). This study concluded that the vast majority of these transcription factors have less access to nucleosomal DNA than to free DNA, but that a subset of TFs are capable of directly binding nucleosomal DNA. Importantly, this study revealed that there are differences in the binding configuration of transcription factors to free or nucleosomal DNA. Taking together, these results uncovered a diverse interaction landscape of affinities between transcription factors and DNA (rather than a binary phenomenon).

B.1.2.2. TFs binding kinetics

Binding kinetics usually refers to the speed of a TF binds to its target and how fast it dissociates from it. We can then assess how long a TF binds to its target, the residence time. To have access to the dynamics of TFs binding, experiments such as Fluorescence Recovery After Photobleaching are usually employed (Sprague et al. 2004; Mueller et al. 2008) Figure 1.10, or single-molecule tracking (SMT) (Izeddin et al. 2014; Liu and Tjian 2018; Gebhardt et al. 2013) (Figure 1.11).

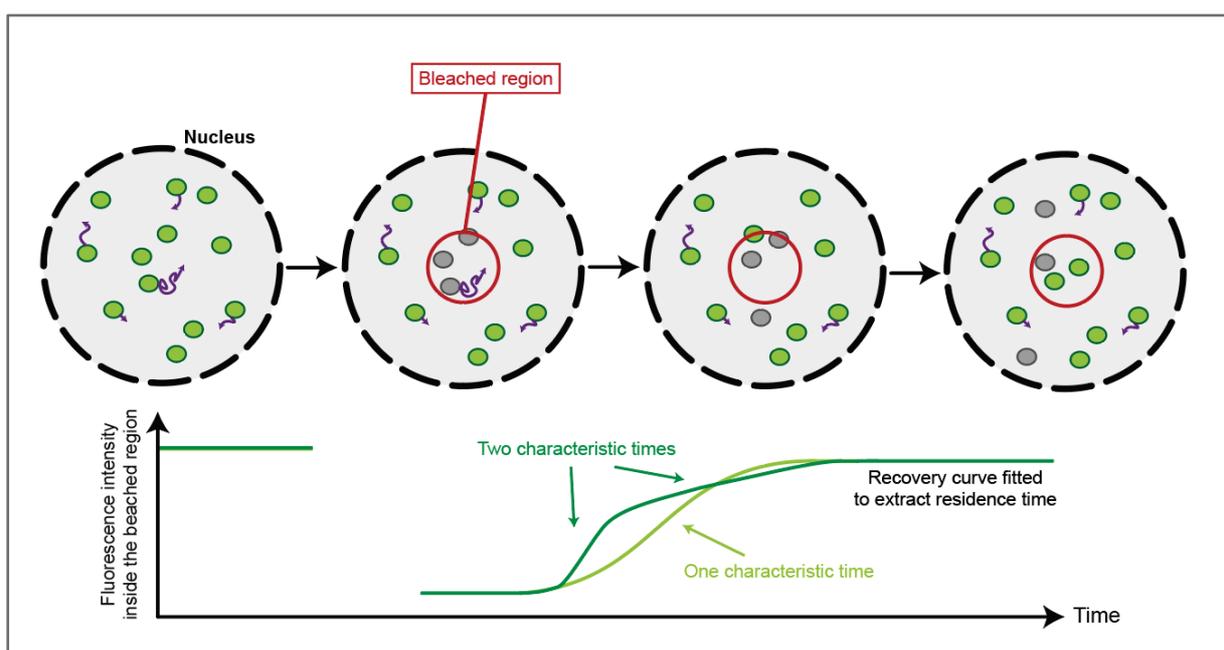


Figure 1. 10 : Scheme of Fluorescence Recovery After Photobleaching (FRAP) experiment.

FRAP consists of photobleaching a sub-region (red circle) of TF proteins coupled with fluorophores (green dots) inside the nucleus. During the recovery phase, bleached molecules (grey dots) move out of the bleach region and unbleached proteins (green dots) move in. Measurement of FRAP is done by averaging the fluorescence intensity in the bleached region (red circle) as a function of time. The speed of the recovery is determined by the dynamic properties of the studied protein. Fast recovery is indicating diffusion while transient binding leads to slow down of the curve. This curve is then fitted with one, two (or more) parameters depending on the shape of the curve (one inflection point = one characteristic time, two inflections = two characteristic times).

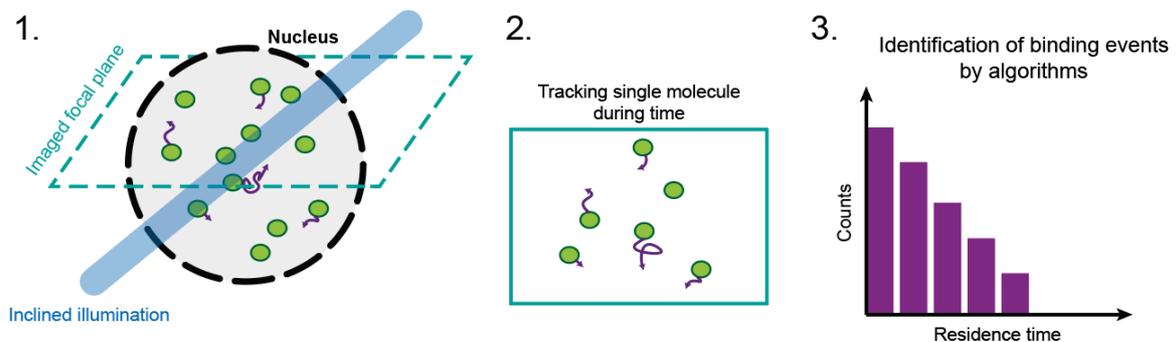


Figure 1. 11 : Scheme of Single Molecule Tracking (SMT) experiment.

Detection of individual fluorescently tagged molecules (usually using HaloTag (Los et al. 2008)) using an inclined illumination beam (HILO illumination, 1.). Imaging data are recorded during time to track single molecule and identify binding events (2.). The duration of each binding event is extracted to obtain residence times of the studied protein (3.).

These techniques allow retrieving the k_{off} = the time for the TFs to unbound its target, which corresponds to the apparent residence time if we take the inverse (Figure 1.12). Generally, TFs (regardless of being pioneer or not) can have apparent residence times from less than a second to the order of tens of seconds. One of the shortest has been measured for the Estrogen Responsive Unit (ERU) with a residence time of 0.65 sec (Gebhardt et al. 2013). In the case of Cohesin in mES cells, its residence time can be very long, in the order of 22 min (Hansen et al. 2017). Yet, it seems that most TFs bind quite transiently (Table 1.1) (Mazzocca

et al. 2021). Usually, apparent long resident times is assumed in many papers to represent sequence-specific binding whereas short ones may reflect non-specific binding (Deluz et al. 2016; Chen et al. 2014b; Mazzocca et al. 2021). However this interpretation is debatable given the technical limitations due to imaging (level of expression in transiently expressing cells, photobleaching over time, binding to non-chromatin related components) (Sprague et al. 2004; Mueller et al. 2008; Mazzocca et al. 2021). Also, it not yet clear whether the timescales of non-specific versus specific binding are well separated. Specific DNA binding can be assessed with DNA binding mutants form of the studied TF. An elegant example is represented by FRAP experiments with FoxA1 mutants in hepatocyte cells (Sekiya et al. 2009). This study revealed that a minimum of two point mutations affects FoxA1 target recognition but not non-specific binding *in vitro*. Conversely, two other point mutations affect only non-specific binding and increase FoxA1 mobility in the nucleus. Furthermore, as pioneer-factors have the ability to directly bind nucleosomes, DNA binding and nucleosome binding can occur in the same time-scale. In that case, a point mutation in the DNA binding could therefore not affect the apparent residence time, corresponding to the nucleosome binding.

However, all of this is based on the theory where specific binding to a sequence would lead to a stronger binding therefore a longer residence time. This is totally logical but remain to be fully demonstrated. We can still conclude that depending on the TFs studied, a TF can interact strongly with DNA but also transiently potentially reflecting the scanning of DNA (Raccaud et al. 2019).

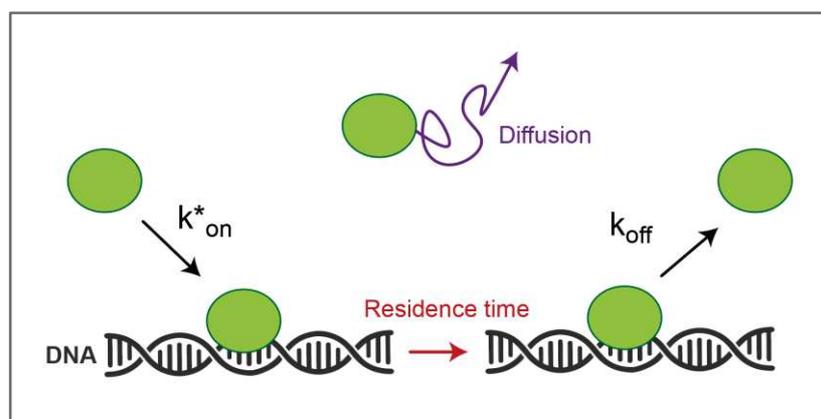


Figure 1. 12 : Transcription factor binding kinetics.

Multiple kinetic parameters can be extracted from live imaging of fluorescent proteins experiments. Diffusion (purple arrow) corresponds to free diffusive molecules. k^*_{on} corresponds to the rate of the TF to bind to DNA. k_{off} represents the rate to which the molecule leaves the binding state. Residence time can be extracted from these parameter ($=1/k_{off}$).

PROTEIN	INTERPHASE	REFERENCE	ORGANISM
TBP	RT : 88s	(Teves et al. 2018)	mES cells
FKH	RT : 30-60s	(Gomez-Lamarca et al. 2018)	Salivary glands
SU(H)	RT : 0.5-2s	(Gomez-Lamarca et al. 2018)	Salivary glands
HAIRLESS	RT : 0.5-2s	(Gomez-Lamarca et al. 2018)	Salivary glands
ZELDA	RT : 2.5-3.3s	(Dufourt et al. 2018)	embryo
BICOID	RT : 2s	(Mir et al. 2018)	embryo
SOX2	RT : 12s	(Chen et al. 2014b)	Es cells
BHLHB8	T1/2 : 10.82s	(Raccaud et al. 2019)	mES cells
CDX2	T1/2 : 14.03s	(Raccaud et al. 2019)	mES cells
FOXA1	T1/2 : 6.47s	(Raccaud et al. 2019)	mES cells
NANOG	T1/2 : 3.55s	(Raccaud et al. 2019)	mES cells
POU5F1	T1/2 : 5.33s	(Raccaud et al. 2019)	mES cells
SOX2	T1/2 : 7.68s	(Raccaud et al. 2019)	mES cells
MAX	T1/2 : 16.21s	(Raccaud et al. 2019)	mES cells
TEAD1	T1/2 : 8.51s	(Raccaud et al. 2019)	mES cells
SOX15	T1/2 : 2.92s	(Raccaud et al. 2019)	mES cells
SOX2	RT : 19.7s	(Teves et al. 2016)	mES cells
GRD	RT : 1.45s	(Gebhardt et al. 2013)	MCF-7 cells
ERE	RT : 3.85s	(Gebhardt et al. 2013)	MCF-7 cells
ERU	RT : 0.65s	(Gebhardt et al. 2013)	MCF-7 cells
CTCF	RT : 1min	(Hansen et al. 2017)	mES cells
COHESIN	RT : 22min	(Hansen et al. 2017)	mES cells
FOXA1	T1/2 : 14s	(Sekiya et al. 2009)	H2.35 cells
GATA-4	T1/2 : 7s	(Sekiya et al. 2009)	H2.35 cells
NF-1	T1/2 : 5s	(Sekiya et al. 2009)	H2.35 cells

C-MYC	T1/2 : 8s	(Sekiya et al. 2009)	H2.35 cells
HMGB1	T1/2 : 1s	(Sekiya et al. 2009)	H2.35 cells

Table 1. 1 : Non-exhaustive list of TFs residence time.

RT: residence time. T_{1/2}: the halftime of recovery is the time from the bleach to the time point where the fluorescence intensity reaches the half (1/2) of the final recovered intensity. H2.35 is an epithelial-like mouse cell line. MCF-7 is a human breast cancer cell line.

B.1.2.3. Consequence of TFs binding on transcription regulation

Transcription activation has been shown to be sensitive to TF concentration. Indeed, increasing TF concentration was reported to be associated with an enhanced mRNA production (Dar et al. 2012; Ko et al. 1990; Larson et al. 2013). This reflects the rate at which the regulatory element, promoter or enhancer, is bound by the TF. Quite recently, TFs residence time has also been reported to influence mRNA production. As for concentration, an increased residence time correlates with an augmentation in mRNA production (Donovan et al. 2019; Callegari et al. 2019; Clauß et al. 2017; Stavreva et al. 2019). Interestingly, by combining artificial activators that could be tuned to play on the DNA binding domain and the MS2/MCP system (Bertrand et al. 1998) in human cell lines, *Popp et al.* reported that TF residence time has even more effect than concentration on transcription activation, thus by increasing burst⁷ frequency and not burst duration or amplitude (Popp et al. 2021). Now multiple studies suggest that TF residence time might be one of the best predictor for gene transcription (Senecal et al. 2014; Loffreda et al. 2017).

To nuance these conclusions, it should be noted that these results were obtained with reporter assays and might depend on TF intrinsic properties and their dimerization (or multimerization) and/or interaction with co-factors. But in any case this highlights that the relationship between TF kinetics and transcription efficiency is probably more complex than just reflecting gene occupancy.

⁷ A burst of transcription is a property of most of the genes in which transcription can occur in "bursts" or "pulses" of mRNA production.

B.1.3. Nuclear microenvironment

As concentration of TF can be important for gene regulation, several laboratories have studied the distribution of a specific factor inside the nucleus. In fact, some TFs have a non-homogeneous distribution and can form what is called 'micro-environments', 'hubs' or 'condensates' (Tsai et al. 2017; Mir et al. 2017; Liu et al. 2014; Sabari et al. 2020; Chen et al. 2014b). This has also been reported for RNA polymerase II (Cisse et al. 2013). I believe that most of the TFs harbor such a distribution. Whether this leads to a functional advantage is not fully understood. The local concentration of the morphogen⁸ Bicoid, was suggested to increase its time-averaged occupancy, therefore favoring recruitment of other factors to activate transcription (Mir et al. 2017). Complementary to this, Huang et al. have shown that Bicoid differential concentration influences the timing of transcription (Huang et al. 2017). This is particularly interesting knowing that Bicoid, and also another factor present in the early embryo Zelda, binds very transiently to the DNA (Mir et al. 2017, 2018; Dufourt et al. 2018; Kent et al. 2020) therefore this local concentration increases the chances for the factor to bind (Figure 1.1).

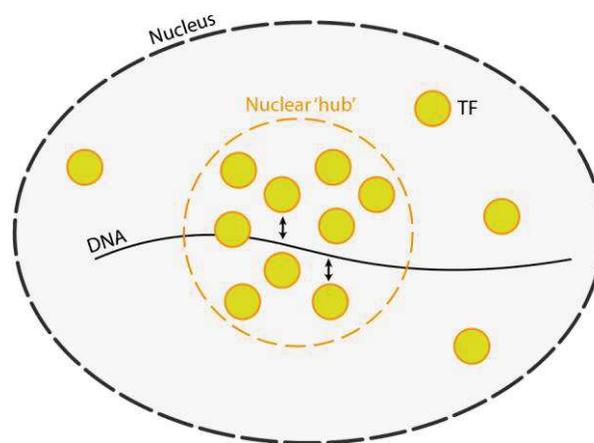


Figure 1. 13 : Scheme of transcriptional factor 'hub' inside nucleus.

TF (yellow dots) are not homogenously distributed inside nucleus and can be more concentrated at certain loci of DNA.

⁸ A morphogen is a signaling molecule (here a transcription factor) that act over long distances to induce fate responses in cells based on the concentration of morphogen.

In terms of transcription, this local concentration has led to relatively conflicting results. Indeed, few studies reported that high local concentration of TF can significantly increase transcription (Trojanowski et al. 2021; Wei et al. 2020; Schneider et al. 2021) but can also decrease transcription (Schneider et al. 2021). However, one recent study has shown that TF activity can be concentration dependent and mediated by their low-complexity domain⁹ (Chong et al. 2021). Indeed, *Chong et al.* reported that when substantially increasing concentration of the EWS TF, with its low-complexity domain, in human cell line, it represses transcription of a target gene. Previously, it has been described that EWS low-complexity domains interactions are necessary for the activation of the microsatellite-associated target genes (Chong et al. 2018) This means that to a certain threshold of concentration of factors containing low-complexity domains, transcription can be locally repressed. To my knowledge this is the first study that shows the effect of fine-tuning TF concentration *in cellulo* on transcription activation.

Importantly, these nuclear 'hubs' might also contain several specific *cis*-regulatory elements such as super enhancers¹⁰. It is the case for super-enhancers bound by the transcription factor BRD4 (Sabari et al. 2018) which reveals a selective partitioning of nuclear condensates and a specific 3D genome organization (Sabari et al. 2020) (Figure 1.14). What I call clustering here is different from looping between two DNA *cis*-regulatory sequences. To my mind, a cluster of *cis*-regulatory sequences contains two or more *cis*-regulatory elements and therefore can contain several loops. This spatial clustering of multiple loci has been shown to participate to gene regulation positively (activation) (Sabari et al. 2018; Allahyar et al. 2018; Oudelaar et al. 2018) but also negatively in case of repressors or heterochromatin associated proteins (Treen et al. 2021; Peng et al. 2020; Strom et al. 2017). The exploration of functional role of such clustering and TF concentration are just emerging and might be TF and DNA loci specific. During embryogenesis, few studied suggested this phenomenon of *cis*-regulatory element

⁹ Protein domains with low sequence complexity of amino acids. Here in the case of transcription factors, I referred to this domain as intrinsically disordered in their structure meaning that it does not have a defined structure.

¹⁰ Region of the genome with a high concentration of enhancers usually bound by multiple transcription factors. They have been shown to drive transcription of specific genes important for cell fate.

clustering in *Drosophila* embryos (Dufourt et al. 2018; Hug et al. 2017a; Cardozo Gizzi et al. 2019; Bantignies et al. 2011). During my thesis, as a collaborative project with the lab of M. Nollmann (Centre de Biochimie Structurale, Montpellier), we investigated the existence of *cis*-regulatory regions spatial clustering with imaging based methods. This work is published in the manuscript by *Espinola et al.* (see Annexe 4: Cis-regulatory chromatin loops during early *Drosophila* development).

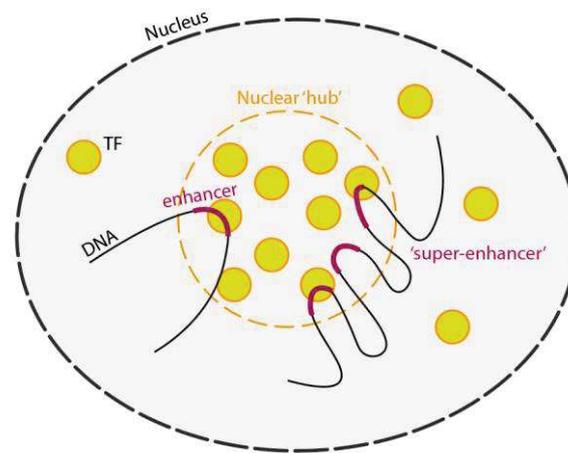


Figure 1. 14 : Scheme of nuclear 'hub' with cis-regulatory elements.

TF nuclear 'hubs' are believed to be clustered with enhancers and 'super-enhancers' to facilitate transient binding of TFs.

B.2. TFs in embryonic development

Development in all metazoan requires a precise orchestration of gene expression in order to give to the cells the right identities. This orchestration is mainly dictated by morphogen gradients and TFs expression in specific tissues (Schep and Adryan 2013). One important aspect that needs to be highlighted is the concept of priming. Priming represents the modification of specific chromatin domains converted from a closed state to an open state. Chromatin priming has first been shown with single locus analysis (Gualdi et al. 1996; Xu et al. 2007; Hoogenkamp et al. 2009) and then genome-wide (Spitz and Furlong 2012; Gifford et al. 2013). This priming consists of chromatin modifications (e.g. on histone tail) and TFs binding (mostly pioneer factors) (Bonifer and Cockerill 2017) relatively a long time before the onset of transcription. For example in *Drosophila* embryos, Zelda binds to its targets genes such as *even-skipped* or *zerknüllt* before the onset of transcription (Nien et al. 2011; Harrison

et al. 2011). This priming is tightly regulated in order to give the right identities to the cells during differentiation, this has been shown for pioneer factors which can exhibit cell-specific actions (Sérandour et al. 2011).

C. Maintenance and inheritance of transcriptional programs

In biology, definitions of certain words can vary between scientists and evolve during years. Classically, the term epigenetic refers to changes in gene regulation, which are not due to a modification in the DNA sequence. However, this notion is also used to define any heritable change, still not due to nucleotide sequence modification. Here, I will not use the term epigenetic for inheritance and memory, because to me epigenetic is also used for mechanisms and specific modifications, whereas inheritance here refers to the concept of transmission.

C.1. Transcriptional memory

The concept of transcriptional memory has been first discovered in yeast (Acar et al. 2005) with galactose induction and inositol starvation (Brickner et al. 2007). It has also been shown in plants with environmental stress (Ding et al. 2012, 2013; Sani et al. 2013) and in mammalian cells during interferon induction (Light et al. 2013; Kamada et al. 2018). During *Drosophila* development, a memorization process has been suggested to be at play for *hunchback* gene highly synchronous expression after mitosis, given the short amount of time for the gene to 'read' the Bicoid TF gradient (Porcher et al. 2010). The main idea of these studies is that cells that has been first stimulated are "primed" during a first induction and will activate faster and stronger after a certain period of time, the "memory time window".

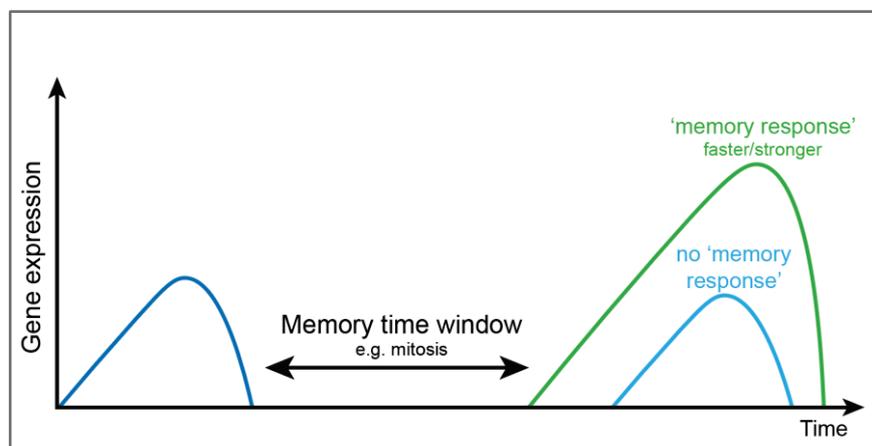


Figure 1. 15 : Representation of the memory time window involved in transcriptional memory.

The first induction of transcription (dark blue curve) lead to a stronger and or faster transcriptional response (green curve) after the memory time window (which could be mitosis or DNA replication for example). No change in transcriptional output represent no 'memory response' (light blue curve).

In these studies, the memory time window is considered as the lag time prior to gene activation after stimulation (Figure 1.15). This time window could be more or less long and we can imagine that in a developing organism this memory time window can be biological processes such as replication or mitosis (Bheda et al. 2020). Indeed, these events are believed to be temporally regulated to avoid any replication-transcription machineries leading to DNA damages for example (García-Muse and Aguilera 2016; Hamperl et al. 2017), or that the compact nature of mitotic chromosomes preclude transcription machinery binding (Martínez-Balbás et al. 1995; Ohta et al. 2010; Parsons and Spencer 1997; Prescott and Bender 1962). But recent studies (Palozola et al. 2017; Wojenski et al. 2021) claim that transcription is not arrested during mitosis by sequencing nascent RNA with EU-RNA-Seq. However, given the experimental system used, these could be questionable. Indeed, *Palozola et al.* employed drug-synchronized cells and do not take into account the potential heterogeneity in mitotic-arrested cell populations. After synchronization, the cells are progressively desynchronized and therefore might not have the same speed of cell cycle progression. In fact, these data have been re-analyzed and it seems that fewer proportion of genes show transcription during mitosis (approx. 2000) presumably around metaphase (Sarnataro et al. 2020). Moreover, these experiments were performed in bulk cell population and therefore do not take into account the behavior of single cells.

C.2. Memory and mitosis

In addition to examples documenting transcriptional memory upon stimulation, live-imaging studies revealed the existence of such a memory bias in wild type, un-stimulated cells. The direct visualization of the inheritance of transcriptionally active states down cell lineages was first monitored in *Dictyostellium* (Muramoto et al. 2010). The authors were the first to observe

a strong similarity in frequency of transcriptional firing between mother and daughter cells has been observed on a reporter gene (Figure 1.16). The existence of a transcriptional mitotic memory has next been shown in *Drosophila* embryos where nuclei deriving from previously transcriptional active nuclei activate on average faster after mitosis than the ones deriving from inactive nuclei (Ferraro et al. 2016) on a reporter transgene (Figure 1.17 and section C.4.3.). Studies in cultured cells have also shown the same concept (Palozola et al. 2017; Teves et al. 2018; Deluz et al. 2016; Phillips et al. 2019). Chromatin states inheritance across cell divisions has been of interest recently and can be active or repressive (Elsherbiny and Dobрева 2021).

There are two very important groups of protein known to contribute to the inheritance of chromatin states: Polycomb Group (PcG) and Trithorax Group (TrxG) proteins. These two groups of proteins function antagonistically to maintain active (TrxG) or silent (PcG) states of gene expression over many generations (Steffen and Ringrose 2014).

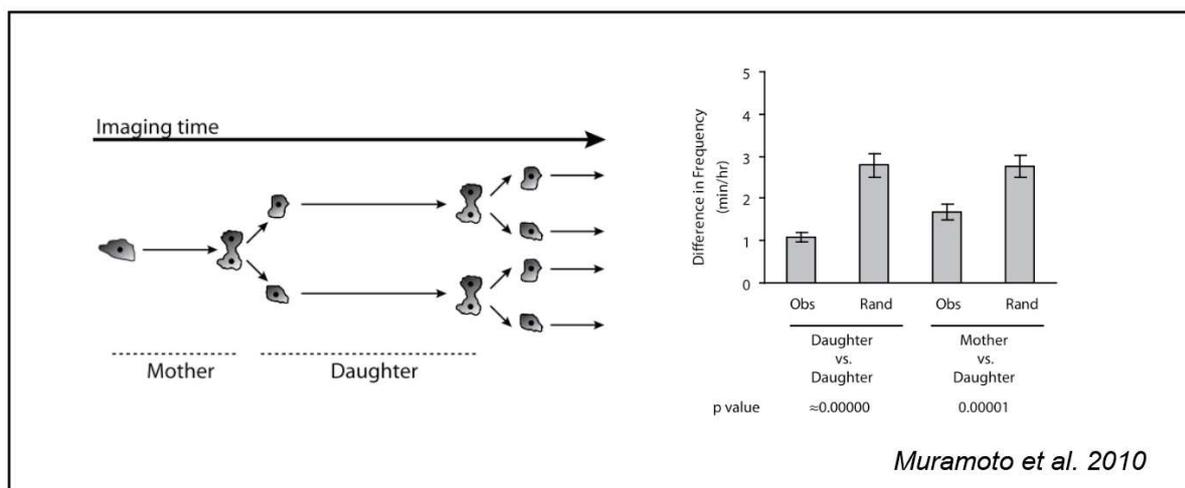


Figure 1. 16 : Transcription memory between mother and daughter cells in *Dictyostellium*.

(Left) Schematic of monitoring transcription down cell lineage from mother to daughter cells.

(Right) Histograms showing the difference of transcription frequency between daughter cells and mother-daughter cells compare to random cells.

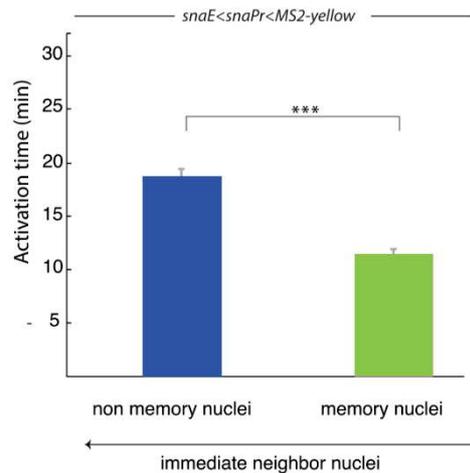


Figure 1. 17 : Transcriptional memory in the *Drosophila* embryo.

Histogram of the activation time of transcription of nuclei coming from inactive (blue: non memory nuclei) or active (green: memory nuclei) nuclei. Data of this graph are from (Ferraro et al. 2016).

C.2.1. Repressed chromatin states

Repressive states are usually maintained via polycomb repressor complexes, associated with histone marks modification, typically trimethylation of lysine 27 and lysine 9 of histone H3 (H3K27me3 and H3K9me3). Polycomb repressive complex 2 (PRC2) catalyzes the deposition of H3K27me3, which leads to the recruitment of polycomb repressive complex 1 (PRC1) and triggers chromatin compaction. H3K27me3 and PRC2 mediate inheritance of repressed states. A loss of PRC1 and/or PRC2 leads to a failure of maintenance of the repressed states of PRC2 targets genes after several cell generations (Jadhav et al. 2020; Pengelly et al. 2013; Saksouk et al. 2015). In *Drosophila*, element bound by PRC2 (polycomb response element, PRE) are known to be important for repression maintenance (Steffen and Ringrose 2014; Coleman and Struhl 2017; Laprell et al. 2017). PREs have been shown to be critical for cell fate maintenance from embryo to larval stage, with the *bx1* PRE required to transfer positional information from embryonic enhancers to larval imaginal enhancers (Chan et al. 1994). Regarding H3K9me3, a similar mechanism has been observed, mediated by methyltransferase proteins such as SUV3-9. After its deposition, H3K9me3 is bound by the heterochromatin protein 1 (HP1) leading to chromatin compaction spreading and heterochromatin formation.

PcG proteins has been shown to not be excluded from mitotic chromatin (Follmer et al. 2012; Steffen et al. 2013) in S2 cells and embryos, indicating a potential maintenance of repressed chromatin states during mitosis.

Moreover, H3K27me₃ has been shown to be maternally inherited and present in early embryonic mitosis (Zenk et al. 2017) in the case of *Drosophila* and H3K9me₃ detected at the two cell stage of mammal development (Wang et al. 2018). This suggests that transgenerational inheritance (through meiosis¹¹) and mitotic inheritance might be linked and might involve similar mechanisms to ensure cell fate maintenance.

C.2.2. Active chromatin states

In contrast to repressed chromatin states, active states are typically maintained with an antagonist protein complex TrxG. These complexes catalyze trimethylation of H3K4 and are associated with histone acetylation marks (Schuettengruber et al. 2007). A recent study established that histones from active genes are not transmitted through cell divisions (Escobar et al. 2019). This is interesting because it suggests that active states are propagated through an active mechanism relying on the constant presence of regulatory factors such as specific TFs.

What could be the support of this memory of active states? To find out we need to have a look at the mitotic landscape. Here, I will mainly describe active states inheritance as my PhD project was focused on transcriptional activation through cell generations.

C.3. Mitotic bookmarking

Condensation of chromatin during mitosis represents a big challenge in chromatin organization in order to faithfully transmit the transcriptional program to daughter cells. This re-organization during mitosis has been proposed to coincide with a large eviction of chromatin associated proteins (Parsons and Spencer 1997). However, it is now well established that some factors are retained during mitosis (Raccaud et al. 2019; Festuccia et al. 2017; Teves et al. 2016) and therefore can index particular genomic regions to facilitate

¹¹ Cell division which gives rise to four daughter cells each with half the number of chromosomes of the parent cell, as in the production of gametes (ovules or spermatozooids).

transcriptional activation upon mitotic exit. This phenomenon is thought to be relevant for maintenance of cellular identity and is termed mitotic bookmarking (Michelotti et al. 1997). Bookmarking could be supported by several mechanisms (Figure 1.18 and more details in Annexe 4) and I will briefly summarize them in the next four paragraphs.

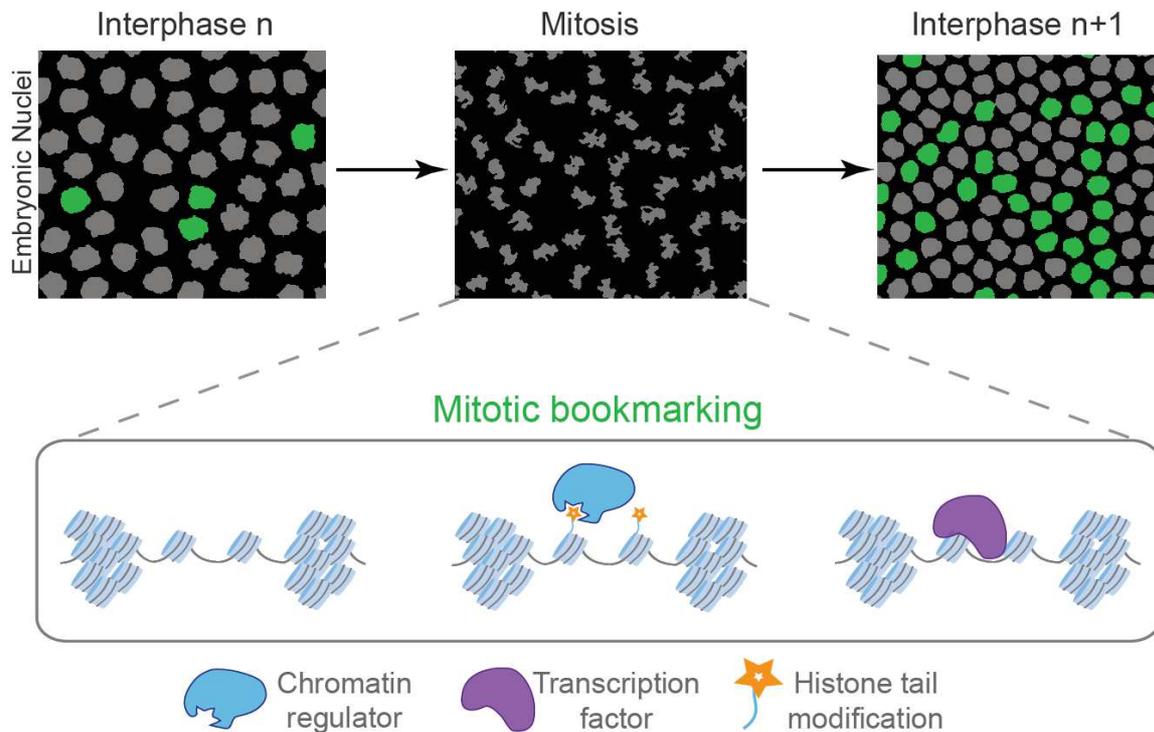


Figure 1. 18 : Potential mechanisms of mitotic inheritance.

Description of different mechanisms to ensure mitotic inheritance at specific gene. (Top) False colored nuclei of a *Drosophila* embryo in interphase and mitosis. Grey nuclei represents transcriptionally inactive nuclei and green transcriptionally active. (Bottom) Mitotic inheritance could happen via open chromatin states, chromatin regulators (blue), transcription factors (purple) and/or histone tail modification (orange).

C.3.1. Open loci during mitosis

To hypothesize about the potential mechanism of mitotic inheritance, one can imagine that the chromatin accessibility is maintained through mitosis and that this confers by itself the transmission of active chromatin states (Figure 1.19). Indeed, most of the open loci in interphase are also open, to a lesser extent, during mitosis. This is true in *Drosophila* (Blythe

and Wieschaus 2016) developing embryos but also in mammalian mouse embryonic stem cells (mESC) (Festuccia et al. 2019; Hsiung et al. 2015). However, in the case of the transgene containing *doc* enhancer and *snail* promoter, the act of a mother nucleus being active does not always lead to a faster transcriptional activation after mitosis (Dufourt et al. 2018). This means that it is not because the gene has been transcribed at a given nuclear cycle, that it will activate faster in the next one. Also, this accessibility during mitosis is likely true for specific developmental gene but also housekeeping gene which has been described for being prioritized over cell specific genes (Palozola et al. 2017). Interestingly, even if accessibility is maintained through mitosis, it has recently been reported that nucleosome positioning is rearranged during mitosis but not at bookmarked loci by Esrrb (Festuccia et al. 2019). Overall, these studies suggest that chromatin accessibility is indeed important for mitotic inheritance, but I would favor that this could be a result of the action of specific TFs.

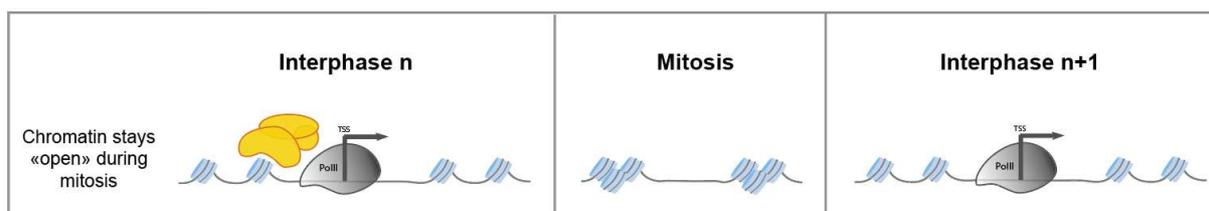


Figure 1. 19 : Mechanism of mitotic inheritance via open chromatin states.

C.3.2. Transcription factors in mitosis

At first, most of TFs were believed to be removed from mitotic chromatin (de Castro et al. 2016; Martínez-Balbás et al. 1995). Various recent studies reported a specific TF type which remain associated with mitotic chromatin, also called bookmarking factors (Festuccia et al. 2017; Elsherbiny and Dobrevá 2021) (Figure 1.20). However, papers reported that formaldehyde fixation can, in some cases, trigger the artifactual eviction of TFs from mitotic chromosome (Pallier et al. 2003; Teves et al. 2016). To circumvent this potential artifact, the use imaging in live cells seems a good complementary method. Whether the sum of fluorescent signal of a specific TF reflects its binding or not needs to be addressed with quantitative imaging techniques such as FRAP. Indeed, despite the technical artefact that fixation can create, it interrogates on the intrinsic nature of the TF binding to mitotic chromatin. Some TF, such as FoxA1 in liver cells can 'decorate' the entire mitotic

chromosome (Caravaca et al. 2013), whereas some such as the GAGA-associated factor or Esrrb only have a small fraction remaining (Raff et al. 1994; Raccaud et al. 2019) (Figure 1.20). This suggests that there are at least two types of mitotic binding, potentially reflecting a dynamic or a more stable binding. A mid-scale screen done in mESc on 501 TFs revealed that more than 70% have a substantial fraction on mitotic chromatin (Raccaud et al. 2019) (Figure 1.21), and that their recovery time in interphase correlates with the one in mitosis as assessed by FRAP. Interestingly, they found that the electrostatic properties of a given TF determine its mitotic bound fraction. This revealed that the intrinsic properties of TF are crucial determinants for its binding during mitosis. They and others (Teves et al. 2016) also showed that there is a wide variety of binding times of TFs during mitosis and interphase (Table 1.2). For instance, the TATA Binding Protein (TBP) has been shown to have a stable binding (residence time of 118s) (Teves et al. 2018) whereas Sox2 appears to be dynamic during mitosis (residence times of 4.3sec in mitosis and 19.7sec in interphase (Teves et al. 2016). In fact, until this recent study about TBP, TF binding appears to be quite dynamic (Elf et al. 2007; Swinstead et al. 2016; Chen et al. 2014b; Gebhardt et al. 2013; Mueller et al. 2008). What is the function of these different modes of binding during mitosis is a question that remains to be elucidated.

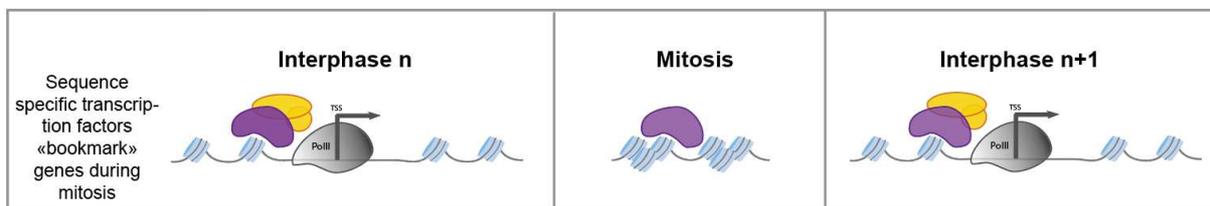


Figure 1. 20 : Mechanism of mitotic inheritance via mitotic retention of TFs.

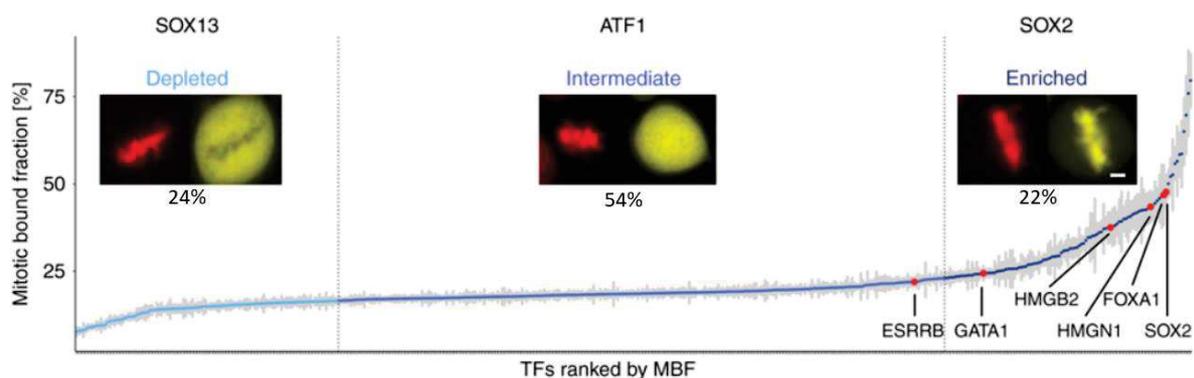


Figure 1. 21 : Mid-scale imaging screen in mESc reveals a high fraction of TF remaining on mitotic chromosome.

501 TFs ranked by their mitotic bound fraction (MBF) value which represent the ratio of the fluorescence of each TF on chromosome (red) compare to nucleoplasm. Figure adapted from (Raccaud et al. 2019).

PROTEIN	INTERPHASE	MITOSIS	REFERENCE	ORGANISM
TBP	RT 88s	118s	(Teves et al. 2018)	mES cells
BHLHB8	T1/2 : 10.82s	T1/2 : 13.48s	(Raccaud et al. 2019)	mES cells
CDX2	T1/2 : 14.03s	T1/2 : 7.8s	(Raccaud et al. 2019)	mES cells
FOXA1	T1/2 : 6.47s	T1/2 : 3.87s	(Raccaud et al. 2019)	mES cells
POU5F1	T1/2 : 5.33s	T1/2 : 4.7s	(Raccaud et al. 2019)	mES cells
SOX2	T1/2 : 7.68s	T1/2 : 9.97s	(Raccaud et al. 2019)	mES cells
MAX	T1/2 : 16.21s	T1/2 : 30.86s	(Raccaud et al. 2019)	mES cells
TEAD1	T1/2 : 8.51s	T1/2 : 4.6s	(Raccaud et al. 2019)	mES cells
SOX15	T1/2 : 2.92s	T1/2 : 2.03s	(Raccaud et al. 2019)	mES cells
SOX2	RT : 19.7s	RT : 4.3s	(Teves et al. 2016)	mES cells
RUNX2	T1/2 : 12s	T1/2 : 7-20s	(Pockwinse et al. 2011)	U2OS cells

Table 1. 2 : List of TF binding kinetics which have been measured in interphase and mitosis.

RT: residence time. T_{1/2}: the halftime of recovery is the time from the bleach to the time point where the fluorescence intensity reaches the half (1/2) of the final recovered intensity. U₂OS: human osteosarcoma cells.

C.3.3. Chromatin associated factors

Another class of factors could play a role in mitotic inheritance: the chromatin associated factors. These factors don't directly bind to DNA but can recognize histone mark modification through specific domain (e.g. bromodomain for acetylation recognition), also called 'readers' and/or can modify histone tail and be called 'writers' (through SET domain for methyltransferase) (Zhang et al. 2015). Chromatin remodelers can also be assigned to this class, such as the SWI/SNF family, as they are known to play crucial role in gene regulation by displacing nucleosome therefore allowing accessibility or not of the DNA (Clapier et al. 2017).

There are few examples of chromatin-associated factors that are present in mitosis and could represent a mechanism of mitotic bookmarking (Figure 1.22). In *Drosophila*, the TrxG member Ash1, a 'writer' of the H₃K₃₆ac mark, has been shown to decorate mitotic chromosome in the early embryo. This was shown with live imaging of a Ash1-GFP transgene (Steffen et al. 2013). Another example is Brd4 which has been shown to bookmark specific promoters of genes in mouse embryonic fibroblast cells using fixed and live imaging (Dey et al. 2009) and present in mitosis in zebrafish embryos (Toyama et al. 2008). In the case of chromatin remodelers, to my knowledge there is no clear evidence for their retention during mitosis. SWI/SNF has been shown to be inactivated by phosphorylation during mitosis of human cell line (Stukenberg et al. 1998) but has been shown in yeast that the same proteins are present in late mitosis (Krebs et al. 2000). However, little is known about the recruitment of such factors during mitosis and particularly *in vivo*.

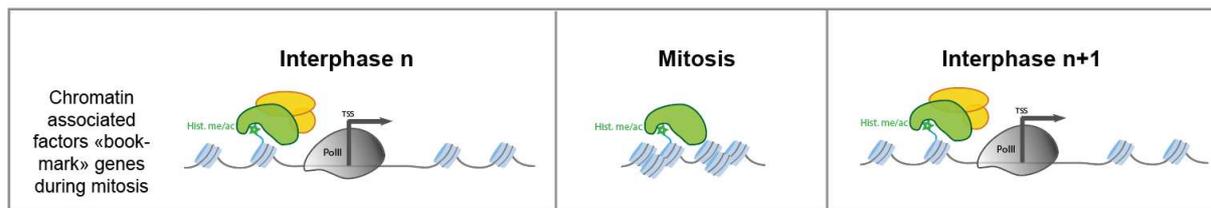


Figure 1. 22 : Mechanism of mitotic inheritance via mitotic retention of chromatin associated factors.

C.3.4. Histone marks

The role of Brd4 as a bona fide mitotic bookmarker has recently been challenged. In the study of *Behera et. al*, mitotic depletion of Brd4 does not lead to a delay of transcriptional activation in G₁ in those studied cells with their timing of analysis, but its associated histone marks, H₃K₁₄ac, H₃K₂₇ac, H₃K₁₂₂ac, and H₄K₁₆ac, are broadly preserved in mitosis and predicts post-mitotic transcription activation patterns (Behera et al. 2019). The fact that bookmarking might not relate to the presence of TF during mitosis interrogates on the role of histone modification as bookmarkers and transmitters of information from mother to daughter cells. Histones undergo many post-translational modifications on multiple residues, including phosphorylation, methylation, acetylation, and ubiquitination (Kouzarides 2007). In general, methylation of histones changes much less than other marks most probably due to its

chemical stability (Wang and Higgins 2013; Byvoet et al. 1972). However, this mark has also been shown to be reversible by the action of histone demethylases (Borun et al. 1972; Bannister et al. 2002). Nevertheless, certain methylation and acetylation has been shown to remain at specific sites during mitosis such as H3K27ac, H3K4me3 and H4K5ac (Kouskouti and Talianidis 2005; Valls et al. 2005; Muramoto et al. 2010; Zhao et al. 2011; Javasky et al. 2018), in mammalian cells or *Dictyostellium*. In embryos, there are few examples of mitotic retention of histone marks, as H3K4me3 in *Drosophila* and *Xenopus* embryos (Black et al. 2016; Hörmanseder et al. 2017) and H4 acetylation in zebrafish embryos (Toyama et al. 2008). Finally, recent studies have shown that the active genes associated H4K16ac mark and repressed genes associated H3K27me3 mark are present in early *Drosophila* mitotic cycles but are also transmitted transgenerationally (Samata et al. 2020; Zenk et al. 2017). This suggests a mechanism, where histone marks would bookmark specific loci to recruit chromatin regulators faster after mitosis (Figure 1.23). Whether this is a general rule for chromatin regulator bookmarking remain to be demonstrated.

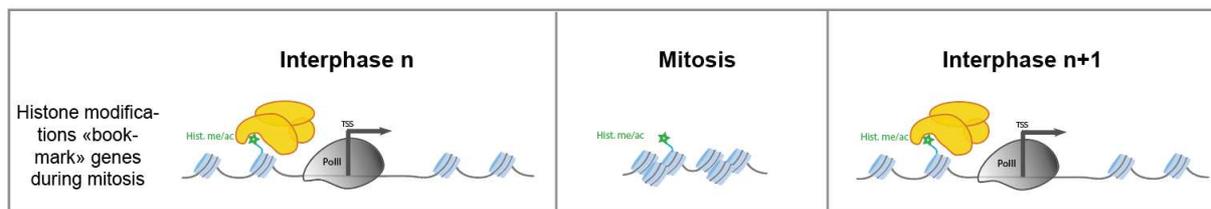


Figure 1. 23 : Mechanism of mitotic inheritance via mitotic retention of histone tail modifications.

Altogether, these different mechanisms have clearly been shown to be at play for mitotic inheritance, but there are not mutually exclusive. Indeed, we can suppose that a combination of these different mechanisms might be in place for specific transcription programs governing cell fate.

C.4. Mitotic memory: towards a faster transcription re-activation after mitosis?

C.4.1. Monitoring transcription activation after mitosis

There are two ways to measure post mitosis transcriptional timing. One is to quantify the timing of activation after mitosis of all the cells in the pattern of the expression of a given

gene or a whole transcriptional program as mostly done in cultured cells (Teves et al. 2018; Palozola et al. 2017; Festuccia et al. 2016; Zhang et al. 2019a). These studies have been using genome-wide approaches such as nascent chromatin associated RNA-seq (Teves et al. 2018), EU-RNA-Seq (Palozola et al. 2017), RNA-polIII ChIP-seq (Zhang et al. 2019a) or gene targeted approach such as RT-qPCR (Kadauke et al. 2012). A second way is to take into account the past of the quantified cells by looking at the kinetics of reactivation after mitosis (Muramoto et al. 2010; Ferraro et al. 2016; Dufourt et al. 2018; Zhao et al. 2011). This requires the capacity to quantify the precise timing of transcriptional activation in live cells, which is allowed by labeling systems such as the MS2/MS2-Coat-Protein (MCP) method (Figure 1.24) (Bertrand et al. 1998; Pichon et al. 2018). This approach is a bipartite system: the MS2 loops will form upon transcription and will be recognized by the MCP fused to a fluorescent protein.

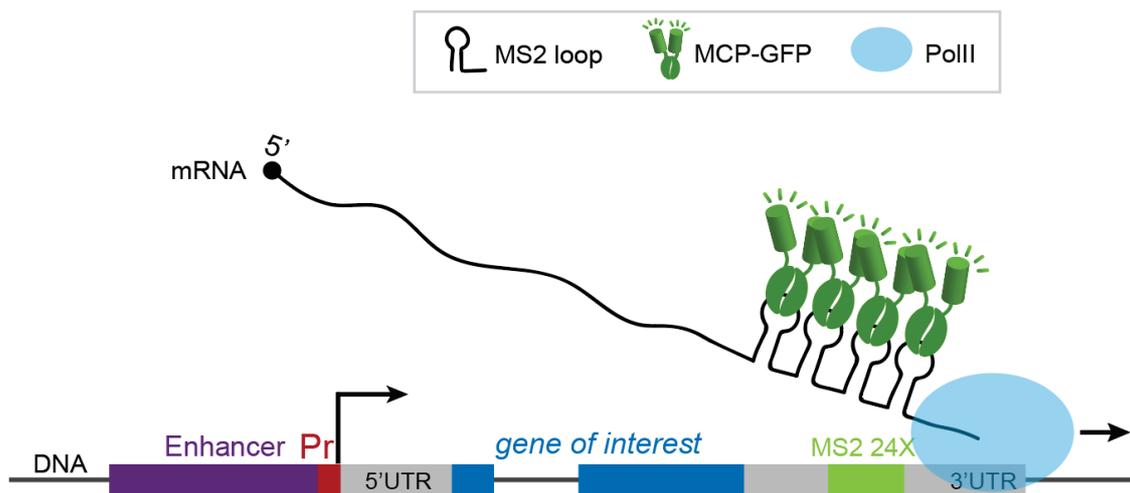


Figure 1. 24 : MS2/MCP system to visualize transcription in live.

Scheme of the MS2/MCP system. Transcription of the gene of interest is under the control of an enhancer (purple) and a promoter (red). Upon transcription by the RNA polymerase II (blue, Pol II), the MS2 loops will be recognized by the MS2 Coat Protein fused to a GFP creating a fluorescent signal above the background.

C.4.2. Transcription is a multi-step process

As transcription is shut down during mitosis and since general TFs (GTFs) are mostly evicted from mitosis (with the exception of TBP (Teves et al. 2018)), transcription re-activation must be re-established after the chromatin condensation of mitosis. Transcription is a multi-step

process, requiring the assembly of several protein complexes at promoters (Cramer 2019). GTFs are sequentially recruited to constitute the pre-initiation complex (PIC) and is followed by the recruitment of RNA Polymerase II (Pol II) at the transcriptional start site (TSS) of the promoter (Gupta et al. 2016) (Figure 1.25). Each of these steps is highly controlled and represents a critical checkpoint for transcription activation after mitosis. Therefore, mitotic bookmarking would represent a layer of regulation which can 'bypass' certain step therefore providing an 'advantage' to the bookmarked loci (Bellec et al. 2018) (Figure 1.26). This could be via reducing the number of steps after mitosis or by accelerating these steps after mitosis or both (Figure 1.26). Modeling these steps of transcription activation led to the finding that one transcription factor (Zelda) accelerates transcription after mitosis. However, Zelda is not retained in mitosis and accelerates the steps but whatever the past transcriptional status of the gene (Dufourt et al. 2018). Thus, the exact role of mitotic bookmarking on steps of transcriptional activation remains to be elucidated.

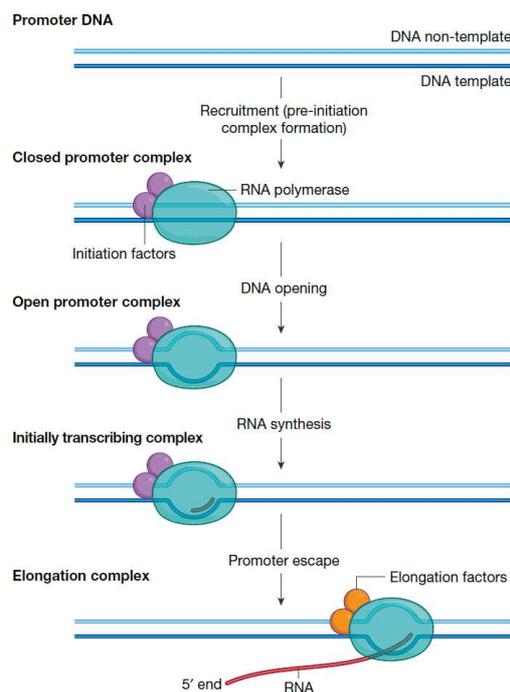


Figure 1. 25 : Steps of transcription.

Schematic of the key steps of transcription with recruitment of the pre-initiation complex to the promoter region of a gene, followed by the recruitment of the RNA polymerase II, opening of the double-strand DNA and RNA synthesis. Figure from (Cramer 2019).

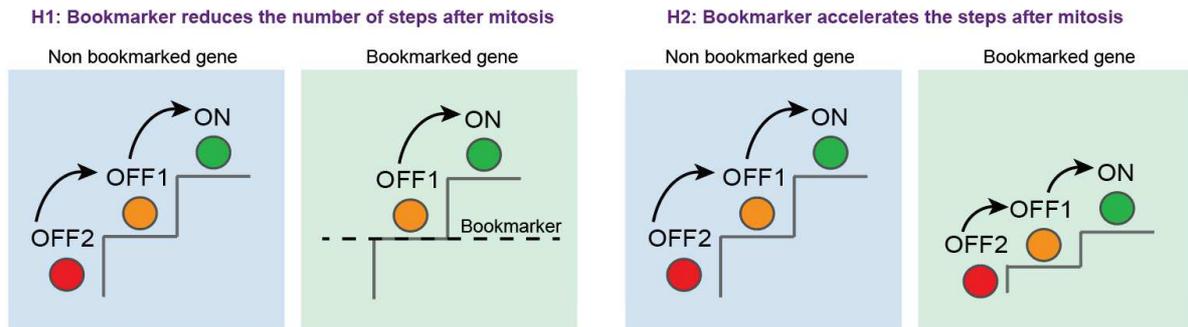


Figure 1.26 : Hypothesis of bookmarking role on post-mitotic transcription activation steps. Dots represents transcriptional states (OFF, inactive or ON for activation). OFF2 is the more 'profound' state of inactivation. Two hypothesis are represented.

C.4.3. Bookmarking as a facilitator for post mitotic re-activation of transcription

Despite numerous examples demonstrating the retention of transcription factors on mitotic chromatin (see section C.3.2.), whether this retention leads to a functional advantage in terms of transcription remains unclear. In fact, a mitotically retained factor could be considered as a bookmarking factor only if its retention has a functional consequence in terms of speed of post-mitotic transcriptional activation. So far, only three studies have been able to identify a role for post-mitotic transcriptional re-activation. Brd4 was the first to be studied and shown to slow down post-mitotic transcription activation when inhibited with JQ1 drug in cultured cells (Zhao et al. 2011). Then GATA1 was shown to be also involved in transcription activation after mitosis (Kadauke et al. 2012) although it might be case specific as GATA1 can act as a repressor (Hsiung et al. 2015). Finally, it is also the case for the general TF TBP which have been depleted using an Auxin-Degron system (Teves et al. 2018). All those studies have shown a role from bookmarking in post-mitotic transcription activation but little is known during development.

The discovery of a transcriptional memory bias in a developing organism, was done using the MS2/MCP system. With this approach it was possible to take into account the past state of the cells, whether it comes from active nuclei or from inactive nuclei (Ferraro et al. 2016) (Figure 1.27). In the lab where I did my PhD, they found that nuclei coming from active nuclei activate faster than the one coming from inactive nuclei on a transgene composed of the *snail* promoter and *snail* core enhancer. However, such a memory bias has never been shown on an endogenous gene.

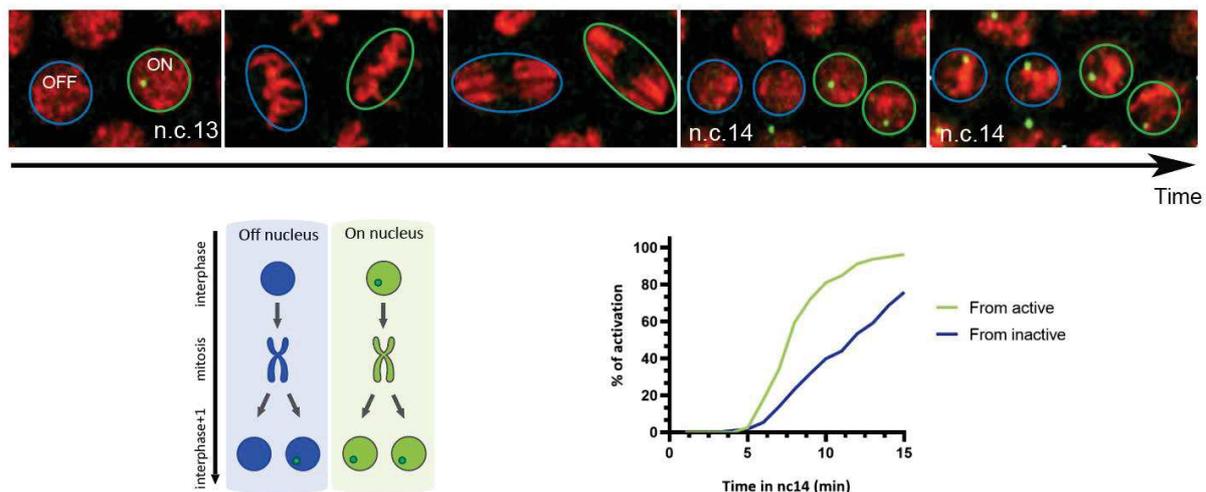


Figure 1. 27 : Scheme of transcriptional memory.

(Top) Screen shot from live imaging of a *Drosophila* embryo expressing *H2Av-RFP* (DNA staining) and *MS2/MCP* representing gene activation (green dots). (Bottom left) Scheme of the two population of nuclei obtained: coming from active (green) and from inactive (blue). Activated nuclei are represented with a green dot inside which schematize RNA detection of a specific gene. (Bottom right) Percentage of activation of nuclei coming from active nuclei (green) and nuclei coming from inactive (blue).

C.5. Mitotic memory: towards a stronger transcription re-activation after mitosis?

Experiments such as live imaging with *MS2/MCP* reporters allow to also measuring the transcriptional activity after mitosis. But very few studies have investigated transcriptional memory in terms of precise level of transcription and protein output. *Muramoto et al.* have shown that sister and mother-daughter *Dictyostellium* cells correlates in frequency of transcriptional firing (Muramoto et al. 2010). In *Zhao et al.* study, the authors showed that *Brd4* mitotic target genes have a faster post-mitotic activation in cultured cells (Zhao et al. 2011). In mammals, few studies looked at gene expression memory at the protein level. In lung carcinoma cells, levels of 20 proteins were measured using fluorescent tagging over cell generation and found to be inherited for few cell generation (Sigal et al. 2006). This phenomenon was also observed in mESCs, where *Nanog* fluorescent protein level was measured and those levels are found to be inherited in subpopulation of cells (Filipczyk et al. 2015). However, those studies rely on protein levels therefore being the result of mRNA and

protein half-lives and are hard to decouple. With the step forward described in Chapter 2, it is now possible to link both mRNA as well as protein levels at the single molecule level.

Little is known about the transcriptional memory in terms of instantaneous activity in phenotypically identical cells. A recent study from *Phillips et al.* investigated the transcriptional fluctuation across cell divisions of a reporter transgene (Phillips et al. 2019). Remarkably, they showed that transcriptional dynamics of sister cells highly correlates, suggesting a similarity of factors contents in the sister cells probably inherited from the mother cell. They also showed that the mean transcriptional activity is correlated between mother and daughter cells. This means that the mother cell can transmit the overall transcriptional output to the daughter cells, whether this is regulated in *cis* and/or *trans* remain to be elucidated.

D. Early *Drosophila* embryogenesis

Early hours of development have been extensively studied in *Drosophila*. I will summarize in this section the specific features of early *Drosophila* development that make it an ideal model of choice to study gene expression regulation of transcription during early embryogenesis in a developing multicellular organism.

D.1. Synchronous and fast first divisions

Drosophila embryogenesis starts with a large egg where the maternal-paternal fused nuclei will divide 13 times in a fast and synchronous manner (Farrell and O'Farrell 2014a). The particularity of this embryonic development is that the earliest nuclear cycles are devoid of gap phases between replication and mitosis. The nuclei alternate between mitosis and S phase (replication) in a very short amount of time. Indeed, the first 14 nuclear cycles occurs in only 1h30, making on average one division every 8.6 minutes (Foe and Alberts 1983; Rabinowitz 1941) (Figure 1.28).

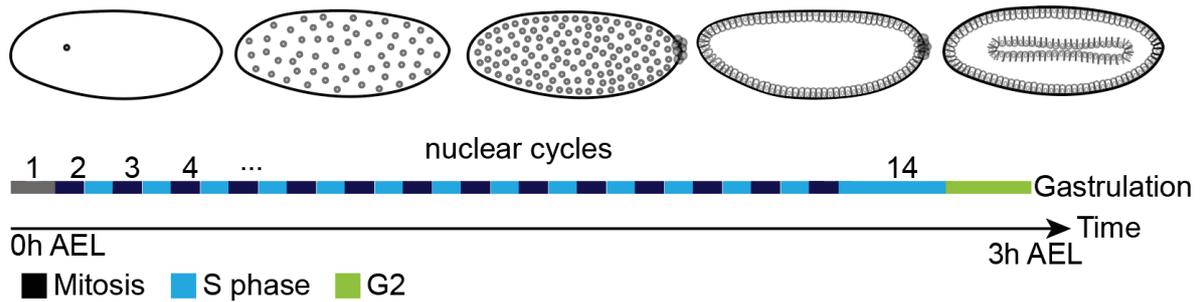


Figure 1. 28 : First hours of *Drosophila* development.

*Scheme of the first 14 nuclear cycles of *Drosophila* embryogenesis with morphological stages illustrated at the top until the first morphological event: gastrulation. Embryos are oriented with anterior on the left and posterior on the right. Each phase of the nuclear cycle is: S phase (light blue), mitosis (dark blue), and G2 (green). The nuclear cycle 14 is the longest and the first to have a G2 phase. AEL: After Egg Laying. Pole cells arise at n.c.8 at the posterior pole of the embryo. These cells will give the future gonads of the fly.*

From the nuclear cycle 8 (n.c.8), the nuclei start to migrate towards the periphery of the egg, near to the plasma and vitelline membrane forming an alignment of nuclei close to the surface of the egg from n.c.10 to n.c.14 also called the blastoderm stage (Farrell and O’Farrell 2014a). Before the n.c.14, the nuclei divide without any membrane between cells, in a common cytoplasm called syncytium. During this period of time, the nuclei will divide synchronously thanks to the periodic action of Cdk1 (Deneke et al. 2016; Fasulo et al. 2012; Royou et al. 2008; Price et al. 2000; Stumpff et al. 2004) (Figure 1.29).

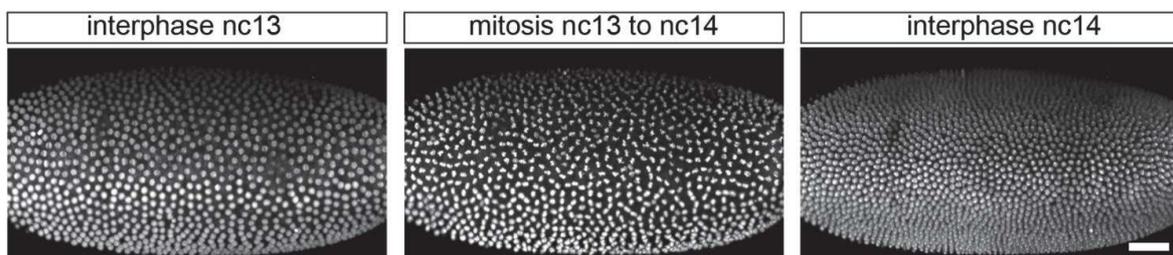


Figure 1. 29 : Synchronized nuclear cycles of a *Drosophila* embryo.

Snap shot images from a H2A-RFP embryo from MuviSpim movie (done by Mathieu Dejean and Sylvain DeRossi), nuclei are in interphase (left panel) or metaphase (middle panel). The embryos are oriented with ventral side in the middle. Scale bar is 50µm.

During the n.c.14, which is longer than the other cycles (45min), the membrane starts to invaginate from the plasma membrane (apical side of nuclei) towards the inside of the embryo (basal side of the embryo) (Figure 1.30). This cellularisation is not linear in time as the last phase (when cellular membrane closes around each cell) is fast compare to the beginning of the process (Falo-Sanjuan and Bray 2021). Following the membrane invagination, nuclei elongate and the genome get more and more organized. Heterochromatin is localized at the apical part of the nuclei (Figure 1.30) (Yuan and O'Farrell 2016).

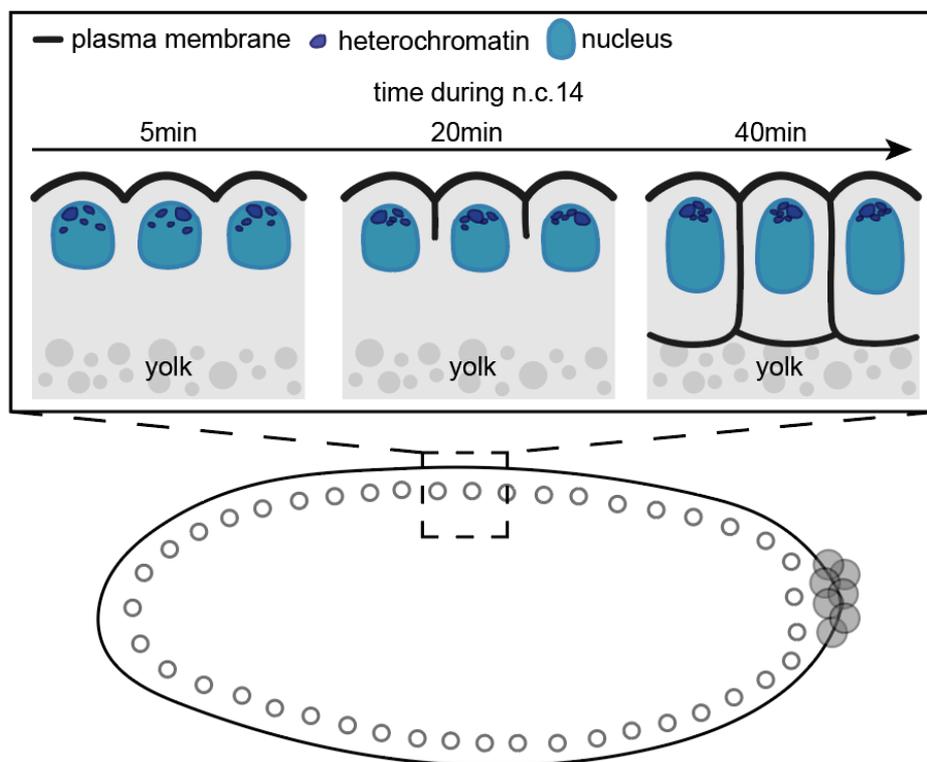


Figure 1.30 : The process of cellularisation during the n.c.14 of *Drosophila* embryogenesis. Scheme of a lateral view of an embryo with square representing the region zoomed in the upper panel.

D.2. Maternal support of early embryogenesis

The fecundated egg contains maternal mRNA and proteins from thousands of genes representing three quarters of the protein-coding transcriptome (Tadros et al. 2007; Thomsen et al. 2010). This maternal support will ensure the correct development of the egg until the zygotic genome starts to be expressed (Figure 1.31, see D.3.). The maternal

deposition comes from oogenesis where the nurse cells (cells located around the oocyte which are polyploid cells providing mRNAs and proteins to the egg) pour out into the oocyte before being fecundated and laid (Spradling 1993). Among those proteins, morphogens such as Bicoid or Nanos are specifically localized in the egg and give rise to gradients which provides a spatial information required for body axes development (Huang et al. 2017; Gurdon and Bourillot 2001; Neumann and Cohen 1997). This maternal deposition is conserved among metazoan and is followed by a clearance of all maternal mRNAs until the zygotic genome get fully activated (Vastenhouw et al. 2019).

D.3. The awakening of the zygotic genome

Concomitantly to maternal products clearance, the zygotic genome starts to be expressed at around 1-1h30 after egg-laying and this process is called the Maternal to Zygotic Transition (MZT) (Figure 1.31). The zygotic expression was first believed to occur in two waves (Tadros and Lipshitz 2009; Chen et al. 2013) but this was revisited with the recent advances and deeper understanding of the zygotic genome activation (ZGA). The current view is that MZT arises gradually rather than via specific waves (Vastenhouw et al. 2019).

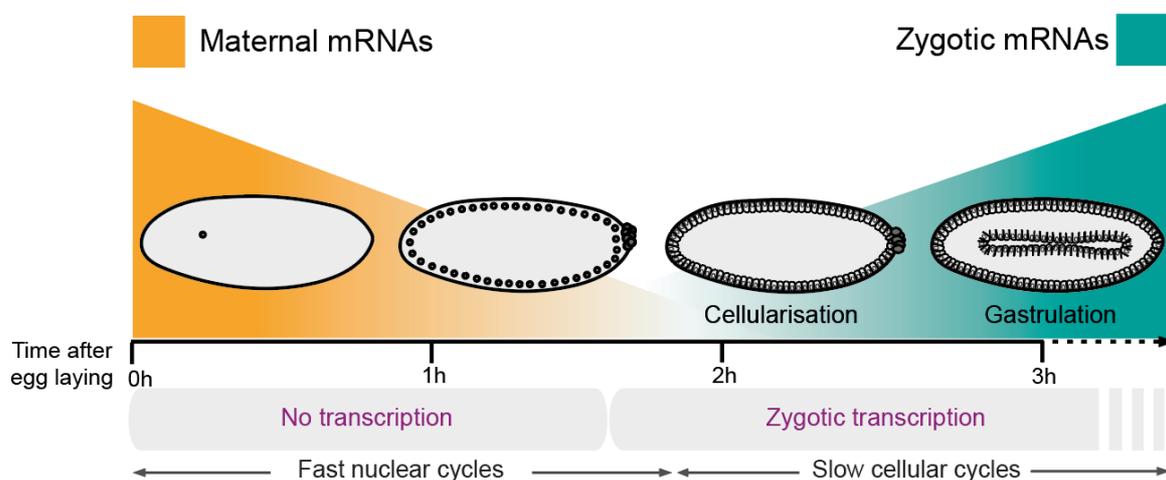


Figure 1. 31 : Scheme of the *Drosophila* Maternal to Zygotic transition (MZT).

Scheme of the MZT. Maternal mRNAs and protein (orange) ensure the development of the embryo until zygotic transcription starts (blue), and this is correlated with longer nuclear cycles. The first morphogenetic event is the gastrulation at around 3hours after egg laying. Figure is from Dufourt et al. 2019.

The first expressed genes are under the control of maternally deposited TF (morphogen, Figure 1.32). They are deposited asymmetrically and lead to the activation of gap genes which themselves activates pair-rule genes in a spatially regulated manner. This generates patterns of gene expression and the formation of the germ layers (Figure 1.33). From these two patterning networks (antero-posterior and dorso-ventral) the embryo is polarized, the cells get specific identities and can develop properly.

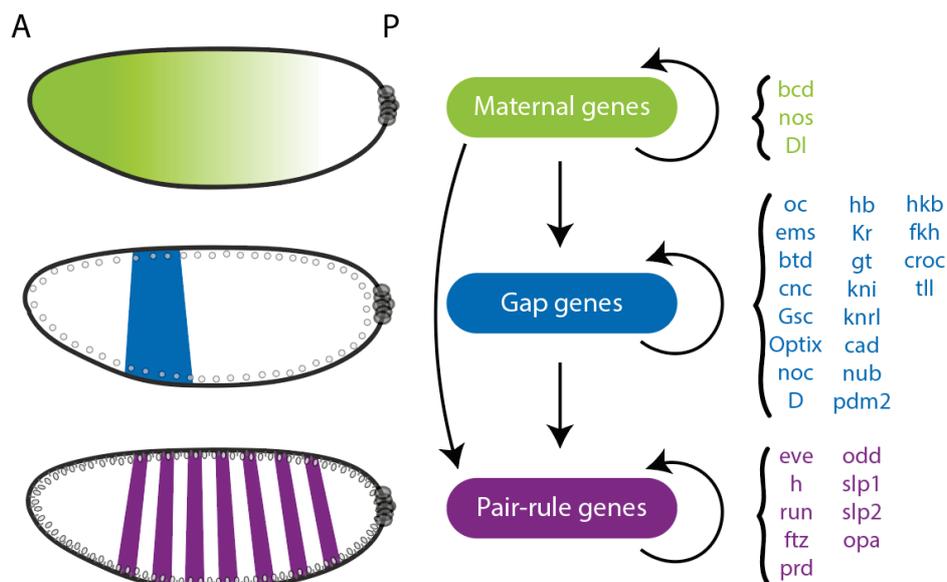


Figure 1. 32 : Cascade of gene activation for antero-posterior patterning in the early *Drosophila* embryo.

*Scheme of pattern of expression of three main classes (maternal in green, gap genes in blue and pair-rule in purple) of genes expressed in the early *Drosophila* embryo. A non-exhaustive list of genes for each class is provided. Embryos are oriented with the anterior to the left (A) and posterior to the right (P).*

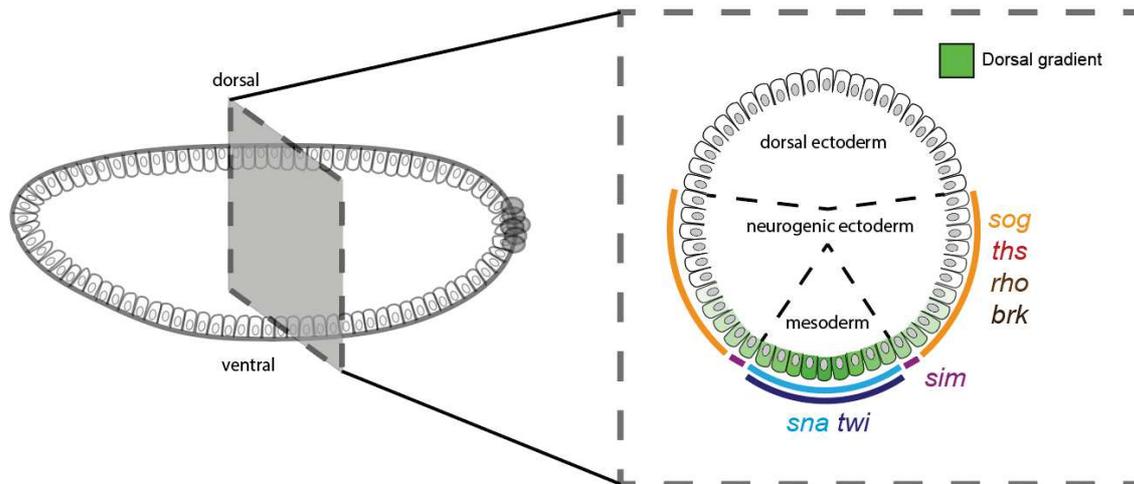


Figure 1. 33 : The formation of germ layers from dorso-ventral patterning in the *Drosophila* embryo.

Scheme of a lateral and transversal view of an embryo with the indicated germ layers. Examples of genes expressed in specific regions of the embryo are indicated.

D.3.1. Major factors at play during ZGA

During ZGA, there is a considerable reshaping of chromatin. Some developmental genes are primed for rapid activation by the modification of their *cis*-regulatory sequences. A particular TF, called Zelda, has the capacity to prepare most of the zygotic genes for transcriptional activation (McDaniel et al. 2019; Schulz et al. 2015a; Harrison et al. 2011). Another factor, called GAGA Associated Factor (GAF) has been shown to be required later for ZGA (Gaskill et al. 2021; Bhat et al. 1996; Moshe and Kaplan 2017). These two factors are crucial to ensure proper ZGA (Liang et al. 2008; Sun et al. 2015; Gaskill et al. 2021; Moshe and Kaplan 2017; Schulz et al. 2015b) and both of them have pioneer factors properties. Zelda has clearly been shown to be able to bind directly nucleosome and be important for transcription (Sun et al. 2015; McDaniel et al. 2019). It is more delicate to qualify GAF as a pioneer factor because one need to test its direct binding to nucleosome *in vitro*. Nevertheless, it has pioneer-like properties such as the ability to open chromatin on naked DNA and able to activate gene expression by recruiting the transcriptional machinery in *Drosophila* S2 cells (Fuda et al. 2015).

D.3.2. Genome organization during ZGA

Another event that occurs concomitantly with the ZGA is the chromatin organization. Indeed, a chromatin architecture starts to emerge clearly at n.c. 12 as shown with Hi-C method¹² (Hug et al. 2017a) Figure 1.34.

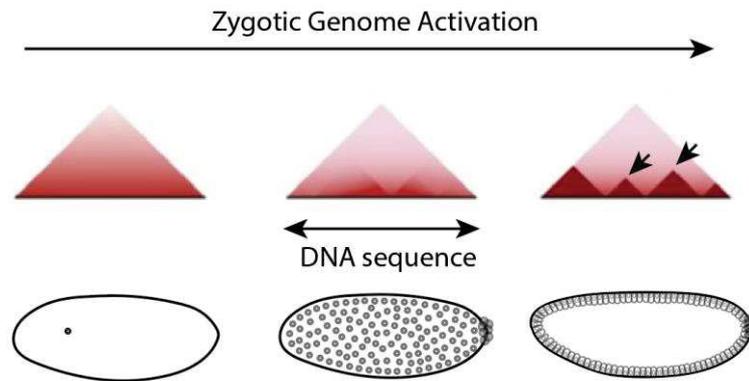


Figure 1. 34 : Genome organization during *Drosophila* ZGA.

Triangles represents HiC contact maps on a specific region of the genome (adapted from hug et al. 2017). Darker red means high probability of contact (black arrowheads).

This chromatin interaction corresponds to cis-regulatory regions that interact together in order to activate or repress genes (Bickmore 2013). The common view is that enhancers and promoters physically get close to regulate transcription activation in a tissue specific manner (Schwarzer and Spitz 2014). But this has been recently challenged by few studies which shows activation of transcription without cis-regulatory element proximity (Benabdallah et al. 2019; Alexander et al. 2019).

During *Drosophila* development and prior to the ZGA, the genome is mostly unstructured. By the time the zygotic genome gets activated, 3D structure starts to emerge with specific DNA sequences getting in contact, forming what is called Topologically Associated Domains (TADs) (Hug et al. 2017a; Ogiyama et al. 2018; Cardozo Gizzi et al. 2019). The stability of these domains is still questioned, but interestingly enough some contact between cis-regulatory

¹² Proximity ligation method that capture the three-dimensional (3D) organizational structure.

regions have been shown to take place before gene transcription activation (Ghavi-Helm et al. 2014).

The study from *Hug et al.* showed that genomic region boundaries in proximity depends on the presence of Zelda (Hug et al. 2017a). Moreover, in *Dufourt et al.* they propose that Zelda proteins organize in 'hubs' (see section B.1.3.) (Dufourt et al. 2018). Collectively, these studies suppose that Zelda 'hubs' could allow physical closeness of regulatory modules bound by Zelda. This is also be discussed in the publication in Annexe 3. Furthermore, GAF has been shown to be able to bring to regulatory element in proximity both in vitro and in cultured cells (Mahmoudi et al. 2002). Mutating GAGA sites within a Polycomb Response Element (PRE) reduces the frequency of looping interaction between PREs (Ogiyama et al. 2018) even if it is not clear is this effect is indirect *via* polycomb group proteins recruitments. Altogether, this provides a hint for this factor to be also involved in 3D genome organization during the ZGA of the *Drosophila* embryo.

E. Mitotic bookmarking during *Drosophila* embryogenesis

Little is known about mitotic bookmarking in a developing embryo. Despite the presence of histone marks during the early cleavages of embryogenesis (Samata et al. 2020; Toyama et al. 2008) and few examples of TFs such as Esrrb decorating mitotic DNA in mouse embryos or Ash1 on mitotic chromosomes of *Drosophila* embryos (Festuccia et al. 2016; Steffen et al. 2013).

In *Drosophila*, there are putative candidates for mitotic bookmarking in the early embryo. I will introduce the few candidates we had when I started my PhD project, and the reasons why we thought it was important to consider them.

E.1. *Female sterile (1) homeotic (Fs(1)h)*

Female sterile (1) homeotic (fs(1)h) gene encodes the dBrd4 protein (*Drosophila* Bromodomain-containing protein 4), a Bromodomain and Extra-Terminal domain (BET) family of protein. dBrd4 is characterized by the presence of two tandem bromodomains (Figure 1.35) which allow the recognition of histone acetylation H3K14, H4K5, or H4K12 (Chiang 2009; Dey et al. 2003). *Fs(1)h* is conserved among vertebrates, called *Brd4*, and is

known to be implicated in a number of cancers and autoimmune diseases (Devaiah et al. 2016b).

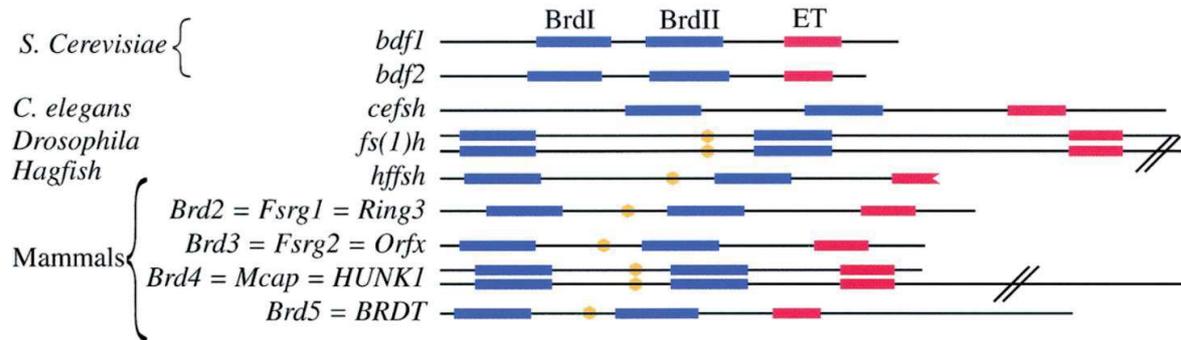


Figure 1. 35 : Conservation of the Bromodomain protein family.

Figure is from (Houzelstein et al. 2002). Scheme of the different domain organization of Brd protein family. Bromodomains are shown in blue and Extra-Terminal domain in pink. Yellow dots represent a highly conserved 10 amino-acid motif which function remains unknown. Histone acetyltransferase (HAT) domain is not represented but is located between the BrdII and ET domains.

This protein is a part of the TrxG protein family therefore an ideal candidate for cell fate maintenance as described in C.2.2. As Brd₄ has been shown to be retained on mitotic chromosomes in zebrafish embryos (Toyama et al. 2008) we can suspect that this is the case in *Drosophila* as well. One of the role of Brd₄ in normal condition (i.e. not in cancer cell lines) was described by Dey et al. who has shown that Brd₄ bookmarks specific genes for transcriptional memory, and upon exiting mitosis, it acts as a signal for rapid transcription in their progeny cells (Dey et al. 2009). Few years later, another study showed that mitotic inhibition of Brd₄ with JQ1 inhibitor delays post-mitotic transcriptional activation (Zhao et al. 2011). As a potential mechanism, Brd₄ has been shown to also acetylate histone tails, through its histone acetyltransferase domain (HAT), of H₃K122 and is able to evict nucleosome (Devaiah et al. 2016a), and this acetyltransferase activity seems to be enhanced or dependent on co-factors such as P300 or CBP to acetylate H₃K27 and H₃K56 (Wu et al. 2018). Brd₄ is also capable of recruiting the transcription elongation factor PTEFb, therefore playing a role in activation and productive transcription (Yang et al. 2005).

Brd4 has been shown regulate transcription notably by binding to super-enhancers (Sabari et al. 2018; Hnisz et al. 2017) and allow contact between enhancer and promoter in cancer (Chapuy et al. 2013; Lovén et al. 2013). It has been proposed that Brd4 could form 'hubs' together with super-enhancers that would favor gene transcription (Sabari et al. 2018).

In *Drosophila*, dBrd4 is maternally deposited and *fs(1)h* mutant is lethal. As its name indicates, null alleles lead to female sterility. Its expression is mostly concentrated in female ovaries, in the early embryo, and larvae nervous system. dBrd4 has two isoforms both expressed in the early embryo (Haynes et al. 1989; Digan et al. 1986) (Figure 1.36).

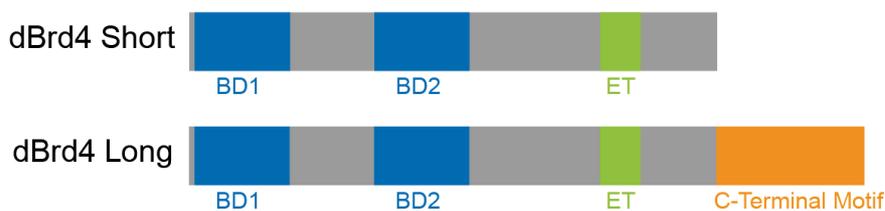


Figure 1. 36 : Scheme of the different isoforms for dBrd4 in *Drosophila*.

Each colored box represents a specific domain of the protein. Bromodomains are in blue (BD) and serve to bind to acetylated histones, extra-terminal domain (ET) in green and carboxy-terminal motif of the long isoform in orange.

E.2. Absent, Small, or Homeotic discs 1 protein (Ash1)

dBrd4 has diverse interactors and partners (such as GAF, CP190, BEAF...) as members of TrxG proteins. One of those members, the histone methyl transferase Ash1 has been shown to interact with dBrd4 physically and genetically (Kockmann et al. 2013) suggesting that they potentially cooperate.

Ash1 catalyzes the dimethylation of H3K36 (Tanaka et al. 2007; Dorigi and Tamkun 2013; Gregory et al. 2007) via its SET domain but is not required for fly survival (Schmähling et al. 2018). This factor targets, among others, *Hox* genes which leads to homeotic transformation¹³ in *ash1* mutants (Tripoulas et al. 1994; Schmähling et al. 2018; Shearn 1989).

¹³ Homeotic genes direct the development of particular body segments or structures during development of any organism. Homeotic transformation will lead to a change in body plan organization.

Ash1 is maternally deposited and known to remain on mitotic chromatin in the early embryo (Steffen et al. 2013) visualized with fused Ash1-GFP transgene. Recently, a study even identified a combination of the Ash1 protein domains required for mitotic retention in the embryo (Steffen et al. 2021). To my knowledge, it is the first time that a domain was identified to be required for mitotic retention only.

Collectively, these features place Ash1 as also a potential mitotic bookmark interacting with dBrd4.

E.3. GAGA associated factor (GAF)

As mentioned in D.3.1., Zelda and GAF are crucial to ensure the proper expression of zygotic genes during the first hours of development. It has previously been reported by the lab that Zelda, using fixed and live imaging, does not stay during mitosis (Dufourt et al. 2018). On the contrary, GAF is partly evicted but remains during mitosis (Raff et al. 1994; Dufourt et al. 2018; Gaskill et al. 2021; Kellum et al. 1995). GAF is encoded by the *Trl* gene and has two isoforms, only the short one being expressed at the early stages (Soeller et al. 1993) (Figure 1.37).

GAF-519 short isoform



GAF-581 long isoform



Figure 1. 37 : Scheme of the different isoforms for GAF.

Each colored box represents a specific domain of the protein. The BTB/POZ domain mediates homomeric dimerization multimerization (Wilkins and Lis 1999). The zinc-finger domain is the DNA-binding domain of the protein.

GAF is known to have a diverse set of functions including the activation and silencing of gene expression, nucleosome organization and remodeling, higher order chromosome architecture and mitosis (Lomaev et al. 2017; Ogiyama et al. 2018; Tsai et al. 2016; Granok et al. 1995). GAF has been described to be necessary for oogenesis (Fedorova et al. 2019) as *Trl*

female null allele mutant don't lay eggs. During early development, depletion of GAF leads to important defects of mitosis (Gaskill et al. 2021; Bhat et al. 1996 and Chapter2) and most of the embryos don't hatch. Moreover, GAF has been shown to play important role in chromatin accessibility and recruitment of PolIII (Gaskill et al. 2021; Fuda et al. 2015; Judd et al. 2021). Taken together, GAF is a good candidate as a mitotic bookmaker in the early *Drosophila* embryo.

Aims

The first part of my PhD project aimed at understanding how is mitotic memory established in the early *Drosophila* embryo. To tackle this question, we subdivided the project in three main objectives:

- 1) Identify a good candidate for mitotic bookmarking in the early embryo

To do this, we employed fixed and live imaging of proteins and look at their distributions during the cell cycle.

- 2) Describe which loci are bound by bookmarking factors in mitotic nuclei in embryos

We therefore wanted to develop a technique to sort interphase and mitotic embryos and perform genome-wide approach to identify those targets.

- 3) Investigate what is the role of this bookmarking on transcriptional memory

For this, we employed live imaging of transcription for genes identified in sub-aim 2) to precisely quantify transcriptional memory and look at the effect of depletion of potential bookmarkers on embryogenesis and transcription.

Chapter 2. Results: Mitotic bookmarking by the pioneer-like factor GAF

This manuscript was recently published as a pre-print on bioRxiv (august 2021) and currently under review at the journal Nature Communications. This article was mostly written by Dr. M. Lagha and myself. It summarizes the results obtained for GAF as a mitotic bookmarkers during early *Drosophila* embryogenesis.

This work is clearly multi-disciplinary, involving live imaging of transcription, fixed imaging DNA and RNA-FISH, image analysis, mathematical modelling, kinetics experiments and genome-wide approaches. Thanks to the great environment at IGMM, I could participate to all those aspects of the project and collaborators could teach me some parts.

When I arrived in M. Lagha laboratory, J. Dufourt has started testing various methods to sort mitotic embryos with a large object flow cytometer. During a year, based on his first attempts, I implemented a solid pipeline that combines both automatic sorting and manual sorting under the microscope to obtain a 99% pure population of mitotic embryos. As this pipeline was based on fixed embryos, treated for immunostaining, my next task was to adapt classical ChIP protocols to this type of embryos. I tested four different types of H3S10ph antibodies and several types of beads (Epoxy beads, Protein A and Protein G beads) to finally choose to couple beads and antibody directly. I then optimized chromatin sonication on mitotic chromatin and systematically checked by qPCR on a handful of targets (snail, hunchback). This again was not trivial, as we did not have any previous positive control.

I then performed libraries (in triplicate, each sample corresponding to ~1000 hand-sorted embryos) and sent these for illumina sequencing. The sequencing results were filtered by A. Zine El Aabidine, a bioinformatician in J.C. Andrau laboratory. I then participated to a training course in ChIPseq analysis during one week and was able to reproduce the analysis from my experiments. Further analysis and plotting the data was done by myself with the help of G. Hunt, a PhD student in M. Mannervik laboratory (The Wenner-Gren Institute, Stockholm University). Imaging for FRAP and FCS of sfGFP-GAF was performed by J.Dufourt and analyzed by C. Favard who chose the equations and fitted the data. A.Trullo coded and developed the software for DNA-FISH analyses and MitoTrack (Trullo et al. 2019). O. Radulescu is a mathematician who analyzed and fitted the memory data from MS2 tagged line movies with a mathematical model that he developed (Dufourt et al. 2018). I designed and performed the cloning of the CRISPR tagged MS2 lines with the help of the intern H. Faure-Gautron that I co-supervised for 6 months. All the live imaging followed by memory

and intensities analysis, as well as fixed imaging were designed and performed by myself with the help of an intern M. Lamarque (that I co-supervised for 6 months). For DNA-FISH experiments, with the advice of F. Bantignies (researcher in G. Cavalli laboratory), I was able to design and set-up the imaging conditions. DNA FISH experiments were primarily performed by the engineer H. Lenden-Hasse. I performed all the genetic crosses used in this study (CRISPR alleles, transgenes for MS2 imaging, RNAi manipulations). I also performed the qPCR, western blots and lethality counts when we assessed RNAi efficiency. I designed and constructed all figures of the manuscript.

The control of transcriptional memory by stable mitotic bookmarking

Maëlle Bellec¹, Jérémy Dufourt¹, George Hunt², Hélène Lenden-Hasse¹, Antonio Trullo¹, Amal Zine El Aabidine¹, Marie Lamarque¹, Marissa M Gaskill³, Heloïse Faure-Gautron¹, Mattias Mannervik², Melissa M Harrison³, Jean-Christophe Andrau¹, Cyril Favard⁴, Ovidiu Radulescu⁵, and Mounia Lagha¹.

Affiliations :

¹ Institut de Génétique Moléculaire de Montpellier, University of Montpellier, CNRS-UMR 5535, 1919 Route de Mende, Montpellier 34293 cedex 5, France.

² Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, 10691 Stockholm, Sweden.

³ Department of Biomolecular Chemistry, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.

⁴ Institut de Recherche en Infectiologie de Montpellier, CNRS UMR 9004, University of Montpellier, 1919 Route de Mende, Montpellier 34293 cedex 5, France.

⁵ LPHI, UMR CNRS 5235, University of Montpellier, Place E. Bataillon – Bât. 24 cc 107, Montpellier 34095 cedex 5, France.

bioRxiv 2021.08.30.458146; doi: <https://doi.org/10.1101/2021.08.30.458146>

Abstract

To maintain cellular identities during development, gene expression profiles must be faithfully propagated through cell generations. The reestablishment of gene expression patterns upon mitotic exit is thought to be mediated, in part, by mitotic bookmarking by transcription factors (TF). However, the mechanisms and functions of TF mitotic bookmarking during early embryogenesis remain poorly understood. In this study, taking advantage of the naturally synchronized mitoses of *Drosophila* early embryos, we provide evidence that the pioneer-like transcription factor GAF acts as stable mitotic bookmarker during zygotic genome activation. We report that GAF remains associated to a large fraction of its interphase targets including at *cis*-regulatory sequences of key developmental genes, with both active and repressive chromatin signatures. GAF mitotic targets are globally accessible during mitosis and are bookmarked via histone acetylation (H₄K8ac). By monitoring the kinetics of transcriptional activation in living embryos, we provide evidence that GAF binding establishes competence for rapid activation upon mitotic exit.

Introduction

Cellular identities are determined by the precise spatio-temporal control of gene expression programs. These programs must be faithfully transmitted during each cellular division. However, with its drastic nuclear reorganization, mitosis represents a major challenge to the propagation of gene expression programs. How cells overcome this mitotic challenge to transmit information to their progeny remains relatively unexplored during embryogenesis (Bellec et al. 2018; Festuccia et al. 2017; Elsherbiny and Dobрева 2021).

Based on live imaging studies and genome-wide profiling experiments on drug-synchronized mitotic cells, it is now well established that a subset of transcription factors (TF), chromatin regulators and histone modifications are retained on their targets during mitosis (Festuccia et al. 2017; Raccaud and Suter 2017; Raccaud et al. 2019) *via* specific, non-specific DNA binding or a combination of both (Cirillo et al. 2002; Raccaud et al. 2019).

When the persistence of TF binding during mitosis is associated with a regulatory role in transcriptional activation upon mitotic exit, TFs can be envisaged as mitotic bookmarkers. The kinetics of post-mitotic re-activation are often examined by whole-genome profiling

experiments of nascent transcription in early G₁ (Palozola et al. 2017; Zhang et al. 2021, 2019a). Combining such approaches with the mitotic depletion of candidate bookmarkers, it was established that some mitotically retained TFs/General TFs/histone marks act as *bona fide* mitotic bookmarkers (Kadauke et al. 2012; Teves et al. 2018; Festuccia et al. 2016).

Parallel to these multi-omics approaches, imaging of transcription in live cells with signal amplifying systems as the MS₂/MCP (Pichon et al. 2018) allows for the direct quantification of the kinetics of transcriptional activation upon mitotic exit. With such approaches, mitotic bookmarking has been associated with an accelerated transcriptional reactivation after mitosis in cultured cells (Zhao et al. 2011). Moreover, this method enabled the visualization of the transmission of active states, referred to as 'transcriptional memory' in *Dictyostellium* and in *Drosophila* embryos (Ferraro et al. 2016; Muramoto et al. 2010). However, how mitotic bookmarking is associated with the transmission of states across mitosis in the context of a developing embryo remains unclear.

This question is particularly important during the first hours of development of all metazoans, when cellular divisions are rapid and frequent. During this period, there is a substantial chromatin reprogramming and transcriptional activation, called Zygotic Genome Activation (ZGA) (Vallot and Tachibana 2020; Schulz and Harrison 2019). The control of this major developmental transition is supervised by key TFs, a subset of which are capable of engaging inaccessible chromatin and foster nucleosome eviction, a defining property of pioneer factors (Zaret and Mango 2016; Iwafuchi-Doi and Zaret 2016; Zaret 2018). Remarkably, many mitotic bookmarking factors have pioneer factor properties (Zaret 2020b).

In *Drosophila melanogaster*, two essential transcription factors with pioneering factor properties, Zelda and GAGA Associated Factor (GAF), orchestrate the reshaping of the genome during ZGA (Liang et al. 2008; Sun et al. 2015; Gaskill et al. 2021; Moshe and Kaplan 2017; Schulz et al. 2015b). Contrary to Zelda, which is not retained during mitosis and is dispensable for transcriptional memory (Dufourt et al. 2018), GAF is known to decorate mitotic chromosomes (Raff et al. 1994; Gaskill et al. 2021; Dufourt et al. 2018). Here we asked whether GAF acts as a mitotic bookmarker during ZGA. GAF, encoded by the *Trithorax-like* gene, binds to repeating (GA)_n sequences and displays a broad set of functions including gene activation or silencing, nucleosome remodeling and chromatin organization (Chetverina et al. 2021; Srivastava et al. 2018). In addition, GAF has been shown to be enriched at paused

promoters (Hendrix et al. 2008; Li and Gilmour 2013) and its manipulation in *Drosophila* S2 cells demonstrated a capacity to rapidly evict nucleosomes, thereby facilitating the recruitment of Pol II at promoters (Fuda et al. 2015; Judd et al. 2021). Together with its mitotic retention, these properties place GAF as a reasonable candidate for mitotic bookmarking during development.

Results

Endogenous GAF is retained during mitosis in early development and stably binds DNA

To investigate the function of GAF during mitosis, we first characterized its distribution during the cell cycle. With immunostaining, we confirmed that GAF is present on chromatin during all stages of mitosis from prophase to telophase (Fig. 1a) (Raff et al. 1994; Dufourt et al. 2018). Next, we examined GAF behavior in living embryos using an endogenously GFP tagged allele of GAF (Gaskill et al. 2021) (Fig. 1b, Movie 1). During mitosis, a large amount of GAF protein is displaced to the cytoplasm, but a clear pool of GAF protein remains associated with mitotic chromosomes (Fig. 1b).

From both live imaging and immunofluorescence data, we observed a strong GAF signal concentrated in large distinct puncta as well as a more diffuse signal within the nucleus. Consistent with previous work²⁹, we found that the majority of large GAF puncta are located at the apical side of the nuclei (Supplementary Fig. 1a, Movie 2), where at this stage, most of centromeric heterochromatin is located (Supplementary Fig. 1b) (Foe et al. 1993). In contrast to GAF apical foci, the rest of the nuclear space contains a homogeneously distributed GAF signal, potentially representing GAF binding to euchromatin (Supplementary Fig. 1a-b, Movie 2). To characterize GAF diffusion and binding kinetics in these regions, we performed Fluorescence Correlation Spectroscopy (FCS) and imaging Fluorescence Recovery After Photobleaching (FRAP) (Auer et al. 2021) on living GAF-GFP embryos during interphase (Fig. 1c, Supplementary Fig. 1c-d). We could not perform FRAP and FCS during mitosis due to their short durations and rapid nuclear movements.

FRAP recovery curves showed two characteristic times, a short one and a surprisingly long one (Fig. 1d). The short recovery time could correspond to fast unbinding or to diffusion. We confirmed that this short recovery time corresponds to diffusion (Fig. 1e) using FCS experiments (Supplementary Fig. 1e-f). The long-lived characteristic time from FRAP data, with a residence time close to a minute (Fig. 1f) is believed to correspond to GAF sequence-specific binding. Such a long-lived binding is an order of magnitude longer than typical TF residence times in *Drosophila* embryos (Mir et al. 2017, 2018; Dufourt et al. 2018). We conclude that GAF protein has the intrinsic capacity to stably bind chromatin. This property could be involved in its capacity to associate to mitotic chromosomes during embryonic divisions.

Capturing GAF mitotic targets genome-wide

Early *Drosophila* embryogenesis provides an ideal system to study mitosis. Indeed, nuclei of the syncytial embryo divide 13 times synchronously before cellularization (Farrell and O'Farrell 2014b). To perform mitotic ChIP, we stained early staged embryos with antibodies against the mitotic specific marker H3S10ph (Supplementary Fig. 2a) (Hendzel et al. 1997; Follmer et al. 2012) and sorted them with a flow cytometer (Fig. 2a and Supplementary Fig. 2b). The pools of embryos were further manually sorted to avoid contamination (Supplementary Fig. 2b). We applied this method to map GAF targets during mitosis and interphase. We retrieved GAF peaks genome-wide in interphase and mitotic samples and classified them into three categories: present only in interphase, only during mitosis, or during both interphase and mitosis, referred to as 'mitotically retained' (Fig. 2b-c''). Remarkably, mitotically retained GAF targets represent 37% of interphase targets, corresponding to a group of ~2000 peaks bound by GAF both in interphase and mitotic embryos (Fig. 2b). The mitotically retained loci comprise many key developmental patterning genes, as exemplified by *snail*, for which the proximal enhancer shows a GAF mitotic peak (Fig. 2c').

Motif search confirmed that GAF peaks are enriched in GAGAG motifs (Fig. 2d), and are centered inside the reads (Supplementary Fig. 2c). However, this consensus GAF binding site did not emerge as a significantly enriched motif in the small sample of GAF mitotic-only targets. We therefore did not analyze in depth this group of GAF targets. Moreover, there was a substantial degree of overlap (~93.5%) when comparing our interphase GAF peaks with published GAF-ChIP-seq data from bulk 2-4h embryos (Koenecke et al. 2017). Thus, we established a pipeline, able to profile mitotic nuclei at a genomic scale, for the first time in a multicellular organism, in the absence of drug synchronization.

Interestingly, the number of GAGAG motifs differs between mitotically retained peaks and interphase only peaks. On average, mitotically retained peaks have 6.2 GAGAG repeats while interphase only bound targets show 2.9 number of motifs (Fig. 2e). Therefore, we conclude that loci with significant number of GAF binding sites are more likely to be bound during mitosis.

Moreover, *de-novo* motif search revealed that while some motifs are present on both categories (interphase only and mitotically retained), a combination of consensus binding

sites is specifically enriched in mitotically retained peaks (e.g. Dorsal, Supplementary Fig. 2d). GAF mitotically retained targets might therefore be regulated by a distinct *cis*-regulatory logic than those from which GAF dissociates during mitosis.

To better characterize GAF-bound loci, we used existing genomic annotations of *cis*-regulatory modules (enhancers, promoters, and insulators) that were previously obtained from whole-genome profiling of the early *Drosophila* embryo (Hug et al. 2017a; Nègre et al. 2010; Koenecke et al. 2017) or validated via reporter transgenes (Kvon et al. 2014) (see Methods, Fig. 2f-f''). This stringent analysis revealed that the majority of GAF mitotically retained regions (65%) corresponds to *cis*-regulatory sequences (Fig. 2f'). This proportion is higher than the interphase only peaks (40%, Fig. 2f).

Mitotically retained GAF marks accessible regions during zygotic genome activation

As GAF displays pioneering properties in many contexts (Moshe and Kaplan 2017; Gaskill et al. 2021; Fuda et al. 2015), we hypothesized that GAF could contribute to chromatin accessibility during mitosis. We therefore determined the degree of chromatin accessibility at GAF-bound loci by using available ATAC-seq data (Blythe and Wieschaus 2016). We observed that GAF mitotically retained regions are globally more open than GAF interphase only or mitotic only targets (Fig. 3a).

More specifically, chromatin accessibility at mitotically retained regions encompasses larger regions than at loci bound by GAF only during interphase. This is in agreement with mitotically retained regions exhibiting a larger number of GAGA binding sites, potentially reflecting an enhanced number of bound GAF proteins able to foster nucleosome eviction (Supplementary Fig. 3a). Moreover, mitotically retained loci open gradually across developmental time windows and remain accessible during mitosis (Fig. 3a). Global chromatin accessibility at GAF mitotically retained targets is mostly linked to accessibility at *cis*-regulatory regions (Supplementary Fig. 3b).

We then asked whether chromatin accessibility at GAF mitotically retained regions required the presence of GAF. For this, we used ATAC-seq data performed on embryos where GAF levels were significantly reduced (Gaskill et al. 2021). From this dataset, we retrieved GAF bound loci for which accessibility was shown to be dependent on GAF. We found that the vast majority of these GAF-dependent regions (96%) correspond to GAF targets that we identified

as mitotically retained (Supplementary Fig. 3d). Interestingly, targets depending on GAF for their accessibility mostly coincide with TSS and enhancer regions but don't overlap TAD boundaries (Hug et al. 2017b) (Supplementary Fig. 3c). Importantly, interphase GAF targets or Zelda-only bound targets (not bound by GAF) did not show such a dependency on GAF for their accessibility (Fig. 3b). Collectively, these results suggest that GAF retention at specific promoters and enhancers during mitosis may foster an accessible chromatin organization, which resists the overall compaction of the genome occurring during mitosis. However, other factors in addition to GAF are likely to foster chromatin accessibility during mitosis.

GAF mitotic-bound regions are enriched with active and repressive histone marks

GAF is known to be present on both active and repressive chromatin regions (Adkins et al. 2006; Chetverina et al. 2021). We therefore assessed the chromatin landscape of GAF mitotically retained regions. For this purpose, we focused on embryonic ChIP-seq profiles of characteristic chromatin marks: H3K27ac for active chromatin state and H3K27me3 for the repressed chromatin state (Li et al. 2014), as well as RNA-seq signal from nc14 embryos (Lott et al. 2011). By clustering GAF mitotically retained regions, we partitioned GAF targets into three distinct clusters (Fig. 3c and Supplementary Fig. 3e). The first cluster (44% of mitotically retained GAF) corresponds to GAF mitotic peaks with significant enrichment in H3K27ac, depleted in H3K27me3 and with a high RNA-seq signal. In contrast, the second cluster (26% of mitotically retained GAF peaks) displayed enrichment for H3K27me3 concomitant with depletion in H3K27ac and low RNA-seq signal. The remaining GAF mitotic targets fall into a third cluster (30% of mitotically retained GAF peaks), which displays no particular epigenetic features with our clustering analysis but shows significantly less chromatin accessibility (Fig. 3d). To examine if additional chromatin modifications mark could discriminate between these three GAF clusters we performed ChIP-seq on the acetylation of lysine 8 of histone H4 (H4K8ac). Indeed among the myriad of chromatin marks labeling active regions, H4K8ac is a prominent mark during initial reshaping of the genome during *Drosophila* ZGA (Li et al. 2014). We used our mitotic ChIP-seq method (Fig. 2a) to map H4K8ac in interphase and mitotic embryos genome-wide (Supplementary Fig. 4a-d). We observed that H4K8ac was particularly enriched in cluster 1 (Fig. 3d and Supplementary Fig. 4e).

Together, these results demonstrate that mitotic GAF retention occurs at genomic regions associated with both active and repressive chromatin states. We propose that the combinatorial action of GAF and histone marks, contribute to the selective mitotic bookmarking of active regions to propagate transcriptional programs across cellular divisions.

GAF mitotic bookmarking is not associated with mitotic loops

Strictly speaking, mitotic occupancy by a TF can be envisaged as a mitotic bookmark only if it leads to a functional 'advantage' upon mitotic exit. Because chromatin loops between *cis*-regulatory regions were observed to be re-established by late anaphase/telophase in mammalian cells (Zhang et al. 2019b) and since GAF is implicated in loop formation in *Drosophila* (Mahmoudi et al. 2002; Ogiyama et al. 2018), we asked if GAF mitotically bound loci could form loops during mitosis in the embryo. We first focused on a specific genomic region containing two developmental genes, *charybde* (*chr*) and *scylla* (*scyl*), separated by 235kb and bound by GAF during both interphase and mitosis (Supplementary Fig. 5a-b). These early expressed genes were previously shown to form a long-range chromatin loop during early development (Ghavi-Helm et al. 2014).

We first confirmed that these loci are physically close and form a loop in nc14 by DNA FISH (Supplementary Fig. 5c). Interestingly, this proximity seems to be reinforced during nc14 progression (Supplementary Fig. 5c). However, while there is an overall genome compaction during mitosis, the distance between *scyl* and *chr* is not different from that of a control locus, in mixed stages of mitosis (Supplementary Fig. 5c). To confirm this result, we examined two other loci using DNA FISH and assessed their potential looping across the cell cycle (Supplementary Fig. 5d). Both *snail* and *escargot* show GAF binding and the H4K8ac mark in interphase and mitosis. While these loci, form a loop in interphase nuclei, this long-range loop is not different from the control locus during mitosis (Supplementary Fig. 5e).

We therefore conclude that, at least for these regions, GAF mitotic binding is not associated with detectable stable mitotic DNA looping.

The GAF bookmarked *scyl* gene harbors transcriptional memory

To test if GAF fosters rapid post-mitotic reactivation, we employed quantitative imaging on a selected GAF mitotically bound target, the zygotically expressed gene *scylla* (*scyl*). This gene is regulated by a promoter/proximal enhancer containing six GAGAG motifs, bound by GAF during interphase and mitosis (cluster 1 of mitotically retained loci) (Fig. 4a). To follow transcription dynamics with high temporal resolution, we utilized the MS2/MCP signal amplification method (Pichon et al. 2018) and quantitative imaging in living embryos. An array of 24X-MS2 repeats was inserted by CRISPR/Cas9 gene editing into the 3'UTR of *scyl* (Fig. 4a). MS2 reporter expression follows *scyl* endogenous expression (Fig. 4b and Supplementary Fig. 6a). Then, we monitored post-mitotic gene reactivation in *nc14* in the ventral (Fig. 4b and Supplementary Fig. 6c) and dorsal side (Supplementary Fig. 6b). In both locations, post-mitotic activation was found to be relatively fast, with a lag time of only 7.5 min and 9 min to reach 50% of the full pattern of activation (t_{50}) in the dorsal ectoderm and mesoderm, respectively (Supplementary Fig. 6d).

In addition to this temporal information within a given interphase, live imaging of transcription in the context the fast-developing *Drosophila* embryo gives access to nuclei genealogy. We assessed whether the transcriptional status of mother nuclei (prior to division) influences that of their descendants (Trullo et al. 2020). Indeed, we have previously shown that within the mesoderm, descendants of active nuclei in *nc13* activate transcription significantly faster than those arising from inactive nuclei, a bias named 'transcriptional memory' (Ferraro et al. 2016). However, this was shown in the context of reporter transgenes and has thus far never been demonstrated at an endogenous locus.

To assess the existence of transcriptional memory at an endogenously mitotically bookmarked locus, we imaged *scyl* expression in the mesoderm. Within this domain, the expression was stochastic in *nc13* (Fig. 4b and Supplementary Fig. 6c, and Movie 3), allowing unambiguous discrimination between active and inactive mother nuclei prior to mitosis. By tracking the timing of activation for daughters arising from active mother nuclei compared to those coming from inactive mother nuclei (Fig. 4c), we observe a clear transcriptional memory bias (Fig. 4d and Supplementary Fig. 6e).

In order to test if this bias was due to a stronger activity of the *scyl* gene in nuclei coming from active mothers, we examined instantaneous intensities of transcriptional sites as they are directly correlated to the mRNA synthesis efficiency. Global transcriptional site intensities

were similar in nuclei coming from active mothers compared to those coming from inactive mothers (Supplementary Fig. 6f).

GAF knock-down delays post-mitotic transcriptional reactivation and affects transcriptional memory

To test whether GAF was involved in the establishment of transcriptional memory, we employed RNAi knock-down (KD) to reduce the pool of maternal GAF. As previous studies reported difficulties to successfully deplete maternal GAF using a specific set of Gal₄ driver (Rieder et al. 2017), we decided to increase the efficiency of our depletion by combining two strong Gal₄ drivers (mat-alpha-Gal₄ and nanos-Gal₄). This strategy induces RNAi at all steps of oogenesis (Dufourt et al. 2014). The level of maternal GAF mRNA KD was estimated to be 88% by qRT-PCR and also confirmed by western blot (Supplementary Fig. 7a), creating a substantial embryonic lethality. However, in this genetic context a few embryos survived until gastrulation, albeit with clear mitotic and patterning defects for GAF targets genes (Movie 4, Supplementary Fig. 7b).

By quantifying post-mitotic reactivation timing of *scyl* in *RNAi-GAF* embryos, we observed a delay of ~6min for t_{50} (Fig. 5a). However, transcription site intensities were not affected upon GAF reduction (Supplementary Fig. 7c). This global trend is also observed when following single nuclei behavior during *nc14* (Supplementary Fig. 7c). We then compared the kinetics of activation in the two subpopulations (from active and from inactive) and found that the transcriptional memory bias was significantly reduced in *RNAi-GAF* embryos (Fig. 5b-c). Such a memory reduction does not occur upon maternal depletion of the pioneer factor Zelda (Dufourt et al. 2018).

Collectively, these data demonstrate that GAF controls the timing of transcriptional activation after mitosis and participates in the establishment of transcriptional memory. Moreover, this temporal effect is not due to a differential promoter activity between neighboring nuclei.

Modeling GAF driven transcriptional memory

We analyzed the statistical distribution of the post-mitotic delay, defined as the lag time between the end of mitosis and the first activation in *nc14*. We have previously developed a simple mathematical model of memory, where this delay was modeled by a mixed gamma

distribution (Dufourt et al. 2018) with two main parameters, the average number of rate-limiting transitions prior to reach the transcription active state (ON) (parameter 'a') and their durations (parameter 'b'). Applying this mathematical model to our live imaging movies of *scyl* transcription dynamics in control (*RNAi-white*) and in GAF depleted embryos (*RNAi-GAF*) revealed that the 'a' parameter was comparable across genotypes (Supplementary Fig. 7d). However, upon GAF KD, the 'b' parameter significantly increased in nuclei coming from active mother nuclei (Supplementary Fig. 7d). Remarkably, this selective decrease in the 'b' parameter within a subpopulation was not observed upon Zelda depletion (Dufourt et al. 2018). In order to be able to compare the effect of various genotypes, subject to distinct *cis*-regulatory codes, we introduced a memory score defined by the ratio $(ab_{\text{inactive}})/(ab_{\text{active}})$. A memory bias exists when this ratio is higher than 1. Using this metric, we observe that endogenous *scyl* exhibits a clear memory bias that vanishes upon GAF depletion (Fig. 5d). Interestingly, a GAF-dependent memory bias was also observed with a second GAF-mitotically bound region (*sna-proximal-enhancer*, cluster 1, Fig. 2c', see Methods) (Fig. 5d and Supplementary Fig. 7e). In all cases, we observe $(ab_{\text{inactive}})/(ab_{\text{active}}) \approx b_{\text{inactive}}/b_{\text{active}}$ (Fig. 5d), suggesting that the primary contribution to the memory bias comes from the transition duration 'b'.

Collectively, these results suggest a model where transcriptional memory bias results from distinct epigenetic paths in nuclei where a given locus is bookmarked by GAF and in nuclei where the same locus is not bound by GAF (Fig. 5e). The preferential bookmarking of active nuclei by GAF could be explained by stochastic GAF binding.

Discussion

We set out to determine how gene regulation by a transcription factor might be propagated through mitosis in a developing embryo. By using a combination of quantitative live imaging and genomics, we provide evidence that the pioneer-like factor GAF acts as a stable mitotic bookmarker during zygotic genome activation in *Drosophila* embryos.

Our results indicate that during mitosis, GAF binds to an important fraction of its interphase targets, largely representing *cis*-regulatory sequences of key developmental genes (Supplementary table 2). We noticed that GAF mitotically retained targets contain a larger number of GAGA repeats than GAF interphase only targets and that this number of GAGA

repeats correlates with the broadness of accessibility. Multiple experiments, with model genes *in vitro* (e.g. *hsp70*, *hsp26*) or from genome-wide approaches clearly demonstrated that GAF contributes to generate nucleosome-free regions (Chetverina et al. 2021). The general view is that this capacity is permitted through the interaction of GAF with nucleosome remodeling factors as PBAP (SWI/SNIF), NURF (ISWI) (Judd et al. 2021) or FACT (Orphanides et al. 1998). Although not yet confirmed with live imaging, immuno-staining data suggest that NURF is removed during metaphase but re-engages chromatin by anaphase (Kwon et al. 2021). If the other partners of GAF implicated in chromatin remodeling are evicted during early mitosis, chromatin accessibility at GAF mitotic targets could be established prior to mitosis onset and then maintained through mitosis owing to the remarkable stability of GAF binding. However, we cannot exclude GAF interactions with other chromatin remodelers (e.g. PBAP) during mitosis and a scenario whereby mitotic accessibility at GAF targets would be dynamically established during mitosis thanks to the coordinated action of GAF and its partners.

We propose that the function of GAF as a mitotic bookmarker is possible because GAF has the intrinsic property to remain bound to chromatin for long periods (residence time in the order of minute). This long engagement of GAF to DNA is in sharp contrast with the binding kinetics of many other TF, such as Zelda or Bicoid in *Drosophila* embryos (Dufourt et al. 2018; Mir et al. 2017). Another particularity of GAF binding, contrasting with other TF, resides in the multimerization of its DNA binding sites as GAGAG repeats in a subset of its targets (76% of mitotically retained peaks display four or more repetitions of GAGAG motifs). Given the known oligomerization of GAF (Espinás et al. 1999) and as GAF is able to regulate transcription in a cooperative manner (Van Steensel et al. 2003), it is tempting to speculate that GAF cooperative binding on long stretches of GAGAG motifs may contribute to a long residence time.

Collectively, we propose that the combination of long residence time and the organization of GAF binding sites in the genome may allow the stable bookmarking of a subset of GAF targets during mitosis.

In this study, we also discovered that a combination of GAF and histone modification could be at play to maintain the chromatin state during mitosis. Indeed, mitotic bookmarking may also be supported by the propagation of histone tail modifications from mother to daughter

cells. Work from mammalian cultured cells revealed widespread mitotic bookmarking by epigenetic modifications, such as H₃K₂₇ac and H₄K₁₆ac (Behera et al. 2019; Liu et al. 2017). Moreover, H₄K₁₆ac transmission from maternal germline to embryos has recently been established (Samata et al. 2020). In the case of GAF, we propose that the combinatorial action of GAF and epigenetic marks, possibly selected via GAF interacting partners, will contribute to the propagation of various epigenetic programs. It would be therefore interesting to employ our established mitotic ChIP method to survey the extent to which *cis*-regulatory regions exhibit different mitotic histone mark modifications during embryogenesis.

A key aspect of mitotic bookmarking is to relate mitotic binding to the rapid transcriptional activation after mitosis. Here we show that GAF plays a role in the timing of re-activation after mitosis. However, we note that GAF binding during mitosis is not the only means to accelerate gene activation. Indeed, we and others have shown that mechanisms such as enhancer priming by Zelda, paused polymerase or redundant enhancers contribute to fast gene activation (Lagha et al. 2012; Bentovim et al. 2017). Moreover, a transcriptional memory bias can occur for a transgene not regulated by GAF (Ferraro et al. 2016). By modeling the transcriptional activation of the gene *scylla*, we reveal that GAF accelerates the epigenetic steps prior to activation selectively in the descendants of active nuclei. We propose a model where GAF binding helps in the decision-making of the post-mitotic epigenetic path. In this model, mitotic bookmarking by GAF would favor an epigenetic path with fast transitions after mitosis (Fig. 5e). In the context of embryogenesis, bookmarking would lead to the fast transmission of select epigenetic states and may contribute to gene expression precision.

Interestingly, GAF vertebrate homolog (vGAF/Th-POK) has recently been implicated in the maintenance of chromatin domains during zebrafish development (Matharu et al. 2021). We therefore suspect that GAF action as a stable bookmarking factor controlling transcriptional memory during *Drosophila* ZGA might be conserved in vertebrates.

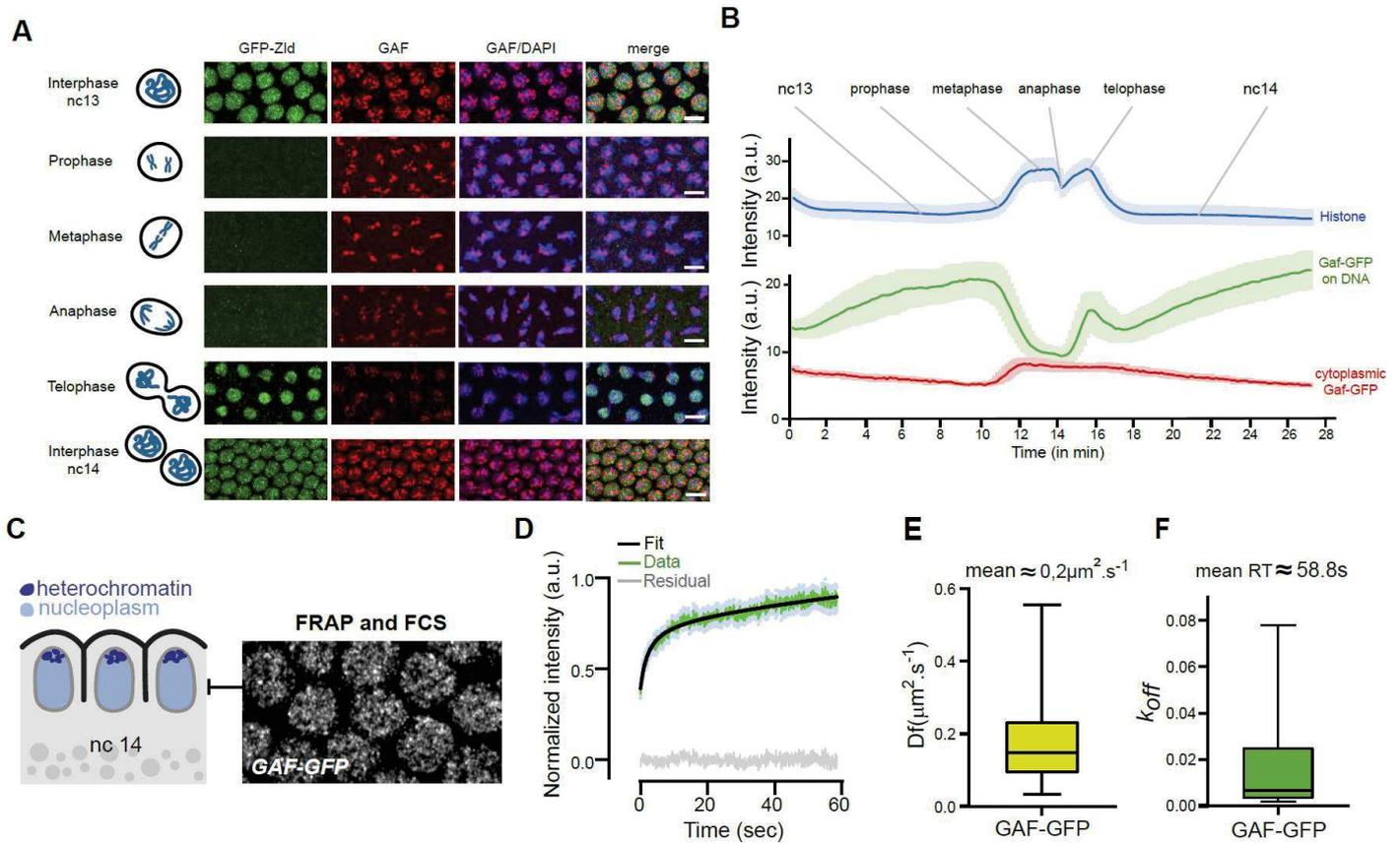


Figure 1: GAF dynamics during nuclear cycles and its kinetic properties

(A) Maximum intensity projected Z-planes of confocal images from immunostaining of Zelda-GFP (green) and GAF (red) on interphase and mitotic embryos at the indicated stages counterstained with DAPI (blue). Scale bar is $5\mu\text{m}$.

(B) Mean fluorescent signal quantifications of GAF-GFP in nucleoplasm (green) and cytoplasm (red), and H2Av-RFP in nucleoplasm during nuclear cycle 13 to 14 extracted from time-lapse movies of embryos expressing GAF-GFP and H2Av-RFP (mean from three movies of three independent embryos).

(C) Schematic of sagittal view of nc14 embryos. Nuclei are represented in light blue and apical heterochromatin regions in dark blue. Right panel shows regions targeted by FRAP and FCS, performed on GAF-GFP embryos.

(D) Mean fluorescence recovery curve (green) from FRAP experiment and fit (black) using a reaction-diffusion model determined at the bleached spot for 23 nuclei from nine nc14 GAF-GFP embryos. Light blue dots represent SEM from different nuclei. Grey curve represents the residual of the fit.

(E) Estimated diffusion coefficient of GAF-GFP. Centered line represents the median and whiskers represent min and max values.

(F) Estimated koff (RT: residence time = $1/k_{off}$) of GAF-GFP. Centered line represents the median and whiskers represent min and max values.

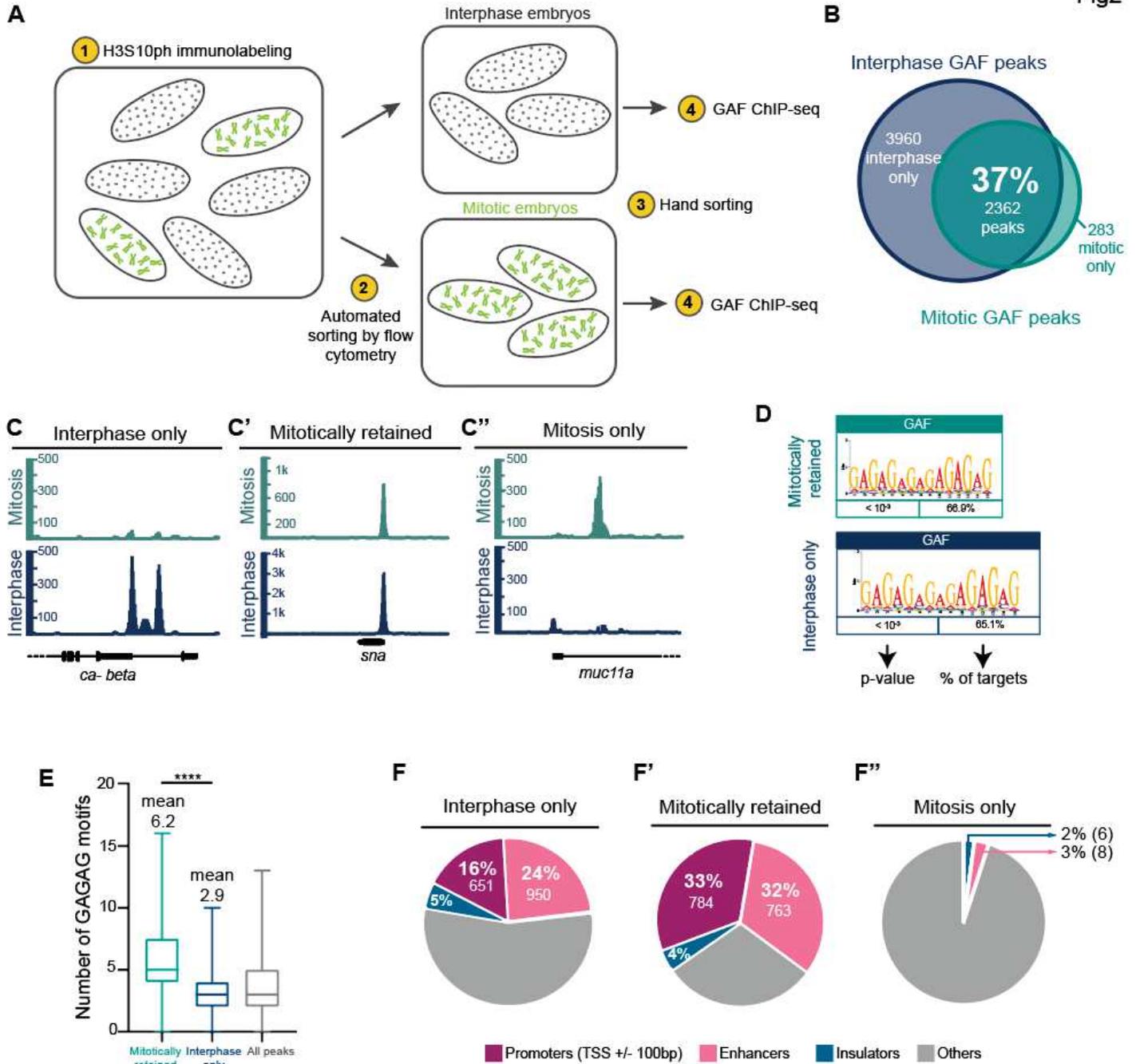


Figure 2: Identification of thousands of mitotically retained GAF loci

(A) Experimental workflow of mitotic embryo sorting followed by GAF-ChIP-seq.

(B) Venn diagram representing the overlap of called GAF-ChIP-seq peaks between interphase and mitotic embryos.

(C) (C') (C'') Genome browser examples of genes from the identified three categories of GAF-ChIP-seq peaks: interphase only, mitotically retained and mitosis only, respectively.

(D) (GA)_n motif enrichment within GAF mitotically retained and interphase only peaks, as reported by MEME.

(E) Box plot representing the number of GAGAG motifs within three different classes of GAF peaks: mitotically retained (light blue), interphase only (dark blue) and all peaks (grey). Centered horizontal line represents the median, whiskers represent min and max values. Two tailed Welch's t-test **** $p < 0.0001$.

(F) (F') (F'') Proportions of GAF-ChIP-seq peaks that overlap diverse cis-regulatory regions in interphase only, mitotically retained and mitosis only GAF-ChIP-seq.

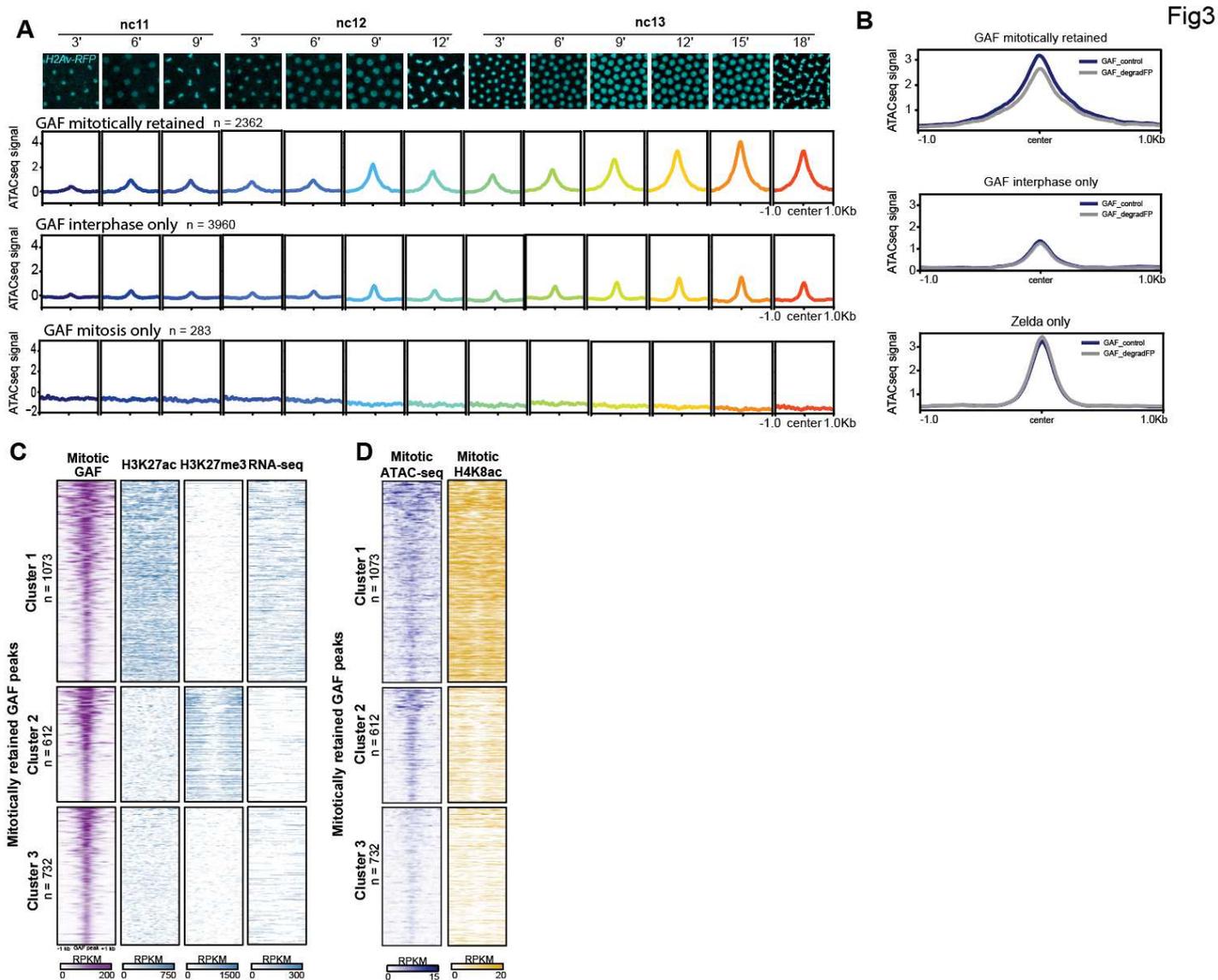


Figure 3: Mitotically retained GAF loci become progressively accessible during Zygotic Genome Activation

(A) Metagene profiles of ATAC-seq signal⁴⁷ centered at mitotically retained, interphase only and mitosis only GAF-ChIP-seq peaks across the indicated stages and represented by the time lapse images from a movie of H2Av-RFP embryos (cyan).

(B) Metagene profiles of ATAC-seq signal in WT (GAF_control, dark blue) and GAF-depleted (GAF_degradFP, grey) embryos²⁵ on GAF mitotically retained, GAF interphase only and Zelda only regions.

(C) Heatmaps of k-means clustered mitotically retained GAF peaks, based on H3K27ac and H3K27me3 ChIP-seq⁴⁹ and RNA-seq⁵⁰.

(D) Heatmaps representing the mitotic ATAC-seq signal ⁴⁷ (dark blue) and the ChIP-seq enrichment of H₄K8ac in mitotic embryos at the clustered mitotically retained GAF peaks from (C).

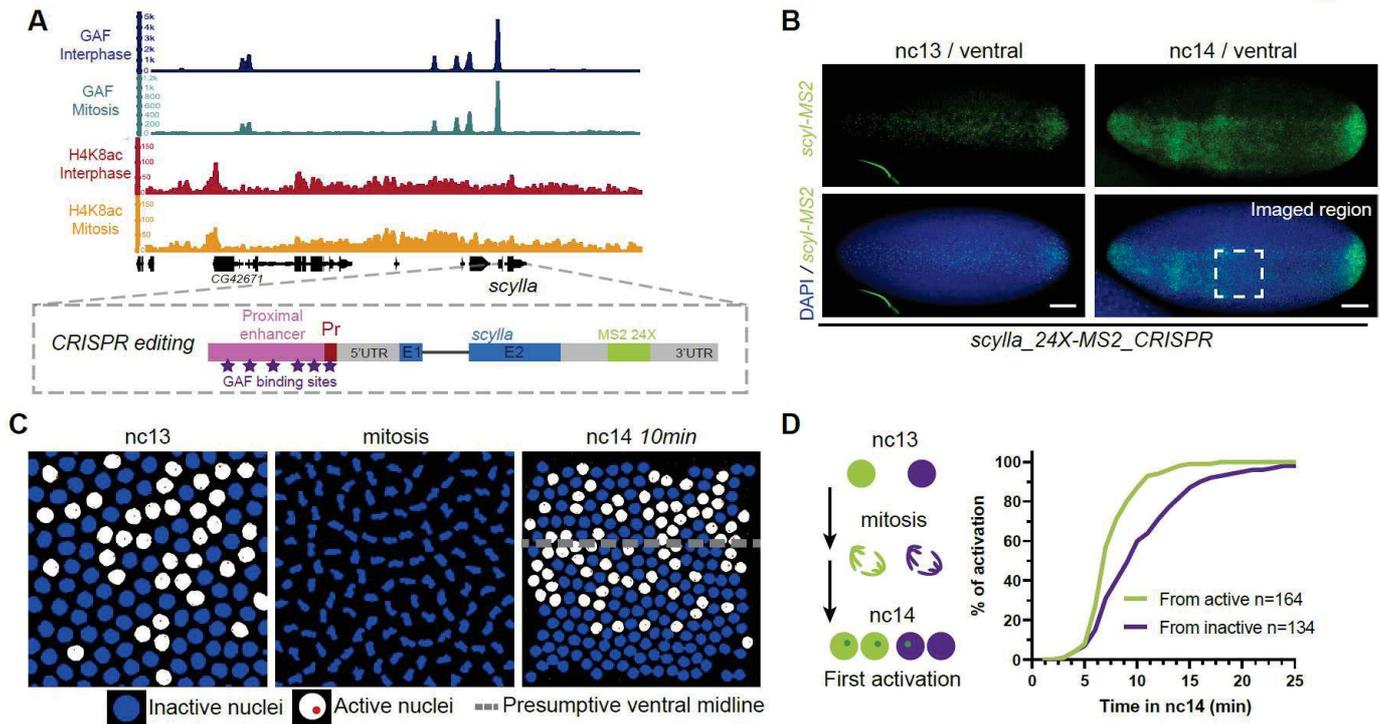


Figure 4: *scylla* gene harbors a transcriptional memory across mitosis.

(A) (Top) Genome browser image of interphase and mitotic GAF (dark blue and turquoise) and H₄K8ac (red and orange) ChIP-seq signal at the *scyl* locus. (Bottom) Schematic of the 24X-MS2 tagging strategy of the *scyl* locus by CRISPR editing.

(B) Maximum intensity projected Z-planes of confocal images from smiFISH with MS2 probes (green) counterstained with DAPI (blue) of *scylla_24X-MS2_CRISPR/+* embryos in nc13 and nc14. Scale bars are 50µm. Dashed box represents the region considered for live imaging experiments.

(C) Snapshots from a representative false-colored movie of *scylla_24X-MS2_CRISPR/+* embryo carrying MCP-eGFP, H2Av-RFP. Active nuclei are represented in white and inactive nuclei in blue. Transcriptional sites are false colored in red. Dashed line represents the presumptive ventral midline.

(D) Quantification of transcriptional memory for *scyl* gene. Left panel: schematic of the two populations of nuclei studied; those derived from active (in green) and those from inactive nuclei (purple). Right panel: cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple). n=number of analyzed pooled nuclei from 4 movies of 4 independent embryos.

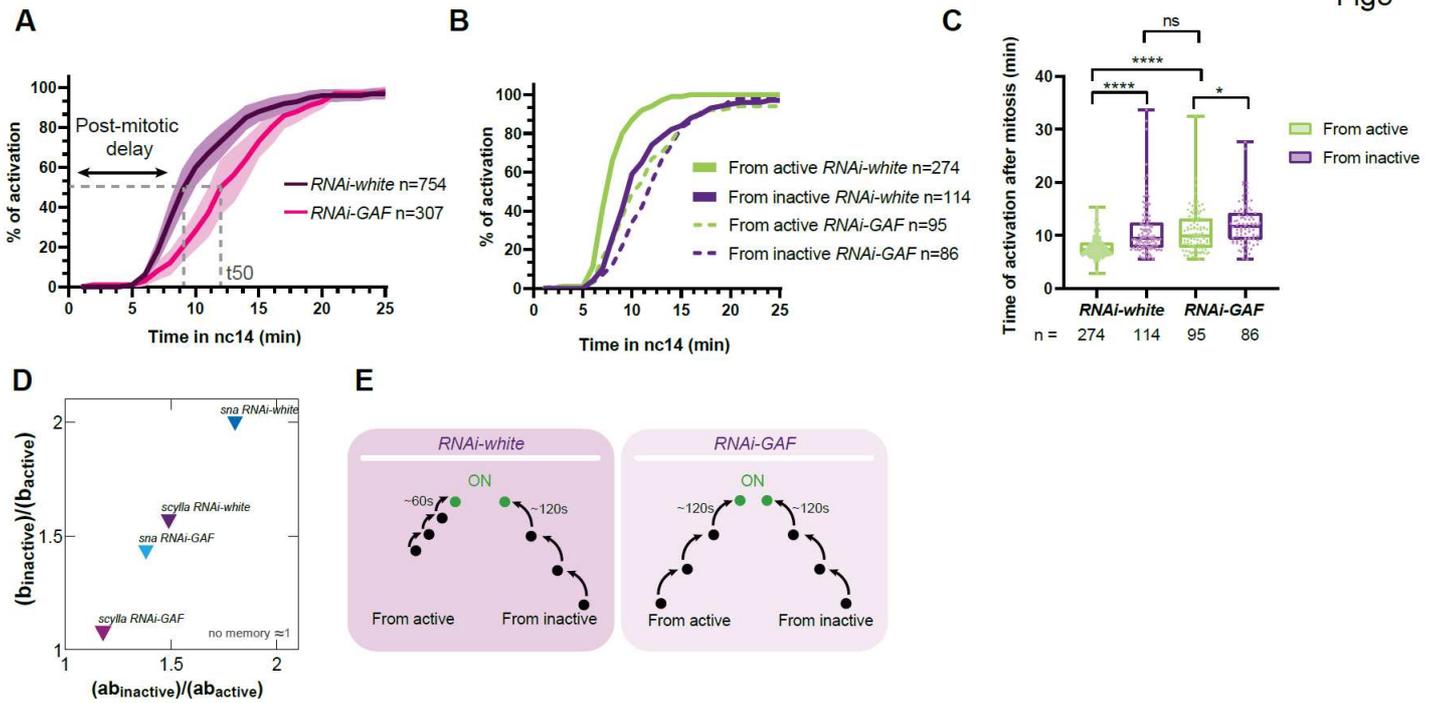


Figure 5: GAF is required for transcriptional memory of *scylla*

(A) Quantification of transcriptional synchrony of *scylla_24X-MS2_CRISPR/+* embryo after mitosis in RNAi-white (control, purple) and *mat-alpha-Gal4/+; nos-Gal4/UASp-shRNA-GAF* embryos (pink). Dashed line represents the t_{50} where 50% of the pattern is activated during nc14. Both of the two daughters derived from each nucleus are quantified. SEM are represented in light purple and light pink. n=number of pooled nuclei analyzed from 4 movies of 4 independent embryos for each condition.

(B) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-white embryos (control, solid curves) and RNAi-GAF *scylla_24X-MS2_CRISPR/+* embryos (dashed curves). n=number of pooled nuclei analyzed from 4 movies of 4 embryos.

(C) Box plot representing the mean time of the first activation after mitosis of nuclei derived from active (green) and inactive (purple) nuclei in RNAi-white embryos and RNAi-GAF *scylla_24X-MS2_CRISPR/+* embryos. Centered horizontal line represents the median. Two tailed Welch's t-test **** $p < 0.0001$, * $p < 0.05$.

(D) Ratios of parameter 'b' and 'ab' in subpopulations from inactive and active nuclei of *scylla_24X-MS2_CRISPR/+* (purples) and *snail-primary-enhancer_MS2* (blues) in RNAi-white or RNAi-GAF embryos. The parameter 'a' corresponds to the average number of transitions

(provided by the sum of weighted probabilities) and the parameter 'b' to the time of each jump from one state to another.

(E) Schematic of the proposed role of GAF in transcriptional memory. In the presence of GAF, nuclei derived from active nuclei have shorter 'b' length than those derived from inactive nuclei whereas in the absence of GAF, both have the same transition times.

Competing interests:

The authors declare that they have no competing interests.

Acknowledgments:

We are grateful to G.Cavalli, B. Schuttengruber, F. Juge, F.Bantignes and J.Chubb for sharing antibodies, advise on DNA FISH and/or insightful discussions.

We thank member of the Lagha lab, Cyril Esnault, Virginia Pimmett and Etienne Schwob for their critical reading of the manuscript. We acknowledge M. Dejean and M. Goussard for technical assistance. We acknowledge the Montpellier Ressources Imagerie facility (France-Biolmaging). Funding: M.B. is a recipient of an FRM fellowship. This work was supported by the ERC SyncDev starting grant to M.L. M.L., J.D., and C.F. are sponsored by CNRS. O.R. acknowledges support from the French National Research Agency (ANR-17-CE40-0036, project SYMBIONT).

Author contributions:

M.L. conceived the project. M.L., M.B. and J.D. designed the experiments. J.D., M.B., H.L-H., M. Lam. and H.F-G. performed experiments. A.T. developed software. C.F., M.B., and J.D. performed kinetic analysis. M.G. and M.H. provided the GAF-GFP strain. O.R. performed mathematical modeling, machine learning and interpreted data. G.H., M.B., A.Z-E-A., M.M. and J-C.A. performed the bioinformatics analysis. M.L. and M.B. wrote the manuscript. All authors discussed, approved, and reviewed the manuscript.

Material and Methods

Fly stocks, handling and genetics

The yw stock was used as a wild type. The germline driver *nos-Gal4:VP16*(BL4937) was previously recombined with a *MCP-eGFP-His2Av-mRFP* fly line (Dufourt et al. 2018). RNAi

were expressed after crossing this recombinant for live imaging (or *nos-Gal4:VP16* for fixed experiments) with *mat-alpha-Gal4* (BL7063), then with *UASp-shRNA-w* (BL35573) or *UASp-shRNA-GAF* (BL41582). Virgin females expressing RNAi, MCP-GFP-His2Av-mRFP and both Gal4 constructs were crossed with MS2 containing CRISPR alleles or transgene-containing males. All experiments were done at 21 °C except RNAi experiments which were done at 25 °C. The C-terminal tagged version of GAF-sfGFP was obtained by CRISPR/Cas9 (Gaskill et al. 2021).

Cloning and transgenesis

The *snail-primary-enhancer_MS2* transgene was obtained by amplification of the *sna* endogenous promoter and primary enhancer using the primers listed in Supplementary table 1. The 128XMS2 tag (Dufourt et al. 2021) was inserted immediately upstream of the yellow reporter gene sequence of the pbphi-yellow plasmid (Ferraro et al. 2016). The transgenic construct was inserted in the VK0033 landing site (BL9750) using PhiC31 targeted insertion (Venken et al. 2006).

The homology arms for the recombination template for CRISPR/Cas9 editing of *scyl* gene to generate *scyl_24X-MS2_CRISPR* were assembled with NEBuilder® HiFi DNA Assembly Master Mix (primers listed in Supplementary table 1) and inserted into pBluescript opened *SpeI/AscI* (for the 5' homology arm) or *XmaI/NheI* (for the 3' homology arm) containing the 24X-MS2 (as in (Dufourt et al. 2018)) inserted after *NotI* digestion. Guide RNA (Supplementary table 1) were cloned into pCFD3-dU6:3gRNA (Addgene 49410) digested by *BbsI* using annealed oligonucleotides (Integrated DNA Technology™). The recombination template and guide RNA plasmids were injected into BDSC#55821 (BestGene Inc.). Transformant flies were screened using a dsRed marker inserted downstream of the 3'UTR of the genes.

Fluorescence recovery after photobleaching

Fluorescence recovery after photobleaching (FRAP) in embryos at nc14 was performed on a Zeiss LSM880 using a 40 × /1.3 Oil objective and a pinhole of 84 μm. Images (256 × 128 pixels, 16bits/pixel, zoom 6x) were acquired every ≈ 53 ms for 1200 frames. GFP was excited with an Argon laser at 488 nm and detected between 492–534 nm. Laser

intensity was kept as low as possible to minimize unintentional photobleaching. A circular ROI (12 × 12 pixels) 0.138 μm/pixel, was bleached using two laser pulses at maximal power during a total of ≈ 110 ms after 10 frames. To discard any source of fluorescence intensity fluctuation other than molecular diffusion, the measured fluorescence recovery in the bleached ROI region (I_{bl}) was corrected by an unbleached ROI (I_{unbl}) of a neighbor's nucleus and another ROI outside of the nucleus (I_{out}) following the simple equation:

$$I_{bl_{corr}}(t) = \frac{I_{bl}(t) - I_{out}(t)}{I_{unbl}(t) - I_{out}(t)} \quad (1)$$

The obtained fluorescence recovery was then normalized to the mean value of fluorescence before the bleaching i.e.

$$I_{bl_{norm}}(t) = \frac{I_{bl_{corr}}(t)}{\frac{1}{N} \sum_{n=1}^{10} I_{bl}(n)} \quad (2)$$

We first employed the pure reaction kinetics model, from the analytical expression of (Spague et al. 2004). With this fit, we were not satisfied as the residuals of the fit were very high. Moreover, we clearly saw from the shape of the curve at least two characteristic times: a short recovery time in the order of seconds and a long one in the order of tens of second to minute.

We then used an equation with two component diffusion and reaction on the first 1100 frames: we started from the analytical expression developed in the Supplementary Equation 35 of (Michelman-Ribeiro et al. 2009a)

$$F(t) = F_{eq}F_D(t) + C_{eq}F_{exc}(t) \quad (3)$$

with C_{eq} defined as above and $F_{eq} = k_{off} / (k_{off} + k_{on}^*)$. $F_D(t)$ is the fluorescence recovery due to diffusion and $F_{exc}(t)$ the fluorescence recovery due to exchange.

Since we used a Gaussian shape illumination profile, $F_D(t)$ is defined using a slightly modified version of the analytical equation of the 20th order limited development of the Axelrod model for Gaussian profile illumination and diffusion (Escoffre et al. 2014; Axelrod et al. 1976). The limited development of Axelrod equation was chosen to be the 20th order. Indeed, diffusion is not fitted with a classical exponential but by the limited development Axelrod equation. A small development (corresponding to the number of order of the polynome) will highly reduce the precision of the fit and higher order of development can add a precision but also can create aberrant fitting. Moreover, the precision we could gain by increasing the order is way below the experimental precision of FRAP experiments in developing organisms, which are particularly noisy:

$$F_D(t) = \frac{1 - e^{-K}}{K} (1 - M) + M \sum_{n=1}^{20} \frac{(-K)^n}{n!} \left(1 + n + 2n \frac{t}{\tau}\right)^{-1} \quad (4)$$

$$M = \frac{I_{(t>30\tau)} - I_0}{1 - I_0} \quad (5)$$

where K is a value proportional to bleaching deepness, estimated for each recovery curve independently from the Axelrod equation, M is the mobile fraction and τ is the half time of recovery. To minimize the effect of mobile fraction on C_{eq} , M was kept between 0.9 and 1.1. Diffusion coefficients of the different molecules were determined according to

$$D = \frac{\beta w^2}{4\tau} \quad (6)$$

with w the value of the radius at $1/e^2$ of the Gaussian beam (in our case, $w=0.83\mu\text{m}$) and β a discrete function of K tabulated in (Yguerabide et al. 1982).

$F_{exc}(t)$ is defined as in (Michelman-Ribeiro et al. 2009), slightly modified with respect to the Gaussian illumination, leading to the following equation:

$$F_{exc}(t) = F_{\infty} - \left(\frac{1-e^{-K}}{K} - F_{\infty} \right) e^{-k_{off}t} \quad (7)$$

with K defined as previously.

Fluorescence correlation spectroscopy

Fluorescence correlation spectroscopy (FCS) experiments were performed on a Zeiss LSM780 microscope using a 40x/1.2 water objective. GFP was excited using the 488 nm line of an Argon laser with a pinhole of 1 airy unit. Intensity fluctuation measured for 10 s were acquired and auto-correlation functions (ACFs) generated by Zen software were loaded in the PyCorrFit program (Müller et al. 2014). Multiple measurements per nucleus in multiple nuclei and embryos at 20 °C were used to generate multiple ACF, used to extract parameters. The FCS measurement volume was calibrated with a Rhodamine6G solution (Dertinger et al. 2008) using $D_f = 414 \mu\text{m}^2.\text{s}^{-1}$. Each time series was fitted with the following generic equation:

$$G(\tau) = 1 + \frac{1}{N} \left(1 + \frac{T e^{-\frac{t}{\tau_r}}}{1-T} \right) \left(\sum_{i=1}^n \frac{f_i}{\left(1 + \frac{t}{\tau_i}\right) \left(1 + \frac{t}{s^2 \tau_i}\right)^{1/2}} \right) + G_{\infty} \quad (8)$$

Using $n=2$ in our fit and where N is the total number of molecules, T is the proportion of the fluorescent molecules N in the triplet state with a triplet state lifetime τ_T (constrained below $10\mu\text{s}$ in our fit), f_i is the proportion of each different diffusing species ($\sum_{i=1}^n f_i = 1$) with a diffusion time $\tau_i = w_{xy}^2 / 4 D$ and $s^2 = w_z / w_{xy}$. We also introduced a G_∞ value to account for long time persistent correlation during the measurements.

We performed FCS to characterize fast kinetics of GAF protein (Supplementary Fig. 1D). FCS measurements were fitted with one or two characteristic times. The quality of the fit was evaluated with the χ^2 function implemented in PyCorrFit in addition to the shape of residuals and the shape of the data measured. We selected two characteristic times, potentially corresponding to two diffusion timings or to a diffusion and binding reaction. To discriminate between these two scenarios, we performed FCS in the cytoplasm, where binding should not occur. Surprisingly, cytoplasmic FCS revealed two characteristic times, on the same order as those retrieved in the nucleoplasm (Supplementary Fig. 1E and F). Therefore, a two-diffusion component model was used to fit the nucleoplasm autocorrelation curves, giving rise to two apparent characteristic diffusion coefficients (D_f) on the order of $22 \mu\text{m}^2 \cdot \text{s}^{-1}$ and $0.45 \mu\text{m}^2 \cdot \text{s}^{-1}$. The fastest D_f corresponds to pure diffusion, as it falls in the range of diffusion of free GFP (Di Bona et al. 2019). The slower diffusion time potentially reflects GAF diffusion within a protein complex.

Immunostaining and RNA in situ hybridization.

A pool of 0-4h after egg-laying (AEL) or 2-4h AEL embryos were dechorionated with bleach for 3 min and thoroughly rinsed with H_2O . They were fixed in 1:1 heptane:formaldehyde-10% for 25 min on a shaker at 450 rpm; formaldehyde was replaced by methanol and embryos were shaken by hand for 1 min. Embryos that sank to the bottom of the tube were rinsed three times with methanol. For immunostaining, embryos were rinsed with methanol two times and washed three times 3 min with PBT (PBS 1x 0.1% triton). Embryos were incubated on a wheel at room temperature for 30 min in PBT, then for 20 min in PBT 1% BSA, and at 4°C overnight in PBT 1% BSA with primary antibodies. Embryos were rinsed three times, washed twice for 20 min in PBT, then incubated in PBT 1% BSA for 20 min, and in PBT 1% BSA with secondary antibodies for 2 h at room temperature. Embryos were rinsed three times then washed three times in PBT for 10 min. DNA staining was performed

using DAPI at 0.5 µg/ml. Primary antibody dilutions for immunostaining were mouse anti-GFP (Roche IgG1 clones 7.1 and 13.1) 1:200; rabbit anti-GAF (gift from Dr. G.Cavalli) 1:250. Secondary antibodies (anti-rabbit Alexa 488-conjugated (Life Technologies, A21206); anti-mouse Alexa 488-conjugated (Life Technologies, A21202); anti-rabbit Alexa 555-conjugated (Life Technologies, A31572)) were used at a dilution 1:500. Fluorescent *in situ* hybridization (FISH) was performed as described in (Dufourt et al. 2018). The digoxigenin-MS2 probe was obtained as (Dufourt et al. 2018) by *in vitro* transcription from a pBluescript plasmid containing the 24X-MS2 sequences, isolated with BamH1/BglII enzymes from the original Addgene MS2 plasmid (#31865). *snail* probe generation was described in (Dufourt et al. 2018). Primary and secondary antibody for FISH were sheep anti-digoxigenin (Roche 11333089001) 1:375; mouse anti-biotin (Life technologies, 03-3700) 1:375; anti-mouse Alexa 488-conjugated (Life Technologies, A21202) and anti-sheep Alexa 555-conjugated (Life Technologies, A21436) 1:500. Mounting was performed in Prolong® Gold.

Images in Supplementary Fig. 1a represent a maximum intensity projection of a stack of 3 z-planes ($\approx 1 \mu\text{m}$). Images in Supplementary Fig. 1b represent a single Z-plane. Images in Fig. 1a represent a maximum intensity projection of a stack of 9 z-planes ($\approx 4,5 \mu\text{m}$).

Single molecule fluorescence *in situ* hybridization (smFISH)

Embryos were fixed as in the previous section, then washed 5 min in 1:1 methanol:ethanol, rinsed twice with ethanol 100%, washed 5 min twice in ethanol 100%, rinsed twice in methanol, washed 5 min once in methanol, rinsed twice in PBT-RNasin (PBS 1x, 0.1% tween, RNasin® Ribonuclease Inhibitors). Next, embryos were washed 4 times for 15 min in PBT-RNasin supplemented with 0.5% ultrapure BSA and then once 20 min in Wash Buffer (10% 20X SCC, 10% Formamide). They were then incubated overnight at 37 °C in Hybridization Buffer (10% Formamide, 10% 20x SSC, 400 µg/ml *E. coli* tRNA (New England Biolabs), 5% dextran sulfate, 1% vanadyl ribonucleoside complex (VRC) and smFISH Stellaris probes against *sna* coupled to Quasar 670 and/or FLAP probes). FLAP-probes against 24X-MS2 and *scyl* were prepared by duplexing 40 pmol of target-specific probes with 100 pmol FLAP-Cy3 oligonucleotides and 1X NEBuffer™ 3 for 3 min at 85 °C, 3 min at 65 °C and 5 min at 25 °C and kept on ice until use. Probe sequences are listed in Supplementary table 1.

Embryos were washed in Wash Buffer at 37 °C and in 2x SCC, 0.1% Tween at room temperature before being mounted in ProLong® Gold antifade reagent. Images were acquired using a Zeiss LSM880 confocal microscope with an Airyscan detector in SR mode with a 40x Plan-Apochromat (1.3 NA) oil objective lens or a 20x Plan-Apochromat (0.8NA) air objective lens. Images were taken with 1024 x 1024 pixels and Z-planes 0.5µm apart. GFP was excited using a 488 nm laser, Cy3 were excited using a 561 nm laser, Quasar670 was excited using a 633 nm laser.

Images in Fig. 4b and Supplementary Fig. 6a-b represent a maximum intensity projection of a stack of 15 z-planes (≈9.5 µm).

H3S10ph immunostaining and mitotic embryos sorting

A pool of 1h30-2h30 AEL embryos were fixed as for immunostaining except the fixation was in 1:1 heptane:1.8% formaldehyde/1X PBS (Thermo Scientific 28906) for exactly 10 min shaking at 450 rpm. Then embryos were rapidly quenched with 125 mM glycine PBS-1x and shaken for 1 min by hand. An anti-phospho-Histone H3 (Ser10) antibody (Cell Signalling #9701) was used at a dilution 1:200. Anti-mouse Alexa 488-conjugated (Life technologies, A21202) was used as a secondary antibody at a dilution 1:500. Embryos were kept in PBT until sorting.

Sorting was done using a COPAS SelectInstrument (Biometrica) with the following parameters: sorting limit low: 1, high: 256; PMT control: Green 650, Yellow 425 and Red 800. A restricted area of sorting (with the highest green signal) was selected representing ≈ 8% of the total population. A container was placed at the output of the non-selected embryos in order to re-pass them through the sorter to collect non-green embryos corresponding to interphase embryos. Right after the sorting, embryos were manually checked under a Leica Z16 APO microscope by placing them on a glass cup and using Drummond Microcaps® micropipettes to remove mis-sorted embryos individually. 1000 embryos per tube were then dried by removing the PBT and kept at -80 °C.

Chromatin Immuno-Precipitation and library preparation

1000 embryos were homogenized in 1 ml of Buffer A (60 mM KCl, 15 mM NaCl, 4 mM MgCl₂, 15 mM HEPES (pH 7.6), 0.5% Triton X100, 0.5 mM DTT, 10 mM Sodium Butyrate and

Protease Inhibitors Roche 04693124001) using a 2 ml Dounce on ice. The solution was then centrifuged 4 min at 2000g at 4 °C. Supernatant was removed and 1 ml of Buffer A was added and this was repeated two times with Buffer A and once with Lysis Buffer without SDS (140 mM NaCl, 15 mM HEPES (pH 7.6), 1 mM EDTA (pH 8), 0.5mM EGTA, 1% Triton X100, 0.5 mM DTT, 0.1% Sodium Deoxycholate, 10 mM Sodium Butyrate and Protease Inhibitors). The pellet was resuspended in 200 µl of Lysis Buffer with 0.1% SDS and 0.5% N-Laurosylsarcosine and incubated 30 min at 4 °C on a rotating wheel. Sonication was done with a Bioruptor® Pico sonication device with 30 sec ON/30 sec OFF cycles for 6-7 min for interphase and 8-9 min for mitotic chromatin. Sonicated chromatin was then centrifuged 5 min at 14000 rpm at 4 °C. The chromatin was then diluted in 1 ml of Lysis Buffer.

Dynabeads® M-270 Epoxy (Invitrogen Life Technologies™, 14301) were prepared in order to directly crosslink antibodies to the beads (anti-GAF, gift from G. Cavalli, or anti-H4K8ac, abcam 15823), avoiding cross reaction with the H3S10ph antibody, following manufacturer protocol. Prior to this, anti-GAF was purified using NAb™ Protein A/G Spin Kit (ThermoScientific). Once the magnetic beads were cross-linked, chromatin was incubated over night at 4 °C on a rotating wheel. Then, beads were washed 7 min at 4 °C once in Lysis Buffer, once in FAT Buffer (1 M TrisHCl pH 8, 0.5 M EDTA pH 8, SDS 10%, 5 M NaCl, 10% Triton), once in FA Buffer (1 M HEPES, pH 7.0-7.6, 5 M NaCl, 0.5 M EDTA pH8, Triton X-100 – 10% NaDeoxycholate) once in LiCl Buffer (1 M Tris-HCl pH 8, 4 M LiCl, 10% Nonidet-P40-Nonidet, 10% NaDeoxycholate and protease inhibitors) and twice in TE (10 mM Tris-HCl pH 8, 0.1 mM EDTA). Elution was done in elution Buffer 1 (10 mM EDTA, 1% SDS, 50 mM Tris-HCl pH 8) for 30 min at 65 °C at 1300 rpm. Eluted chromatin was removed and a second elution step with Elution Buffer 2 (TE, 0.67% SDS) was performed. The two elutions were pooled. Chromatin was then reverse-crosslinked by heating overnight at 65 °C. Next, chromatin was incubated 3 h at 50 °C with ProteinaseK (Thermo Scientific™ EO0491) and RNaseA (Thermo Scientific™ EN0531). DNA was then extracted with phenol/chloroform purification. Biological duplicates were performed for each sample.

Libraries were then prepared using the NEBNext Ultrall DNA Library Prep Kit for Illumina, following the manufacturer's instructions. Sequencing was performed on Illumina HiSeq 4000 on pair-end 75 bp.

ChIP-seq analysis

Both reads from ChIP-seq and Input experiments were trimmed for quality using a threshold of 20 and filtered for adapters using Cutadapt (v1.16). Reads shorter than 30 bp after trimming were removed. Reads were mapped to *Drosophila melanogaster* genome (dm6 release) using Bowtie2 (Langmead and Salzberg 2012). Aligned sequences were processed with the R package PASHA to generate the used wiggle files (Fenouil et al. 2016). Pasha elongates *in silico* the aligned reads using the DNA fragment size estimated from paired-reads. Then, the resulting elongated reads were used to calculate the coverage score at each nucleotide in the genome. Wiggle files representing average enrichment score every 50 bp were generated. In order to normalize the enrichment scores to reads per million, we rescaled the wiggle files using PASHA package. Besides, in order to reduce the over-enrichment of some genomic regions due to biased sonication and DNA sequencing, we subtracted from ChIP sample wiggle files the signal present in Input sample wiggle files. The Rescaled and Input subtracted wiggle files from biological replicate were then used to generate the final wiggle file representing the mean signal.

In order to call the enriched peaks from the final wiggle files, we used *Thresholding* function of the Integrated Genome Browser (IGB) to define the signal value over which we consider a genomic region to be enriched compared to background noise (*Threshold*). We used also the minimum number of consecutive enriched bins to be considered an enriched region (*Min.Run*) as well as the minimum gap above which two enriched regions were considered to be distinct (*Max.Gap*). The three parameters were then used with an in-house script that realizes peak calling by using algorithm employed by *Thresholding* function of IGB.

Peaks calling was done with a threshold of 100 for GAF-ChIP-seq and 22 for H₄K8ac-ChIP-seq, minimum run of 50 bp and maximum gap of 200bp. Interphase only peaks correspond to peaks from interphase ChIP-seq with no overlap with peaks from mitotic ChIP-seq. Mitotically retained correspond to interphase peaks with an overlap (min 1 base pair) with peaks from mitotic ChIP-seq. Mitotic only peaks correspond to peaks from mitotic ChIP-seq with no overlap with peaks from interphase ChIPseq.

Motif search was done with the MEME ChIP tool (MEME suite 5.1.1).

Peaks were considered as promoter if overlap with the region defined by 100 bp around TSS. Peaks were considered as enhancers if overlapping with identified enhancer (Kvon et al. 2014) and/or overlapping with a H3K27ac peak (Koenecke et al. 2016a).

ATAC-seq data are from (Blythe and Wieschaus 2016)(GSE83851). Wig files were converted to BigWig using Wig/BedGraph-to-bigWig converter (Galaxy Version 1.1.1). ATACseq mean signal was then plotted on regions of interest (mitotically retained peak coordinates and Interphase only coordinates) using computeMatrix by centering ATAC-seq signal to the center of the regions (and +/- 1 kb) followed by plotProfile (Galaxy Version 3.3.2.0.0).

Mitotically retained GAF peaks were subdivided by *k*-means clustering based on chromatin state (H3K27ac and H3K27me3 ChIP-seq (Li et al. 2014)) and transcriptional status (nc14 RNA-seq (Lott et al. 2011)) using deepTools (Ramírez et al. 2016). Peaks were partitioned into three clusters: cluster 1, n=1073, cluster 2, n=612 and cluster 3, n=732. To further characterize mitotically retained clusters we plotted heatmaps using deepTools (Ramirez et al., 2016) for publicly available ChIP-seq data for H3K27ac (Li et al. 2014), H3K27me3 (Li et al. 2014), Pc (Koenecke et al. 2016b) and ATAC-seq (Blythe and Wieschaus 2016).

GAF bound loci for which accessibility are dependent on GAF were taken from (Gaskill et al. 2021).

Whole-genome data used in this study

Experiment	Stage	Reference
TSS	-	GCF_000001215.4 (Release_6)
GAF-ChIP-seq	2-4 hr AEL, WT embryos	<i>Koenecke et al. 2017</i>
H3K27ac-ChIP-seq	2-4 hr AEL, tol10b embryos	<i>Koenecke et al. 2016</i>
H3K27ac-ChIP-seq	2-4 hr AEL, gd7 embryos	<i>Koenecke et al. 2016</i>
Enhancer reporter transgenes	Embryos	<i>Kvon et al. 2014</i>
Hi-C	nc12, nc13, nc14	<i>Hug et al. 2017</i>
Insulators	Embryos	<i>Negre et al. 2010</i>
ATAC-seq	nc8, 9, 10...13 every 3min	<i>Blythe & Wieschaus 2016</i>
ATAC-seq degrad-FP	2-2.5 hr AEL	<i>Gaskill et al. 2021</i>
H3K27ac	nc14	<i>Li et al. 2014</i>

H3K27me3	nc14	<i>Li et al. 2014</i>
RNA-seq	nc14	<i>Lott et al. 2011</i>

Live imaging

Movies of *His2Av-mRFP; sfGFP-GAF_CRISPR* (related to Movie 1, Movie 2 and Fig. 1b) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mRFP were excited using a 488 nm and 561 nm laser respectively with the following settings: 256 x 256-pixel images, 15 z-planes 1 µm apart and zoom 4x, resulting in a time resolution of 9.5 sec per Z-stack. Average intensity profiles were measured for histones, nucleoplasmic GAF and cytoplasmic GAF from three movies of embryos transitioning from nc13 into nc14. An automatic tracking of maximum intensity projected images fluorescence was done using a home-made software as in (Dufourt et al. 2018). First a detection of nuclei is made using His2Av-mRFP allowing the monitoring of histone intensity fluctuation, then a mask of His2Av-mRFP detected nuclei was projected on the sfGFP-GAF channel allowing the recovery of sf-GFP-GAF present on histones. Finally, five ROI in each movie corresponding to cytoplasmic regions were tracked for sfGFP-GAF intensity in the cytoplasm.

Movies of *MCP-eGFP-His2Av-mRFP>snail-primary-enhancer_MS2/+* embryos (related to Supplementary Fig. 7e) were acquired using a Zeiss LSM780 with confocal microscope with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mRFP were excited using a 488 nm and 561 nm laser respectively with the following settings: 512 x 512-pixel images, 21 z-planes 0.5µm apart and zoom 2.1x, resulting in a time resolution of 22 sec per frame. Movies were subjected to filtering steps to track transcription foci as 128XMS2 loops result in signal retention during mitosis.

Movies of *MCP-eGFP-His2Av-mRFP>scyl_MS2_CRISPR/+* in *RNAi-White* and *RNAi-GAF* background (related to Movie 3 and 4 and to Fig. 4 and 5 and Supplementary Fig. 6c-f and 7c) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mRFP were excited using a 488 nm and 561 nm laser respectively with the following settings: 552 x 552-pixel images, 21 z-planes 0.5 µm apart and zoom 2.1x, resulting in a time resolution of 5.45 sec per frame. As we observed that GAF knock-down was not complete (some *RNAi-GAF* embryos gastrulate and

develop), movies showing visible developmental defects, such as nuclear dropout, anaphase bridges or failure to gastrulate, were kept for analysis.

Memory movies analysis

Movies were analyzed using Mitotrack (Trullo et al. 2020) as in (Dufourt et al. 2018) with newly implemented tools to filter mitotic 128XMS2 foci in movies of *MCP-eGFP-His2Av-mRFP>snail-primary-enhancer_MS2/+* embryos (mitotic foci are now detected with the 24MS2 array). Briefly, using a custom-made algorithm developed in Python™ and implemented in the MitoTrack software, nuclei were segmented and tracked in 2D, working on the maximum intensity projected stack. Transcription spots were detected and tracked in 3D. All the spots present during mitosis were removed in the successive cycle such that only *de novo* appearing MS2 punctae were analyzed.

For intensity analysis (related to Supplementary Fig. 6f and 7c) the intensity of detected spots was collected for each frame to study the transcriptional intensity behavior throughout nuclear cycle 14. Nuclei coming from inactive and nuclei coming from active were separated for Supplementary Fig. 6f and pooled for Supplementary Fig. 7c.

qRT-PCR in RNAi embryos

Total RNA from 0-2h AEL *RNAi-white* or *RNAi-GAF* driven by *nos-Gal4* and *mat-alphaTub-Gal4* embryos was extracted with TRIzol following the manufacturer's instructions. RNA was DNase-treated. 1 µg of RNA extracted from ~300 embryos per replicate was reverse transcribed using SuperScript IV and random primers. Quantitative PCR analyses were performed with the LightCycler480 SYBR Green I Master system (primers used listed in Supplementary table 1, targeting both isoforms of *GAF*). RNA levels were calculated using the *RpL32* housekeeping gene as reference and not bound by *GAF* according to the *GAF*-ChIP-seq. Each experiment was performed with biological triplicates and technical triplicates.

Western blot analysis

Fifty embryos from *RNAi-white* or *RNAi-GAF* driven by *nos-Gal4* and *mat-alphaTub-Gal4* 0-2h AEL embryos were collected and crushed in 100µl of NuPAGE™ LDS sample buffer and reducing agent. Samples were heated 10min at 70°C, and the volume-equivalent of 5

embryos was loaded per well on a 4-12% Bis-Tris NuPAGE™ Novex™ gel and ran at 180V. Protein transfer was done for 1h10 at 110V to a nitrocellulose membrane, 0.2 μm (Invitrogen, LC2000). Membrane was blocked in 5% milk-PBT (PBS 1X 0.1% Tween 20) for 40 min and incubated overnight at 4°C with primary antibodies 1/2000 mouse anti-GAF or 1/2000 mouse anti-Tubulin in PBT. Anti-mouse and -rabbit IgG-HRP (Cell Signaling #7076 and #7074) secondary antibody were used at 1/4000 and incubated 1hour at room temperature. Chemiluminescent detection was done using Pierce™ ECL Plus (ThermoFisher) kit.

DNA probe preparation and DNA-FISH

Probes were generated using 4 to 6 consecutive PCR fragments of 1.2 to 1.5 kb from *Drosophila* genomic DNA, covering approximately a 10 kb region. Primers are listed in Supplementary table 1. Probes were labeled using the FISH Tag DNA Kit (Invitrogen Life Technologies, F32951) with Alexa Fluor 488, 555, and 647 dyes following manufacturer's protocol. Probes for satellite regions (related to Supplementary Fig. 1b) are from (Garavís et al. 2015).

DNA FISH was performed on 0-4 h AEL yw embryos adapted from (Bantignies and Cavalli 2014). Briefly, embryos were fixed as described above and were rehydrated with successive 3-5 min 1 ml washes on a rotating wheel with the following solutions: (1) 90% MeOH, 10% PBT; (2) 70% MeOH, 30% PBT; (3) 50% MeOH, 50% PBT; (4) 30% MeOH, 70% PBT; (5) 100 % PBT. Embryos were subsequently incubated in 200 μg RNase A (Thermo Scientific, EN0531) in 1 ml PBT for 2 h then 1 h at room temperature on a rotating wheel. Embryos were then slowly transferred to 100% pHM buffer (50% Formamide, 4x SSC, 100 mM NaH₂PO₄, pH 7.0, 0.1% Tween 20) in rotating wheel 20 min per solution 1 (1) 20% pHM, 80% PBS-Triton; (2) 50% pHM, 50% PBS-Triton; 80% pHM, 20% PBS-Triton; 100% pHM.

Cellular DNA and probes were respectively denaturated in pHM and FHB (50% Formamide, 10% Dextran sulfate, 2x SSC, 0.05% Salmon Sperm DNA) for 15 min at 80 °C.

Probes and embryos were quickly pooled in the same PCR tube and slowly hybridized together with the temperature decreasing 1 °C every 10 min to reach 37 °C in a thermocycler. Washes were performed in pre-warmed solution (1 to 4) at 37 °C for 20 min under 900 rpm agitation (1) 50% Formamide, 10% CHAPS 3%, 10% SSC; (2) 40% Formamide, 10% CHAPS

3%, 10% SSC; (3) 30% Formamide; (4) 20% Formamide; then 20 min on a rotating wheel at room temperature using (5) 10% Formamide; (6) PBT; (7) PBS-Triton. Embryos were stained with DAPI at 0.5µg/ml, washed in PBT and mounted between slide and coverslip.

Images were acquired using a Zeiss LSM880 with confocal microscope in Airyscan mode with a Plan-Apochromat 63x/1.4 oil objective lens with the following settings: zoom 3.0x, z-planes 0.3 µm apart, 1024x1024 pixels.

Distance measurements

To measure the distances between probes (*scyl-chrb* and *chrb-ctrl*, or *esg-sna* and *sna-ctrl*), we used a custom-made software developed in Python™. This software is available through this link: https://github.com/ant-trullo/DNA_FishAnalyzer.

All the probes channels were treated with a 3D Laplacian of Gaussian filter (with kernel size 1) and then detected in 3D with manual thresholding on the filtered matrices; for each of the detected spots, the center of mass was determined. DAPI signal was treated with a 3D Gaussian filter (with user-defined kernel size) and the logarithm of the resulting matrix is thresholded with an Otsu algorithm, the threshold value being adjusted separately in each frame. The logarithm was used in order to compensate for non-homogeneous intensity inside nuclei. In order to generate distances, all the spots outside nuclei were removed. Then, nearest mutual neighbor spots were selected by calculating the distances of all the possible couples of spots and picking the smallest set. The distances were calculated with respect to the center of mass and using the Euclidean distance, taking into account the different pixel size on the z axis. A minimum of 10 images from 5 different embryos were analyzed for each condition. Aberrant distances (superior to 1µm) were not considered.

Mathematical modeling of mitotic memory

The analysis of the lag time between mitosis and first initiation is 'non-stationary' and was previously modeled with mixed gamma distributions for the time from mitosis to transcriptional activation in n.c. 14 (Dufourt et al. 2018, Bellec et al. 2018). This mathematical framework describes the waiting times prior to the first detected transcriptional initiation event in n.c. 14, as a sequence of discrete transitions. This model considers that the activation of transcription follows a stepwise progression through a series of non-productive OFF events

that precede activation (ON event). In this model, the average number of rate-limiting transitions is provided by the parameter 'a' while the duration of the transitions by parameter 'b' (expressed in seconds and considered the same for all transitions).

We are interested in the post-mitotic delay, defined as the time needed for post-mitotic transcription (re)activation. We model this time as the sum of two variables as in (Dufourt et al. 2018):

$$T_a = T_0 + T_r \quad (9)$$

where T_0 is a deterministic incompressible lag time, the same for all nuclei, and T_r is a random variable whose value fluctuates from one nucleus to another. The decomposition in Eq. 9 can be justified by the experimental observation that all the reactivation curves (Fig. 5a, Supp7c) start with a nonzero length interval during which no nuclei are activated. Furthermore, T_r is defined such that it takes values close to zero with non-zero probability. This property allows us to set T_0 to the instant when the first nucleus initiates transcription, in order to determine T_r . The random variable T_r is modeled using a finite state, continuous time, Markov chain. The states of the process are $A_1, A_2, \dots, A_{n-1}, A_n$. The states $A_i, 1 \leq i \leq n-1$ are OFF, i.e. not transcribing. The state A_n is ON, i.e. transcribing. Each OFF state has a given lifetime, defined as the waiting time before leaving the state and going elsewhere. Like in (Dufourt et al. 2018), we considered that each of the states has the same lifetime denoted τ . Also, the transitions are considered linear and irreversible: in order to go to A_{i+1} one has to visit A_i and once there, no return is possible. The time T_r is the time needed to reach A_n starting from one of the OFF states. The predictions of these models were compared to the empirical survival function $S_{exp}(t)$ defined as the probability that $T_r > t$, obtained using the Meier-Kaplan method from the values T_r for all the analyzed nuclei. Following Occam's razor principle, we based our analysis on the simplest model that is compatible with the data, which is a model with $n=4$ and homogeneous lifetimes. For this model the theoretical survival function is a mixture of Gamma distributions:

$$S(t) = p_1 \exp(-t/\tau) + p_2 \left(1 - \frac{1}{\Gamma(2)} \gamma(2, t/\tau) \right) + p_3 \left(1 - \frac{1}{\Gamma(3)} \gamma(3, t/\tau) \right) \quad (10)$$

where γ , Γ are the complete and incomplete gamma functions and p_1, p_2, p_3 (satisfying $p_1+p_2+p_3=1$) are the probabilities to reach ON after one, two or three jumps, respectively.

We have also tested more complex models, with uneven lifetimes, more states and therefore, more parameters. However, the complex models did not provide a sensibly better fit with data and generated overfitting identified as large parametric uncertainty.

The parameters 'a' and 'b', summarizing the statistics of the post-mitotic reactivation are defined as

$$a = p_1 + 2p_2 + 3p_3, b = \tau \tag{11}$$

The uncertainty intervals shown in Supplementary Figure 7D express how constraining is the cost function and how sensitive are the parameters. If an interval is close to zero, this means that the parameter is very sensitive, a change of this parameter has a strong influence on the cost function. In general, there are no methods to find parameter statistical confidence intervals in frequentist approaches; it is possible in Bayesian approaches that we did not use. The relative uncertainty due to multiple local optima is provided in percentage, for each parameter. The goodness of fit is given by the sum of squares distance O.

This model was previously developed for transcriptional memory of snail transgenes in Dufourt et al. 2018. We wanted to be able to compare our transcriptional timing data with the ones that has been done in the lab. Specifically, in Dufourt et al. study, the model suggested that descendants of active nuclei have fewer and shorter transitions prior to activation than descendants of inactive nuclei. Moreover, they showed that transcriptional memory can be modeled but was not dependent on Zelda. We tested this model on my GAF related data, and found that transcriptional memory is dependent on the duration of the transition between descendants of active and inactive rather than the number of transitions. We therefore were able to estimate on which parameter transcriptional memory is from.

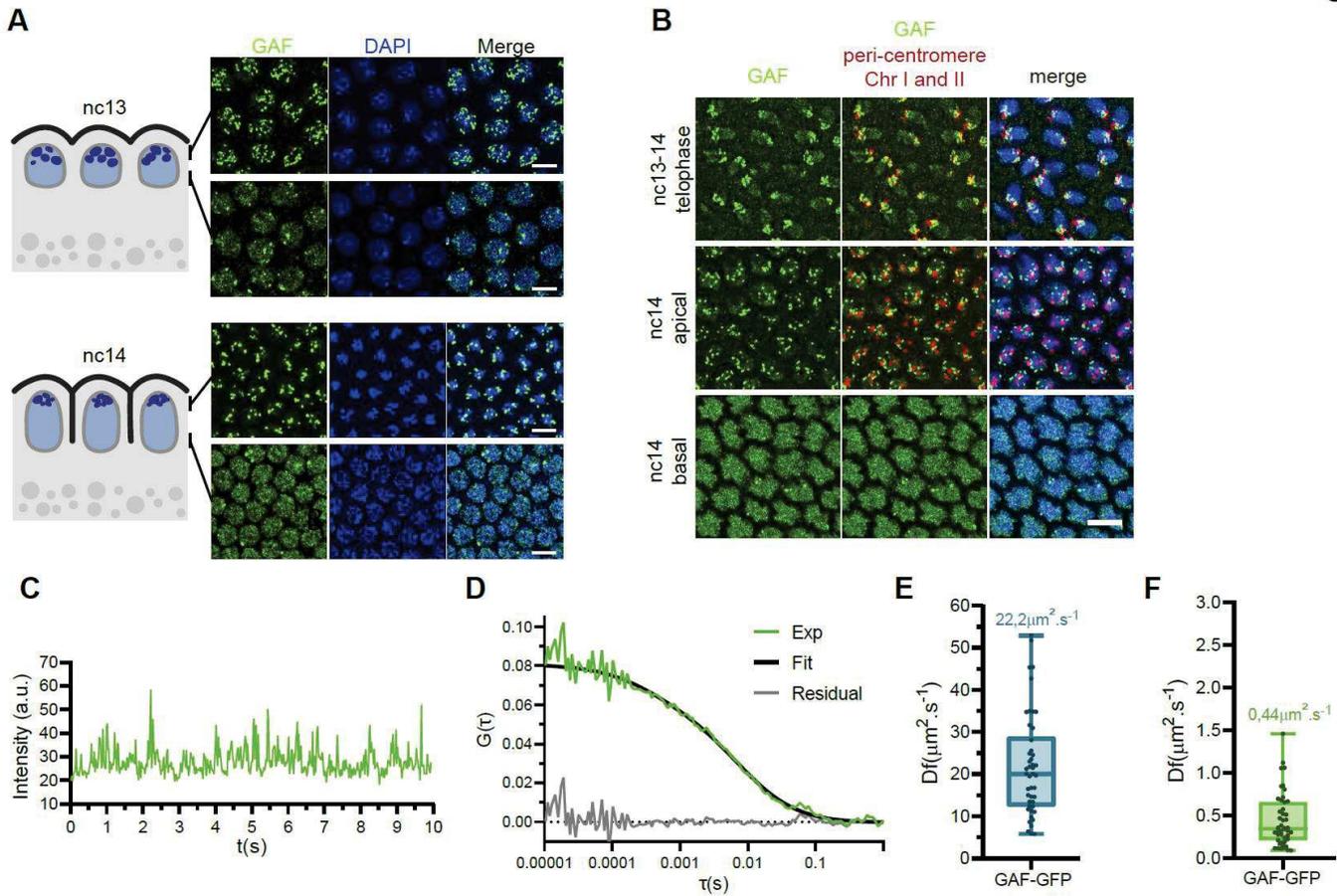


Figure Sup1: GAF puncta localize to heterochromatin nuclear regions

(A) Maximum intensity projected Z-planes of confocal images showing GAF immunostaining (green) in wild type embryos at the indicated stages counterstained with DAPI (blue). Upper panels are images taken at the apical side of the nuclei, bottom panels at the basal side. Scale bars are 5 μ m.

(B) Maximum intensity projected Z-planes of confocal images from DNA-immunoFISH with peri-centromeric probes ⁶⁷ (red) and anti-GAF (green) on wild type embryos and at the indicated nuclear cycles. Scale bar is 5 μ m.

(C) Example of an intensity time trace obtained from FCS in a GAF-GFP nc14 embryo.

(D) Example of autocorrelation function (green curve) related to (c) (black curve represents fitting using a double-diffusion model).

(E) Estimated fast diffusion time for GAF-GFP extracted after fitting FCS data with a double-diffusion model performed in nucleoplasm. Centered horizontal line of the box plot represents the median.

(F) Estimated slow diffusion time for GAF-GFP extracted after fitting FCS data with a double-diffusion model performed in nucleoplasm. Centered horizontal line of the box plot represents the median.

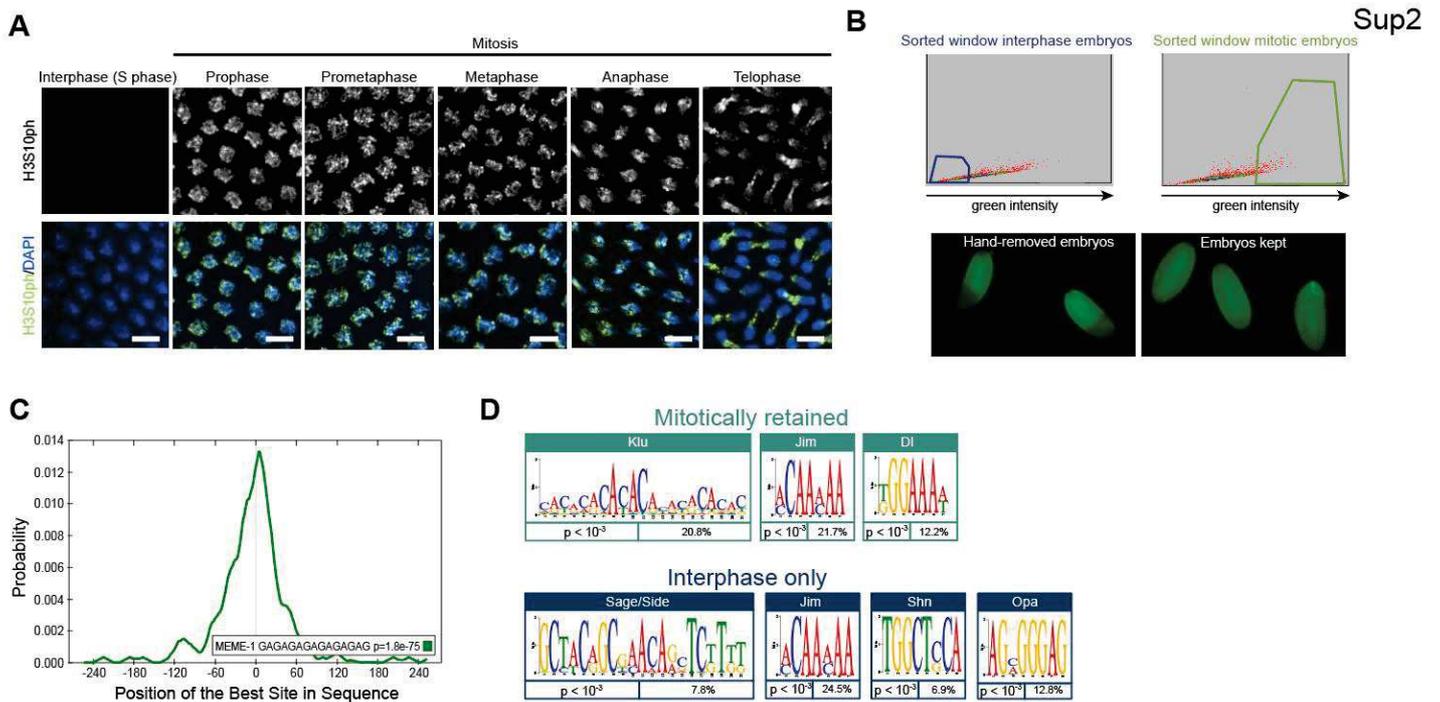


Figure Sup2: Characterization of GAF mitotically retained regions

(A) Maximum intensity projected Z-planes of confocal images of H₃S10ph immunostaining (grey and green) in wild type embryos at the indicated mitotic stages counterstained with DAPI (blue). Scale bars are 5µm.

(B) (Top panels) Isolation of interphase and mitotic embryos previously stained with a H₃S10ph antibody with a flow cytometer. Highlighted squares show the range for selected embryos for the interphase population (blue) and the mitotic population (green). (Bottom panels) Representative embryos stained with an H₃S10ph antibody, undergoing mitotic waves (left image) (removed during the hand sorting) and fully mitotic embryos (right image).

(C) Profile of probability for the GAGAG motif reported by MEME within sequences from GAF-ChIP-seq peaks.

(D) De-novo motif enrichment within GAF mitotically retained and interphase only peaks, as reported by MEME.

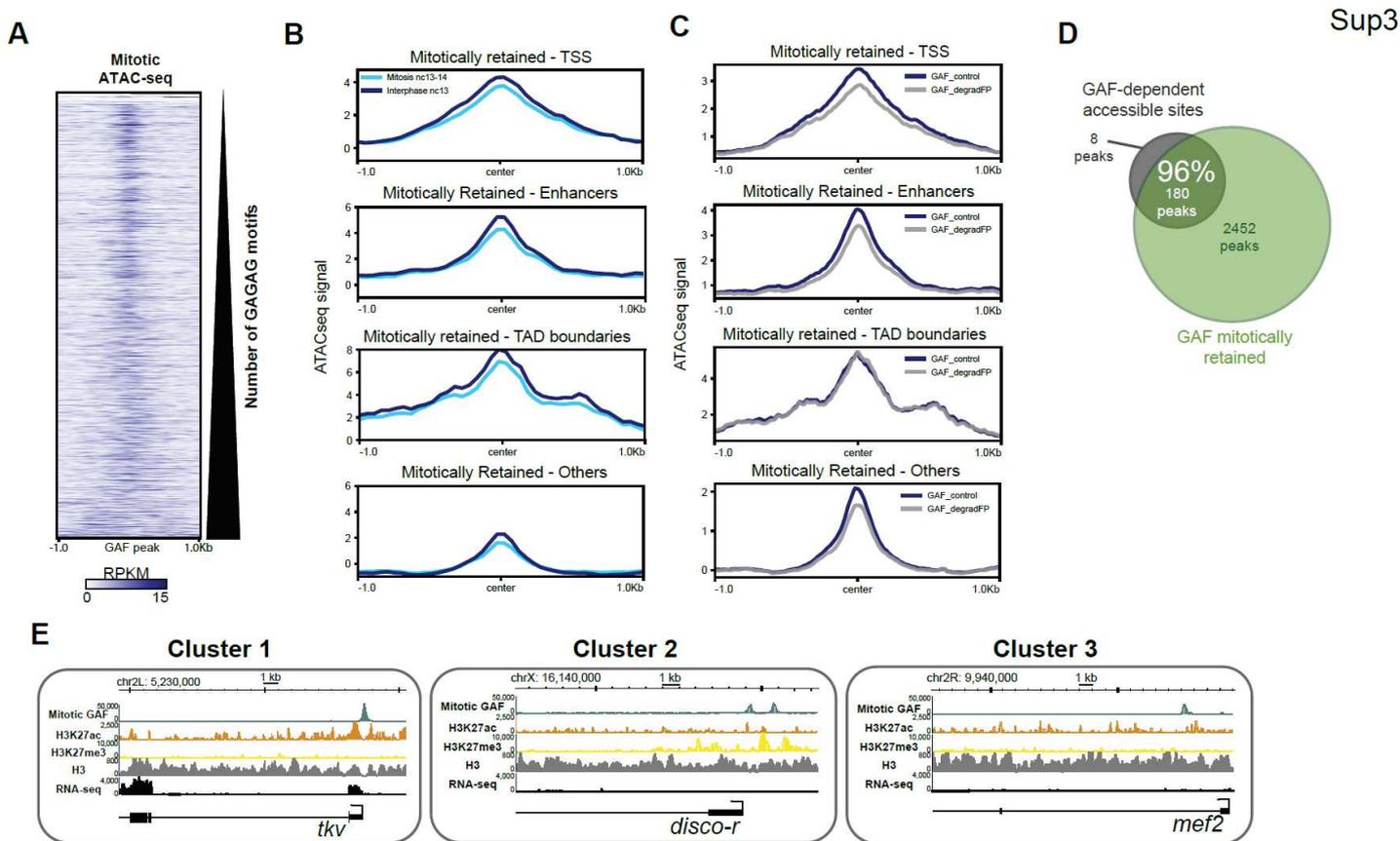


Figure Sup3: Chromatin states of GAF mitotically retained regions

(A) Heatmap of ATAC-seq signal⁴⁷ at GAF mitotically retained regions, partitioned by the number of GAGAG motifs they contain.

(B) Metagene profiles of the ATAC-seq signal⁴⁷ at GAF mitotically retained regions partitioned by the type of cis-regulatory element they overlap. Plots show peaks located at TSS, TAD boundaries⁴⁴, enhancers^{46,49} and others that were unassigned.

(C) Metagene profiles of the ATAC-seq signal in control (GAF control, blue) or GAF depleted (GAF_degradFP, grey)²⁵ embryos in different categories of GAF mitotically retained regions defined in Fig. 2b.

(D) Venn diagram of the overlap between GAF dependent accessible loci²⁵ (grey) and GAF mitotically retained peaks (green).

(E) Genome browser images showing ChIP-seq signals of mitotic GAF, H3K27ac, H3K27me₃ and H₃⁴⁹, alongside RNA-seq⁵⁰ at three loci containing GAF mitotically retained peaks. These three examples represent each of the three clusters defined by distinct chromatin states.

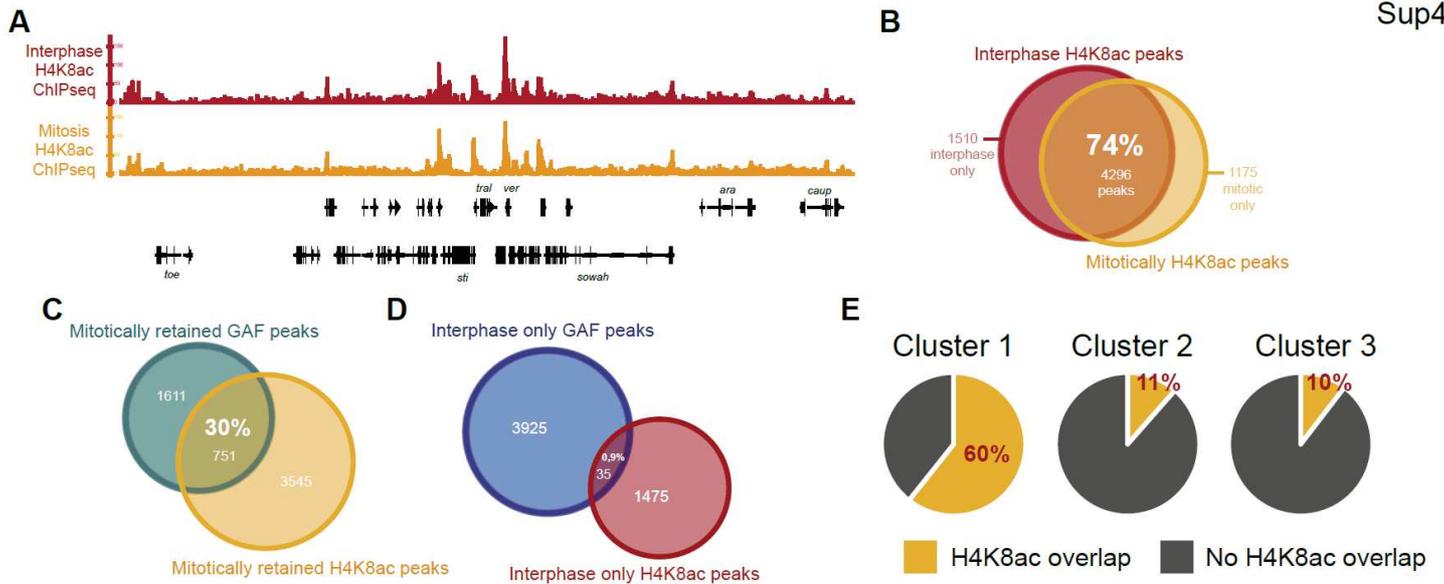


Figure Sup4: H₄K8ac-ChIP-seq in interphase and mitotic embryos

(A) Genome browser image showing H₄K8ac-ChIP-seq profiles of interphase and mitotic embryos at a representative genomic region containing mitotically retained H₄K8ac peaks.

(B) Venn diagram of the overlap between H₄K8ac interphase (red) and mitotic (orange) ChIP-seq peaks.

(C-D) Venn diagram of the overlap between mitotically retained and interphase only (c) or H₄K8ac and GAF-ChIP-seq peaks (D).

(E) Pie charts of the overlap between H₄K8ac-ChIP-seq mitotic peaks and GAF mitotically retained peaks, after their partitioning into three clusters.

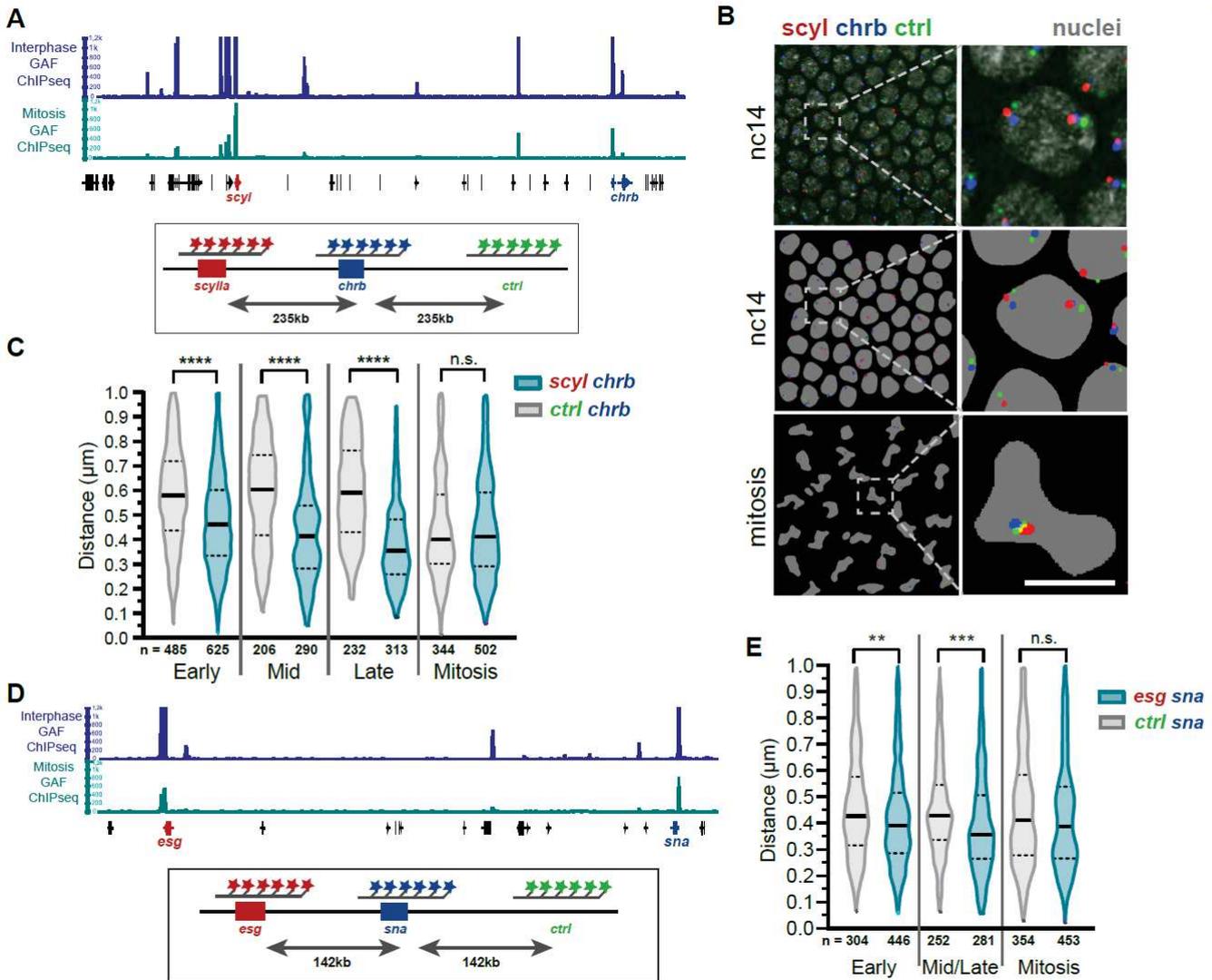


Figure Sup5: Absence of mitotic loops for two GAF mitotically retained loci

(A) (Top) Genome browser image of the interphase and mitotic GAF- and H₄K8ac-ChIP-seq profiles at the *scyl*-*chrB* locus. (Bottom) Schematic of the *scyl*-*chrB* locus with the indicated designed probes for DNA-FISH. The spatial distance between *scyl* and *chrB* was compared to that of a non-bookmarked region located at an equivalent distance (235kb).

(B) DNA-FISH image in *nc14* wild type embryo with *scyl* labeled in red, *chrB* in blue and control region in green. Bottom images represent the same image false colored after analysis as well as nuclei in mitosis. Scale bar is 5µm.

(C) Violin plot representing the distance between scyl-chrb and chrb-ctrl from images taken in early, middle, late n.c.14 interphase or mitotic wild type embryos. Two tailed Welch's t-test **** $p < 0.0001$.

(D) (Top) Genome browser image of the interphase and mitotic GAF-ChIP-seq profiles at the sna-esg locus. (Bottom) schematic of the sna-esg locus with the positions of designed probes for DNA-FISH indicated.

(E) Violin plot representing the distance between sna-esg and sna-ctrl from images taken in nc14 interphase or mitotic wild type embryos. Two tailed Welch's t-test *** $p < 0.001$, ** $p < 0.01$.

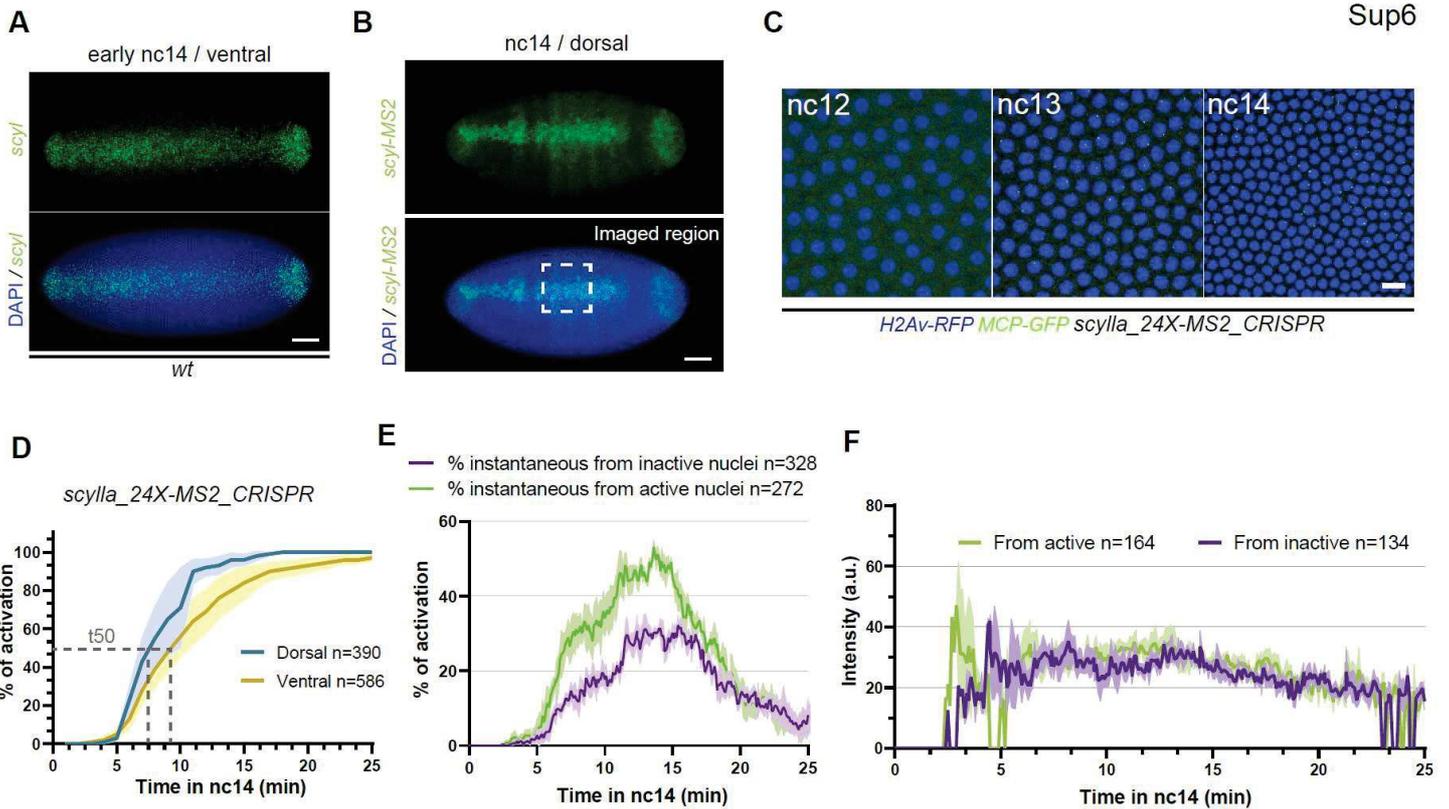


Figure Sup6: *scylla* transcription in the early embryo

(A) Maximum intensity projected Z-planes of confocal images from smiFISH with *scyl* probes (green) counterstained with DAPI (blue) on wild type embryos in early nc14 on the ventral side. Scale bars are 50 μ m.

(B) Maximum intensity projected Z-planes of confocal images from smiFISH with *scyl* probes (green) counterstained with DAPI (blue) on wild type embryos in nc14 on the dorsal side. Dashed square represents the regions imaged for the quantifications shown in (D). Scale bars are 50 μ m.

(C) Snapshots from a representative movie of *scylla_24X-MS2_CRISPR/+* embryo carrying MCP-eGFP, H2Av-RFP. Nuclei are visualized in blue and transcription sites in green. Scale bar corresponds to 10 μ m.

(D) Quantification of the transcriptional synchrony after mitosis in *scylla_24X-MS2_CRISPR* embryos in dorsal (blue) and ventral (yellow) regions. Both daughters of each nucleus are quantified. SEM are represented in light blue and light yellow respectively. n=number of nuclei analyzed from 3 movies of 3 embryos for each condition.

(E) Instantaneous percentage of activation after mitosis of nuclei coming from active nuclei (green) and from inactive (purple). SEM are represented in light purple and light green respectively. n=number of nuclei analyzed from 3 movies of 3 embryos.

(F) Mean intensity of scyl transcriptional site of nuclei coming from active nuclei (green) and from inactive nuclei (purple). SEM are represented in light green and light purple. Both of the two daughters of each nucleus are quantified. n=number of nuclei analyzed from 4 movies of 4 embryos.

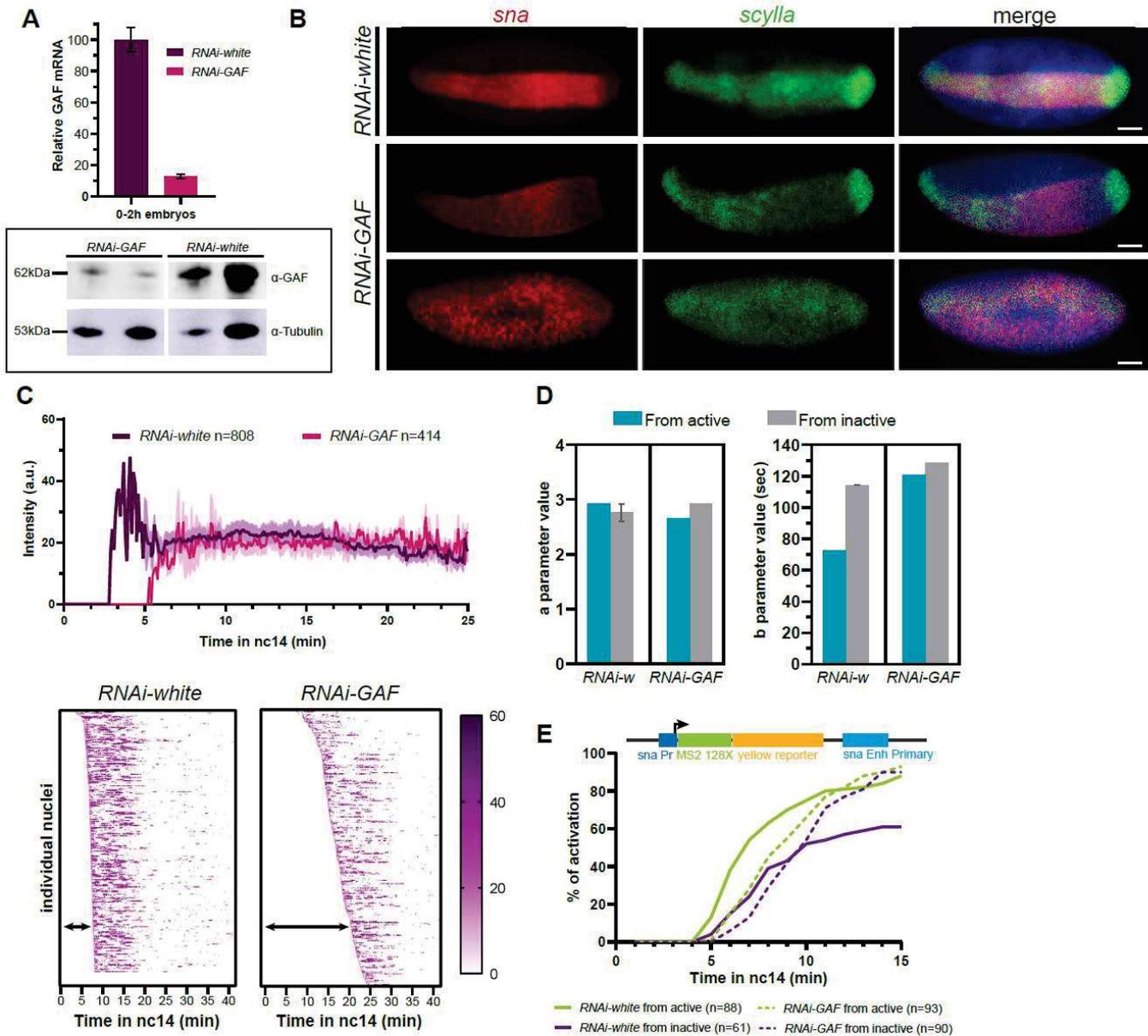


Figure Sup7: Effect of GAF reduction on *scylla* transcription

(A) (Top) Histogram of the relative amount of RPL32 transcripts normalized GAF mRNA in RNAi-white and RNAi-GAF 0-2h embryos quantified by RT-qPCR. (Bottom) Two examples of western blot analysis of RNAi-white and RNAi-GAF 0-2h embryos, with the indicated antibodies. Each experiment was performed in biological triplicates.

(B) Maximum intensity projected Z-planes from confocal images from a smFISH with *scylla* probes (red) and *sna* probes (green) counterstained with DAPI (blue) on RNAi-white and RNAi-GAF embryos, showing different types of phenotypes. Scale bars are 50µm.

(C) (Top) Mean intensity of *scyl* transcriptional site (*scylla_24X-MS2_CRISPR*) of nuclei in RNAi-white embryos (control, purple) and RNAi-GAF embryos (pink). SD are represented in light purple and light pink. Both of the two daughters of each nucleus are quantified. n=number of nuclei analyzed from 4 movies of 4 embryos. (Bottom) Heatmaps of *scyl* (*scylla_24X-MS2_CRISPR*) transcriptional site intensity of individual nuclei in RNAi-white and RNAi-GAF embryos sorted by their first activation time.

(D) Histograms of 'a' and 'b' parameters extracted from mathematical modeling for the *scyl* (*scylla_24X-MS2_CRISPR*) gene in RNAi-white and RNAi-GAF *scylla_24X-MS2_CRISPR* embryos. Error intervals correspond to variation among optimal and best suboptimal fits.

(E) (Top) Schematic representing the snail-primary-enhancer_MS2 transgene. (Bottom) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-white embryos (control, solid curves) and RNAi-GAF embryos (dashed curves). n=number of nuclei analyzed from 4 movies of 4 embryos.

Movie Legends

Movie1: Imaging GAF behavior during the cell cycle

Maximum intensity projection of confocal live imaging of a developing *His2Av-mRFP;GAF-GFP* embryo. Scale bar is 5µm.

Movie2: GAF subnuclear localization

Maximum intensity projection of confocal live imaging of *His2Av-mRFP;GAF-GFP* embryo. Top movie comprises a six Z-planes projected images at the apical side of nuclei, bottom movie is a six Z-planes projected images at the basal side of nuclei. Time is in minutes. Scale bar is 5µm.

Movie3: Transcription of *scylla* in a control embryo

Maximum intensity projection of confocal live imaging of a *mat-alpha-Gal4/+; nos-Gal4, MCP-eGFP, H2Av-RFP/UASp-shRNA-white > scylla_24X-MS2_CRISPR/+* embryo.

Nuclei are visualized in red and transcriptional sites in green. Scale bar is 10µm.

Movie4: Transcription of *scylla* in a GAF maternally depleted embryo

Maximum intensity projection of confocal live imaging of a *mat-alpha-Gal4/+; nos-Gal4, MCP-eGFP, H2Av-RFP/UASp-shRNA-GAF > scylla_24X-MS2_CRISPR/+* embryo.

Nuclei are visualized in red and transcriptional sites in green. Scale bar is 10µm.

Supplementary Tables Legends

Supplementary table 1: Primers sequences for cloning, DNA-FISH and smFISH probes sequences used in this study.

Supplementary table 2: Identified GAF mitotically retained, interphase only and mitotic only peak coordinates with the nearest gene identified and its distance in base pair from the TSS.

Supplementary table 3: Identified GAF mitotically retained, interphase only and mitotic only peak coordinates with their respective features.

Supplementary table 4: Whole-genome data used in this study.

References

1. Bellec, M., Radulescu, O. & Lagha, M. Remembering the past: Mitotic bookmarking in a developing embryo. *Curr. Opin. Syst. Biol.* 11, 41–49 (2018).
2. Festuccia, N., Gonzalez, I., Owens, N. & Navarro, P. Mitotic bookmarking in development and stem cells. *Development* 144, 3633–3645 (2017).
3. Elsherbiny, A. & Dobрева, G. Epigenetic memory of cell fate commitment. *Curr. Opin. Cell Biol.* 69, 80–87 (2021).
4. Raccaud, M. & Suter, D. M. Transcription factor retention on mitotic chromosomes: regulatory mechanisms and impact on cell fate decisions. *FEBS Lett.* 592, 878–887 (2017).
5. Raccaud, M. et al. Mitotic chromosome binding predicts transcription factor properties in interphase. *Nat. Commun.* 10, 1–16 (2019).
6. Cirillo, L. A. et al. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9, 279–89 (2002).
7. Palozola, K. C. et al. Mitotic transcription and waves of gene reactivation during mitotic exit. *Science* (80-.). 358, 119–122 (2017).
8. Zhang, H. et al. CTCF and transcription influence chromatin structure re-configuration after mitosis. *bioRxiv* (2021).
9. Zhang, H. et al. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* 576, 158–162 (2019).
10. Kadauke, S. et al. Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell* 150, 725–737 (2012).

11. Teves, S. S. et al. A stable mode of bookmarking by TBP recruits RNA polymerase II to mitotic chromosomes. *Elife* 7, 1–22 (2018).
12. Festuccia, N. et al. Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* 18, 1139–1148 (2016).
13. Pichon, X., Lagha, M., Mueller, F. & Bertrand, E. A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges. *Mol. Cell* 71, 468–480 (2018).
14. Zhao, R., Nakamura, T., Fu, Y., Lazar, Z. & Spector, D. L. Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nat. Cell Biol.* 13, 1295–1304 (2011).
15. Ferraro, T. et al. Transcriptional Memory in the *Drosophila* Embryo. *Curr. Biol.* 26, 212–218 (2016).
16. Muramoto, T., Müller, I., Thomas, G., Melvin, A. & Chubb, J. R. Methylation of H3K4 Is Required for Inheritance of Active Transcriptional States. *Curr. Biol.* 20, 397–406 (2010).
17. Vallot, A. & Tachibana, K. The emergence of genome architecture and zygotic genome activation. *Curr. Opin. Cell Biol.* 64, 50–57 (2020).
18. Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat. Rev. Genet.* 20, 221–234 (2019).
19. Zaret, K. S. & Mango, S. E. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr. Opin. Genet. Dev.* 37, 76–81 (2016).
20. Iwafuchi-Doi, M. & Zaret, K. S. Cell fate control by pioneer transcription factors. *Development* 143, 1833–1837 (2016).
21. Zaret, K. S. Pioneering the chromatin landscape. *Nature Genetics* vol. 50 167–169 (2018).
22. Zaret, K. S. Pioneer Transcription Factors Initiating Gene Network Changes. *Annu. Rev. Genet.* 54, 367–385 (2020).
23. Liang, H. L. et al. The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. *Nature* 456, 400–403 (2008).
24. Sun, Y. et al. Zelda overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. *Genome Res.* 25, 1703–1714 (2015).

25. Gaskill, M. M., Gibson, T. J. ., Larson, E. D. . & Harrison, M. M. . GAF is essential for zygotic genome activation and chromatin accessibility in the early *Drosophila* embryo. *Elife* 0–43 (2021).
26. Moshe, A. & Kaplan, T. Genome-wide search for Zelda-like chromatin signatures identifies GAF as a pioneer factor in early fly development. *Epigenetics and Chromatin* 10, 1–14 (2017).
27. Schulz, K. N. et al. Zelda is differentially required for chromatin accessibility , transcription factor binding , and gene expression in the early *Drosophila* embryo. 1715–1726 (2015)
28. Dufourt, J. et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun.* 9, 1–13 (2018).
29. Raff, J. W., Kellum, R. & Alberts, B. The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle. *EMBO J.* 13, 5977–5983 (1994).
30. Chetverina, D., Erokhin, M. & Schedl, P. GAGA factor: a multifunctional pioneering chromatin protein. *Cell. Mol. Life Sci.* 78, 4125–4141 (2021).
31. Srivastava, A., Kumar, A. S. & Mishra, R. K. Vertebrate GAF/ThPOK: emerging functions in chromatin architecture and transcriptional regulation. *Cell. Mol. Life Sci.* Vol. 75, 623–633 (2018).
32. Hendrix, D. A., Hong, J. W., Zeitlinger, J., Rokhsar, D. S. & Levine, M. S. Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7762–7767 (2008).
33. Li, J. & Gilmour, D. S. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J.* 32, 1829–1841 (2013).
34. Fuda, N. J. et al. GAGA Factor Maintains Nucleosome-Free Regions and Has a Role in RNA Polymerase II Recruitment to Promoters. *PLoS Genet.* 11, 1–22 (2015).
35. Judd, J., Duarte, F. M. & Lis, J. T. Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. *Genes Dev.* 35, 147–156 (2021).
36. Foe, V. E., Odell, G. M. & Edgar, B. A. Mitosis and morphogenesis in the *Drosophila* embryo: Point and counterpoint. In: *The Development of Drosophila melanogaster*. Cold Spring Harb. Lab. Press 149–300 (1993).

37. Auer, J. M. T. et al. Of numbers and movement – Understanding transcription factor pathogenesis by advanced microscopy. *DMM Dis. Model. Mech.* 13, (2021).
38. Mir, M. et al. Dense bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev.* 31, 1784–1794 (2017).
39. Mir, M. et al. Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos. *Elife* 7, 1–27 (2018).
40. Farrell, J. A. & O’Farrell, P. H. From egg to gastrula: How the cell cycle is remodeled during the drosophila mid-blastula transition. *Annu. Rev. Genet.* 48, 269–294 (2014).
41. Hendzel, M. J. et al. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* 106, 348–360 (1997).
42. Follmer, N. E., Wani, A. H. & Francis, N. J. A Polycomb Group Protein Is Retained at Specific Sites on Chromatin in Mitosis. *PLoS Genet.* 8, (2012).
43. Koenecke, N., Johnston, J., He, Q., Meier, S. & Zeitlinger, J. Drosophila poised enhancers are generated during tissue patterning with the help of repression. *Genome Res.* 27, 64–74 (2017).
44. 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).
45. Nègre, N. et al. A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet.* 6, (2010).
46. Kvon, E. Z. et al. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature* 512, 91–95 (2014).
47. Blythe, S. A. & Wieschaus, E. F. Establishment and maintenance of heritable chromatin structure during early drosophila embryogenesis. *Elife* 5, 1–21 (2016).
48. Adkins, N. L., Hagerman, T. A. & Georgel, P. GAGA protein: A multi-faceted transcription factor. *Biochem. Cell Biol.* 84, 559–567 (2006).
49. Li, X. Y., Harrison, M. M., Villalta, J. E., Kaplan, T. & Eisen, M. B. Establishment of regions of genomic activity during the Drosophila maternal to zygotic transition. *Elife* 3, 1–20 (2014).

50. Lott, S. E. et al. Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-Seq. *PLoS Biol.* 9, (2011).
51. Mahmoudi, T., Katsani, K. R. & Verrijzer, C. P. GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J.* 21, 1775–1781 (2002).
52. 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).
53. Ghavi-Helm, Y. et al. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512, 96–100 (2014).
54. Trullo, A., Dufourt, J. & Lagha, M. MitoTrack, a user-friendly semi-automatic software for lineage tracking in living embryos. *Bioinformatics* 36, 1300–1302 (2020).
55. Rieder, L. E. et al. Histone locus regulation by the *Drosophila* dosage compensation adaptor protein CLAMP. *Genes Dev.* 31, 1494–1508 (2017).
56. Dufourt, J. et al. Spatio-temporal requirements for transposable element piRNA-mediated silencing during *Drosophila* oogenesis. *Nucleic Acids Res.* 42, 2512–2524 (2014).
57. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S. & Reinberg, D. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92, 105–116 (1998).
58. Kwon, S. Y., Jang, B. & Badenhorst, P. The ISWI chromatin remodelling factor NURF is not required for mitotic male X chromosome organisation. *microPublication Biol.* 2–7 (2021).
59. Espinás, M. L. et al. The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. *J. Biol. Chem.* 274, 16461–16469 (1999).
60. Van Steensel, B., Delrow, J. & Bussemaker, H. J. Genomewide analysis of *Drosophila* GAGA factor target genes reveals context-dependent DNA binding. *Proc. Natl. Acad. Sci. U. S. A.* 100, 2580–2585 (2003).
61. Behera, V. et al. Interrogating Histone Acetylation and BRD4 as Mitotic Bookmarks of Transcription. *Cell Rep.* 27, 400-415.e5 (2019).
62. Liu, Y. et al. Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells. *Cell Rep.* 19, 1283–1293 (2017).

63. Samata, M. et al. Intergenerationally Maintained Histone H₄ Lysine 16 Acetylation Is Instructive for Future Gene Activation. *Cell* 182, 127–144.e23 (2020).
64. Lagha, M., Bothma, J. P. & Levine, M. Mechanisms of transcriptional precision in animal development. *Trends Genet.* 28, 409–416 (2012).
65. Bentovim, L., Harden, T. T. & DePace, A. H. Transcriptional precision and accuracy in development: From measurements to models and mechanisms. *Dev.* 144, 3855–3866 (2017).
66. Matharu, N. K., Yadav, S., Kumar, M. & Mishra, R. K. Role of vertebrate GAGA associated factor (vGAF) in early development of zebrafish. *Cells Dev.* 166, 203682 (2021).
67. Garavís, M. et al. The structure of an endogenous *Drosophila* centromere reveals the prevalence of tandemly repeated sequences able to form i-motifs. *Sci. Rep.* 5, 1–10 (2015).
68. Dufourt, J. et al. Imaging translation dynamics in live embryos reveals spatial heterogeneities. *Science (80-.).* 372, 840–844 (2021).
69. Venken, K. J. T., He, Y., Hoskins, R. A. & Bellen, H. J. P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science (80-.).* 314, 1747–1751 (2006).
70. Michelman-Ribeiro, A. et al. Direct measurement of association and dissociation rates of DNA binding in live cells by fluorescence correlation spectroscopy. *Biophys. J.* 97, 337–346 (2009).
71. Escoffre, J. M., Hubert, M., Teissié, J., Rols, M. P. & Favard, C. Evidence for electro-induced membrane defects assessed by lateral mobility measurement of a gpi anchored protein. *Eur. Biophys. J.* 43, 277–286 (2014).
72. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. & Webb, W. W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys. J.* 16, 1055–1069 (1976).
73. Yguerabide, J., Schmidt, J. A. & Yguerabide, E. E. Lateral mobility in membranes as detected by fluorescence recovery after photobleaching. *Biophys. J.* 40, 69–75 (1982).
74. Müller, P., Schwille, P. & Weidemann, T. PyCorrFit-generic data evaluation for fluorescence correlation spectroscopy. *Bioinformatics* 30, 2532–2533 (2014).
75. Dertinger, T. et al. The optics and performance of dual-focus fluorescence correlation spectroscopy. *Opt. Express* 16, 14353 (2008).

76. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359 (2012).
77. Fenouil, R. et al. Pasha: a versatile R package for piling chromatin HTS data. *Bioinformatics* 15, 2528–30 (2016).
78. Koenecke, N., Johnston, J., Gaertner, B., Natarajan, M. & Zeitlinger, J. Genome-wide identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol.* 17, 1–19 (2016).
79. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165 (2016).
80. Bantignies, F. & Cavalli, G. Topological Organization of *Drosophila* Hox Genes Using DNA Fluorescent In Situ Hybridization. *Methods Mol. Biol.* 1196, 103–120 (2014).

Chapter 3. Results: Trithorax Group proteins and mitotic bookmarking

Parallel to my work on GAF, I investigated if dBRD₄ and Ash1 could act as mitotic bookmarkers during *Drosophila* early development. In this chapter, I will describe the results obtained and describe additional experiments required prior to publication. Therefore, I did not present these results as an advanced manuscript.

For both Ash1 and dBRD₄, I attempted to identify the mitotic bound targets genome-wide. Unfortunately, we did not obtain an efficient antibody against the protein Ash1 either for ChIP or immunostaining. However, we used an anti-dBrd₄ (kind gift from Igor Dawid, National Institute of Health, Maryland) and could obtain convincing results (two ChIP-seq replicates for interphase chromatin, one for mitotic chromatin and one sample is in progress). In addition, to retrieve binding kinetics of dBrd₄, I generated a CRISPR GFP-tagged allele. Moreover to characterize dBrd₄ functions, I generated a knock-down and assed 3D genome organization by HiC and RNA-seq (in collaboration with Juanma Vaquerizas Lab, Max Planck Institute for Molecular Biomedicine, Germany). Finally, I performed genetic epistatic experiments in which I depleted both Ash1 and dBrd₄.

Localization of dBRD₄ and Ash1 during the cell cycle

To first examine whether dBrd₄ was retained in mitosis during *Drosophila* embryogenesis, I created a transgene expressing dBrd₄ fused to super-folder GFP (sfGFP) under the maternally expressed driver *αTubulin* (Steffen et al. 2013) and performed live imaging. The *αTubulin* driver is strongly expressed during oogenesis leading to an important deposition in the laid egg. We observed that dBrd₄ is retained throughout the cell cycle (Figure 3.1 and 3.2). We also confirmed the observations made by the L. Ringrose lab (Steffen et al. 2013), that Ash1 is retained on mitotic chromosomes (Figure 3.1). Mitotic retention of dBrd₄ and Ash1 is high and these two proteins seem to 'decorate' chromatin.

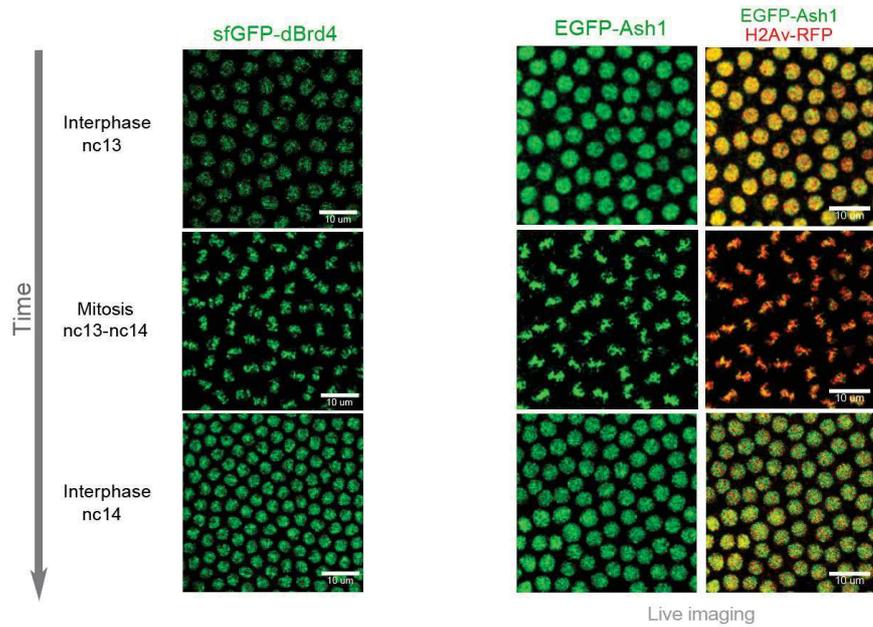


Figure 3. 1 : Maternally deposited *dBrd4* and *Ash1* are retained during mitosis in the *Drosophila* embryo.

Snap shots from maximum intensity projected live imaging of sfGFP-dBrd4/+ and EGFP-Ash1/HistH2Av-RFP embryos during interphases n.c.13 and 14 and mitosis n.c.13 to 14.

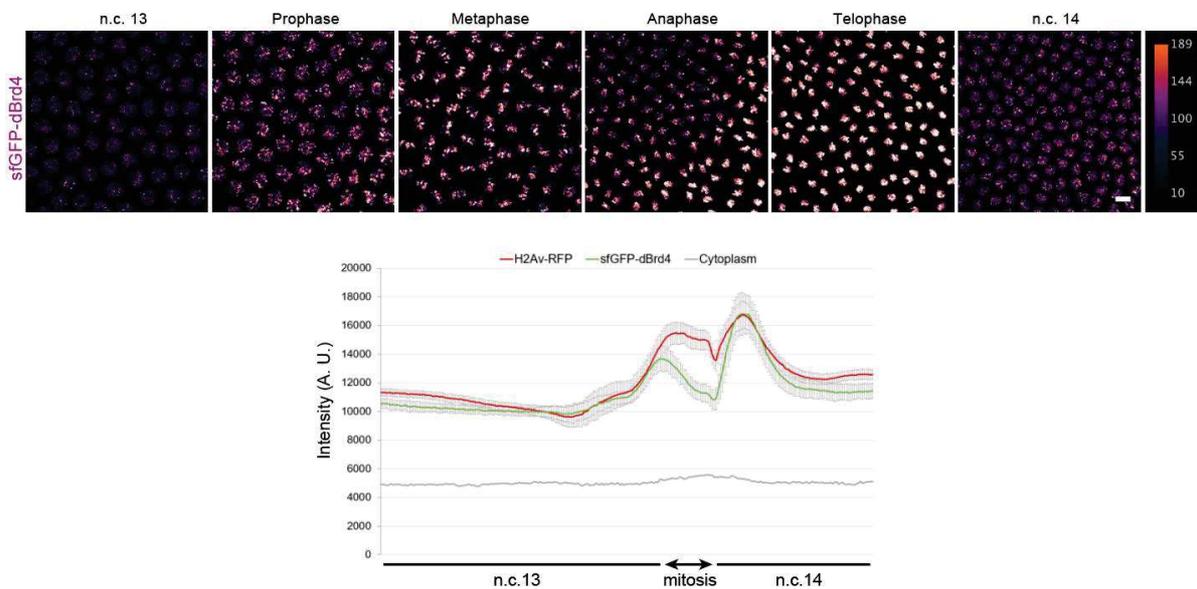


Figure 3. 2 : Maternally deposited *dBrd4* is retained during all phases of mitosis.

(Top) Snap shots from a movie of a transgenic sfGFP-BRD4/+ embryo at each step of the mitosis nuclear cycle 13 to nuclear cycle 14. Scale bar is 5μm. Note an increase in intensity at telophase.

(Bottom) Mean fluorescent signal quantifications of sfGFP-dBrd4 in nucleoplasm (green) and cytoplasm (grey), and H2Av-RFP in nucleoplasm during nuclear cycle 13 to 14 extracted from time-lapse movies of embryos expressing sfGFP-dBrd4 and H2Av-RFP (mean from three movies of three independent embryos).

Interestingly, we noticed that dBrd4 signal (assessed from a dBrd4-sfGFP transgene) was not uniformly distributed within nuclei but instead formed nuclear puncta at each interphase (Figure 3.3). This distribution is reminiscent of Brd4 nuclear distribution in mESCs (Sabari et al. 2018). Nuclear dBrd4 foci are also observed with the sfGFP-dBrd4 CRISPR allele that I generated (Figure 3.3) as well as in fixed wild type embryos (Figure 3.4). The sfGFP-dBrd4 CRISPR line was difficult to obtain (after several attempts of recombination matrix cloning and injections) but reinforce the observation of nuclear foci of dBrd4. Further characterizations of dBrd4 binding dynamics are currently under investigation in the laboratory.

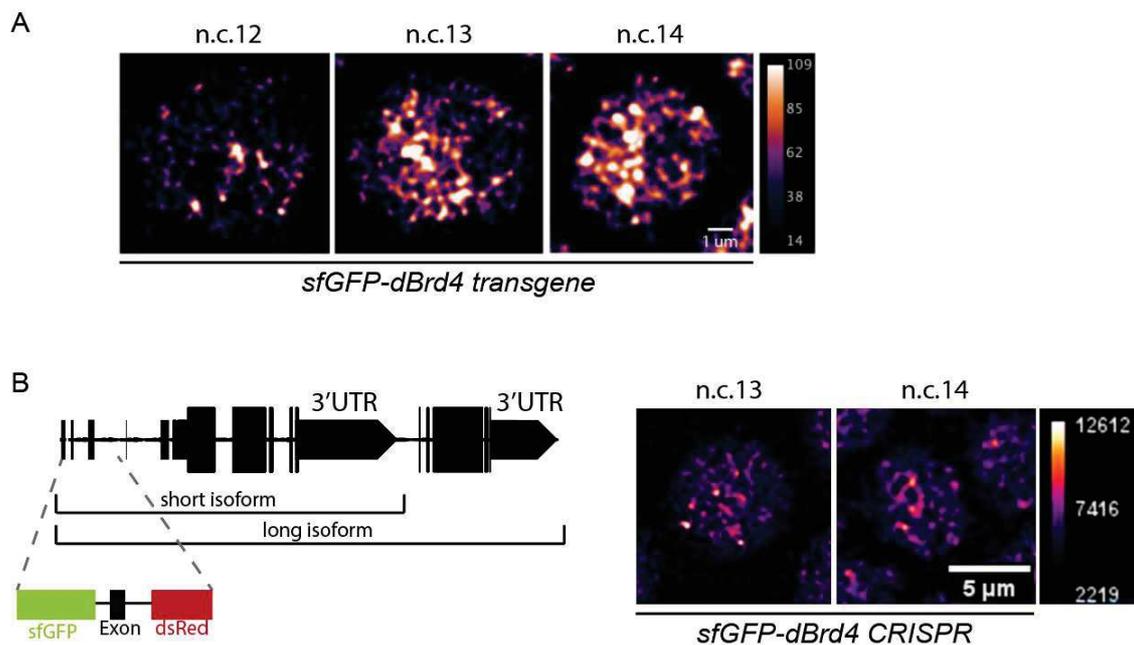


Figure 3.3 : dBrd4 forms nuclear clusters.

(A) Snap shots from a movie of a transgenic sfGFP-Brd4/+ embryo of a nucleus at nuclear cycle 12, 13 and 14. (B) Scheme of the CRISPR strategy used to insert sfGFP (green box) after the ATG of dBrd4 coding sequence. A dsRed marker was inserted in an intronic region. Right, snap shots from a movie of a sfGFP-Brd4 CRISPR/+ embryo of a nucleus at nuclear cycle 13 and 14.

Knowing that Zelda also forms nuclear 'hubs' (Dufourt et al. 2018; Mir et al. 2018), we looked if those hubs could share the same microenvironment with dBrd4. Some of the clusters overlap but not all of them (Figure 3.4).

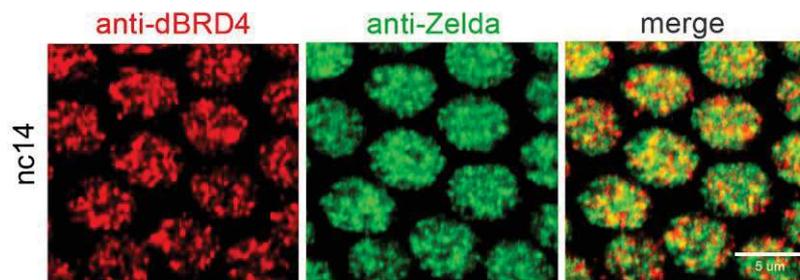


Figure 3. 4 : dBrd4 puncta do not exclusively co-localize with Zelda.

Microcopy image of one Z-plane from n.c. 14 sfGFP-Zelda embryo immuno-stained with anti-dBrd4 (red) and anti-GFP (green).

We can speculate that dBrd4 nuclear foci could contain several *cis*-regulatory regions such as super-enhancers, as described in Chapter 1 E.1 (Sabari et al. 2018). Such a clustering might have a role on transcription regulation by concentrating dBrd4 and increasing the rate of dBrd4 binding to its targets. In the future, to ask whether these nuclear foci contain *cis*-regulatory regions, we could perform DNA-FISH against dBrd4 target enhancers (Kellner et al. 2013) (either with few probes or with a multiplex approach as Hi-M, as described in Annexe 3), combined with immunostaining against dBrd4. One way to assess the kinetics of dBrd4 binding would be to perform FRAP in the nucleoplasm and in the foci. Such measurements would inform on the biophysical nature of dBRD4 nuclear hubs. For instance we could learn if they represent phase-separated condensates that dynamically exchange dBrd4 molecules as it has been describe for Brd4 in mESCs (Sabari et al. 2018).

Role of dBrd4 in ZGA and 3D genome organization

To characterize the role of dBrd4 during ZGA, we first investigated a potential function in spatial genome organization. Specifically we questioned how a maternal depletion of dBRD4 would affect 3D genome organization and zygotic transcription in the early embryo. To do this, I generated dBrd4 knockdown embryos using RNA interference against dBrd4 and the *maternal-αTubulin* driver. This approach leads to developmental defects and none of the

embryos hatched. Furthermore, I attempted to visualize transcription using MS2/MCP reporter but transcription seems too affected to be detected. We collected embryos depleted in dBrd4 and now plan to perform single embryo RNA-seq experiments to assess the impact of dBrd4 at a genome-wide transcriptional level (Figure. 3.5)

In collaboration with Noelia Diaz and Elizabeth Ing-Simmons from Juanma Vaquerizas laboratory (Max Planck Institute for Molecular Biomedicine, Germany), we collected staged dBrd4 knockdown and control embryos (RNAi-white) and performed HiC-seq experiments to obtain genome-wide chromosomal contact maps. Elizabeth Ing-Simmons did the analysis on one replicate and more replicates will be analyzed in the future. Interestingly, we did not observe an overall loss of genomic contacts upon dBrd4 depletion (Figure 3.5). In the contrary, we observed some regions that contact more frequently on chromosome 2L and R of RNAi-dBrd4 embryos than of control embryos (Figure 3.5, black arrows). Surprisingly, the control embryo chromosomal contact map does not resemble to the map of n.c.14 embryos from Hug et al. 2017. We suspect that this is due to the staging of embryos. For the RNAi-white and RNAi-dBrd4 n.c.11 to n.c.13 embryos were used because usually RNAi-dBrd4 embryos do not reach n.c.14 due to developmental defects. Therefore, we will compare our data to the earlier stages from Hug et al. 2017. Although these results are preliminary, we can suppose that depletion of dBrd4 has an effect on chromosome 3D organization of the early embryo and might help to organize the genome during ZGA.

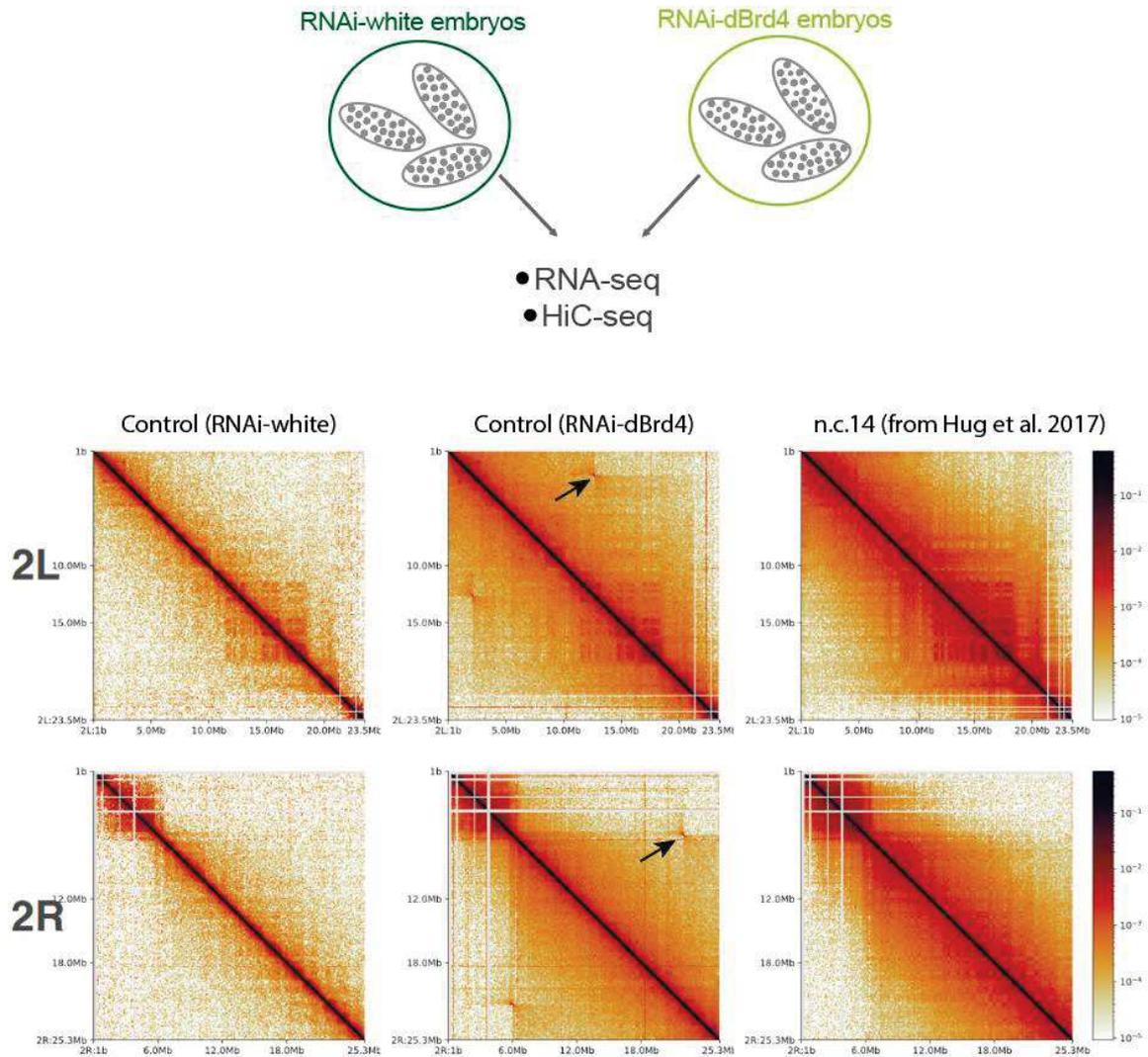


Figure 3. 5 : dBrd4 depletion generate chromosome conformation rearrangements.

(Top) Scheme of the pools of embryos used for RNA-seq and HiC-seq experiments. (Bottom) Contact maps of the chromosomes 2L and 2R in control embryos (RNAi-white), RNAi-dBrd4 and wild type n.c.14 embryos from Hug et al. 2017 HiC experiments. Black arrows point a high frequency of contacts between two genomic loci that were not present in the control sample.

dBrd4 mitotically retained regions

We next sought to identify dBrd4 mitotically retained loci genome-wide. To do this, I performed ChIP-seq on mitotic and interphase embryos as in Chapter 2, Material and Methods. We succeeded to obtain two replicates for interphase but only one replicate for mitotic embryos gave satisfactory ChIP signal. dBRD4 has never been ChIP profiled in

Drosophila embryos (only in S2 cells (Kockmann et al. 2013)) and was difficult to ChIP in our hands. We attribute this poor ChIP signal to the fact that this protein does not directly bind to DNA and that during mitosis chromatin is more compact and thus harder to sonicate. We also have a limited amount of chromatin after sorting the embryos (compare to S2 cells for example). We are currently performing another experiment to reach at least two replicates. Overall, mitotic dBrd4 signal was lower than in interphase. Moreover, dBrd4 signal is broader than for GAF, which could be expected as dBrd4 does not directly bind DNA but recognize histone modifications.

The ChIP analysis is still in progress, however we could identify mitotically retained regions as exemplified by the *Spi/msbl1* locus (Figure 3. 6).

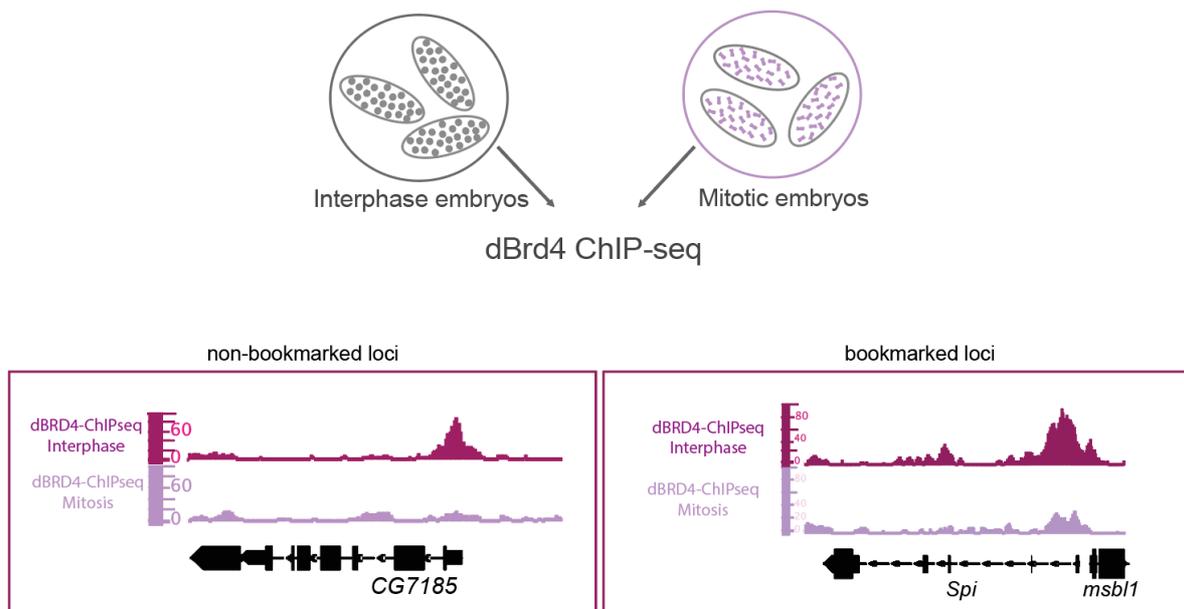


Figure 3. 6 : dBrd4 is retained on specific genes during mitosis.

(Top) Scheme of the pools of embryos used for dBrd4 ChIP-seq experiment. (Bottom left) dBrd4 ChIP-seq profile on interphase (pink) and mitotic (light pink) embryos at the non-bookmarked *CG7185* gene. (Bottom right) dBrd4 ChIP-seq profile on interphase (pink) and mitotic (light pink) embryos at the bookmarked *Spi* and *msbl1* genes.

Even if more bioinformatic analyses are required (in progress in the lab), these results indicate that dBrd4 is retained on some of its target genes.

Genetic approaches to study the impact of dBrd4 and Ash1 on the timing of transcriptional activation

To investigate a potential role of dBrd4 and Ash1 on transcriptional activation and mitotic memory, we used MS2/MCP live imaging approach with a reporter gene (as in Ferraro et al. 2016). Specifically, I used a transgene carrying a *snail* minimal enhancer with *snail* core promoter (*sna-transgene*) (see C.4.3.). This transgene is expressed in the ventral side, the mesoderm, of the embryo. To ensure non-limiting levels of the *sna* activators Dorsal and Twist (Ip et al. 1992), we studied the transcription activation in a spatial manner and used the ventral furrow¹⁴ as an indicator to define dorso-ventral coordinates (as described in Dufourt et al. 2018). We studied temporal dynamics of *sna-transgene* activation in a region of 50µm from either side of the ventral furrow and defined internal and external regions (Figure 3.7).

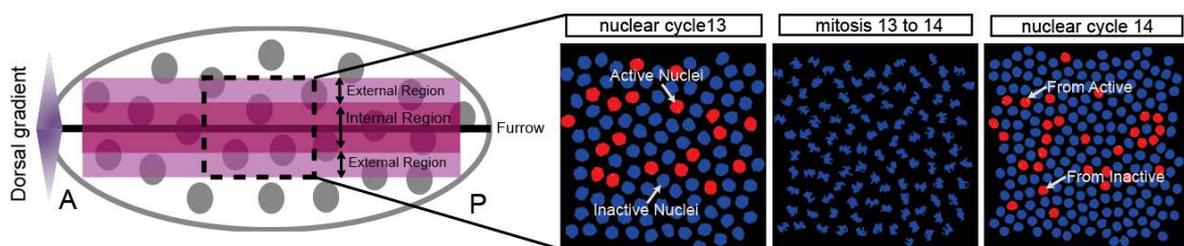


Figure 3.7 : Experimental set-up of transcription activation quantification.

(Left) Scheme of a ventral side of an embryo with the region imaged in the dashed square. (Right) Snap shots of a false-colored movie with active nuclei in red and inactive in blue. Tracking nuclei captures the lineage information for each nucleus.

We used RNA-interference against *ash1* with the maternal driver *nanos-Gal4* expressed in *Drosophila* ovaries. As assessed by qRT-PCR, we successfully depleted *Ash1* mRNA in early embryos (Figure 3.8.A). These embryos could still gastrulate and no major defects were observed. Hence, *Ash1* seems to not be critically required during oogenesis and early embryonic development as the RNAi was induced only in the ovaries and before ZGA. We quantified post-mitotic transcriptional activation of the *sna-transgene* in control (*RNAi-white*) and *Ash1* depleted (*RNAi-Ash1*) embryos. We see an overall delay in post-mitotic synchrony

¹⁴ Line of invaginating nuclei during gastrulation process.

(Figure 3.8.B) and a reduction in the probability of nuclei to be active in n.c.13 (however not significant, Figure 3.8.C).

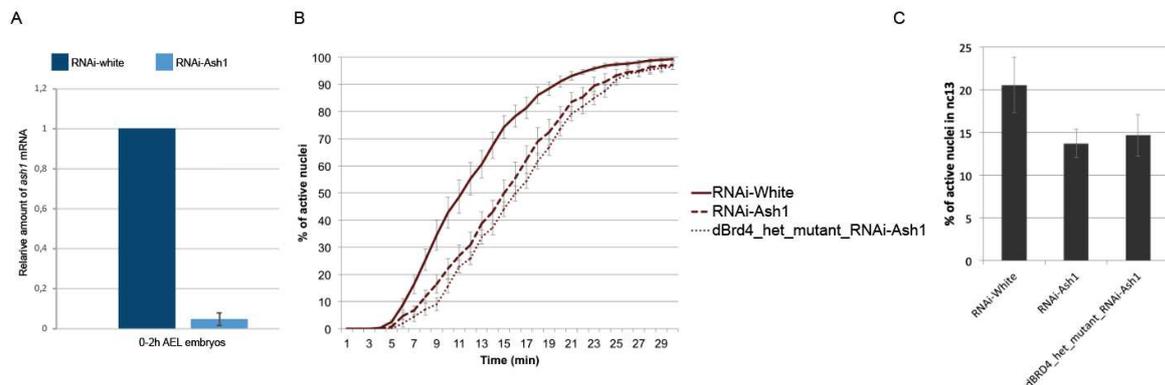


Figure 3. 8 : Maternal *Ash1* reduction delay post-mitotic transcription activation of the *sna*-transgene.

(A) Histogram of the relative amount of *RPL32* transcripts normalized *Ash1* mRNA in RNAi-white and RNAi-Ash1 0-2h embryos quantified by RT-qPCR. (B) Quantification of transcriptional synchrony of *sna*-transgene/+ embryo after mitosis in RNAi-white (control) $n=$ from 9 movies, *nos-Gal4/UASp-shRNA-Ash1* (dashed line) from 9 movies and *dBrd4^{-/+};nos-Gal4/UASp-shRNA-Ash1* (dot line) from 6 movies embryos. SEM are represented in grey. $n=$ number of nuclei analyzed from 4 movies of 4 independent embryos for each condition. (C) Histogram of the percentage of active nuclei in n.c.13 of RNAi-white, *nos-Gal4/UASp-shRNA-Ash1* and *dBrd4^{-/+};nos-Gal4/UASp-shRNA-Ash1* movies.

Despite the delay of transcriptional activation after mitosis, the transcriptional memory bias seems not to be affected upon *Ash1* depletion (Figure 3.9.A and B). Indeed, nuclei coming from active nuclei and inactive nuclei are delayed in terms of transcriptional timing but nuclei coming from active activate still faster than nuclei coming from inactive nuclei. We observe a memory bias in *Ash1* RNAi in both the central part of the pattern as well as in the external regions (Figure 3.9.A and B). We note a higher delay of the activation of the nuclei derived from active nuclei in the external region of the pattern (Figure 3.9.B), suggesting a spatial dependence for post-mitotic timing of activation. It is known that the TF activator *Dorsal* is in higher concentration at the ventral midline (Kanodia et al. 2009; Rushlow and Shvartsman 2012) and could potentially compensate for the loss of *Ash1*.

Overall we conclude that Ash1 is important for post-mitotic transcriptional synchrony but does not play a role by itself in transcriptional memory.

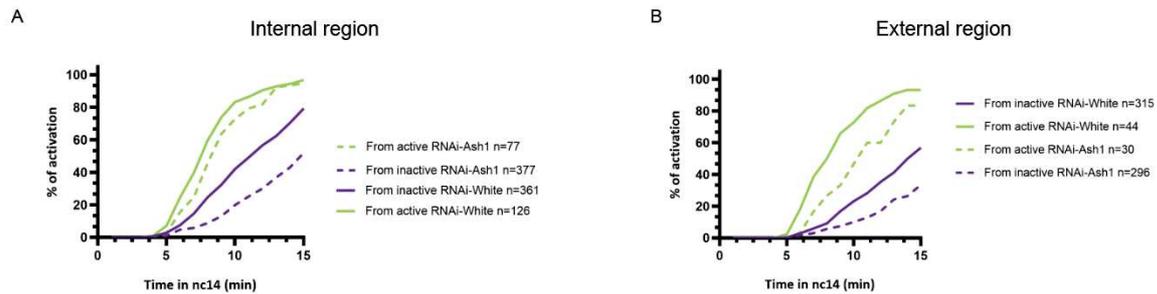


Figure 3.9 : Depleting maternal Ash1 do not alter the transcriptional memory bias.

(A) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-white embryos (control, solid curves) and RNAi-Ash1 *sna-transgene*/⁺ embryos (dashed curves). *n*=number of nuclei analyzed from 9 movies of 9 embryos for each genotype, in the internal part of the pattern (\pm 25 μ m from the ventral midline, see Figure 3.5). (B) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-white embryos (control, solid curves) and RNAi-Ash1 *sna-transgene*/⁺ embryos (dashed curves). *n*=number of nuclei analyzed from 9 movies of 9 embryos for each genotype, in the external part of the pattern (25 to 50 μ m from the ventral midline, see Figure 3.5).

dBrd4 (*fs(1)h*) mutants are lethal and female do not lay eggs (Digan et al. 1986; Florence and Faller 2008). We therefore attempted to study heterozygous mutant for dBrd4 (*fs(1)h¹¹¹²*). The *fs(1)h¹¹¹²* is an amorphic allele¹⁵ (Bagley et al. 2014) which was created in a previous genetic screen of lethal X-chromosome mutations (Zheng et al. 2008). *sna-transgene* post-mitotic transcriptional activation is not affected when reducing dBrd4 amount in the embryo (Figure 3.10). It is the case for nuclei coming from active nuclei and inactive nuclei (Figure 3.10).

¹⁵ An allele shown by molecular evidence to completely lack function, producing either a completely inactive gene product or none at all.

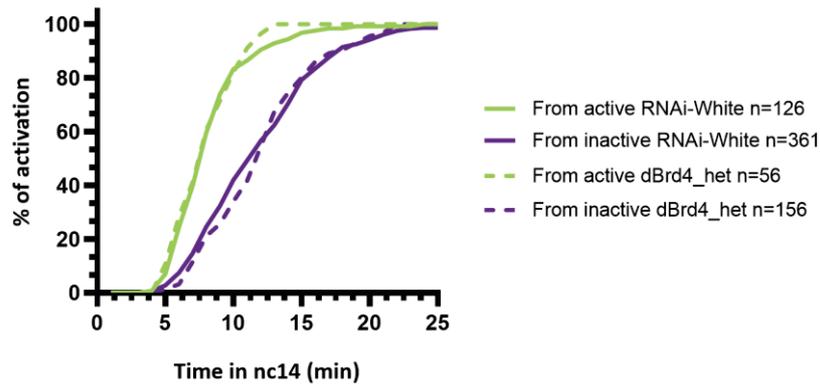


Figure 3. 10 : Reducing *dBrd4* level does not affect post-mitotic transcriptional timing of activation.

Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-white embryos (control, solid curves) and *dBrd4* heterozygous mutant *sna-transgene/+* embryos (dashed curves). *n*=number of pooled nuclei analyzed from 9 movies of 9 embryos for RNAi-white and 3 movies for *dBrd4* heterozygous mutant.

Thus, reducing *dBrd4* is not sufficient to affect transcription activation, potentially because of a compensatory mechanism by other factors such as *Ash1*. Indeed, *Ash1* has been shown to physically and genetically interact with *dBrd4* (Kockmann et al. 2013). Consequently, we combined *Ash1* knock-down experiments to a genetic context where *dBrd4* maternal dose was reduced, *dBrd4_het_RNAi-Ash1*.

To look at the effect of *dBrd4* only, we compared the *RNAi-Ash1* in *dBrd4* heterozygous mutant embryos to *RNAi-Ash1* embryos. In that case, only the level of *dBrd4* is different from these two genetic contexts. Even if the memory bias is still present in embryos with a single copy of maternal *dBrd4* in an *RNAi-Ash1* context (Figure 3.11), the kinetics of post-mitotic transcriptional reactivation appears affected in nuclei derived from active mothers. In such background, a slight reduction in the memory bias is observed in both regions, central part of the pattern and external part (Figure 3.11, green arrows). This effect is relatively small, but we need to remember that *dBrd4* was reduced by only half dose (in the maternal pool only). This trend may hint for a potential role of *dBrd4* on mitotic memory.

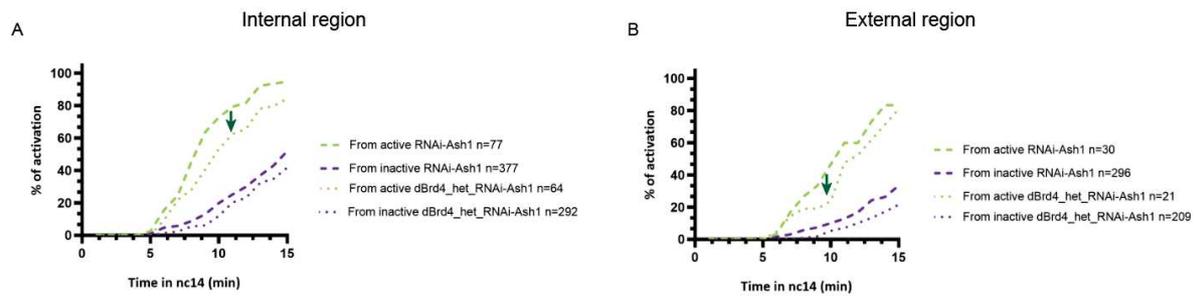


Figure 3. 11 : Reducing maternal dBrd4 affects the transcriptional memory.

(A) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-Ash1 embryos (dashed curves) and RNAi-Ash1 dBrd4 heterozygous mutant *sna-transgene/+* embryos (dot curves). *n*=number of pooled nuclei analyzed from 9 movies of 9 embryos for RNAi-white and 6 movies from 6 embryos for RNAi-Ash1 dBrd4 heterozygous mutant, in the internal part of the pattern ($\pm 25\mu\text{m}$ from the ventral midline, see Figure 3.5). (B) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) in RNAi-Ash1 embryos (dashed curves) and RNAi-Ash1 dBrd4 heterozygous mutant *sna-transgene/+* embryos (dot curves). *n*=number of pooled nuclei analyzed from 9 movies of 9 embryos for RNAi-white and 6 movies from 6 embryos for RNAi-Ash1 dBrd4 heterozygous mutant, in the external part of the pattern (25 to $50\mu\text{m}$ from the ventral midline, see Figure 3.5).

The next step would be to deplete dBrd4 totally but as the female don't lay eggs, we wanted to deplete dBrd4 specifically in the embryo and in a reversible manner. To reach such reversibility, I turned to optogenetics.

I chose the LEXY (Niopek et al. 2016) system because to me the readout was the easiest. With this system, upon blue light exposure the protein of interest is exported from nuclei thanks to the nuclear export signal in the LEXY domain being exposed upon activation. Therefore, you have access to the localization in real time of the targeted factor and its depletion in the nucleus allows being certain of its non-binding to the DNA. Moreover, a recent study showed the efficiency of the system in the *Drosophila* embryo (Kogler et al. 2021).

I fused the LEXY domain to the dBrd4 coding sequence and the fluorescent protein mCitrine (Figure 3.12). The goal was to trigger the inactivation of dBrd4 only during mitosis to decipher its role in mitosis from interphase. I first tested this construct in *Drosophila* S2 cultured cells

and was able to detect mCitrine-dBrd₄ nuclear signal (Figure 3.13). After 1min30 of blue light exposure, we were able to considerably reduce dBrd₄ nuclear localization without significant bleaching (control area). However, no complete nuclear signal recovery was observed at least in 20min. Therefore, this optogenetic system might be a good option but needed to be tested and might behave differently between cells and tagged protein. I attempted to tag endogenous dBrd₄ in *Drosophila* flies but unfortunately without success. This tool is still under optimization in the laboratory. However I could obtain a control fly strain carrying the LEXY domain fused to a mCherry fluorescent protein with a Nuclear Localization Signal (NLS). As a proof of concept, I could efficiently deplete mCherry-LEXY protein from a nuclei in a developing *Drosophila* embryo (Figure 3.14).

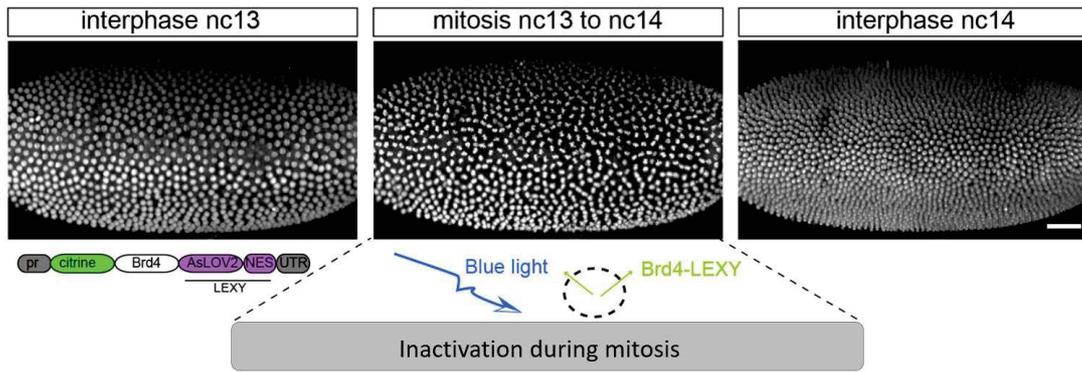


Figure 3. 12 : Scheme of the dBrd4 optogenetic tool.

Snap shot images from a H2A-RFP embryo from MuviSpim movie (done by Mathieu Dejean), nuclei are in interphase (left panel) or metaphase (middle panel). The embryos are oriented with ventral side in the middle. Scale bar is 50 μ m. Scheme of the construct build : pr = promoter region of *maternal- α tubulin*, asLOV2 = domain which will change conformation upon blue light activation, NES = nuclear export signal which will be available for recognition upon light exposure, UTR = 3'UTR of the *maternal- α tubulin*.

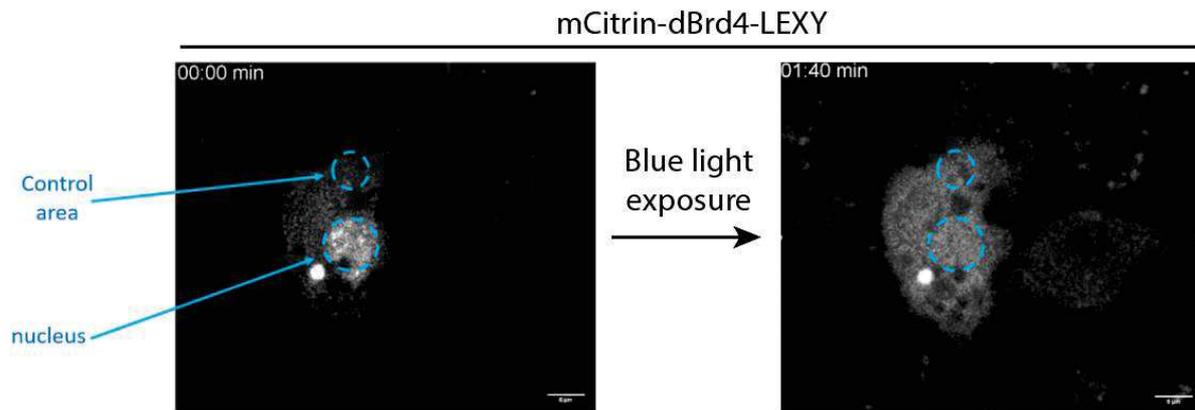


Figure 3. 13 : Optogenetic experiment in S2 cells.

Snap shots images from confocal movie of S2 cell. Cells were transfected with mCitrine-dBrd4-LEXY construct. The nucleus and control area (blue circles) were illuminated with 458nm wave length laser during 1min30. Scale bars are 5 μ m.

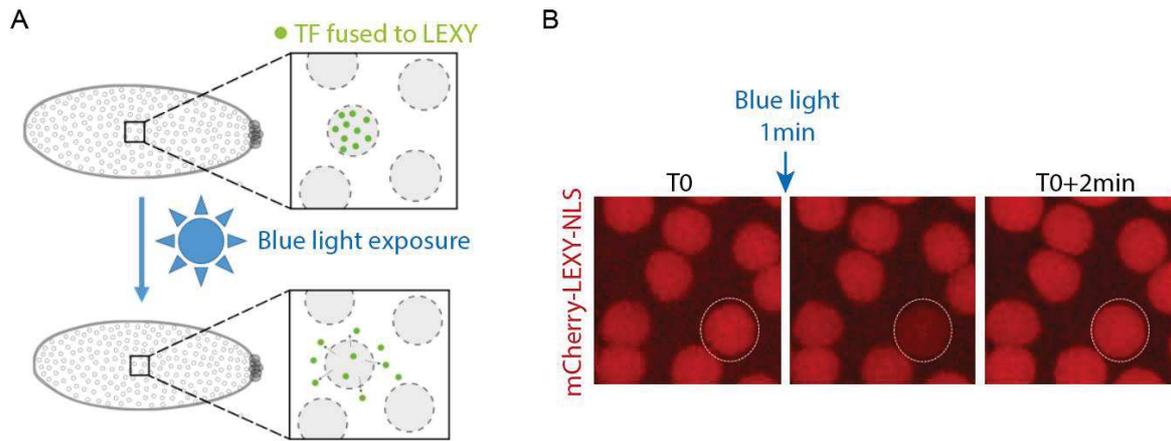


Figure 3.14 : mCherry-LEXY optogenetic system in the *Drosophila* embryo.

(A) Schematic of the experiment principle of the mCherry-LEXY nuclear depletion in an embryo.

(B) Nuclear depletion of mCherry-LEXY upon blue light exposure (dashed white circle) in a n.c.13 embryo.

Chapter 4. Discussion

During my PhD, I sought to determine how cell identity is kept during the early cleavage stages of embryogenesis. By cell identity, I imply transcriptional programs. By developing a method for mitotic ChIP-seq, I could determine the profiles of a TF (GAF), a chromatin modifier (dBrd4) and a histone modification (H₄K8ac) during mitosis of early *Drosophila* embryos. We identified GAF as a stable mitotic bookmark during zygotic genome activation, and I will primarily discuss here the results obtained for this TF. However, many aspects of the discussion here could be envisaged for chromatin modifiers and histone marks.

Towards a better understanding of TF action during mitosis and interphase

Mapping GAF bound targets during mitosis led to identify the existence of a transcriptional memory bias at an endogenous developmental gene. To the best of my knowledge, my work was the first to show a transcriptional memory at an endogenous locus. Measuring such a transcriptional memory is not trivial, as a stochastic activation in n.c.13 is required to obtain nuclei deriving from active and from inactive mothers in the next nuclear cycle. Reducing maternal GAF protein amount has shown that this memory is dependent, at least in part, of its presence in the early embryo. To clearly conclude about the mitotic role of GAF, disentangling the role of GAF during mitosis from its role during interphase will be necessary. One approach would consist in performing a loss of function of GAF specifically during mitosis with optogenetics methods (see result section 2.). However, my attempts to edit GAF with the LEXY domain were unsuccessful. But we have to consider the limitations of this approach. Indeed, I observed in transfected S2 cells that the recovery time to which GAF-LEXY get back to the nucleus can be in the order of minutes. As in the *Drosophila* embryo the nuclear cycles last from 10 to 40 min this might be not fast enough to disentangle the impact of mitotic depletion and interphase. Another approach would be to *de novo* elicit GAF binding during mitosis, with a tethering strategy (gain of function). As an attempt to perform such tethering and in collaboration with Dr. Andreas Möglich and Dr. Jeremy Dufourt, I generated a transgenic strain with GAF fused to an opto-activable dCas9 (LOV domain fused to a deactivated Cas9). The rationale is that upon light activation, the GAF-dCas9 would bind to specific targets, guided by specific gRNA (provided from another strain), therefore bringing GAF ectopically to its targets, but only during mitosis and only at specific loci. With this, we

can create an artificial bookmarked gene to see if GAF is sufficient for mitotic memory. I obtained *snail* enhancer and *doc* enhancer (gRNA lines) and I am currently testing these stocks. In parallel, I also generated the same tool with a deacetylase to specifically deacetylate histone marks during mitosis. Indeed my work suggested that many *cis*-regulatory regions exhibit histone acetylation during mitosis. I specifically examined H₄k8ac in Chapter 2. I believe that this will reveal new role of the epigenetic code during mitosis.

The concept of transcriptional memory

Conceptually, mitotic memory can be seen as a mechanism to ensure robustness in gene expression during development. We can imagine development like an object hold by several ropes. If you cut one rope, the object will not fall and will still be stable. You can cut several ropes, with a different combination but the object will still not fall. You can weaken the stability of the object until you cut enough ropes to fall it down. I see the mitotic memory as one of this rope. It might not be an essential mechanism by itself, but contribute to the proper development of an organism.

This mitotic memory means that a mother cell can pass its identity information to the daughters as early as in an undifferentiated embryo. Yet, the cells will have to adopt a different fate even if they derive from the original same mother. Identities are the results of a combination of different factors; indeed different cells can express the same single gene but give different specification when expressed in a combination with different genes. In that sense, memory from mother to daughter cells can be seen as the basis of cell identity that will then be influenced by external signals depending, for instance, on the position within the embryo. It would therefore be interesting to test for how many cell cycles this mitotic memory is kept, or if this memory is transient and can change during time. One would need to track cells during many cell cycles.

In our case, the mitotic memory bias is observed on the timing of activation of each nucleus. In fact, nuclei coming from inactive mothers will also be activated at some point. Thus the transcriptional memory here is defined as precision in timing of activation rather than cell identity as all the cells will end up being mesodermal cells (but still remain progenitors).

The 'memory' cells expressing earlier would potentially produce more mRNAs as it is the case for the *snail* transgene described in Ferraro et al. 2016. As also discussed in Chubb 2016, we

could speculate that this timing of activation is important for producing the right amount of mRNA to ensure the proper development.

The case of the *scylla* gene

We identified *scylla* as a target of GAF during mitosis and that this gene harbors a mitotic transcriptional memory in an endogenous context. But what could be the function of such a memory in the case of *scylla*?

The gene *scylla*, together with the gene *charybde*, have been shown to be required for head involution, a morphogenetic event happening just after gastrulation (Scuderi et al. 2006) and inhibit cell growth (Reiling and Hafen 2004). It can act as cell death activator, probably giving its name of the Greek mythology 'monster'. Interestingly, *scylla* expression can be upregulated under the stress condition of hypoxia at the larvae stage (Reiling and Hafen 2004). This upregulation provides a better survival of larvae development suggesting a critical role of *scylla* (and *charybde*) under hypoxia condition. It is tempting to hypothesize that the transcriptional memory of this gene might be important in particular during stress induction and that its timing of activation might be crucial for the organism to be able to respond to environmental stress, participating to robustness during development.

What defines the presence of transcriptional memory or not?

Results from my thesis work show that GAF controls, in part, the kinetics of post-mitotic reactivation. But we also showed that GAF is present on active and repressed gene. We revealed the presence of transcriptional memory for one gene but this might not be the case for other genes. So what could dictate the presence of a transcriptional memory at specific loci?

Given the pleiotropic role of GAF, I expect that this factor is not alone playing a role in mitotic bookmarking. Memory might be supported by a combination of factors that bookmark specific genes for activation but also repression. Given the observation that GAF is not sufficient to trigger memory on a reporter transgene (Annexe 1) in the dorsal part of the embryo but that it controls memory in other parts of the embryo (mesoderm), we can imagine that several factors, recruited to specific loci, are required to trigger memory. Intriguingly, none of the studied genes and transgenes (*doc* transgenes, *tom* gene, Annexe 1

and 2) expressed in the dorsal part, show any strong memory bias between nuclei derived from active and from inactive mother nuclei. I suggest that this memory could be spatially dependent, probably due to the presence of other transcription factors such as Dorsal. This factor has been shown to not clearly stay on mitotic chromatin in live (Reeves et al. 2012) as well as in fixed embryos (Esposito et al. 2016, and data not shown). Therefore this factor might help for a rapid post-mitotic reactivation by being quickly recruited by the help of GAF for instance.

Cooperative binding

By doing GAF ChIP-seq in mitotic embryos, we found that mitotic retained loci contain significantly more GAGAG motifs than interphase only loci. Moreover, GAF is known to be able to form oligomers (Espinás et al. 1999) *in vitro* and in cultured S2 cells. It is possible that GAF oligomerization on long GAGAG repeats would favor its retention during mitosis and maybe change its binding kinetics. To test this, extracting binding kinetics by FRAP or FCS in a GAF POZ domain mutant would be key. Indeed, the POZ domain has been shown to mediate oligomerization of GAF (Espinás et al. 1999). Alternatively, a new technique is capable of measuring oligomer formation (Hinde et al. 2016) and could potentially be applied to GAF-GFP. We can imagine that on several GAF binding motifs (typically the mitotic retained regions) one GAF protein can bind and nucleates oligomerization of several proteins. If this oligomerization is efficient, then the residence time would increase and provides a longer binding for mitotic retained regions. This hypothesis is difficult to test, but measuring GAF binding kinetics at different concentration of GAF *in vitro* on a fixed amount of targets, combined with oligomerization assay, would provide information regarding the effect of TF concentration on the k_{off} .

As illustrated in Chapter 1 E.3., a polyQ domain is present in both isoforms of the protein GAF. These polyQ domains have been showed to have prion¹⁶-like properties and trigger protein aggregation in yeast cells (Tariq et al. 2013). Moreover, its polyQ domains has been shown to

¹⁶ A prion is a protein that can become pathogenic by changing the conformation in space of other proteins in contact.

facilitates multimerization of GAF *in vitro* (Wilkins and Lis 1999). Therefore, polyQ domains might have a role on GAF binding and on its function as a TF.

Together with the observation that GAF forms big clusters in embryonic nuclei, such data could highlight a regulation of TF binding by oligomerization, potentially affecting mitotic retention and transcription regulation.

Bookmarking factors binding kinetics

By imaging the tagged sfGFP-GAF and with immunostaining, we saw that GAF is present in all embryonic cells (Chapter 2). However, it is only the nuclei deriving from active nuclei that harbor a memory bias. This could be explained by a differential of binding of GAF (binding in the active nuclei but not in the inactive of nuclear cycle 13) at specific loci, or that the concentration of GAF is different. To test this, although challenging, we could measure the binding of this factor on a single locus by for example labeling a specific DNA locus in live with synthetic zinc-finger proteins with anti-FLAG and anti-HA frankenbodies (Liu et al. 2021), expressing GAF-GFP and performing FRAP on this locus on different cells in the embryo. An orthogonal approach would be to use TetO/TetR system (Normanno et al. 2015) or ParS/ParB (Dubarry et al. 2006; Chen et al. 2018) to label DNA loci in live. This will give us information on the possible different types of binding of the same protein at different loci and/or localization within an embryo.

So far, no kinetic experiments such as FRAP has been performed in mitosis of living embryos. This is probably due to the challenge of the important movement of mitotic chromosomes specially in *Drosophila*. A good option would be to perform kinetics measurements in organisms with longer mitosis such as mammals. A good option would be to perform kinetics measurements in organisms with longer mitosis such as mammals. Therefore, even though challenging, mouse embryos or embryo-like structures such as human embryoids¹⁷ might be considered. A difference in binding kinetics between interphase and mitosis could be due to a different concentration of the concerned TF (for instance if its binding depends on

¹⁷ A multilayered cluster of differentiating Embryonic Stem Cells that resembles to an embryo at a specific stage of early development. One example is a blastoid, a structure that contains the same cell types and tissue topology as a blastocyst.

cooperativity) or its association with other factors that could stabilize or destabilize its binding.

Overall many examples reveal that acute depletion of TF help to understand different functions heretofore unknown. Similar approaches for mitotic versus interphase action will certainly unveil new functions of TFs once the tools established.

Chapter 5. Introduction: Translation dynamics during embryogenesis

Recent technological advances have been developed for a better understanding of the regulation of gene expression in space and time. As introduced in the chapter 1, the MS₂/MCP system has been a real breakthrough for the deepness of our knowledge concerning dynamics of transcription (Pichon et al. 2018). However, we know that the central dogma of molecular biology, stated by Francis Crick in 1957, consists of two main steps: transcription and translation. Yet, little is known about the dynamics of translation, and this is particularly true in the context of a developing organism. This lack of knowledge is primarily due to technical limitations. While transcription live imaging with the MS₂/MCP system exists since 1998, there was no equivalent method to image translation until 2016.

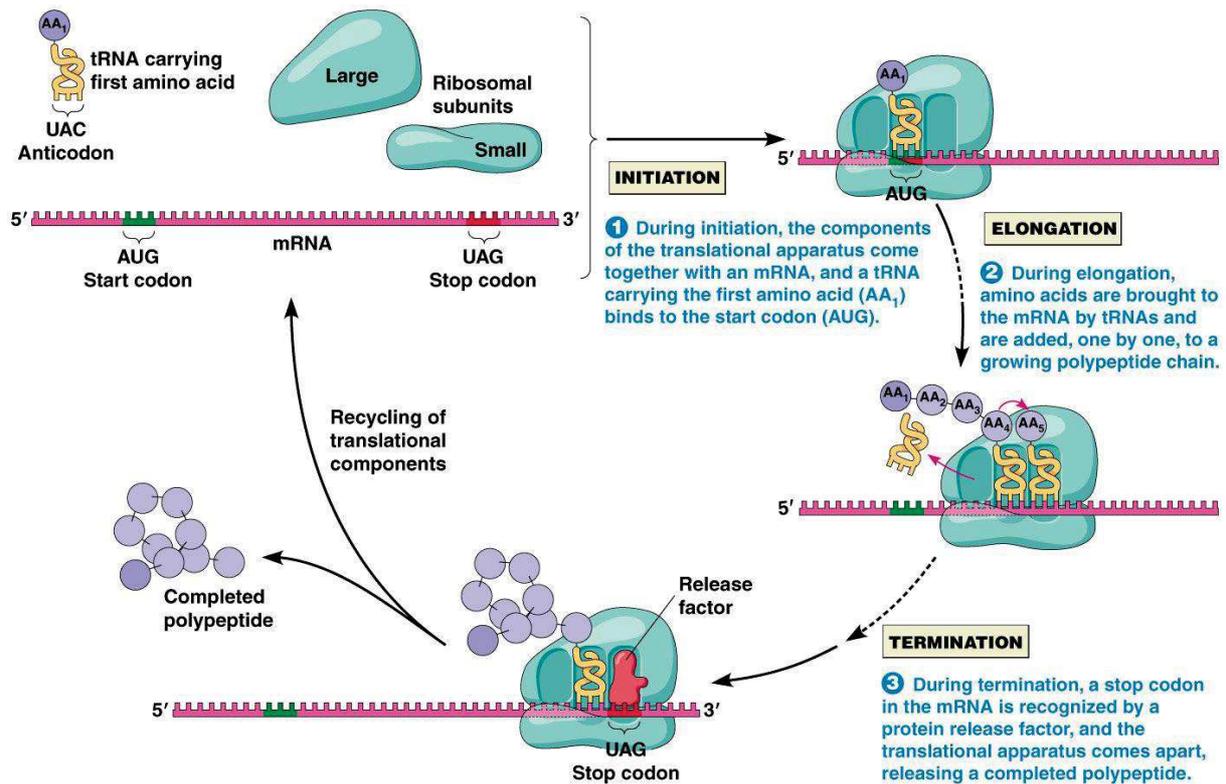
During my PhD, I was involved in a parallel project, initiated by Dr. Jérémy Dufourt, aiming at establishing a method to image translation in space and time in developing *Drosophila* embryos. As this project does not represent my main PhD project, I am intentionally summering it in a concise manner.

Initially, I contributed to this work by performing Single Particle Tracking analysis and I was second author of the manuscript for the first submission. But my implication was extensively increased during the revision process, taking place during Covid-19 lockdown periods, prior to publication of this work. I spent more than 6 months full time participating to the design and execution of experiments and interpretation of data. To account for this contribution, and in agreement with all co-authors, I am signing this article as a co-first author.

In the following chapters, I will introduce this methodology and the kind of research projects it can lead to in *Drosophila* but also in other model organisms, in particular when combined to imaging transcription and mRNA molecules.

A. Mechanisms of translation

In eukaryotes, when an mRNA is transcribed, matured and transported outside the nucleus, it is translated by the translation machinery. The process of translation is composed of three phases: initiation, elongation and termination (Kasinath et al. 2006) (Figure 5.1). Each of these phases can be highly regulated in order to give the right amount of protein at the right place and time.



© 2012 Pearson Education, Inc.

Figure 5.1 : Three main steps of mRNA translation.

Translation is composed of three main steps represented here: 1: initiation; 2: elongation and 3: termination.

A.1. Initiation phase of translation

There are at least two ways to initiate translation on a given mRNA by bringing the 80S subunit (composed of a considerable amount of nucleic acids and proteins) of the ribosome to the start codon (AUG). The first known mechanism is cap-dependent (Hellen and Sarnow 2001). The cap is a methylated guanosine triphosphate located at the 5' end of the mRNA and allows the recruitment of the eukaryotic translation Initiation Factor 4 E (eIF4E) protein which will lead to the assembly of ribosomal subunit 40S then the 60S (which forms the 80S, Figure 5. 2). The second mechanism is when the subunits directly bind to a specific sequence in the mRNA called Internal ribosomal entry site (IRES) (Komar and Hatzoglou 2005). This IRES-driven translation is not fully understood but may involve several eukaryotic Initiation Factors (eIFs) (Thoma et al. 2004). It seems that 10% of the mRNAs go through IRES-driven translation (Komar and Hatzoglou 2005).

In order to locate start codons during the initiation phase, ribosomal scanning of mRNA is generally accepted as the predominant mechanism (Kozak 1978). Scanning consists of the progression of the small ribosomal subunit in a 5' to 3' direction to probe for the initiation codon (reviewed in Shirokikh et al. 2019).

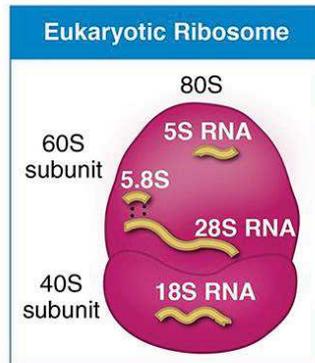


Figure 5. 2 : Scheme of the subunits which forms a eukaryotic ribosome.

Image is from www.quora.com.

A.2. Elongation phase of translation

Peptide chain formation comes from the addition of amino acids in accordance with the codon sequence in the mRNA (Browne and Proud 2002). The elongation phase involves the mRNA, the amino acyl tRNA, the 80S ribosome (composed of 40S and 60S subunits) and eukaryotic Elongation Factors (eEFs). The tRNA serves as recognition for a codon sequence and it allows a peptide bond between the new amino acid with the previous one to finally form a chain of amino acid (Figure 5. 3).

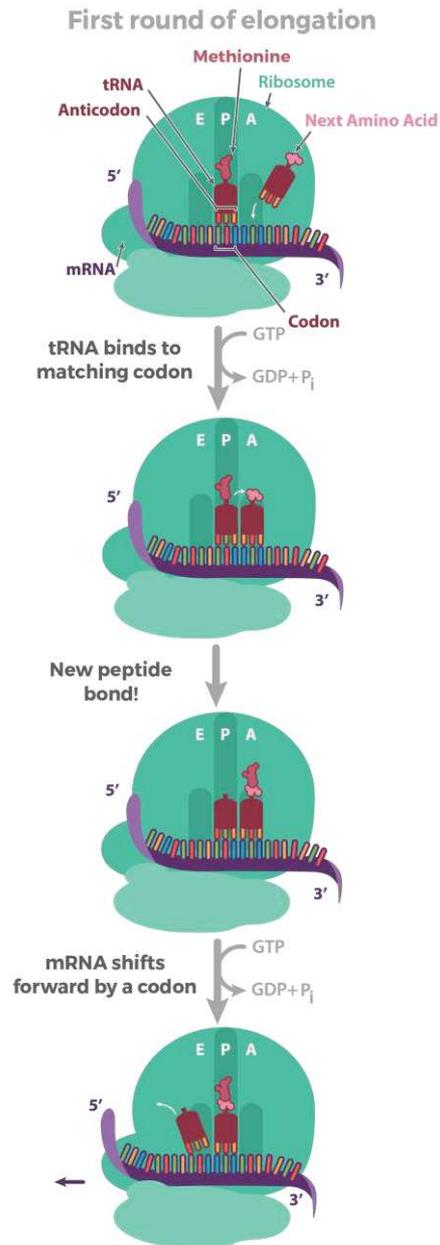


Figure 5.3 : Scheme of the elongation phase of translation.

mRNA (purple) stand in the groove between the 60S and 40S ribosomal units. Specific amino acyl tRNA bring amino acids, and peptide formation occurs in accordance with codon sequence of the mRNA. Exit (E), peptidyl (P), and amino acyl (A) sites on the ribosomes are shown. Scheme is from www.khanacademy.org.

A.3. Termination phase of translation

The termination phase occurs when the 80S ribosome arrives to a stop codon (e.g. UGA) of the mRNA. The 80S ribosome is released and split into 60S and 40S and helped by the

ribosomal eukaryotic Release Factor 1 (eRF1) (Fu et al. 2019). These subunits can be recycled for other rounds of translation.

All of these three phases can represent potential steps of regulation and rate limitations. For instance eIF6 has been shown to inhibit the assembly of the 80S ribosome and can be released under a stimulus (Ceci et al. 2003). Translation elongation can be inhibited by the RNA binding protein FMRP which binds to 60S subunit impeding tRNA recruitment as well as translation elongation factors (Chen et al. 2014a).

B. Regulation of translation in space

B.1. mRNA localization

How the encoded genomic information is differentially interpreted by the cells to get the right identity is *via* the expression of specific genes (i.e. transcriptional program) but also how these mRNAs are differentially translated. The localization of mRNA is suspected to be one important mechanism to address a localized production of proteins. In this case, mRNAs can be addressed to specific cellular compartment to be translationally repressed or activated (Das et al. 2021). This mRNA localization is a highly conserved process known from decades, present in yeast as well as in *Xenopus* and *Drosophila* embryo and has been extensively studied in neuronal cells (Martin and Ephrussi 2009) (Figure 5. 4).

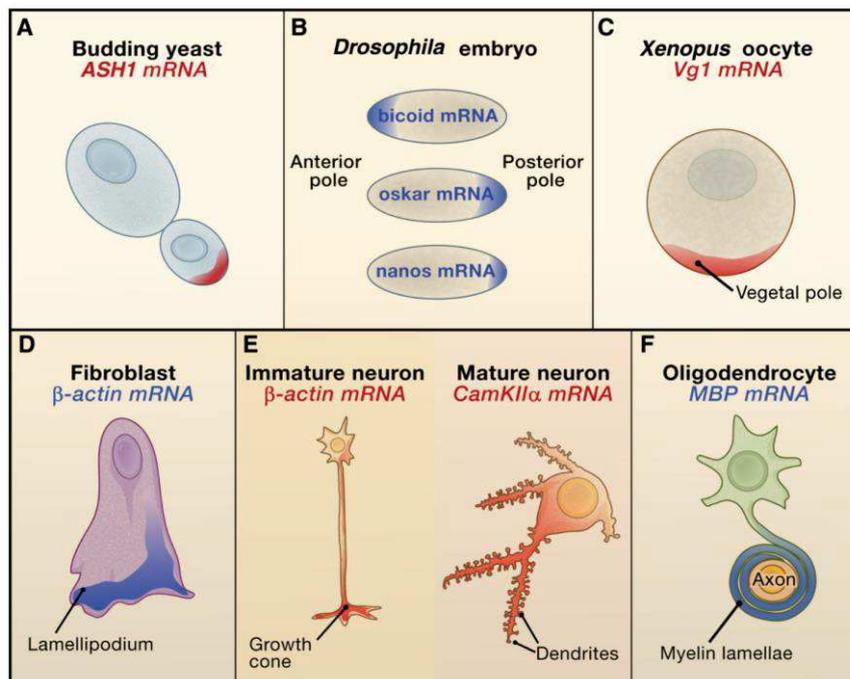


Figure 5. 4 : Examples of localized mRNAs.

Figure is from (Martin and Ephrussi 2009). (A) *ASH1* mRNA localizes to the bud tip in budding yeast. (B) *bicoid* mRNA localizes to the anterior pole, *oskar* and *nanos* mRNAs to the posterior pole of *Drosophila* embryos. (C) *Vg1* mRNA localizes to the vegetal pole in *Xenopus* oocytes. (D) In chick and mammalian fibroblasts, β -actin mRNA localizes to lamellipodia. (E) In immature mammalian neurons, β -actin mRNA localizes to growth cones and in mature, *CamKII α* mRNA is present in distal dendrites. (F) In mammalian oligodendrocytes, myelin basic protein (MBP) mRNA localizes to myelin lamellae.

The major mechanism of mRNA localization is through active transport on actin filament or microtubules. In this case, some RNA-binding protein recognize specific cis-regulatory sequences located generally in the 3' or 5' UTR of the mRNA, also called 'zipcode' (Engel et al. 2020; Buxbaum et al. 2015).

B.2. Local translation

It is established that mRNA localization regulates gene expression (Rongo et al. 1995, also see Chapter 1 A.2.1), but mechanisms underlying localized mRNA translation are still mostly unknown.

Specific inhibition of translation can lead to localized translation as it is the case in neurons where β -actin mRNA is translationally silent during transport until it gets activated when reaching its final localization to active spines (distal dendrites) (Hüttelmaier et al. 2005). Recent advances in live imaging of translation challenged this idea by demonstrating that mRNA with the β -actin 3'UTR are actively transcribing at distal but also at proximal dendrites (Wu et al. 2016).

In some circumstances, mRNA can be spatially protected from degradation as it is the case for *nanos* mRNA in *Drosophila* embryos where the protein Oskar 'protects' the mRNA specially at the posterior pole of the embryo from degradation complexes (Zaessinger et al. 2006).

Another way, not mutually exclusive, to control translatability of an mRNA in space is the different composition of the translation machinery, which revealed to be the rule rather than the exception (Shi and Barna 2015) (Figure 5. 5). Different composition of translation component can lead to more or less efficient translation, which might be important during development, but is for now poorly understood.

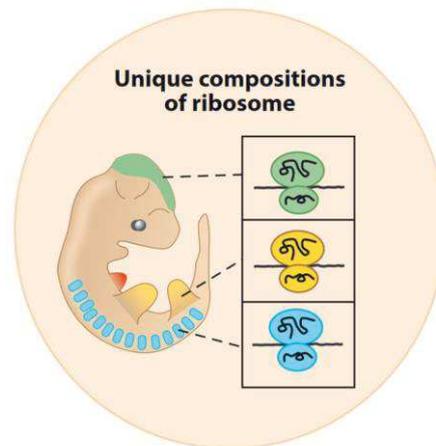


Figure 5. 5 : Different ribosomal composition in a mouse embryo.

Scheme representing unique compositions of ribosomes in different cell and tissue types of a mouse embryo. Image from (Shi and Barna 2015).

C. Regulation of translation in time

Since 2009, the technique of ribosome profiling followed by next generation sequencing (Ingolia et al. 2009) have been employed to explore the translation of mRNA genome-wide

(Vardi-Oknin and Arava 2019; Williams et al. 2014; Shiber et al. 2018). Although this technique revolutionized the field of quantification of translation, it lacks spatial and precise temporal resolution as well as access to single mRNA molecules translation dynamics. To analyze the translation regulation in time, the best way is to measure its dynamics in live. There are few techniques which allow the visualization of nascent protein production (Lyon and Stasevich 2017). One of them is the SunTag system which has been first developed in 2014 for full protein visualization (Tanenbaum et al. 2014). In 2016 four labs implemented this technique to be able to measure translation dynamics in living cells (Yan et al. 2016; Pichon et al. 2016; Wu et al. 2016; Wang et al. 2016). This technique relies on the recruitment of detector composed of a fluorescent protein fused to an antibody (Fab, scFv, nanobody...) and its association to a Tag encoded by a specific peptide disposed in arrays (Figure 5. 6). Upon translation, the nascent peptide will be bound by the fluorescent antibody and give an increased signal upon translation.

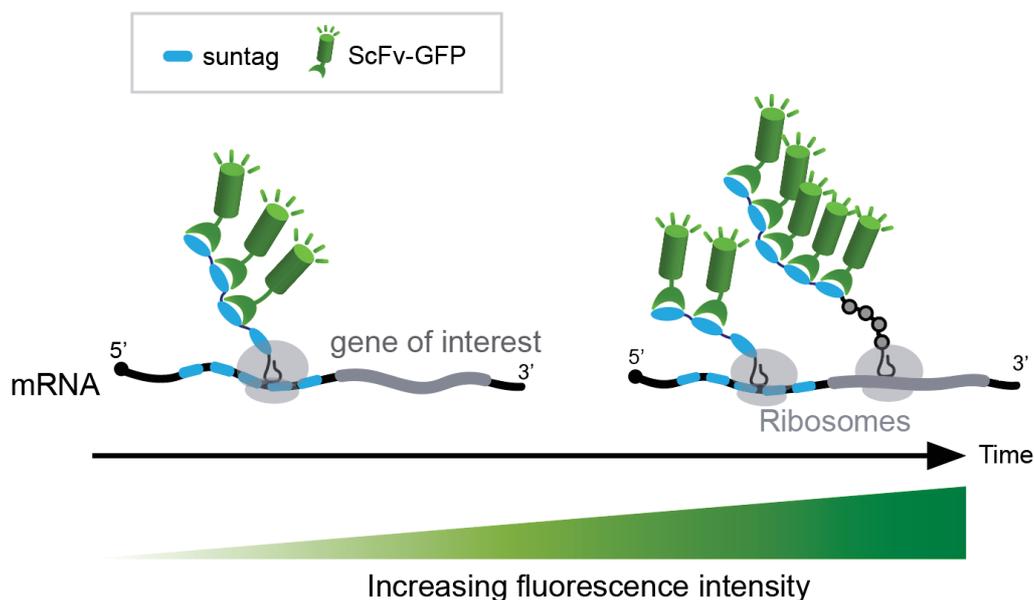


Figure 5. 6 : Scheme of the SunTag system principle.

SunTag tandem sequences (blue) are bound by the scFv-GFP (green) upon translation.

This allowed the measurements of translation rates at the single mRNA molecules at the subcellular level. However, until now we didn't have access to the information of regulation

of translation in a multicellular organism. Only the first round of translation of the *oskar* mRNA was captured in a living organism using the Translating RNA Imaging by Coat protein Knockoff (TRICK) method (Halstead et al. 2015) in *Drosophila ovaries*.

Using live imaging techniques, some studies revealed a dynamic regulation of translation (see also Chapter 6. Results). For example, *Formicola et al.* reported that in *Drosophila* neurons, ribonucleoprotein (RNP) granules are remodeled upon external stimuli and activates translation of associated mRNAs very rapidly (Formicola et al. 2021). In the study of *Wu et al.*, they have shown that neuronal mRNA translation appears in burst (Wu et al. 2016).

All those technical advances led to the discovery of another layer of gene regulation and will allow a better understanding of regulation of translation in time as shown in *Neurospora crassa* where translation is regulated in a rhythmic manner following the circadian clock, until now believed to be only transcriptionally regulated (Caster et al. 2016). Another example of spatio-temporal regulation of translation is in mammalian oocytes in which translational hotspots were discovered at the onset of meiosis surrounding the nucleus (Susor et al. 2015).

D. Translation in the early *Drosophila* embryo

A mid-scale screen in *Drosophila* embryos revealed that $\approx 70\%$ of mRNAs have a particular subcellular localization (Lécuyer et al. 2007; Wilk et al. 2016) (Figure 5. 7). However, the consequences of such localization in terms of translatability of these mRNA is way less understood. This suggest a potential huge heterogeneity between cells, not only in terms of mRNA but also in protein output.

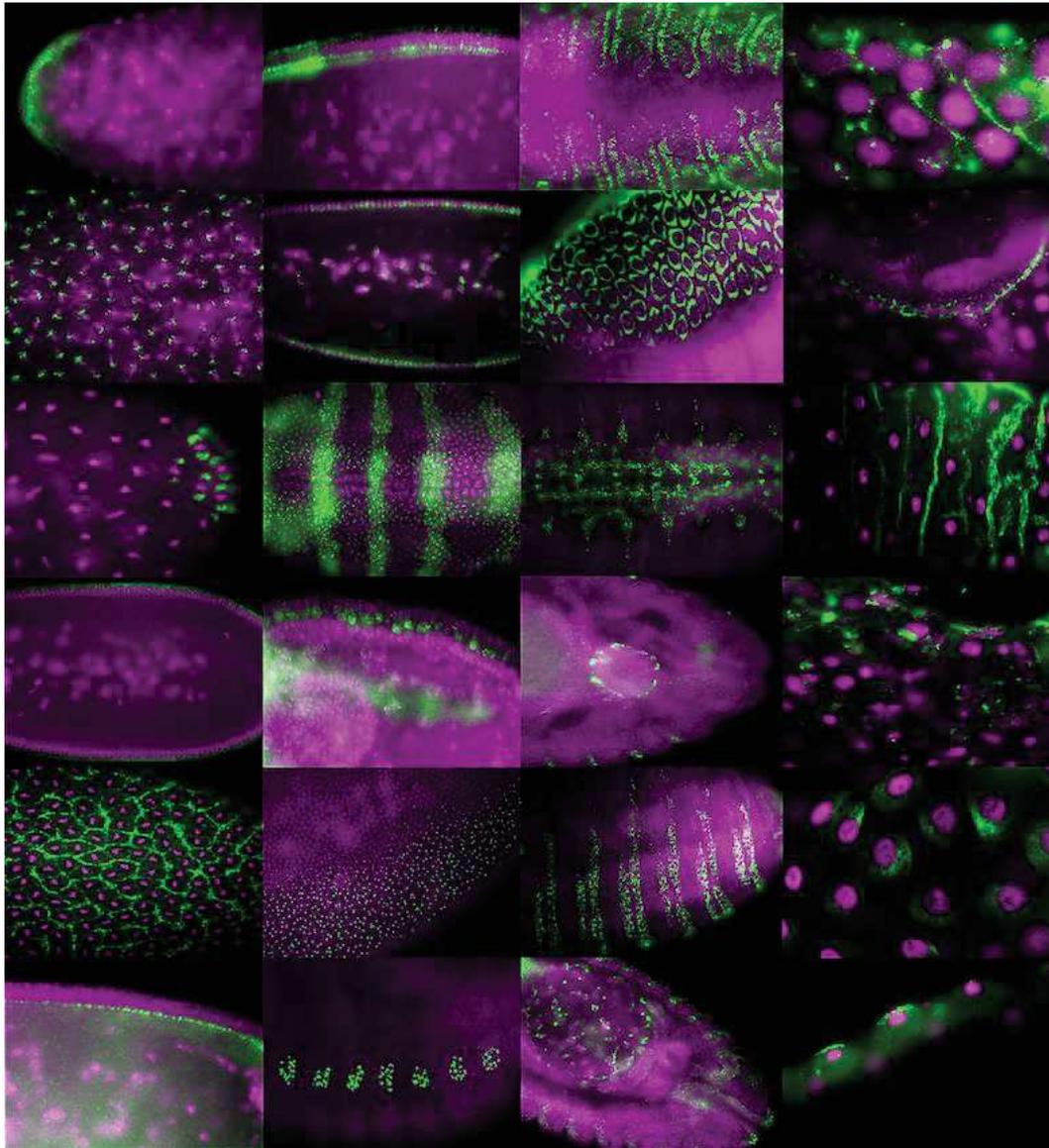


Figure 5.7 : In-situ hybridization of diverse mRNAs in *Drosophila* embryos.

DNA is represented in pink and mRNA in green. Images are from fly-fish.ccb.utoronto.ca.

Most of our knowledge about mRNA translation in the early embryo comes from staged proteomics (Fabre et al. 2016) but this does not provide information about the dynamics, in time and space, of translation regulation. We also know that transcription speed and efficiency are very optimized during the first hours of development, to ensure robustness in front of numerous morphological variations (Boettiger and Levine 2009; Lagha et al. 2012). But is it also the case for translation or is there a consequence on translation?

Aims

The main goal of this second part was to develop the SunTag method in a multicellular living organism. The goal of this deployment was to be able to answer to biological questions so far unreachable in an embryo: are all RNA molecules engaged in translation? Is there a spatial and/or temporal regulation? What are the main quantitative features we can extract of translation of developmental genes?

To answer this, we needed to develop tools to image translation dynamics in the *Drosophila* embryo, and precisely quantify its fluctuation during time and space.

Chapter 6. Results: Imaging translation dynamics of *twist*

TRANSLATION

Imaging translation dynamics in live embryos reveals spatial heterogeneities

Jeremy Dufourt^{1*†}, Maelle Bellec^{1†}, Antonio Trullo¹, Matthieu Dejean¹, Sylvain De Rossi², Cyril Favard³, Mounia Lagha^{1*}

Much is known about the factors involved in the translation of messenger RNA (mRNA) into protein; however, this multistep process has not been imaged in living multicellular organisms. Here, we deploy the SunTag method to visualize and quantify the timing, location, and kinetics of the translation of single mRNAs in living *Drosophila* embryos. By focusing on the translation of the conserved major epithelial-mesenchymal transition–inducing transcription factor Twist, we identify spatial heterogeneity in mRNA translation efficiency and reveal the existence of translation factories, where clustered mRNAs are cotranslated preferentially at basal perinuclear regions. Observing the location and dynamics of mRNA translation in a living multicellular organism opens avenues for understanding gene regulation during development.

More than 60 years ago, it was established that mRNA is translated to make protein. However, studies have revealed that the level of a given mRNA and the amount of protein it encodes do not directly correlate (1). This lack of colinearity may partially result from differential translational regulation in subcellular compartments where mRNAs are targeted (2, 3). To quantitate and compare mRNA and nascent protein, methods are needed to visualize these molecules *in vivo*. Live imaging of mRNA has been possible since 1998 (4), but

a similar method to image many cycles of translation was only established in 2016 in cultured cells (5–9) and has yet to be established in an intact developing organism. With its rapid development and the simple arrangement of nuclei in the syncytial blastoderm stage, the *Drosophila melanogaster* embryo represents a model organism to image gene expression.

To visualize translation using a reporter transgene, we used the SunTag system, whereby repetitions of an epitope (named *suntag*) are added to the protein of interest and are detected with a genetically encoded single-

chain antibody (called scFv) fused to a fluorescent protein (10) (Fig. 1A). To implement the SunTag method in *Drosophila* embryos, we focused on the gene *twist*, which encodes a conserved transcriptional activator of the mesodermal program in metazoans (11). In *Drosophila* early embryos, this gene is expressed during the activation of the zygotic genome in a specific ventral domain. We created a *twi_suntag* transgene (fig. S1F and supplementary text) that enables the labeling of Twi protein with 32 *suntag* repeats. Additionally, we created scFv-fluorescent lines to detect *suntag* peptides (fig. S1, A to E; movie S8; and supplementary text). In the presence of the *twi_suntag* transgene and scFv–green fluorescent protein (GFP) detector protein, distinct spots were detected within the presumptive mesoderm of living embryos (figs. S1 and S2, movie S9, and supplementary text). However, *twi_suntag* expression appeared stochastically in this domain (fig. S1G; fig. S2, B, E, and F; and supplementary text).

Having demonstrated our ability to observe translation with a reporter transgene, we then

¹Institut de Génétique Moléculaire de Montpellier, University of Montpellier, CNRS-UMR 5535, Montpellier 34293 cedex 5, France. ²MRI, BioCampus Montpellier, CNRS, INSERM, University of Montpellier, Montpellier, France. ³Institut de Recherche en Infectiologie de Montpellier, CNRS UMR 9004, University of Montpellier, Montpellier 34293 cedex 5, France. *Corresponding author. Email: jeremy.dufourt@igmm.cnrs.fr (J.D.); mounia.lagha@igmm.cnrs.fr (M.L.)
†These authors contributed equally to this work.

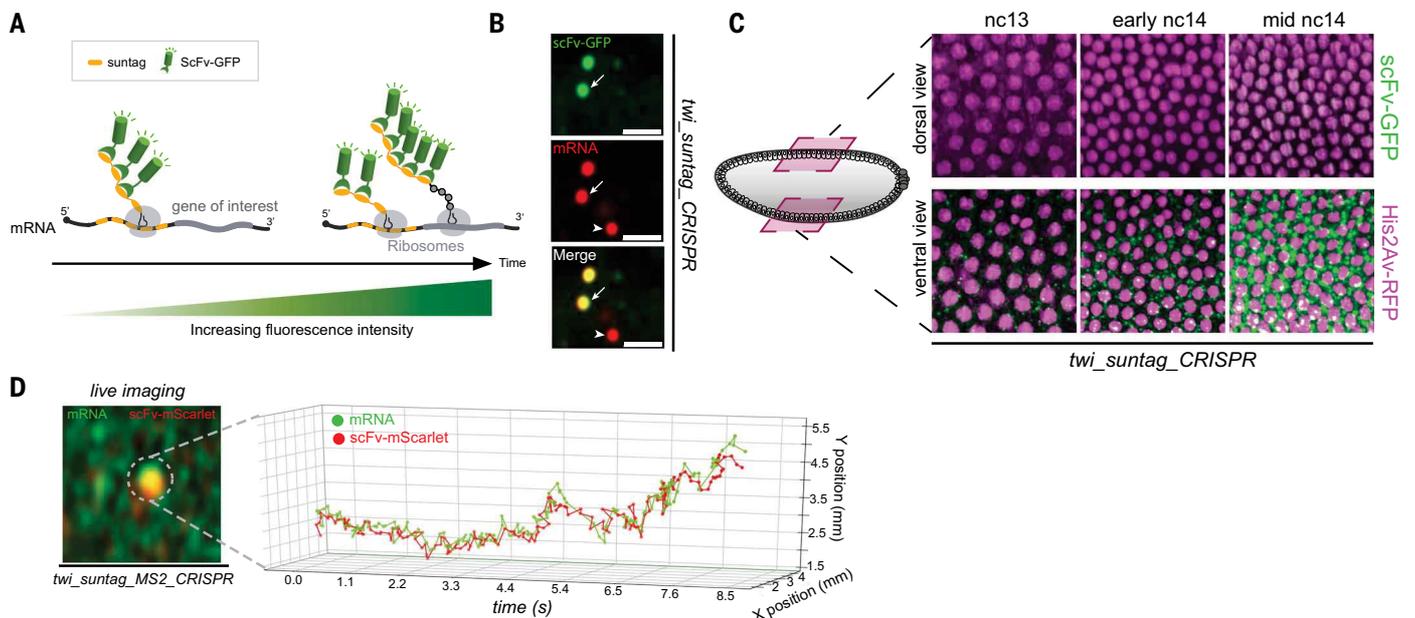


Fig. 1. Imaging translation of mRNAs in *Drosophila* embryos. (A) Principle of the SunTag system. Repetitions of *suntag* epitopes are added to the protein of interest and are detected with a single-chain antibody (scFv) fused to GFP. (B) Zoomed-in confocal images of n.c.14 *twi_suntag_CRISPR* *Drosophila* embryos expressing scFv-GFP (green) stained with *suntag* probes (red), exhibiting two groups of mRNA molecules: colocalizing with scFv-GFP signal (arrows) and not

colocalizing with a GFP signal (arrowheads). Scale bars, 1 μ m. (C) Live imaging of a *His2Av-mRFP/+;scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo by MuViSPIM (multiview selective plane illumination microscope) (images from movie S1). (D) Spatiotemporal tracking of mRNA and translation signal from an *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+* embryo (image from movie S3).

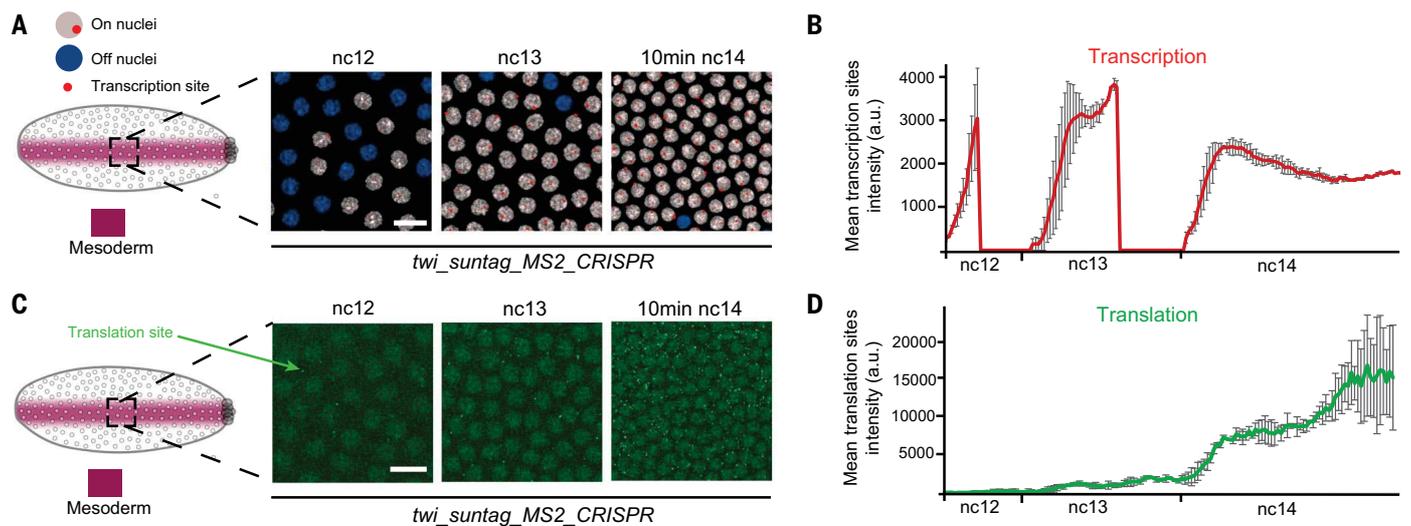


Fig. 2. Dynamics of transcription and translation. (A and B) Live imaging of an *MCP-eGFP-His2Av-mRFP>twi_suntag_MS2_CRISPR/+* embryo showing transcription sites (TSs) in red (A) and quantification of TS intensities (B) ($n = 2$ movies). Scale bar, 10 μm . a.u., arbitrary units. (C and D) Live imaging of a *scFv-GFP/+>twi_suntag_MS2_CRISPR/+* embryo (C) and quantification of translation site intensities over time (D) ($n = 2$ movies). Scale bar, 10 μm .

monitored *twi* translational dynamics from its endogenous locus with a *twi_suntag_CRISPR* allele (fig. S3, A and B). By performing single-molecule mRNA labeling [single-molecule fluorescence in situ hybridization (smFISH)] with the simultaneous detection of native scFv-GFP, we could detect two populations of cytoplasmic mRNA molecules: (i) those colocalizing with a bright GFP signal—i.e., $69 \pm 3\%$ in nuclear cycle 14 (n.c.14) ($n = 5$ embryos)—presumably corresponding to mRNAs being translated and (ii) those devoid of a GFP signal (Fig. 1B). Next, we questioned whether these bright scFv foci could be detected in living embryos with light sheet microscopy, and we found that *twi* translation was strongly induced during n.c.14 (Fig. 1C and movie S1) and was specific to the mesoderm. Bright but rare scFv-GFP foci appeared as early as n.c.12 and persisted during mitoses (fig. S3C).

To determine whether scFv-GFP spots correspond to nascent sites of translation, we imaged *twi_suntag_CRISPR* embryos injected with puromycin, a translation inhibitor. We did not observe scFv-GFP spots close to the injection site (fig. S3D and movie S10C). To observe nascent translation of single mRNA particles in live embryos, we engineered a *twi_suntag_MS2_CRISPR* and combined a *scFv-mScarlet* with an *MCP-GFP* transgenic line (fig. S4 and supplementary text). For this dual cytoplasmic imaging, single mRNA molecules are labeled with an MS2 array, visualized using the coat protein of bacteriophage MS2 (MCP) fused to GFP (12), while nascent proteins are labeled with the suntag peptides, recognized by the scFv antibody fused to mScarlet. Confocal imaging revealed distinct molecules of cytoplasmic mRNAs with, in

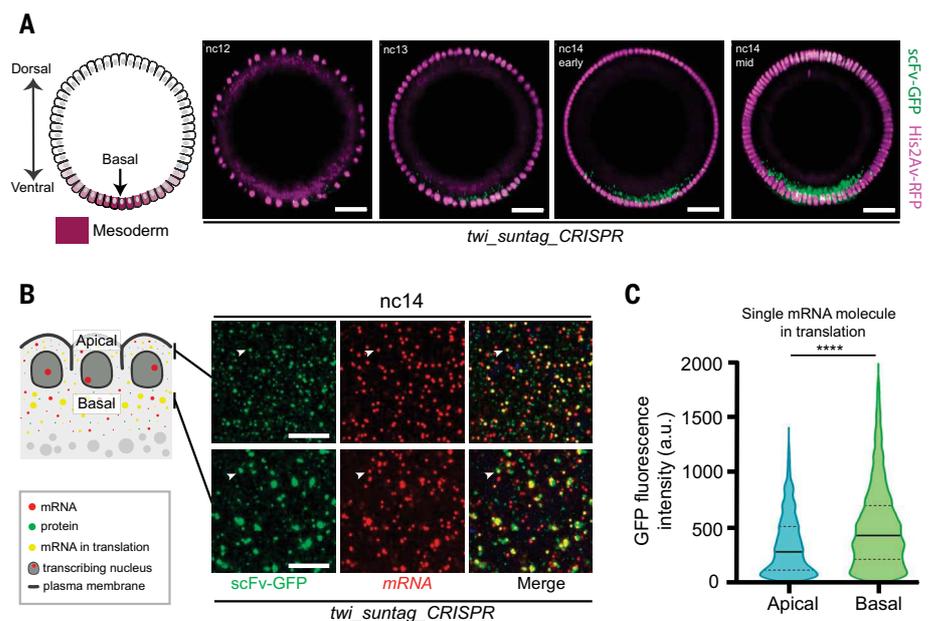


Fig. 3. Spatial heterogeneity of translation. (A) Live imaging of a *His2Av-mRFP/+;scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo (MuviSpim cross section) (images from movie S4). Scale bars, 30 μm . (B and C) Representative confocal image (apical and basal z-stacks shown separately) of an *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo expressing scFv-GFP (green) labeled with suntag probes (red) (scale bars, 5 μm) (B) and quantification shown as a violin plot of the distribution of scFv-GFP intensities colocalizing with single mRNA molecules located apically ($n = 2380$; blue) and basally ($n = 4202$; green) (C). Two-tailed Welch's t test; **** $P < 0.0001$.

some cases, a red scFv-mScarlet signal on top (fig. S4G and movie S2). This dual-color live imaging confirms the existence of two mRNA pools, with a subset of *twi* mRNA undergoing translation. It further shows that these mRNA and nascent proteins move together (Fig. 1D and movie S3), revealing that mRNAs in translation are not static.

By combining SunTag and MS2 labeling, it is possible to image transcription and translation and quantify their degree of correlation. In the case of *twi*, the timing of translation is consistent with its mRNA production (fig. S5A). Live imaging of the *twi_suntag_MS2_CRISPR* reveals that translation peaks in n.c.13 (Fig. 2, A and B) (13). Thus, the largest

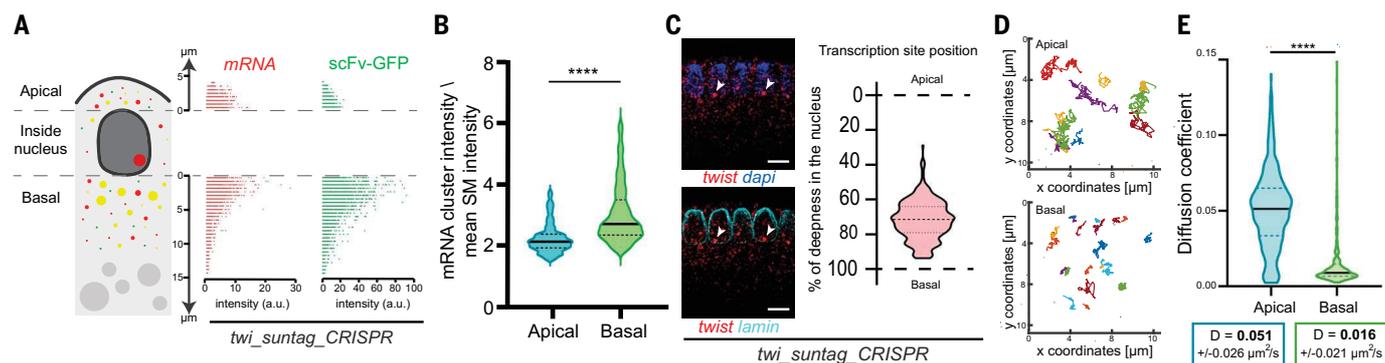


Fig. 4. Local translation of *twi* mRNA. (A) Intensity quantification of *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryos labeled with *suntag* probes showing mRNA spots (red) and *scFv-GFP* spots (green) in 0.5- μm -spaced Z-planes of four n.c.14 embryos. (B) Distribution of the intensity of mRNA clusters in apical ($n = 523$; blue) and basal ($n = 1384$; green), normalized by the mean intensity of a single mRNA molecule. Two-tailed Welch's t test; **** $P < 0.0001$. (C) Sagittal views of

twi_suntag_CRISPR/+ embryos labeled with *suntag* probes, anti-lamin, and 4',6-diamidino-2-phenylindole (DAPI) (scale bars, 5 μm) and corresponding quantification of TS positions along the apico-basal nuclear axis. (D and E) Single-particle tracking on *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryos. (D) Examples of color-coded translation foci trajectories. (E) Violin plots of the estimated diffusion coefficient distributions of apical and basal particles. Two-tailed Welch's t test; **** $P < 0.0001$.

wave of mRNA production precedes the timing of the largest burst of *twi* translation (Fig. 2, C and D; fig. S4D; and movies S1 and S12A). Further, the timing of *twi* translation is consistent with the timing of nuclear Twi protein emergence (fig. S5B).

To gain more insight into the dynamics of *twi* translation, we used the SunTag method to reveal translation kinetics (5, 7–9). We determined that SunTag-Twi fusion protein was fully translated (fig. S5C). Then, by correlating temporal intensity fluctuations of single spot *scFv-GFP* (5, 7–9), elongation and initiation rates were estimated to be in the order of 35 amino acids per second and 13 s, respectively (fig. S5, D to H). These rates are probably upper estimates and do not reflect the variability between mRNAs. Nonetheless, these rates lead to an overall translation efficiency of seven ribosomes per mRNA (fig. S5I), consistent with ribosome profiling experiments (14). Collectively these data suggest that the relatively late timing of *twist* translational activation could be partly compensated by its fast translation kinetics.

Using a transverse view of a developing embryo, the sites of translation in n.c.14 appeared much more prominent in the basal perinuclear region (i.e., toward the interior of the embryo), although translation was also observed in the apical perinuclear space (Fig. 3A, fig. S6A, and movies S4 and S12B). To further investigate this apparent spatial bias, we quantified the *scFv-GFP* signal in these two compartments (fig. S6, B and C, and fig. S7). In contrast to earlier developmental stages—where translation is equivalent in the apical and basal cytoplasmic spaces—in n.c.14, the largest and brightest spots of *twi* translation appeared mainly in the basal cytoplasm. To estimate translation efficiency, we extracted the intensity of the *scFv-GFP* signal overlapping individual mRNA molecules (see materials and meth-

ods). We found that in the basal perinuclear space, a single molecule of mRNA is on average 50% more intense in the *scFv-GFP* channel than a single molecule located apically, which suggests an enhanced efficiency of translation (Fig. 3, B and C). This bias is also observed with *twi_suntag* transgene (fig. S6, D and E). Collectively, these data demonstrate that translation efficiency of identical mRNA molecules depends on their subcellular localization. This spatial heterogeneity does not seem to rely on a differential distribution of ribosomes and might be supported by a higher basal availability of mitochondria (fig. S6, F and G, and movie S5).

Live imaging data revealed the existence of large *scFv-GFP* foci predominantly present in the basal cytoplasm. Simultaneous detection of mRNA and translation foci shows that these large size translation foci overlap large mRNA foci (Fig. 3B and fig. S6D). To better characterize these large foci, we quantified mRNA densities and *scFv-GFP* signal. Although mRNA molecules were present along the entire depth of a cell volume, their intensity was clearly enhanced at the level of the basal perinuclear space (Fig. 4A), where they tend to assemble in clusters (fig. S8, A and B). These mRNA clusters were of varying sizes and were larger in the basal perinuclear cytoplasm (Fig. 4, A and B, and fig. S8, B and C). In total, $94 \pm 3\%$ of these mRNA clusters were engaged in translation ($n = 4$ embryos). Thus, we consider them as translation factories, echoing what has been shown in mammalian cells (7, 15, 16). Similar translation factories are observed with an *ilp4-suntag* transgenic reporter (fig. S9 and supplementary text).

Twi translation factories are distinct from germ plasm granules and processing bodies (P-bodies) (fig. S10, A and B, and movie S15). Clustering of *twi* mRNA in the basal cytoplasm is also observed in wild-type as well as

in *twi* hemizygous embryos, albeit with a reduced frequency, which suggests that clustering partly depends on mRNA concentration (fig. S8D). Basal mRNA clustering is also detected for other mRNAs (fig. S11A). However, clustering of mRNAs is not a specific feature of the basal cytoplasm, as it is also observed apically (fig. S11, B and C) and largely documented for pair-rule genes (17, 18).

The site of *twi* mRNA major clustering might be, in part, dictated by the localization of its site of transcription (Fig. 4C; fig. S12, A and B; and movie S6). A preferential export of mRNA toward the basal cytoplasm would favor basal *twi* mRNA clustering, which would be rapidly cotranslated in factories. In the case of a nuclear protein like Twi, its translation in factories nearby the nuclear periphery could favor rapid nuclear import of newly formed proteins, as suggested by Twi protein stainings (fig. S12C).

twi translation occurs before complete cellularization. Consequently, its messenger ribonucleoproteins (mRNPs) could theoretically diffuse between neighboring pseudocells. To gain insight into *twi* mRNP mobilities, we tracked *twi_suntag_CRISPR* mRNPs in different cytoplasmic locations (Fig. 4D and movie S7). The trajectories and the mean square displacement (MSD) revealed clear, distinct properties of apical versus basal particles (fig. S13 and supplementary text). For example, the diffusion coefficient of mRNPs is one-third as fast in the basal compartment compared with the apical (Fig. 4E). The sublinear growth of the MSD curves suggests subdiffusive behavior in both compartments (fig. S11, C to G). Thus, we conclude that, in the basal perinuclear cytoplasm, *twi* translation sites diffuse slower because of their larger size.

By focusing on *twi* mRNAs as a paradigm for transcription factor encoding transcripts, we have uncovered fundamental features of

translation in a living organism such as heterogeneity in translation efficiencies of identical mRNAs and the existence of translation factories. Local translation of multiple mRNAs could have several benefits. First, it could favor the assembly of newly synthesized proteins in complexes. This is potentially the case for *Twi*, known to homodimerize (19). Second, localized protein synthesis could favor fast delivery of newly formed proteins to their destinations. Correlation between mRNA localization and protein function is well documented (2). The SunTag method now allows us to bridge the gap between mRNA and protein localization. In the case of *twi*, we propose that local and enhanced translation close to the nuclear envelope favor rapid nuclear import of neosynthesized *Twi* protein. This might be generalizable to other transcription factors, as proposed for pair-rule proteins (18).

Finally, the clustering of mRNAs and their cotranslation restricts the diffusion capacities of mRNPs. In the context of a syncytial embryo, this property could be exploited to limit the diffusion and allow spatial precision in cell fate decisions. As cellularization proceeds with an apico-basal directionality, apical anchoring of mRNAs represents an optimal strategy to limit diffusion. However, translation dynamics of these apical mRNAs remain to be demonstrated. In contrast, for mRNAs located basally in a compartment, where short-range diffu-

sion lasts for a relatively long period of time, we propose that clustering and rapid local translation restrict the diffusion capacities of mRNPs. Thus, precision in the establishment of developmental patterns cannot only be attributed to precision in transcriptional activation. We anticipate that our approaches will pave the way to investigating previously inaccessible translation modalities during development and differentiation.

REFERENCES AND NOTES

1. K. Becker *et al.*, *Nat. Commun.* **9**, 4970 (2018).
2. E. Lécuyer *et al.*, *Cell* **131**, 174–187 (2007).
3. A. Chin, E. Lécuyer, *Biochim. Biophys. Acta Gen. Subj.* **1861**, 2956–2970 (2017).
4. E. Bertrand *et al.*, *Mol. Cell* **2**, 437–445 (1998).
5. X. Yan, T. A. Hoek, R. D. Vale, M. E. Tanenbaum, *Cell* **165**, 976–989 (2016).
6. T. Morisaki *et al.*, *Science* **352**, 1425–1429 (2016).
7. X. Pichon *et al.*, *J. Cell Biol.* **214**, 769–781 (2016).
8. B. Wu, C. Eliscovich, Y. J. Yoon, R. H. Singer, *Science* **352**, 1430–1435 (2016).
9. C. Wang, B. Han, R. Zhou, X. Zhuang, *Cell* **165**, 990–1001 (2016).
10. M. E. Tanenbaum, L. A. Gilbert, L. S. Qi, J. S. Weissman, R. D. Vale, *Cell* **159**, 635–646 (2014).
11. T. Sandmann *et al.*, *Genes Dev.* **21**, 436–449 (2007).
12. K. Tantale *et al.*, *Nat. Commun.* **7**, 12248 (2016).
13. J. E. Sandler, A. Stathopoulos, *Genetics* **202**, 1575–1584 (2016).
14. S. W. Eichhorn *et al.*, *eLife* **5**, e16955 (2016).
15. L. Chang, Y. Shav-Tal, T. Trcek, R. H. Singer, R. D. Goldman, *J. Cell Biol.* **172**, 747–758 (2006).
16. R. Chouaib *et al.*, *Dev. Cell* **54**, 773–791.e5 (2020).
17. I. Davis, D. Ish-Horowicz, *Cell* **67**, 927–940 (1991).
18. S. L. Bullock *et al.*, *Development* **131**, 4251–4261 (2004).
19. I. Castanon, S. Von Stetina, J. Kass, M. K. Baylies, *Development* **128**, 3145–3159 (2001).

ACKNOWLEDGMENTS

We are grateful to E. Bertrand, R. Zinzen, I. Izzedin, P. Lasko, T. Hurd, and X. Pichon for sharing flies, reagents, and software. We thank C. Desplan, J. Chubb, R. Bordonne, F. Besse, T. E. Saunders, J. Dejardin, and V. L. Pimmett, for their critical reading of the manuscript. We acknowledge L. Bellec, H. Lenden, and M. Goussard for technical assistance. We acknowledge the Montpellier Ressources Imagerie facility (France-Biologymaging). **Funding:** M.B. is a recipient of an FRM fellowship. This work was supported by the ERC SyncDev starting grant and a HFSP-CDA grant to M.L. M.L., J.D., and C.F. are sponsored by CNRS. S.D.R. is sponsored by INSERM. **Author contributions:** M.L. conceived the project. M.L. and J.D. designed the experiments. J.D., M.B., and M.D. performed experiments. A.T. developed software. C.F., M.B., and J.D. performed kinetic analysis. S.D.R. and J.D. performed MuViSPIM imaging. J.D., M.L., A.T., M.B., and C.F. analyzed the data. J.D., M.L., and M.B. interpreted the results. M.B. created artwork. M.L. wrote the manuscript with help from J.D. and M.B. All authors discussed, approved, and reviewed the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text and/or the supplementary materials, and fly stocks will be deposited at Vienna Drosophila Resource Center.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6544/840/suppl/DC1
Materials and Methods
Supplementary Text
Supplementary Sequences
Figs. S1 to S14
Table S1
References (20–37)
MDAR Reproducibility Checklist
Movies S1 to S15

[View/request a protocol for this paper from Bio-protocol.](#)

29 April 2020; accepted 13 April 2021
Published online 29 April 2021
10.1126/science.abc3483

Imaging translation dynamics in live embryos reveals spatial heterogeneities

Jeremy Dufourt, Maelle Bellec, Antonio Trullo, Matthieu Dejean, Sylvain De Rossi, Cyril Favard and Mounia Lagha

Science **372** (6544), 840-844.

DOI: 10.1126/science.abc3483 originally published online April 29, 2021

Quantifying translation in space and time

During development, precise control of gene expression establishes reproducible patterns, leading to the formation of organs at the right time and place. The emergence of developmental patterns has been primarily studied at the transcriptional level, but the fate of these transcripts has received little attention. Dufourt *et al.* used the SunTag labeling method to image the dynamics of translation of individual messenger RNA (mRNA) molecules in living fruit fly embryos. This work revealed "translation factories"—clusters of mRNA and translation machinery—and heterogeneities in the efficiency of translation between identical mRNAs.

Science, abc3483, this issue p. 840

ARTICLE TOOLS

<http://science.sciencemag.org/content/372/6544/840>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2021/04/28/science.abc3483.DC1>

REFERENCES

This article cites 37 articles, 10 of which you can access for free
<http://science.sciencemag.org/content/372/6544/840#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works



science.sciencemag.org/cgi/content/full/science.abc3483/DC1

Supplementary Materials for

Imaging translation dynamics in live embryos reveals spatial heterogeneities

Jeremy Dufourt*†, Maelle Bellec†, Antonio Trullo, Matthieu Dejean, Sylvain De Rossi, Cyril Favard, Mounia Lagha*

*Corresponding author. Email: jeremy.dufourt@igmm.cnrs.fr (J.D.); mounia.lagha@igmm.cnrs.fr (M.L.)

†These authors contributed equally to this work.

Published 29 April 2021 on *Science* First Release

DOI: 10.1126/science.abc3483

This PDF file includes:

Materials and Methods
Supplementary Text
Supplementary Sequences
Figs. S1 to S14
Table S1
Captions for Movies S1 to S15
References

Other Supplementary Material for this manuscript includes the following:

(available at science.sciencemag.org/cgi/content/full/science.abc3483/DC1)

MDAR Reproducibility Checklist (.pdf)
Movies S1 to S15 (.mp4)

Materials and Methods

***Drosophila* stocks and genetics**

The *yw* stock was used as a control. *His2av-mRFP* (BL23651), *UASp-me31B-GFP* (BL51530), *nos-Gal4:VP16* (BL4937) stocks come from Bloomington. The *UASp-mito-mCherry* strain was gift from T.Hurd. The germline driver *nos-Gal4* was recombined with the *scFv-sfGFP-NLS* and the *scFv-mScarlet-NLS*. *MCP-eGFP-His2Av-mRFP* comes from(20).

Cloning and Transgenesis

The *twi-suntag* transgene was synthesized (GenScript Biotech) (Supplementary text) with 32x *suntag* repeats(7) into pUC57-simple. The *twi_suntag_MS2* transgene was generated based on the *twi_suntag* transgene with 128x MS2 repeats(12) inserted in the *XbaI* restriction site. Constructs were inserted into pbPHi(21) using *PmeI* and *FseI* and injected into BL9750 using PhiC31 targeted insertion(22)(BestGene, Inc.). All *scFv-sfGFP*, *scFv-TagRFPT* and *scFv-mScarlet* lines were generated using NEBuilder® HiFi DNA Assembly Master Mix with primers listed in supplementary table 1 and inserted into pNosPE_MCP-eGFP (Supplementary text) after removal of MCP-eGFP.

The recombination templates for CRISPR/Cas9 editing of *twist* gene to generate *twi_suntag_CRISPR* and *twi_suntag_MS2_CRISPR* were assembled with NEBuilder® HiFi DNA Assembly Master Mix (primers listed in supplementary table 1) and inserted into pBluescript opened *KpnI* and *SacI*.

The *twi-suntag* transgene was digested with *FseI* and *SacII* and inserted into the CRISPR recombination template opened with *FseI* and *SacII*. 128x MS2 repeats(12) were inserted in an *XbaI* restriction site. Guide RNA (supplementary table 1) were cloned into pCFD3-dU6:3gRNA (Addgene 49410) digested by *BbsI* using annealed oligonucleotides (Integrated DNA Technology™). The recombination template and guide RNA plasmids were injected into BDSC#55821 (BestGene Inc.). Transformant flies were screened using dsRed marker inserted downstream of the 3'UTR of *twi_suntag_CRISPR* and *twi_suntag_MS2_CRISPR*. *MCP-TagRFPT* constructs were assembled by replacing the eGFP fragment of pNosPE_MCP-eGFP using

NheI/BamHI with the TagRFPT coding sequence amplified by PCR (supplementary table 1) from TagRFP-T-Rabenosyn-5 (Addgene 37537).

MCP-TagRFPT-NLS was generated by insertion of the TagRFPT-NLS coding sequence into pNosPE_MCP-eGFP with NEBuilder® HiFi DNA Assembly Master Mix (primers listed in supplementary table 1).

Twipe_TwiPr_24XMS2 was generated by insertion of the *Twipe_TwiPr* sequence into *Not1 BamHI* opened snaE enhancer plasmid(20, 21) using NEBuilder® HiFi DNA Assembly Master Mix (primers listed in supplementary table 1).

The *ilp4-suntag* transgene was synthesized (GenScript Biotech) (Supplementary text) into pUC57-simple. The 32x suntag repeats were inserted into *Kpn1* and *EcoRV* restriction sites. Constructs were inserted into pbPhi(21) using *PmeI* and *Nhe1* and injected into BL9750 using PhiC31 targeted insertion(22) (BestGene, Inc.).

Western blot analysis

Fifty embryos from wild type (*yw*), *twi_suntag_CRISPR* and *twi_suntag_MS2_CRISPR* were collected 1 to 5 hours after egg laying and crushed in 100µl of NuPAGE™ LDS sample buffer and reducing agent. Samples were heated 10min at 70°C, and the volume-equivalent of 5 embryos was loaded per well on a 4-12% Bis-Tris NuPAGE™ Novex™ gel and ran at 180V. Protein transfer was done for 1h10 at 110V to a nitrocellulose membrane, 0.2 µm (Invitrogen, LC2000). Membrane was blocked in 5% milk-PBT (PBS 1X 0.1% Tween 20) for 40 min and incubated overnight at 4°C with primary antibodies 1/1000 mouse anti-GCN4 (Novus Biologicals, C11L34, NBP2-81274) or mouse 1/2000 anti-Tubulin in PBT. Anti-mouse and -rabbit IgG-HRP (Cell Signaling #7076 and #7074) secondary antibody was used at a 1/4000 dilution and incubated 1hour at room temperature. Chemiluminescent detection was done using Pierce™ ECL Plus (ThermoFisher) kit. After the detection of the first antibody (anti-GCN4), the membrane was incubated 10min in Thermo Scientific™ Restore™ PLUS stripping buffer and blocked as before in order to hybridize with anti-Tubulin.

Single molecule fluorescence *in situ* hybridization (smFISH) and Immunostaining

Embryos were dechorionated with bleach for 3 min, thoroughly rinsed with H₂O and fixed in a 1:1 solution of 10% formaldehyde: 100% heptane for 25 min with

shaking. Formaldehyde was replaced by methanol and embryos were vortexed for 1min. Embryos that sank to the bottom of the tube were rinsed three times with methanol before storage at -20°C. For immunostaining, embryos were rinsed with methanol and washed three times with PBT (PBS 1X 0.1% Triton X-100). Embryos were incubated on a rotating wheel at room temperature twice for 30 min in PBT, once for 20 min in PBT+ 1% BSA, and at 4 °C overnight in PBT 1% BSA with guinea-pig anti-Twist 1/200 (gift from Robert Zinzen), rabbit anti-RpS5a, RpS5b (gifts from Paul Lasko, both 1/200), rabbit anti-RpS6 (Cell Signaling #2217, 1/200), rabbit anti-Aub (gift from Paul Lasko, 1/200), rat anti-Vasa (DSHB 760351, 1/200), mouse anti-Lamin (DSHB ADL67.10, 1/200), mouse anti-Dorsal (DSHB 7A4, 1/50). Embryos were rinsed three times and washed twice for 30 min in PBT, then incubated in PBT+ 1% BSA for 30 min, and in PBT +1% BSA with secondary antibodies anti-guinea-pig Alexa 555-conjugated (Life technologies, A21435), anti-mouse Alexa 647 (Invitrogen, A32728), anti-rabbit 488 (Invitrogen, A21206) 1/500 for 2 h at room temperature. Embryos were rinsed three times then washed three times in PBT for 10 min. DNA staining was performed using DAPI at 0.5 µg/mL.

smFISH was performed as follows: wash 5min in 1:1 methanol:ethanol, rinse twice with ethanol 100%, wash 5min twice in ethanol 100%, rinse twice in methanol, wash 5min once in methanol, rinse twice in PBT-RNa (PBS 1x, 0.1% tween, RNasin® Ribonuclease Inhibitors). Then, embryos were washed 4 times for 15 min in PBT-RNa supplemented with 0.5% ultrapure BSA and then once 20 min in Wash Buffer (10% 20X SCC, 10% Formamide). They were then incubated overnight at 37°C in Hybridization Buffer (10% Formamide, 10% 20X SSC, 400µg/ml *E.coli* tRNA (New England Biolabs), 5% dextran sulfate, 1% vanadyl ribonucleoside complex (VRC) and smFISH Stellaris probes against *suntag* coupled to Quasar 570 or Quasar 670 and/or against *twist* coupled to Quasar 570 and/or against *snail* coupled to Quasar 670 and/or against *htl* coupled to Quasar 570. Probe sequences are listed in supplementary table 1. Probes against 32X MS2 coupled to Cy3 were a kind gift from Edouard Bertrand. FLAP-probes against *ush* and *pnr* (sequences provided upon request) were prepared by duplexing 40pmol of target-specific probes with 100pmol FLAP-Cy3 oligonucleotides and 1X NEBuffer™ 3.1 for 3min at 85°C, 3min at 65°C and 5min at 25°C and kept on ice until use.

Embryos were washed in Wash Buffer at 37°C and in 2X SCC, 0.1% Tween at room temperature before being mounted in Pro-Long® Diamond antifade reagent. Images

were acquired using a Zeiss LSM880 confocal microscope with an Airyscan detector in SR mode with a 40x Plan-Apochromat (1.3NA) oil objective lens, a 63X Plan-Apochromat (1.4NA) oil objective lens or a 20x Plan-Apochromat (0.8NA) air objective lens. GFP was excited using a 488nm laser, Cy3 and Quasar570 were excited using a 561nm laser, Quasar670 was excited using a 633nm laser.

Images in Figure 3B represent a maximum intensity projection of a stack of 3 z-planes ($\approx 1\mu\text{m}$). Images in Figure 4C represent a maximum intensity projection of a stack of 2 z-planes ($\approx 0.5\mu\text{m}$). Images in Figure S11C represent a maximum intensity projection of a stack of 4 z-planes ($\approx 1.5\mu\text{m}$). Images in Figure S12C represent a maximum intensity projection of a stack of 2 z-planes ($\approx 0.5\mu\text{m}$).

Image analysis from smFISH experiments

Analysis of smFISH data was accomplished by custom-made algorithms developed in PythonTM. Briefly, a blob detection was performed on the scFv-GFP and mRNA channels (green and red respectively) separately. Raw data were filtered frame by frame with a two-dimensional Difference of Gaussian Filter with kernel sizes determined as in(23) and the filtered images were thresholded with a user-defined threshold value; the choice of the threshold was driven by visual inspection via an interactive graphical tool. All 3D connected components of the resulting binary images were considered as spots, which were then filtered in size with a volume threshold of 10 pixels. We removed spots with centroids lying in a frame where nuclei were detected. This allowed us to restrict our analysis to mRNAs located in the apical or basal perinuclear cytoplasmic regions. Within the detected mRNA pool, three populations could be distinguished by volume or intensity.

Analysis of smFISH data allowed detection of single molecules of mRNA, mRNAs in clusters and mRNAs undergoing decay. By looking at the mRNA spots not engaged in translation (red spots not overlapping green spots), a separable population can be characterized by a very low spot volume. It was theorized that these small spots are mRNA in decay, that are unable to be translated. To remove these small mRNA spots possibly in decay, the volume threshold of a single molecule mRNA was refined by looking at the histogram of the volume of the mRNA spots in translation (red spots overlapping with at least one pixel of a green spot). This led to the detection of two major populations on the histogram: single molecules of mRNA, representing the larger population shown in the histogram, and mRNA clusters, characterized by a smaller

population and a larger volume. A double Gaussian fitting was performed on the histogram to differentiate the single molecule and cluster populations. Then a threshold value was defined for the lower bound of the single molecule volume as $\mu - 3\sigma$, where μ and σ are, respectively, the mean and the standard deviation of the Gaussian function fitting single molecule mRNA spots.

Verifying the shape of single molecules and clusters further validated the segmentation. A single molecule is a spot-like object, and so its intensity pattern has the shape of a Point Spread Function (PSF) that can be well approximated as a 3D Gaussian function. In order to check that the detected spots are single molecules, the intensity profile of the detected spots was tested against a theoretical PSF for agreement. For each spot, a 3D Gaussian fitting was performed and the goodness of the fit estimated with the R^2 value given as

$$R^2 = 1 - \left(\sum_i (y_i - f_i)^2 \right) / \left(\sum_i (y_i - \bar{y})^2 \right)$$

where y_i are the observed values, f_i are the fitted values and \bar{y} the averaged observed value. R^2 is defined so that it approaches 1 if the fitting is perfect and close to 0 if the fitting is the average value of the distribution. Outliers from fitting were removed using the ROUT method(24), embedded in GraphPad Prism 8 software, with Q set to 0.1%.

Having obtained the correct threshold value by image analysis for the volume, the mRNA channel was reanalyzed to connect the spot detection results from both channels. Two clear populations of mRNAs were obtained that were divided into two subpopulations based on position above or below the nucleus.

The first parameter extracted was the percentage of mRNA in translation, given by the ratio between the number of mRNA spots overlapping a translation spot, divided by the total number of mRNA spots.

The efficiency of translation was determined by estimating mRNA single molecule intensity separately for above and below the nucleus, by performing a Gaussian fitting on the histogram of intensities of all detected mRNA spots. Single molecules were, by far, the largest population, so a single Gaussian fitting gives an estimate of the single molecule intensity not influenced by the arbitrarily small population of mRNA clusters, also present in the histogram. The single molecule upper intensity bound was defined as $\mu + 3\sigma$, where μ and σ are the mean and the standard

deviation of the fitting respectively. Single mRNA molecules and clusters of *twist* endogenous mRNAs (related to Figure S8D) were quantified in 3D in the basal region. Detected spots were split into single molecules and clusters using an intensity threshold. The intensity threshold was determined starting from the histogram of the intensities of all the detected objects: we took the central value of the bin with the higher population as mean value (μ) and then we calculated the data dispersion (DD) as:

$$DD = \frac{1}{N} \sum_1^N |x_i - \mu|$$

where N is the number of detected objects and x_i is the intensity of each object. The threshold intensity was defined as $\mu + DD$. All the objects with a higher intensity were considered as clusters, the others as single molecules.

For translation spots, the intensity was rescaled, independently for each spot, by dividing by the background to account for the free pool of GFP provided by the *nos-scFv-GFP-NLS* transgene. The background was calculated for each spot as the average intensity value of the pixels surrounding the spot itself. This allowed us to rescale by the background, because in this case there is a bath of free diffusing GFP.

To score differential clustering of mRNAs, above and below nuclei, all mRNA spots intensities above $\mu + 3\sigma$ (mRNA clusters) were rescaled by the mean intensity of a single mRNA molecule extracted from the Gaussian fitting, which allowed each mRNA entity (cluster of varying intensities) to be expressed as a function of a single molecule intensity. For translation efficiency and clustering of mRNA, outliers were removed using the ROUT method(24) embedded in GraphPad Prism 8 software, with Q set to 0.1%. For smFISH of the CRISPR alleles, analysis of the apical region was performed using between 5 and 8 confocal Z-planes spaced 0.5 μ m apart, and analysis of the basal region using 17 to 28 Z-planes spaced 0.5 μ m apart and only during the initial twenty minutes of n.c.14. At late n.c.14, green foci not overlapping mRNA are observed, representing potential aggregates. For smFISH analysis with the *twi_suntag* transgene, only one filtering with a volume threshold of 10 pixels was performed as only two populations appeared (single molecule mRNA and mRNA clusters). Analysis for the apical region was performed using 6 to 12 confocal Z-planes spaced 0.5 μ m apart and analysis for the basal region using 36 to 41 confocal Z-planes spaced 0.5 μ m apart.

Intensity Spatial Correlation analysis

The variation of the aggregation level of fluorophores was studied as a function of the apico-basal coordinate using Intensity Spatial Correlation analysis(25). For each z-frame of our Z-stack, the spatial correlation was calculated and then fitted with a 2D Gaussian function:

$$g(c, n) = g(0,0) * \exp \{-(c^2 + n^2)/w_0^2\} + g_0.$$

From fitting we learn $g(0,0)$, correlation value at the origin, w_0 which is the waist of the Gaussian and g_0 which is the average background value. The cluster density (CD) is then given by:

$$CD = \langle n \rangle / (\pi w_0^2) \approx 1 / (g(0,0) * w_0^2)$$

where $\langle n \rangle$ is the density of the fluorophores and $\langle n \rangle \approx 1/g(0,0)$ (26). The degree of aggregation (DA) is given by:

$$DA = \langle i \rangle / CD \rightarrow DA \approx \langle i \rangle * g(0,0) * w_0^2$$

where $\langle i \rangle$ is the average intensity of the z-frame. The estimate of the degree of aggregation is only proportional to the true value due to the approximations used, and so can only be employed ratiometrically between two frames of the same stack.

Transcription Site apico-basal coordinate

To position the TS within the apico-basal volume of the nucleus, nuclei were first established using a 3D reconstruction. For each frame of a Z-stack, a 2D segmentation was performed on the DAPI channel. These sections were compiled in 3D to form nuclear volumes, which were then fit with an ellipsoid shape to approximate the 3D shape of the nucleus. The apicobasal positioning of the TS (identified with a blob detection algorithm and a high-volume threshold to be separated by mRNA molecules using our smFISH custom-made algorithms developed in Python™) could

then be defined as the Z position of the TS center of mass relative to the Z position of the first frame of the ellipsoidal fitting.

Light-Sheet Microscopy

For light-sheet imaging (related to Movie S1, S4 and S9, S12, S14 and Figures 1C, 3A, S4D and S6A), we employed the MuViSPIM(27) (Luxendo, Brüker company GMBH). This setup provides two-sided illumination with two Nikon 10x/0.3 water objectives and two-sided detection with two Olympus 20x/1.0 W objectives. The light sheet is created by scanning a Gaussian beam, as in digital scanned laser light-sheet microscopy (DSLM). We used the line mode which is an in-built synchronization between rolling shutter readout mode of sCMOS cameras and digital light sheet scanning. This allows the rejection of out-of-focus and scattered light, thus improving the signal to noise ratio. Images are acquired by two ORCA Flash 4.0 (C91440) from Hamamatsu and processed by LuxControl v1.10.2. A 50ms exposure time was used for the green and red channel with a 488nm and 561nm laser excitation respectively. Maximum intensity projections were processed with Fiji(28) (Movie S1, Movie S9 and Movie S12A). Fusion is processed by Image Processor v2.9.0 from Luxendo (GMBH). Deconvolution was performed after the fusion process using Huygens Professional v 19.10 (Scientific Volume Imaging B.V). The Gaussian multi-view light sheet parameters were used for processing of 3D+t images. 3D reconstruction was done with Imaris v9.5.0 (Bitplane, Oxford company). The ortho slicer tool was used to show cross sections of embryos with $\approx 5\mu\text{m}$ extended section thickness for Movie S4 and Movie S12B and $\approx 10\mu\text{m}$ extended section thickness for Movie S14 and Figure S5J.

Live imaging of the scFv genetic tools

For all live imaging experiments, embryos were dechorionated with tape and mounted between a hydrophobic membrane and a coverslip as described previously(20).

Movies for *scFv-GFP-noNLS* (related to Movie S8A) and *scFv-GFP-NLS* (named *scFv-GFP* in the text) (related to Movie S8B) were acquired using a Zeiss LSM780 confocal microscope with a Plan-Apochromat 40x/1.3 oil objective lens, with GFP excitation using a 488nm laser. A GaAsP detector was used to detect GFP fluorescence with the

following settings: 1024x 1024 pixels, with each Z-stacks comprised of 6 planes spaced 2 μ m apart, 16-bit and zoom 2.0.

Movies for *scFv-GFP-Bcd3'UTR-NLS* (related to Figure S1E) were acquired using a Zeiss LSM780 confocal microscope with a Plan-Apochromat 40x/1.3 Oil objective and the following settings: 512x 512 pixels, 16-bit.

Movies for *scFv-mScarlet* (related to Movie S8D) were acquired using a Zeiss LSM880 confocal microscope with a Plan-Apochromat 40x/1.3 oil objective lens, with mScarlet excitation using a 561nm laser. A GaAsP detector was used to detect mScarlet fluorescence with the following settings: 1024x 1024 pixels, each Z-stack comprised of 14 planes.

Movies for *scFv-TagRFPT* (related to Movie S8C) were acquired using a Zeiss LSM880 confocal microscope with a Plan-Apochromat 40x/1.3 oil objective lens, with TagRFPT excitation using a 561nm laser. A GaAsP detector was used to detect TagRFPT fluorescence with the following settings: 1024x 1024 pixels, with each Z-stack comprised of 14 planes.

mRNA and SunTag dual-color live imaging

Movies for *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_transgene/+* (related to Figure S2D and Movie S11) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mScarlet were excited using a 488nm and 561nm laser respectively, with the following settings: 220x 220-pixel images and zoom 20x. For visual representation we took the square value of each pixel.

Movies for *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+* (related to Figure S4G and Movie S2) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mScarlet were excited using a 488nm and 561nm laser respectively with the following settings: 220x 220-pixel images and zoom 20x. Under these conditions a frame was acquired every \approx 103ms. For visual representation we took the square value of each pixel.

Movies for *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+* (related to Figure 1D and Movie S3) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mScarlet were excited using a 488nm and 561nm laser respectively, with the

following settings: 292x 292-pixel images and zoom 15x. Under these conditions a frame was acquired every ≈ 46 ms. To follow motions of translating foci and their associated RNA, the two-color images were then split into 2 channels and analyzed independently in Icy(11). The x-y positions of the RNA and translation foci centers were determined using the active contour plugin of Icy. RNA and translation foci were then tracked over time to determine the changes in these center positions using the same plugin.

Movie for *MCP-eGFP-His2Av-mRFP/+>twi_suntag_MS2_CRISPR/+* (related to Figure S4E-F and Movie S13) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and RFP were excited using a 488nm and 561nm laser respectively with the following settings: 276x 276-pixel images, with each Z-stack comprised of 3 planes spaced 1 μ m apart and zoom 8x. To enhance the difference between signal and background for visual representation we took the square value of each pixel and processed stacks containing 3 Z-planes ($\approx 2\mu$ m) were maximum intensity projected.

Live imaging of transcription

Movies of *MCP-eGFP-His2Av-mRFP>twi_suntag_MS2_CRISPR/+* and *MCP-eGFP-His2Av-mRFP>TwiPE_TwiPr_24XMS2/+* (related to Figure 2A-B and Figure S2F) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mRFP were excited using a 488nm and 561nm laser respectively with the following settings: 892x 892-pixel images, 35 Z-planes 0.5 μ m apart and zoom 1.8x.

Movies were subject to filtering steps to track transcription foci as 128x MS2 loops result in signal retention during mitosis. Briefly, using a custom-made algorithm developed in PythonTM, nuclei were segmented and tracked in 2D, working on the maximum intensity projected stack.

Transcription spots were detected and tracked in 3D: all the spots present during mitosis were removed in the successive cycle such that only *de novo* appearing MS2 punctae were analyzed. Once the mitotic spots were removed, the intensity of detected spots was collected for each frame to study the transcriptional intensity behavior throughout each nuclear cycle.

Movie for *MCP-eGFP-His2Av-mRFP>twi_suntag_MS2_CRISPR/+* (related to Figure S12B and Movie S6) were acquired using a Zeiss LSM880 with confocal

microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and RFP were excited using a 488nm and 561nm laser respectively with the following settings: 220x 220-pixel images, each Z-stacks comprised 17 planes spaced 1 μ m apart and zoom 10x. Figure S12B and Movie S6 show a maximum intensity projection of 8 Z-planes (\approx 7 μ m).

Live imaging of translation

Movies of *scFv-GFP-NLS/+>twi_suntag_MS2_CRISPR/+* (related to Figure 2C and D) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP was excited using a 488nm laser with the following settings: 640x 640-pixel images, each Z-stack comprised of 50-60 planes spaced 0.5 μ m apart and zoom 2.5x. Movies were then processed to remove frame outside of the embryos or containing the membrane signal to correct drifting, and processed stacks containing 20 Z-planes (\approx 9.5 μ m) were maximum intensity projected (using a custom-made software, developed in PythonTM). The sum of translation intensity was extracted frame by frame in 2D from maximum intensity projected movies. A non-linear (logarithmic) transformation was first performed along with a pre-smoothing on the data to prepare the movie for spots detection. The time series was then filtered with a Difference of Gaussian filter and the resulting time series was thresholded with a threshold value expressed as a function of the mean and standard deviation of the filtered stack. The thresholded time series was used to mask the raw data, so that the translation intensity could be defined (frame by frame) as the sum of the pixel intensities of the masked stack.

Movies of *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* (related to Figure S6B and C) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP was excited using a 488nm laser with the following settings: 640x 640-pixel images, each Z-stack comprised of 50-60 planes spaced 0.5 μ m apart and zoom 2.5x.

Movies were then processed to remove frame outside of the embryos or containing the membrane signal to correct drifting, and processed stacks were maximum intensity projected using custom-made software, developed in PythonTM.

Snapshots related to Figure S6B show: Upper panels a maximum intensity projection of four Z-planes located apically (\approx 1.5 μ m). Lower panels show a maximum intensity projection of 21 Z-planes located basally (\approx 10 μ m).

Movies for *UASp-mito-mCherry/+;nosGal4,scFv-GFP-NLS/+>twi_suntag_CRISPR/+* (related to Figure S6G and Movie S5) were acquired using a Zeiss LSM880 confocal microscope with a Plan-Apochromat 40x/1.3 oil objective lens, with GFP excitation using a 488nm laser. A GaAsP detector was used to detect GFP and a PMT for mCherry fluorescence with the following settings GFP and mCherry were excited using a 488nm and 561nm laser respectively with the following settings: 580x590-pixel images, 8 Z-planes 0.5 μ m apart and zoom 4x. Live imaging snapshots in Figure S6G show a maximum intensity projection of 8 Z-planes (\approx 3.5 μ m) taken \approx 10min after nc13-nc14 mitosis.

Movies for *UASp-me31B-GFP/+;nosGal4,scFv-mScarlet-NLS/+>twi_suntag_CRISPR/+* (related to Figure S10B and Movie S15) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP and mScarlet were excited using a 488nm and 561nm laser respectively with the following settings: 568x568-pixel images, each Z-stack comprised of 10 planes spaced 0.5 μ m apart and zoom 2x.

Laser power was measured regularly with 10x Plan-Apochromat air objective lens and was maintained in the range of \approx 28 μ W for the 488nm laser, \approx 19 μ W for the 561nm laser and \approx 17 μ W for the 633nm laser.

Translation Kinetics Analysis

Movies for extraction of translation kinetics from *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* (related to Figure S5D-I) embryos were acquired during the first 20min of n.c.14 using a Zeiss LSM880 with confocal microscope in fast Airyscan mode with a Plan-Apochromat 40x/1.3 oil objective lens. GFP was excited using a 488nm laser with the following settings: 132 x 132-pixel images, each Z-stack comprised of 15 planes spaced 0.5 μ m apart and zoom 8x. Under these conditions a Z-stack was acquired every \approx 0,75s.

Fluorescence intensity fluctuations of translation sites were obtained with Icy image software(29), using spot detector and tracking plugins. In order to generate the tracked intensity fluctuation, each detected spot was fitted to a Gaussian from which background was subtracted. Tracks were then selected if their time length was superior to 150 s, and were discarded when bleaching was observed during the time course. Moreover, as performed in (6, 8), tracks were kept only if the decorrelation time was above 5s.

Temporal autocorrelation of the intensity fluctuation was performed in Matlab using the multi-tau algorithm developed in Python by Paul Müller (2012) *Python multiple-tau algorithm*. The multiple tau algorithm allows to avoid overweighting of the end of the correlation versus the beginning.

The elongation dwell time T was obtained by fitting the measured auto-correlation, with the following analytical expression as in(6, 8).

$$G(\tau) = \frac{(T-\tau)}{cT^2} \cdot H(T - \tau) + G_{\infty}$$

where c is the translation initiation rate and $H(t)$ is the Heaviside step function equal to 1 for $t>0$ and zero otherwise. In order to take into account the disappearance of translation sites and consequently the shortness of the tracks we observed here compared to(6, 8), a new term was added, namely G_{∞} , fitting the non-zero decorrelation at long time. Autocorrelation curves were fitted with a Levenberg-Marquadt algorithm using the `lsqcurvefit` built-in function of Matlab 2015b (Mathworks Inc., USA).

Calibrations of optical properties at different positions of the embryo

We calibrate changes in the optical properties of microscope, generated by the embryo sample, by estimating the average fluorescence per molecule and the possible change in the excitation PSF at the basal and the apical part of the embryo. For this, we performed classical FCS experiments (related to Figure S7) on a Zeiss LSM780 microscope using a 40X/ 1.2 NA water objective. scFv-GFP was excited using the 488nm line of an Argon laser and detected through a pinhole of 1 airy unit. The subsequent FCS measurement volume was calibrated with a Rhodamine6G solution using $D=414 \mu\text{m}^2/\text{s}$. Intensity fluctuations were acquired for 10s and auto-correlation functions (ACFs) generated by Zen software were loaded in the `PyCorrFit(30)` and fit with the following analytical expression:

$$G_D(\tau) = \frac{1}{N} \left(1 + \frac{T}{1-T} e^{-\tau/\tau_T} \right) (F_1 G_{D_1} + F_2 G_{D_2}) + G_{\infty}$$

with

$$G_{D_x} = \left(1 + \frac{\tau}{\tau_{D_x}} \right)^{-1} \left(1 + \frac{\tau}{\omega^2 \tau_{D_x}} \right)^{-1/2}$$

where $\tau_{Dx} = w_{xy}^2 / 4 D_x$ and $\omega = w_z / w_{xy}$, T the triplet fraction, F1 and F2 the respective proportion of fast (free) and slowly diffusing scFv-GFP. G_{Dx} being the dynamic part of the decorrelation, that allows for estimation of τ_{Dx} . The triplet state lifetime (τ_T) was arbitrarily fixed to 5 μ s(31). A G_∞ value was introduced to account for possible long-time persistent correlation during the measurements as in the translation kinetics analysis. For each correlogram the value of N and D_1 (number of molecules and diffusion coefficient of the fast diffusing fraction) were estimated. The average molecular brightness was also determined as the average fluorescence brightness ($\langle I_f \rangle$) divided by the number of molecules N. Multiple measurements were done on the apical and basal side of nuclei from different *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryos.

Puromycin and HEPES Injection

Embryos were dechorionated with tape, lined up on a hydrophobic membrane covered in heptane glue with desiccation at room temperature for 10 minutes prior to being covered in Voltalef Oil 10S (VWR), and injected with 10mg/mL puromycin (InvivoGen) or HEPES using a FemtoJet ***5427*** (Eppendorf) micro-injector and associated Femtotips® I (Eppendorf) needles. Embryos were injected in the lateral region immediately prior to coverslip positioning. Time lapse images (related to Movie S10) were acquired using a confocal LSM780 (Zeiss) microscope. A GaAsP detector was used to detect the GFP fluorescence excited using a 488nm laser.

In order to graphically represent the behavior of the impact of puromycin drug treatment, translation bright spots were segmented by a simple intensity threshold of the maximum intensity projection of the raw data: external z frames were removed from the projection and fake spot detections were manually deleted. Each image was divided into 16-pixel square edges (1024 squares per image). Then, the number of detected spots was counted in each square, to estimate their concentration. The concentration was visualized frame-by-frame thanks to a color-map (related Movie S10).

Detection and tracking of single particles

For single particle tracking (related to Movie S7A and S7B), movies were acquired with a Zeiss LSM880 confocal microscope in fast Airyscan mode with a Plan-

Apochromat 40x/1.3 oil objective lens, GFP was excited using a 488nm laser and the following settings: 132 x 132 pixels, each Z-stack comprised of 10 planes 0.5 μ m apart, Zoom 8x. Under these conditions a Z-stack was acquired every \approx 500ms.

Each movie was maximum intensity Z-projected using Fiji(28). Single particle trajectories were generated with MatLab 2015b (Mathworks Inc., USA) using SLIMfast, which implements the Multiple-Target-Tracing algorithm(32, 33). Individual spots were localized using a blob diameter of 9 pixels. They were then tracked using maximal frame gap of 0 frame and a maximal expected diffusion coefficient set to $D_{\max} = 0,5 \mu\text{m}^2.\text{s}^{-1}$. Tracked movies were then evaluated using the script evalSPT(34). Each tracked spot was visually checked and poorly tracked spots were manually removed using evalSPT. In total 237 traces were obtained from three movies apically and 263 traces from three movies basally. 1D displacement in x and y were obtained from evalSPT tool for each movie. Mean square displacements (MSD) were calculated for tracks present for at least 10 consecutive frames using the MSDanalyzer MatLab script(35). Coefficients diffusion were first obtained by fitting 0 to 20 seconds of the MSD of individual trajectories with a linear model as follows:

$$y = a * x + b$$

To estimate MSD, according to Einstein's theory, the MSD of Brownian motion is described as

$$MSD(t) = 2dDt$$

with d representing dimensionality, in this case, $d = 2$. D represents the diffusion coefficient.

A linear fitting was considered unsatisfactory when $R^2 < 0.8$ (Figure S13) and proceed to a non-linear model fitting.

For this, MSD of individual trajectories were fitted with a non-linear model, as follows: If $x = \log(\Delta t)$ and $y = \log(\text{MSD})$ where ' Δt ' are the delays at which the MSD is calculated, then this method fits $y = f(x)$ by a straight line

$$y = \alpha * x + \gamma$$

so that the MSD curves were approximated by $\text{MSD} = \gamma * \Delta t^\alpha$.

According to a previous modeling study(36), an α value of 1 indicates free diffusive movement, a value of α close to 2 indicates movement by active transport, and a value of $\alpha < 1$ indicates a motion constrained in space.

Supplementary Text

Imaging translation dynamics from a *twi-suntag* transgene

In order to implement the SunTag method we created various scFv-fluorescent lines for maternal deposition of the scFv detector (scFvGFP-NLS, referred to as scFv-GFP in the text, scFv-GFP-noNLS, scFv-GFP-NLS-bcd-3'UTR, (Figure S1A-C and E). A diffuse, soluble GFP signal is detected in all scFv lines, showing that scFv can be genetically encoded in *Drosophila* without forming aggregates (Figure S1B, C and movie S8A and S8B). By performing mRNA labeling using single molecule fluorescent *in situ* hybridization (smFISH) with the simultaneous detection of native scFv-GFP, on the *twi_suntag* transgene containing 32x suntag repeats(7) (Figure S1F-H), we could detect two populations of cytoplasmic mRNA molecules: those co-localizing with a bright GFP signal, $\approx 56\%$ of total mRNA in nc14 (n=6231 from three smFISH), corresponding to mRNAs in translation, and those devoid of a GFP signal and corresponding to untranslated mRNAs (Figure S1H). By live imaging with light sheet microscopy(27), distinct spots were detected clearly above background within the presumptive mesoderm of *twi_suntag* transgenic embryos (Figure S1I and Movie S9). These bright spots correspond to nascent sites of translation, as they do not appear upon puromycin injection into live embryos, while they appear upon HEPES injection (Figure S1J and Movie S10A-B).

Imaging mRNA and their translatability with a *twi-suntag-MS2* transgene

To visualize both mRNA and their translation in living embryos, we created another transgene where, in addition to the suntag sequences, MS2 repeats were inserted in the 3'UTR region (Figure S2A). To amplify the fluorescent signal produced by single molecules of mRNA, we employed a 128-loop array of a new generation of optimized MS2 loops(12). Similar to the *twi_suntag* transgene, *twi_suntag_MS2* transcripts were expressed in the presumptive mesoderm (Figure S2B) and we could detect two populations of cytoplasmic mRNAs, the first corresponding to mRNAs in translation (arrow) and the second to mRNAs not in translation (arrowhead) (Figure S2C). Using two-color live imaging of the *twi_suntag_MS2* transgene (with scFv-GFP and MCP-TagRFPT detector transgenes), we occasionally detected mRNA spots co-localizing and traveling with suntag signal in the cytoplasm of living embryos (data not shown).

However, fast diffusion of mRNA molecules and very rapid photobleaching of the fluorescent tag precluded their tracking. To try to bypass these weaknesses, we hypothesized that translation could be less sensitive to photobleaching (because of binding of new scFv to neosynthesized suntag proteins). We generated a scFv-TagRFPT-NLS but this fluorophore undergoes rapid photobleaching and forms aggregates (white arrow heads left panel Figure S1D, and Movie S8C). Next, we developed a scFv-mScarlet-NLS which has a stronger signal and doesn't form aggregates during the developmental timing that we studied (right panel Figure S1D and Movie S8D). Combined to MCP-GFP, this new fly line allowed us to track, for a short time, mRNA in translation in live embryos (Figure S2D, Movie S11). We note however that rapid photobleaching prevents the quantitative analysis of putative translation bursts, as it has been achieved in cultured cells(5-9). In conclusion, we have developed and validated the genetic tools to image translation in intact live *Drosophila* embryos. We can thus image translation in live embryos with scFv-GFP and use smFISH to detect and quantify cytoplasmic mRNAs at high resolution.

While the suntag signal was restricted specifically to the mesoderm, *twi_suntag* expression appeared stochastically in cells within this domain (Figure S1G, I and S2B). Our *twi* transgenes did not fully recapitulate *twi* canonical expression (Figure S2E) due to the presence of only the *twi* proximal enhancer in the transgene, as evidenced with a *twiPE_twipr_MS2* transgene (Figure S2F).

Imaging simultaneously transcription site and translation with a *twi-suntag-MS2* CRISPR

By combining suntag and MS2 labeling, it is possible to image transcription and translation from an endogenous gene within a whole embryo (Figure S5J and Movie S14). Thus, double labeling with suntag/MS2 repeats provides a method to visualize the precise chronology of events and possible flow of information from transcription to translation. However, to be more quantitative, we decided to image translation and transcription separately using detector protein fused to a GFP that is brighter and is less subject to photobleaching as compared to red emitted fluorescent proteins.

Imaging mRNA and their translatability with an *ilp4-suntag* transgene

To test if local translation, could be extended to basally-located mesodermal mRNAs other than *twi*, we generated an *ilp4-suntag* transgene. The choice of this gene stems from its expression and the characterization of its *cis*-regulatory sequences. Indeed, from the very useful FlyFISH database images(2), we observed that *ilp4 (mes3)* mRNAs were enriched in the basal cytoplasm of nc14 embryos. We note however that the FlyFISH database is not of single molecule resolution. Moreover, nc14 ventral expression of this gene is, in part, dictated by a characterized *cis*-regulatory element(37). Thus, we created an *ilp4-suntag* transgene under the control of this *cis*-regulatory element (Figure S9A) (see methods). By performing single molecule mRNA labeling using single molecule fluorescent *in situ* hybridization (smFISH) with *suntag* probes and with the simultaneous detection of native scFv-GFP, we could detect two populations of cytoplasmic mRNA molecules: those co-localizing with scFv-GFP spots (mRNA in translation) (Figure S9B arrows) and those not co-localizing with scFv-GFP signal (mRNA not in translation) (Figure S9B stars). In the basal cytoplasm, we observe large scFv-GFP foci, co-localizing with large mRNA clusters (Figure S9B arrowheads). Thus, similarly to *twi* mRNA, *ilp4* mRNA can be co-translated basally.

Tracking of translating mRNPs reveals distinct mobilities not due to cellular movements

To quantitatively characterize the diffusive properties of *twi_suntag_CRISPR* mRNPs, we imaged the scFv-GFP signal in the apical and basal perinuclear region in the mesoderm at high speed and employed single particle tracking (Figure 4D-E). We first observed that the mRNPs located above nuclei tend to displace faster than those located below as quantified by 1D displacement (Figure S13A) and without preferential drift (Figure S13B). Therefore, the displacement measured is not due to cellular movement.

Extended main Figure legends

Figure1

C: Snapshot of a live imaging of a His2Av-mRFP/+;scFv-GFP-NLS/+>*twi_suntag_CRISPR*/+ embryo by MuViSPIM (Movie S1). A 50ms exposure time was used for the green and red channel with a 488nm and 561nm laser excitation

respectively. Maximum intensity projections of half of the embryo ($\approx 100 \mu\text{m}$) for each side were processed with Fiji. Time resolution 2min between Z-stacks.

D: Snapshot from movie for MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+ (Movie S3) were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode, GFP and mScarlet were excited using a 488nm and 561nm laser respectively and a frame was acquired every $\approx 46\text{ms}$.

Figure 2

A: Snapshots of maximum intensity projected 35z-planes ($\approx 17\mu\text{m}$) live imaging of a MCP-eGFP-His2Av-mRFP>twi_suntag_MS2_CRISPR/+ embryo showing false colored transcription sites (TS) in red, transcriptionally active nuclei in grey and transcriptionally inactive nuclei in blue. Movies were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode. GFP and mRFP were excited using a 488nm and 561nm laser respectively and a stack was acquired every 19.6s.

C: Snapshots of maximum intensity projected 20 Z-planes ($\approx 9.5\mu\text{m}$) live imaging of a scFv-GFP/+>twi_suntag_MS2_CRISPR/+ embryo. Movies were acquired using a Zeiss LSM880 with confocal microscope in fast Airyscan mode. GFP and mRFP were excited using a 488nm and 561nm laser respectively and a stack was acquired every 22.5s.

Figure 3

A: Snapshots of maximum intensity projected reconstructed cross sections of $\approx 5\mu\text{m}$ from a live imaging of a MuViSPIM of a His2Av-mRFP/+;scFv-GFP-NLS/+>twi_suntag_CRISPR/+ embryo (Movie S4). Nuclei are detected using His2Av-mRFP and suntag using scFv-GFP. A 50ms exposure time was used for the green and red channel with a 488nm and 561nm laser excitation respectively. Time resolution 2min between Z-stacks.

B: Representative confocal image of apical and basal maximum intensity projected 3 Z-planes ($\approx 1\mu\text{m}$) shown separately of an scFv-GFP-NLS/+>twi_suntag_CRISPR/+ embryo expressing scFv-GFP (green) labelled with suntag smFISH probes (red). Images were acquired using a Zeiss LSM880 with confocal microscope in super resolution Airyscan mode.

Figure 4

D: Maximum intensity projection of 2 z-planes ($\approx 0.5\mu\text{m}$) of sagittal views of *twi_suntag_CRISPR/+* embryos labelled with *suntag smFISH* probes, anti-lamin and DAPI. Images were acquired using a Zeiss LSM880 with confocal microscope in super resolution Airyscan mode.

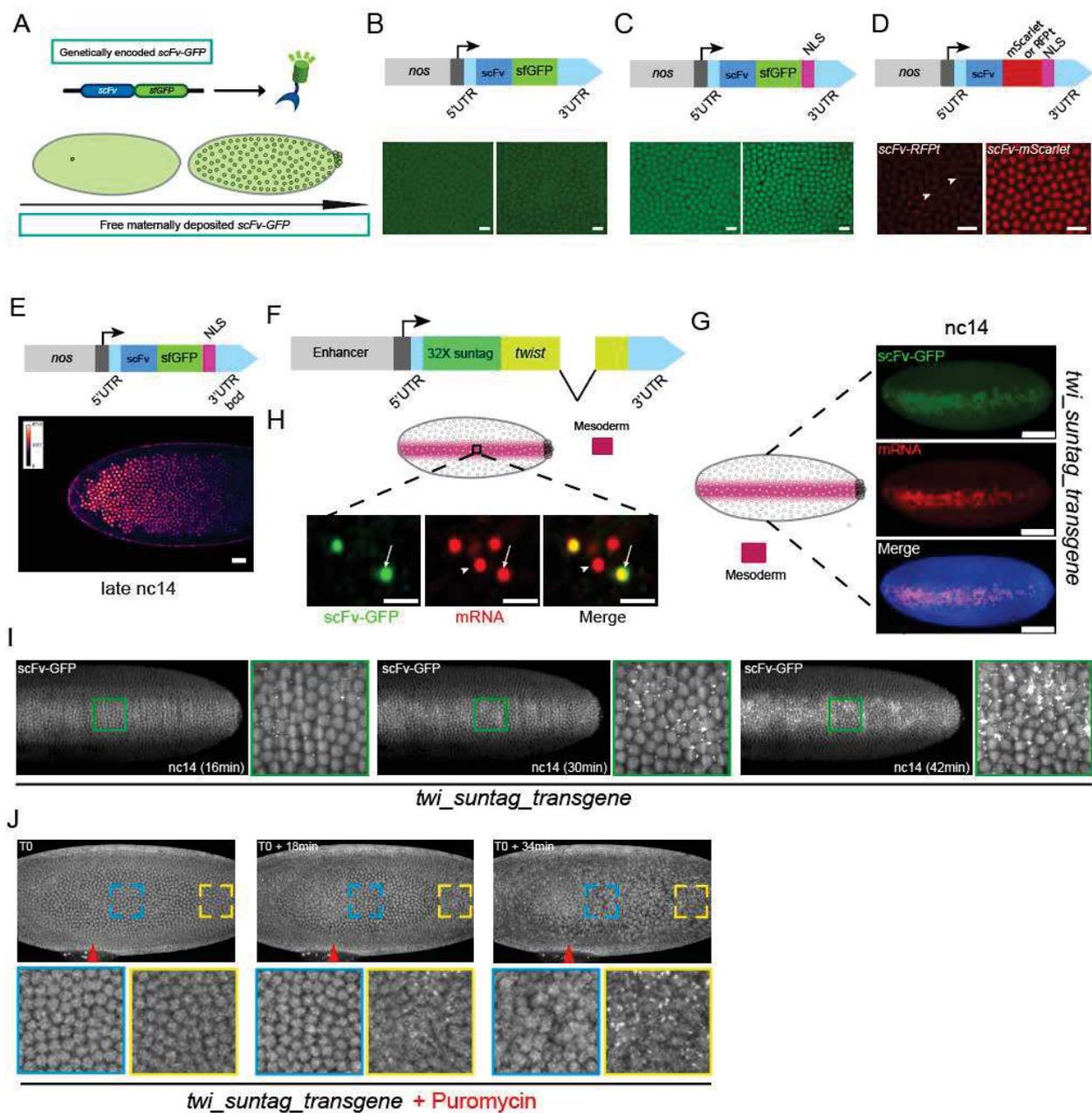


Figure S1: Imaging translation in drosophila embryos from a *twi-suntag* transgene

A, Schematic representation of the genetically encoded and maternally deposited *scFv-GFP* detector lines created in this study.

B-D, Snapshots of maximum intensity projection of confocal time lapse movie of *Drosophila* embryos expressing the *scFv-GFP* (**B**), the *scFv-GFP-NLS* (**C**), the *scFv-RFPt* or *scFv-mScarlet* (**D**) transgene during the n.c.14. Arrowheads point to *scFv-RFPt* aggregates. NLS: nuclear localization signal, scale bars 10 μ m. Related Movie S8.

E, Heatmap of GFP fluorescence at late n.c.14, showing an enhanced fluorescence at the anterior part of the *scFv-GFP-NLS-bcd3'UTR* transgenic embryo, scale bars 20 μ m.

F, Schematic representation of *twist_suntag* transgene.

G, Maximum intensity projection of a ventral view of a *scFv-GFP-NLS/+>twi_suntag/+* transgenic embryo in n.c.14 stained with endogenous *scFv-GFP* (green), smFISH probes against *suntag* repeats (red) and DAPI (blue). Scale bars 100 μ m.

H, Zoomed confocal single plane within the mesoderm of a *scFv-GFP-NLS/+>twi_suntag/+* transgenic embryo stained with endogenous scFv-GFP (green) and probes against *suntag* repeats (red). scFv-GFP foci colocalizing with *suntag* probes reveal mRNAs in translation (arrow), while spots that are only stained with *suntag* probes correspond to non-translating mRNAs (arrowhead), scale bars 1 μ m.

I, Snapshots of dorsal and ventral views of a maximum intensity projection from a *scFv-GFP-NLS/+>twi_suntag/+* live transgenic embryo imaged with a MultiView Selective Plane Illumination Microscope, MuViSPIM, (related Movie S9). Associated zoomed images are shown on the right. All embryos are ventral views, oriented with anterior to the left.

J, Maximum intensity projection of frames taken from live imaging (related Movie S10A) of a *scFv-GFP-NLS/+>twi_suntag/+* transgenic embryo injected with puromycin (related Movie S10B). Red arrowheads indicate the site of drug injection. T0 corresponds to \approx 25-30min after puromycin injection. Zoomed images from the color-coded indicated regions are provided in the lower panels.

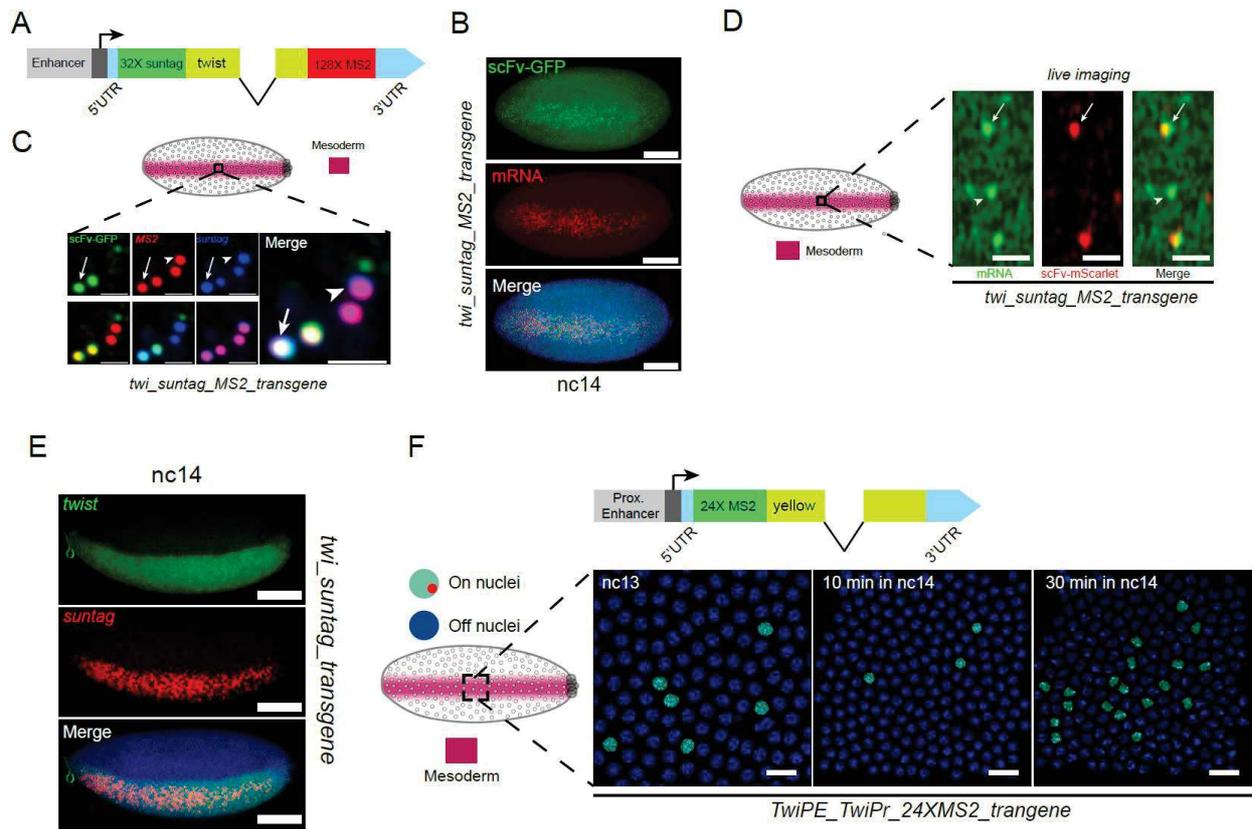


Figure S2: Imaging translation in drosophila embryos from an endogenous locus

A, Schematic of the *twi_suntag_MS2* transgene.
B, Maximum intensity projection of confocal images from a n.c.14 *scFv-GFP-NLS/+>twi_suntag_MS2/+* transgenic embryo, stained with endogenous scFv-GFP (green), smFISH probes against *suntag* repeats (red) and DAPI (blue). Scale bars 100µm.

C, Zoomed confocal plane of a *scFv-GFP-NLS/+>twi_suntag_MS2/+* transgenic embryo stained with endogenous scFv-GFP (green) and probes against *suntag* (blue) and *MS2* (red) repeats. scFv-GFP foci colocalizing with *suntag* and *MS2* probes reveal mRNAs in translation (arrow), while spots that are only stained with *suntag* and *MS2* probes represent mRNAs not in translation (arrowhead), scale bars 1µm.

D, Snapshots from live imaging of a *MCP-GFP/+; scFv-mScarlet-NLS/+>twi_suntag_MS2/+* transgenic embryo (related Movie S11) exhibiting two groups of mRNAs indicated by arrows (engaged in translation) and arrowheads (untranslated). Scale bars 1µm.

E, Maximum intensity projection of confocal images from a *scFv-GFP-NLS/+>twi_suntag/+* transgenic embryo fixed at n.c.14, stained with probes against *twist* (green), *suntag* repeats (red) and DAPI (blue). Scale bars 100µm.

F, Live imaging of *MCP-GFP,H2Av-RFP/+>TwiPE_TwiPr_24XMS2/+* transgenic embryo to detect transcriptionally active nuclei. False-colored snapshot images from the movie, showing transcriptionally active and inactive nuclei in light green and dark blue respectively and transcription sites in red. scale bar 10 µm.

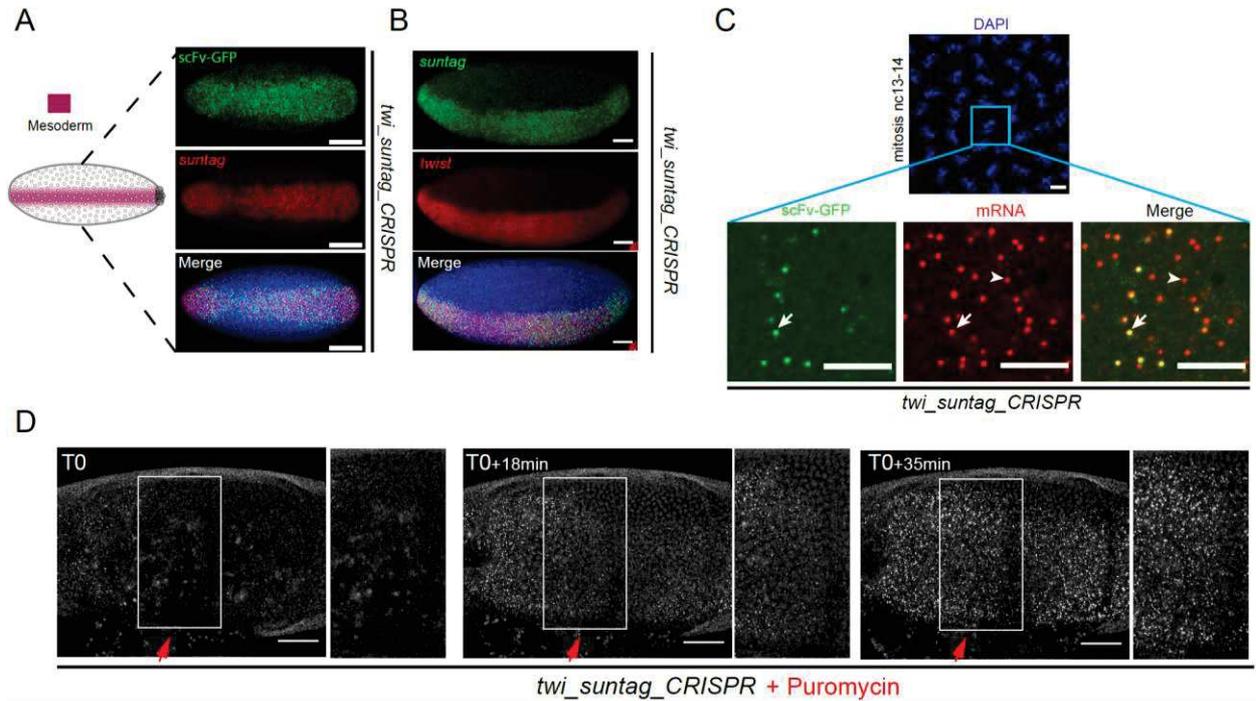


Figure S3: Capturing the timing of translation of endogenous *twist* mRNAs

A-B, Maximum intensity projection of confocal images of n.c.14 *twi_suntag_CRISPR/+ Drosophila* embryos expressing scFv-GFP (green) stained with DAPI (blue) and a suntag smFISH probe (red) in **A** and with suntag and *twi* smFISH probes in **B**. Scale bars 100 μ m.

C, Maximum intensity projection of three z-planes ($\approx 1 \mu$ m) of confocal images of mitosis from n.c.13 to n.c.14 of a *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo expressing scFv-GFP, labelled with suntag smFISH probes and DAPI (blue). Right panels are zoomed views exhibiting two groups of mRNAs (red) molecules: those co-localizing with scFv-GFP signal (arrows) and those not co-localizing with a GFP signal (arrowheads), scale bars 5 μ m.

D, Maximum intensity projection of frames taken from live imaging (related Movie S10C) of a *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo injected with puromycin. Red arrowheads indicate the site of drug injection. T1 corresponds to ≈ 25 -30min after puromycin injection. Zoomed images (white rectangle) are provided in right panels for each time point. Lag time between each time point are ≈ 18 min.

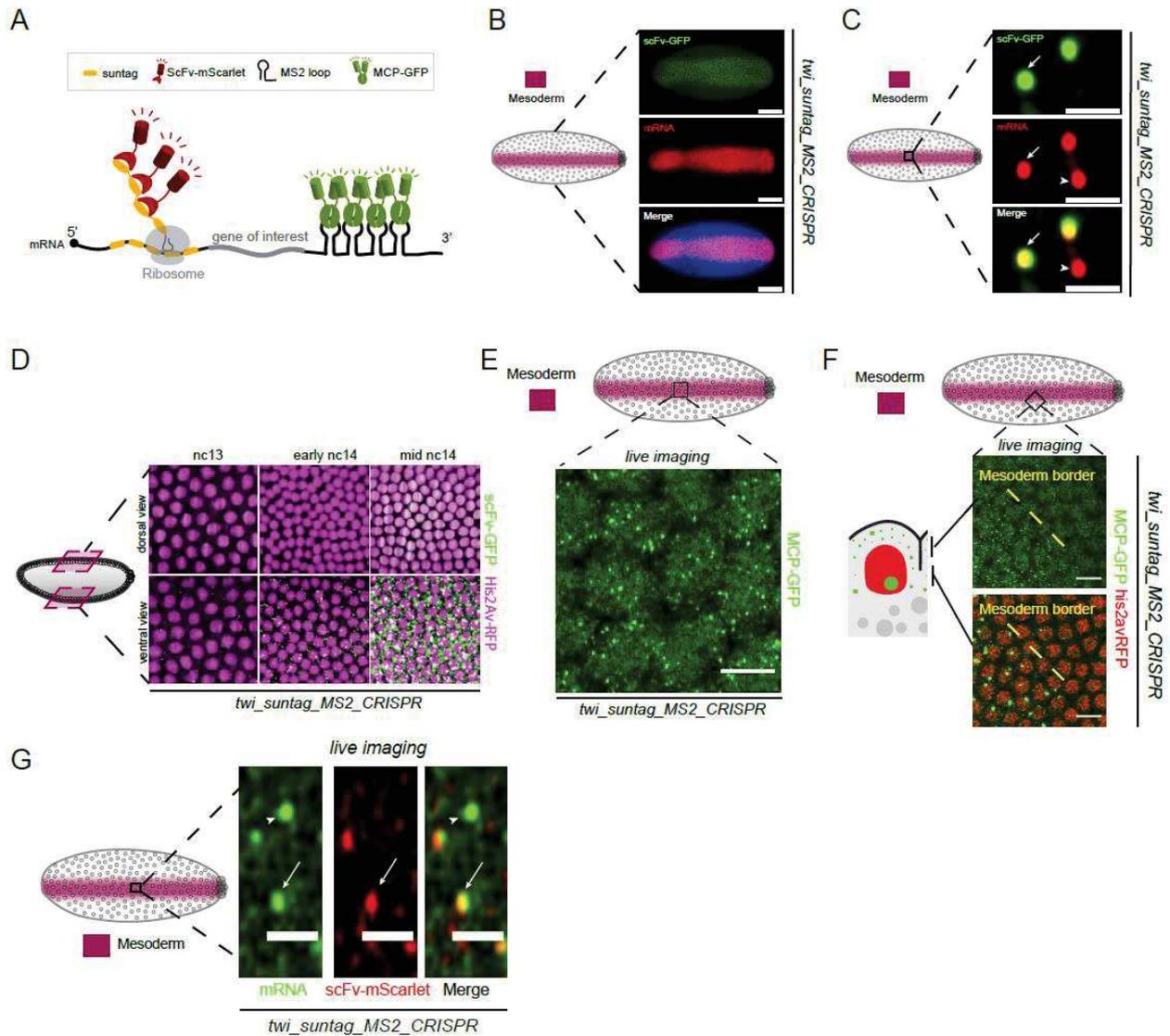


Figure S4: Dual-color imaging of mRNA and its translation

A, Principle of dual color imaging with the Suntag and the MS2 systems. Labeled antibody probes, here single-chain variable fragments coupled to mScarlet (scFv-mScarlet), bind to repeated suntag epitopes, allowing visualization of nascent translation. Tandem array of suntag peptides are shown in orange at the 5' end of the gene, scFv-GFP molecules in green and ribosomes in light grey. MS2 loops (in black) are inserted at the 3' end of the gene and upon transcription will be bound by the MS2 Coat Protein (MCP) coupled to GFP (MCP-GFP).

B-C, Maximum intensity projection of confocal images of n.c.14 *scFv-GFP-NLS/+>twi_suntag_MS2_CRISPR/+* *Drosophila* embryos expressing scFv-GFP (green) labelled with *suntag* probes (red) and DAPI (blue), scale bars 100µm. **C**, zoomed view within the mesoderm shows two distinct mRNA (red) pools: those engaged in translation (co-localizing with scFv-GFP signal, arrows) and untranslated single molecules (red, arrowheads). Scale bars 1µm.

D, Maximum intensity projection snapshots of *His2Av-mRFP/+>scFv-GFP-NLS/+>twi_suntag_MS2_CRISPR/+* embryo imaged with a MuVi-SPIM, (related Movie S12). Upper panels show the dorsal part of the embryo while the lower panels show the ventral mesoderm part of the embryo, where *twi* is expressed. Nuclei are shown in magenta and sites of translation in green.

E-F, Representative maximum intensity projection of three z-planes ($\approx 1 \mu\text{m}$) image of a *MCP-GFP,His2Av-mRFP/+>twi_suntag_MS2_CRISPR/+* time-lapse movie embryo within the *twi* expression pattern (related Movie S13) and **F**, at the border part of the *twi* expression pattern.

G, Snapshot image of a *MCP-eGFP/+; scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+* live embryo (related Movie S2) exhibiting two groups of mRNAs indicated by arrows (engaged in translation) and arrowheads (untranslated). Scale bars $1 \mu\text{m}$.

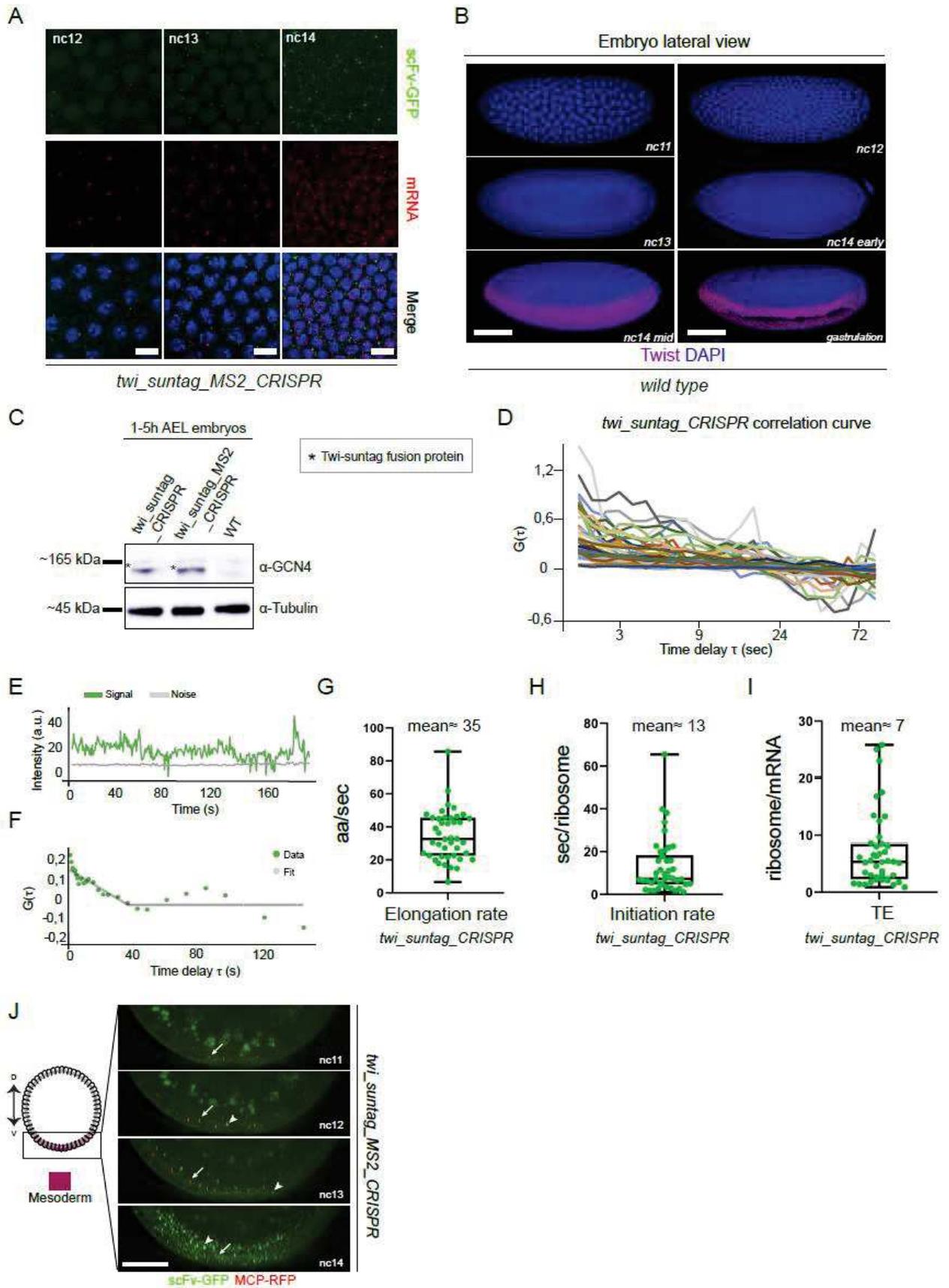


Figure S5: Revealing the timing of *twi* translation in the mesoderm

A, Maximum intensity projection of confocal images from n.c.12, n.c.13 and n.c.14 *scFv-GFP-NLS/+>twi_suntag_MS2_CRISPR/+* embryos expressing scFv-GFP (green) and labelled with *suntag* probes (red) and DAPI (blue). Scale bar, 10 μ m.

B, Wild type fixed embryos at the indicated developmental stage, stained with a Twi antibody (magenta) and DAPI (blue). All embryos are oriented with anterior to the left and ventral on the bottom. Scale bars, 100 μ m.

C, Western blot analysis of 1-5h after egg laying (AEL) *twi_suntag_CRISPR*, *twi_suntag_MS2_CRISPR* and wild type embryos. Hybridization with anti-GCN4 shows a 165kDa band corresponding to the Twist fused to *suntag* repeats translated protein (black star).

D, Autocorrelation curves (n=42 traces from 9 movies) of *twi_suntag_CRISPR* polysomes from *scFv-GFP/+>twi_suntag_CRISPR/+* embryos showing the variability in translation dynamics related to **G-I**.

E, The intensity of a tracked *scFv-GFP/+>twi_suntag_CRISPR/+* polysome can be measured as a function of time.

F Autocorrelation curves (green) calculated from intensity fluctuation data in **E** and its fit (grey).

G-I, Kinetic parameters extracted from the fit of 42 traces from 9 movies with elongation rates in **G**, initiation rates in **H** and Translation Efficiency (TE) in **I**.

J, Representative ventral cross section of a transverse plane from reconstructed MuViSPIM images of a *scFv-GFP-NLS/MCP TagRFPT>twi_suntag_MS2_CRISPR/+* embryo (related Movie S14). Transcription sites are visualized with the MCP-TagRFPT (red dots, arrow) and sites of translation with the scFV-GFP (green dots, arrowhead). Scale bar 30 μ m. Left panel: schematic of a sagittal section throughout an embryo, with the dorso-ventral axis (D, V) and the presumptive mesoderm (purple) indicated.

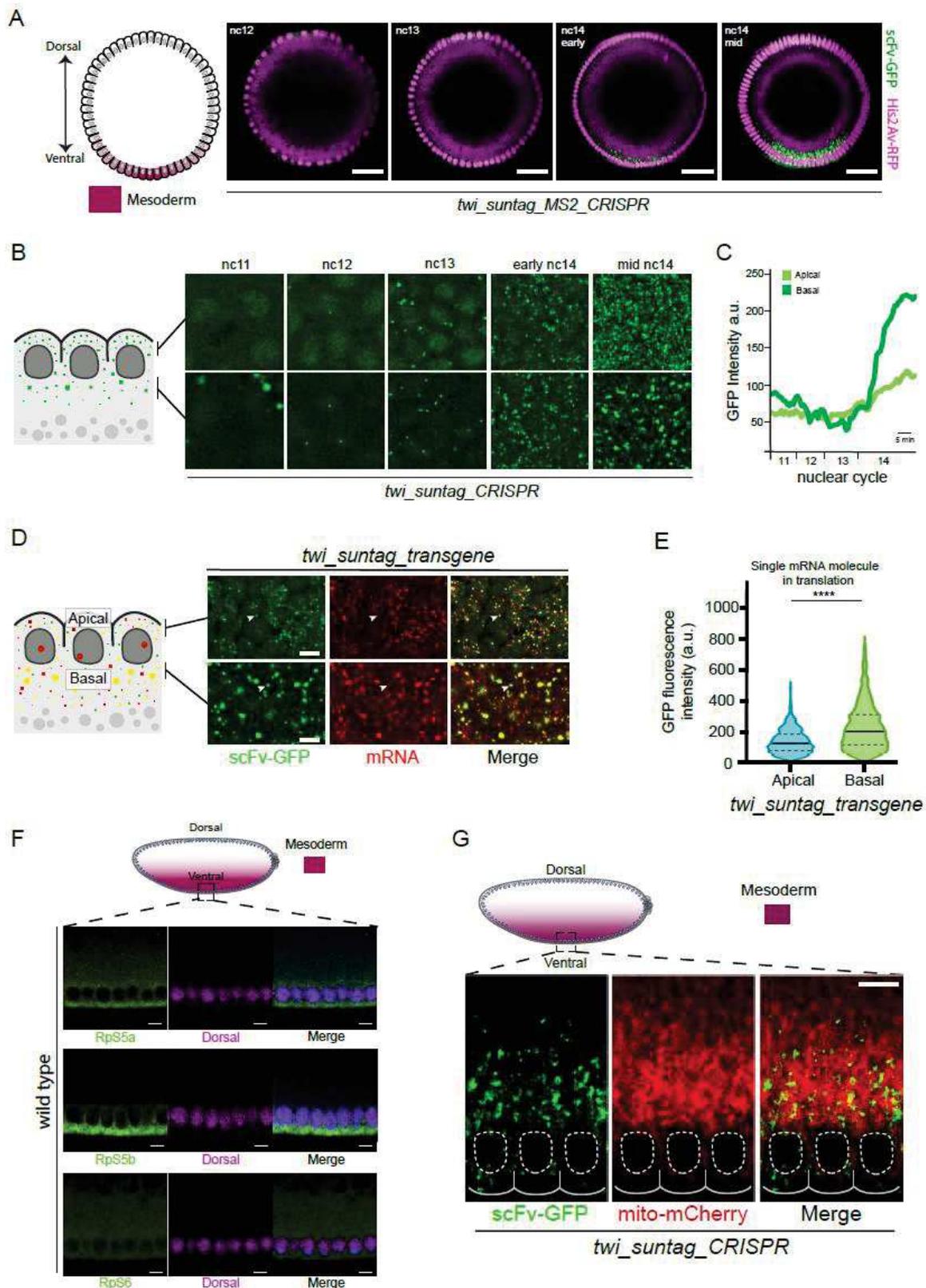


Figure S6: Revealing a spatial heterogeneity in translation sites in the mesoderm
A, Representative cross section of a transverse plane from reconstructed MuViSPIM images of a *His2Av-mRFP/+; scFv-GFP-NLS/+>twi_suntag_MS2_CRISPR/+* embryo (related Movie S12B). Ventral mesodermal nuclei are located at the bottom. Nuclei are observed with a *His2Av-mRFP* transgene (magenta) and sites of translation with the scFv-GFP (green dots). Scale bars 30 μ m.

B-C, Live imaging of a *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo, with representative images of apically or basally located Z-planes in **B**, and their quantification (total intensities of GFP) in **C**.

D, Maximum intensity projection of three z-planes ($\approx 1 \mu\text{m}$) of confocal images of n.c.14 *scFv-GFP-NLS/+>twi_suntag/+* transgenic *Drosophila* embryos expressing scFv-GFP (green) labelled with *suntag* probes (red) and DAPI (blue), scale bars $5 \mu\text{m}$.

E, Violin plots of the distribution of scFv-GFP intensities colocalizing with single mRNA molecules located apically (n=681; blue) and basally (n=2404; green) from three n.c.14 *scFv-GFP-NLS/+>twi_suntag/+* transgenic embryos labelled with *suntag* smFISH probes (see methods section). Centered black bar represents the median, dashed black lines represent quartiles. **** $p < 0.0001$ with a two-tailed Welch's t-test.

F, Maximum intensity projection of three Z-planes ($\approx 1 \mu\text{m}$) of confocal images of wild type embryos stained with antibodies against ribosomal proteins subunits RpS5a, Rps5b and RpS6 (green) together with Dorsal antibody (magenta) and DAPI (blue). All images are taken on n.c.14 aged embryos. Scale bars $5 \mu\text{m}$.

G, Live imaging of mitochondria (Movie S5) with *UASp-mito-mCherry/+;nosGal4,scFv-GFP-NLS>twi_suntag_CRISPR/+* embryo. Scale bar, $5 \mu\text{m}$.

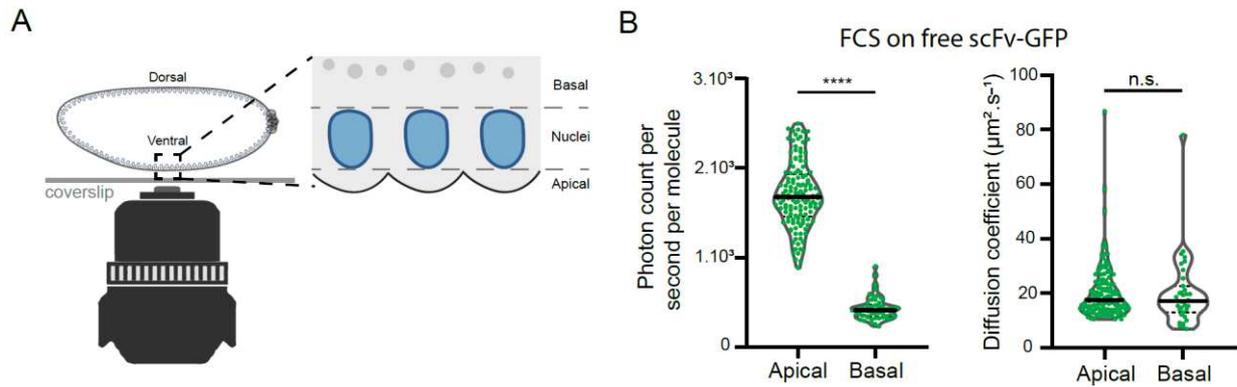


Figure S7: Fluorescence Correlation Spectroscopy on free scFv-GFP molecules

A, Illustration of the imaging setup employed in this study. The apical side of the embryo is close to the coverslip, while the basal side, deeper in the embryo, is more distant from the coverslip.

B, Violin plots of the brightness (left panel) and diffusion coefficient (right panel) of the free scFv-GFP of the measured by FCS apically ($\approx 2 \mu\text{m}$ inside the embryo) and basally ($\approx 20 \mu\text{m}$ inside the embryo) in count per second per molecule and $\mu\text{m}^2 \cdot \text{s}^{-1}$ respectively. Measurements have been performed on three different *scFv-sfGFP-NLS/+* embryos. **** $p < 0.0001$ and not significant (n.s.) $p > 0.05$ with a two-tailed Welch's t-test.

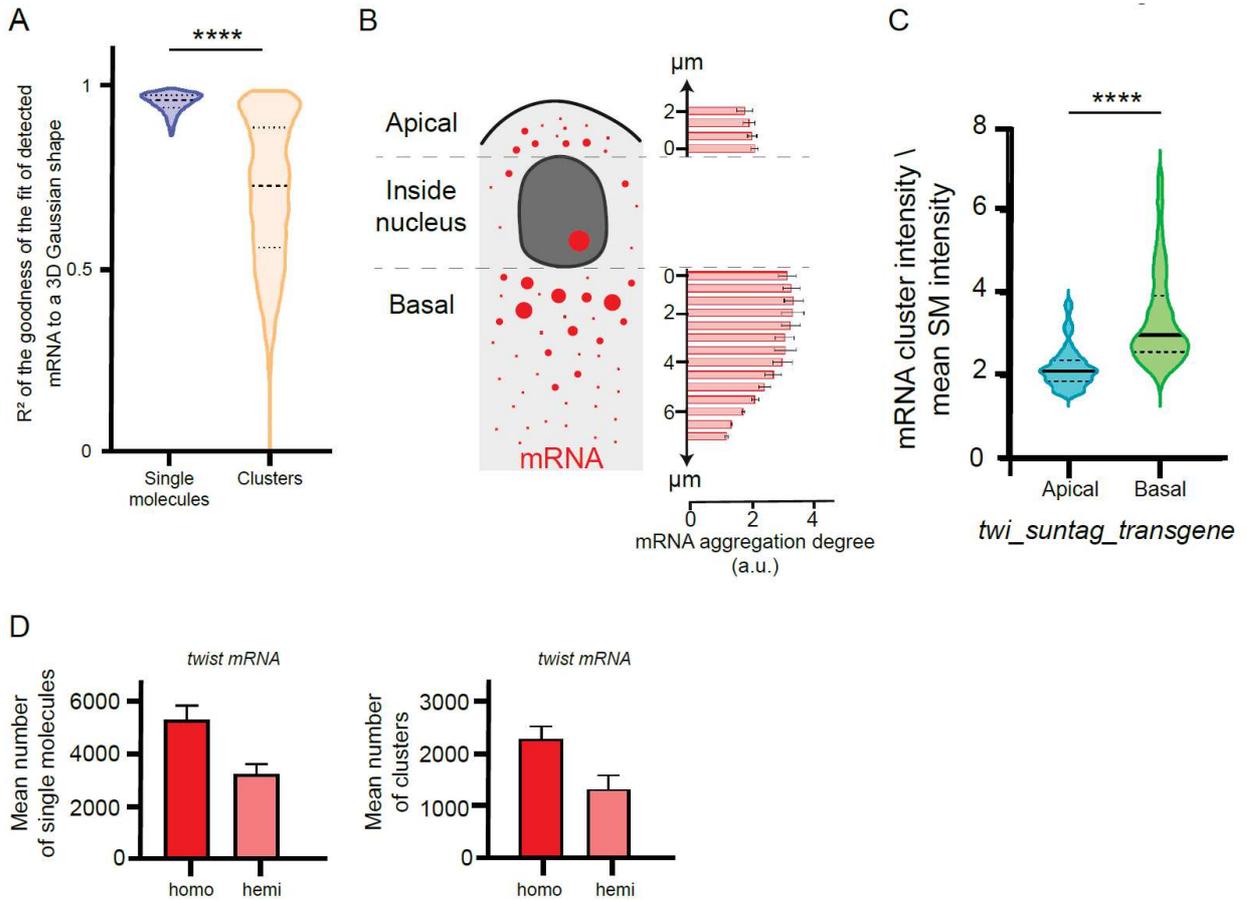


Figure S8: Quantification and characterization of *twi* mRNA clusters

A, Violin plots showing the distribution of the R² coefficient of the Gaussian 3D fitting of *twi_suntag_CRISPR* single molecules (purple) and clusters (orange).

B, Quantification of the degree of aggregation with spatial correlation analysis of *twi_suntag_CRISPR* mRNAs along the z axis for four n.c.14 embryos *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* labelled with *suntag* smFISH probes. Mean and SD are represented.

C Violin plots of the distribution of the number of mRNA cluster intensity divided by the mean single molecule intensity quantified from images taken apically (n=207; blue) and basally (n=688; green) from three n.c.14 embryos *scFv-GFP-NLS/+>twi_suntag/+* transgene labelled with *suntag* smFISH probes. Centered black bar represent median, dashed black lines represent quartiles. **** p < 0.0001 with a two-tailed Welch's t-test.

D, Histograms of mean and SD numbers of basal *twist* mRNA single molecule (upper panel) and number of clusters (lower panel) from wild type embryos homozygous for *twist* (dark red) and hemizygous for *twist* (light red).

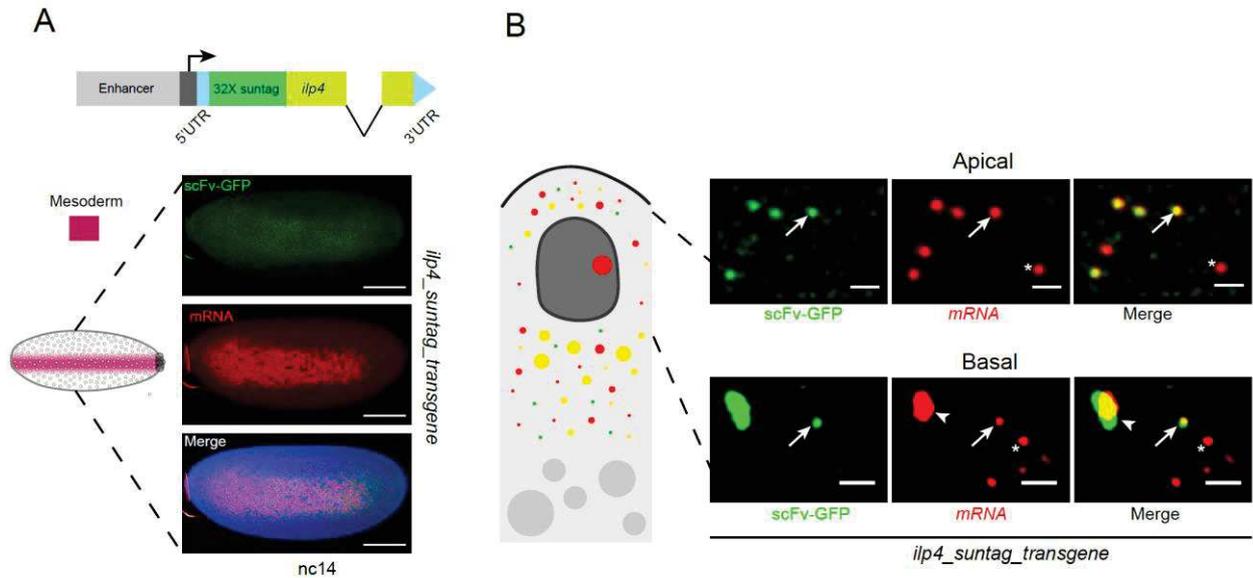


Figure S9: *ilp4* mRNAs form basal mRNA clusters locally translated

A-B *scFv-GFP-NLS/+>ilp4_suntag* transgenic embryos expressing scFv-GFP, labelled with suntag smFISH probes and DAPI (blue). **A**, Schematic representation of *ilp4_suntag* transgene and maximum intensity projection of a ventral view of a whole embryo in n.c.14, scale bars 100µm **B**, Representative zoomed images of 2 planes (≈0.5 µm) maximum intensity projection, apical (upper panels) and basal (lower panels) showing three types of mRNAs: untranslated (stars), in translation (arrows) and large GFP foci colocalizing with cluster of mRNAs (arrowhead). Scale bars 1 µm.

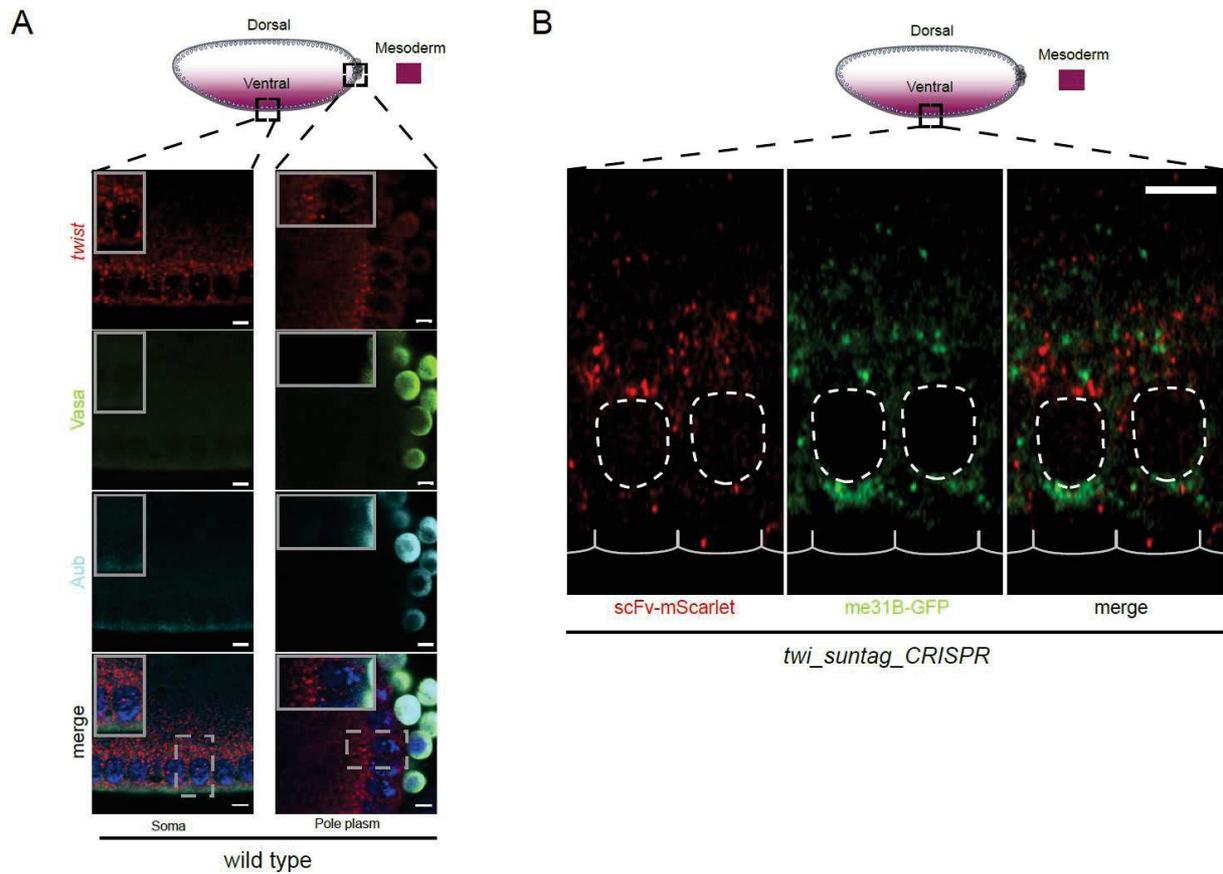


Figure S10: *twi* mRNA clusters are distinct from germ granules and p-bodies

A, Maximum intensity projection of three z-planes ($\approx 1 \mu\text{m}$) of confocal images of immuno staining against Vasa and Aubergine proteins combined with smFISH using *twist* probes on wild type embryos. Left panels are images taken within the center of the *twi* expression pattern, right panels are taken at the posterior extremity of *twi* expression pattern. Scale bars $5 \mu\text{m}$.

B, Snapshots of one z-planes of confocal time lapse movie of *UASp-me31B-GFP/+; scFv-mScarlet-NLS,nosGal4/+>twi_suntag_CRISPR/+* embryos (related Movie S15). Dashed circles represent nuclear membrane and solid lines the plasma membrane. Scale bar $5 \mu\text{m}$.

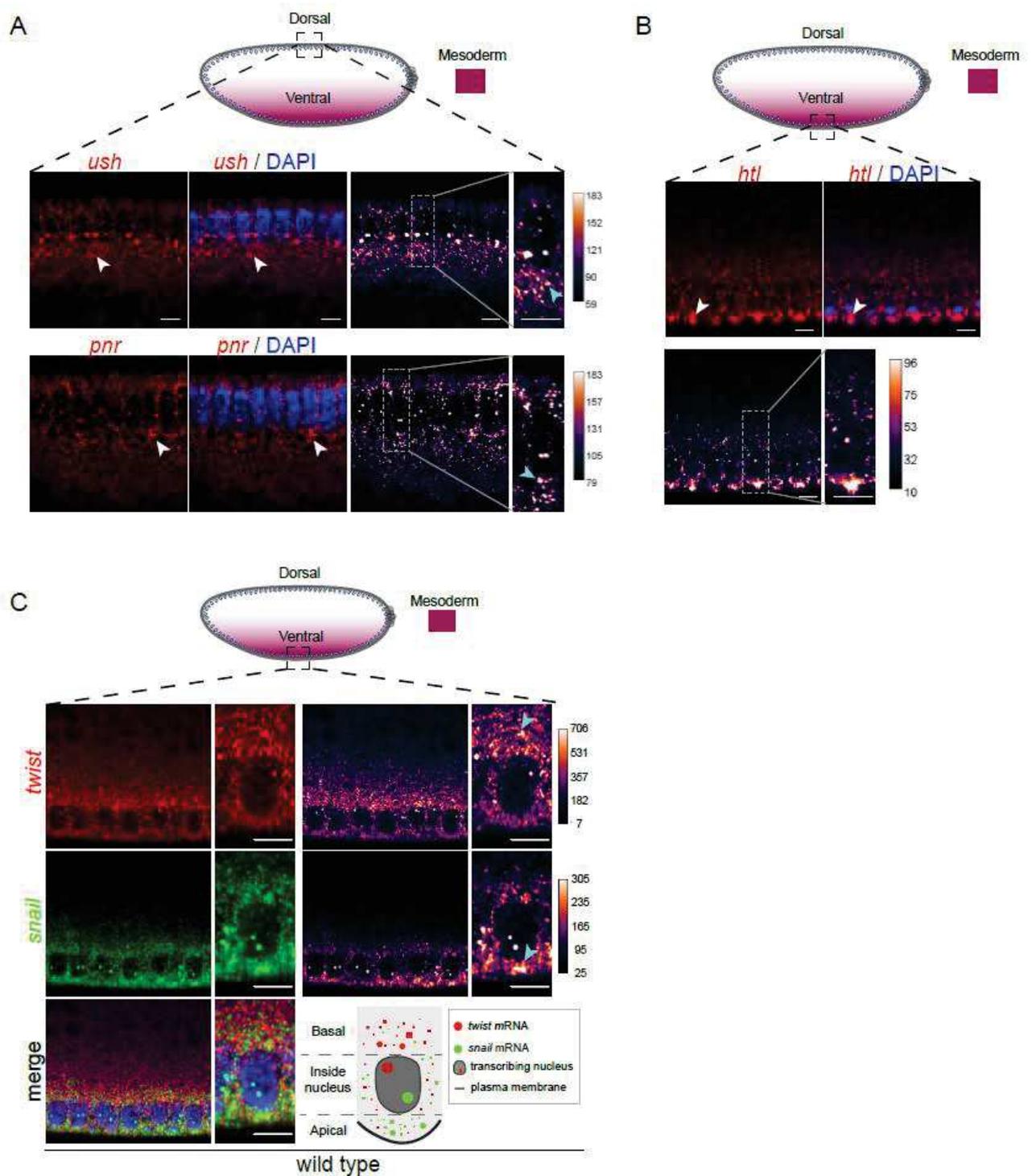


Figure S11: Examples of localized mRNA clusters

A, Maximum intensity projection of three z-planes ($\approx 1 \mu\text{m}$) of confocal images of n.c.14 wild type embryos labelled with *ush* or *pnr* FLAP-probes (red) together with DAPI (blue). Right panels are the same images, shown as false colored. Arrowheads point clusters of mRNA located basally. Scale bars $5 \mu\text{m}$. Note the basal position of the *ush* transcription site.

B, Maximum intensity projection of three z-planes ($\approx 1 \mu\text{m}$) of confocal images of n.c.14 wild type embryos labelled with *htl* smFISH probes (red) and DAPI (blue). Arrowheads point clusters of mRNA located apically. Lower panels are the same images, shown as false colored. Scale bars $5 \mu\text{m}$.

C, Maximum intensity projection of four z-planes ($\approx 1.5 \mu\text{m}$) of confocal images of n.c.14 wild type embryos labelled with *twi* (red) and *sna* (green) probes and DAPI (blue). Scale bars $5\mu\text{m}$, (arrowheads: mRNA clusters). Right panels are the same images, shown as false colored.

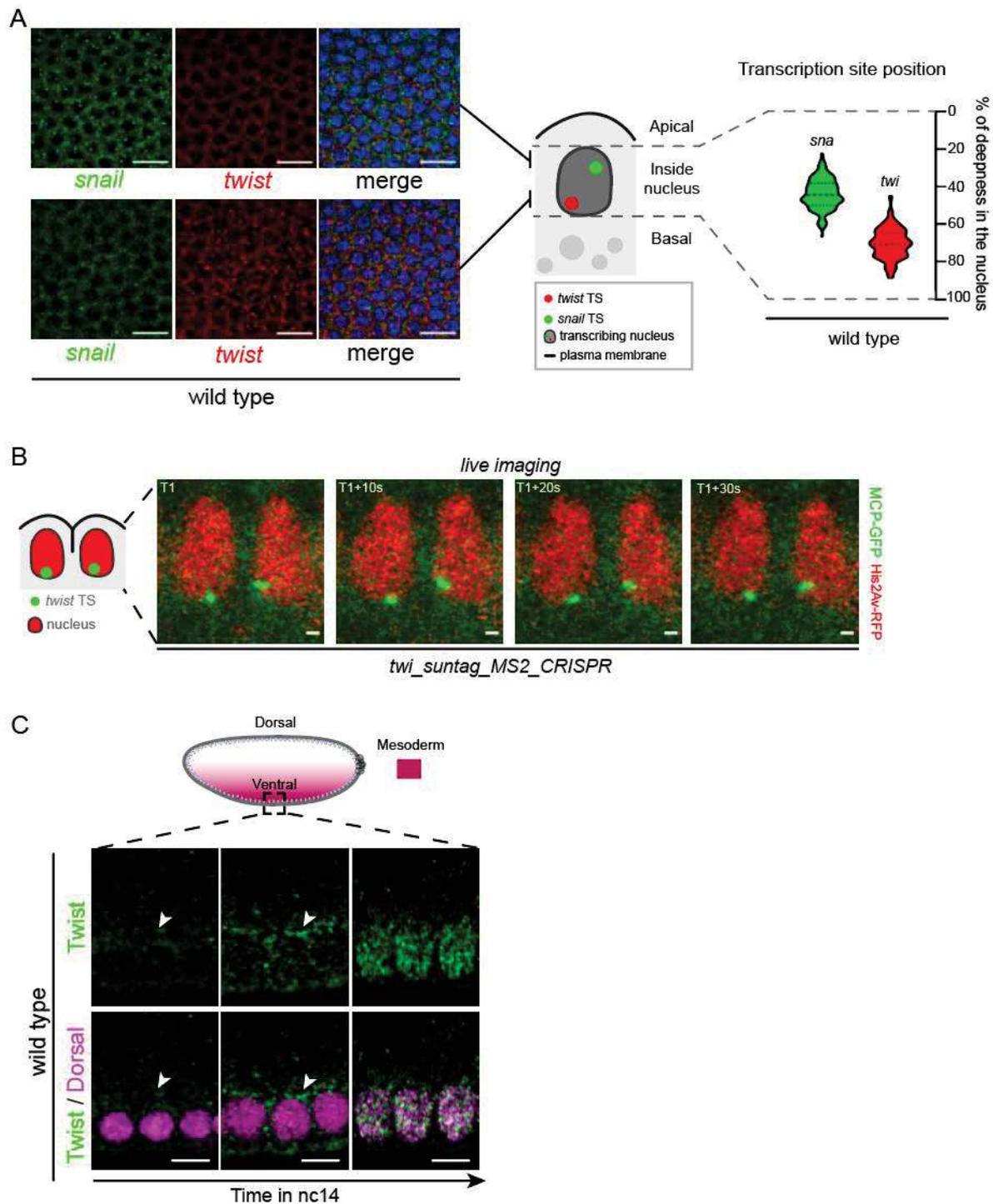


Figure S12: Apico-basal distribution of transcriptional sites for *sna* and *twi* genes

A, Maximum intensity projection of three z-planes ($\approx 1 \mu\text{m}$) of confocal images of n.c.14 wild type embryos labelled with *sna* (green) and *twi* (red) smFISH probes together with DAPI (blue). Upper left panels show images taken in the apical part of the nuclei and lower left panels in the basal part. Scale bars $10 \mu\text{m}$. Right panel represents quantification of *sna* and *twi* transcription site positions along nuclei apico-basal axis.

B, Snapshots of maximum intensity projection of 8 z-planes ($\approx 8 \mu\text{m}$) of confocal time lapse movie of *MCP-GFP, H2Av-RFP/+>twi_suntag_MS2_CRISPR/+* embryos (related Movie S6). Scale bars $1 \mu\text{m}$.

C, Immuno-labeled n.c.14 staged embryos with anti-Dorsal and anti-Twist antibodies.
Scale bars 5 μ m, (arrowhead: Twi foci).

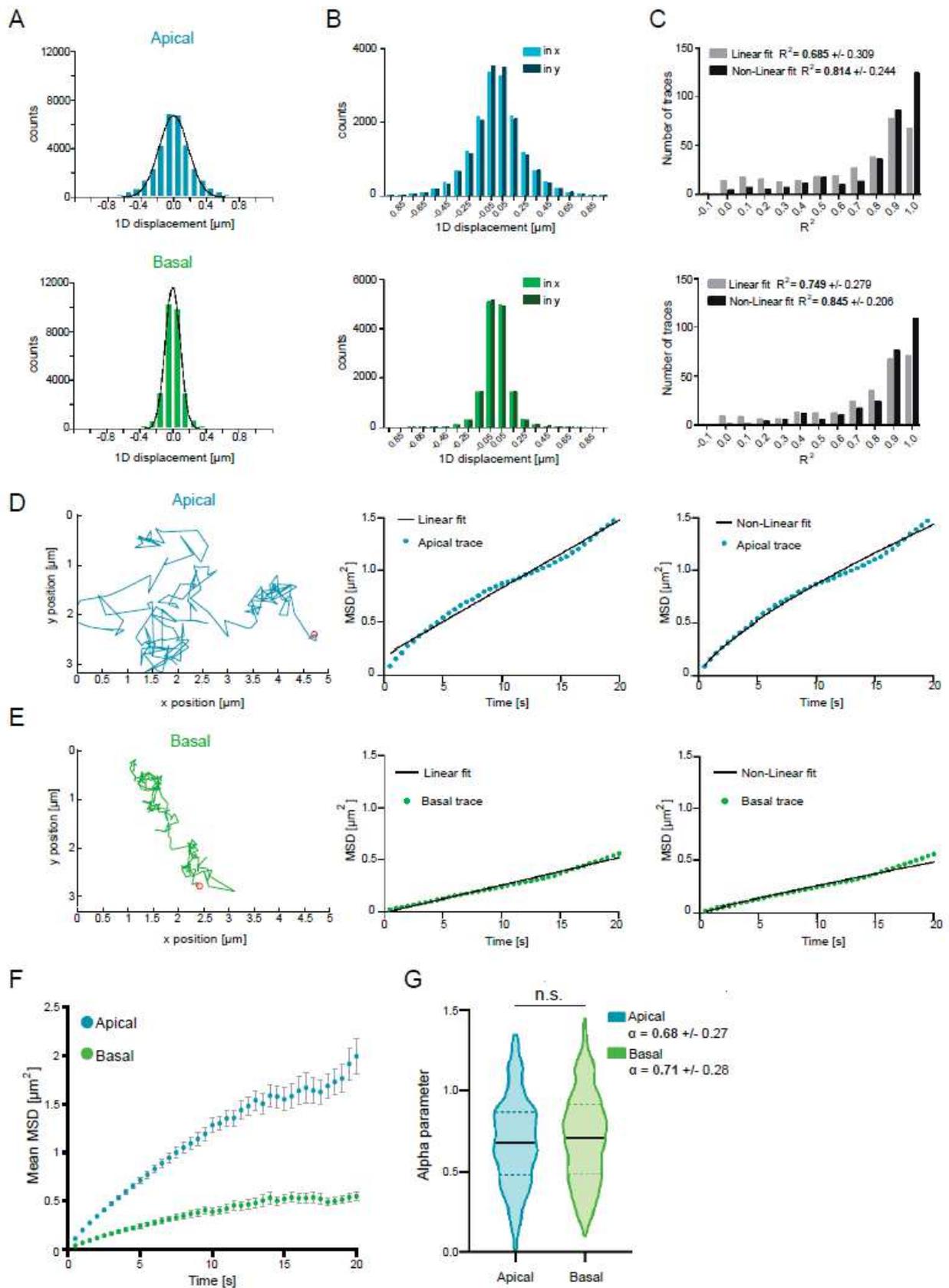


Figure S13: Quantifying the mobilities of *twi* mRNPs translation sites
A, 1D displacement distribution in x and y merged measured between two consecutive frames of traces located in the apical (blue) and basal (green) perinuclear cytoplasm

of *scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryos. Gaussian fitting is shown in black.

B, 1D displacement distribution in x and in y separately measured between two consecutive frames of traces located in the apical perinuclear cytoplasm (top histogram, x (light blue) and y (dark blue)) and basal perinuclear cytoplasm (bottom histogram, x (light green) and y (dark green)). Displacements are equally distributed in x and y showing no drift due to cellular movements.

C, Distribution of R^2 as a goodness of the linear fit (light grey) and non-linear fit (dark grey) of apical traces (top histogram) and basal traces (bottom histogram). A good fit is considered when $R^2 > 0.8$.

D, One representative trajectory of a translation particle located apically (left panel) as well as linear (middle panel) and non-linear (right panel) fitting of its MSD over time.

E, One representative trajectory of a translation particle located basally (left panel) as well as linear (middle panel) and non-linear (left panel) fitting of its MSD over time.

F, Graph representing the mean Mean Square Displacement (MSD) as a function of time for particles located apically (n=237 traces, three movies; blue) or basally (n=263 traces, three movies; green) nuclei. Error bars represent SEM.

G, Violin plots of the alpha parameter for the particles located apically (blue) and basally (green) using anomalous fitting (centered black bar represent median, dashed lines represent quartiles). Not significant (n.s.) $p > 0.05$ with a two-tailed Welch's t-test.

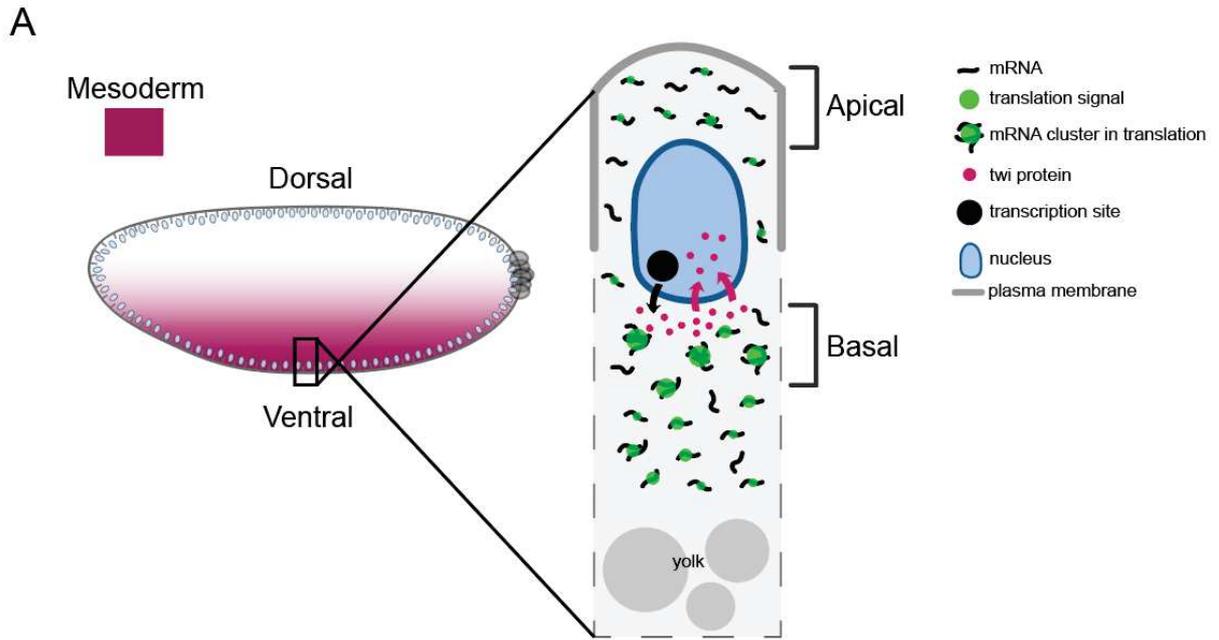


Figure S14: Graphical summary

A, Schematic summarizing the spatial heterogeneity in *twist* translation.

Captions for Movies S1 to S15

Movie S1: Maximum intensity projection of LightSheet Z-stack live imaging of *His2Av-mRFP/+;scFv-sfGFP-NLS/+>twi_suntag_CRISPR/+* embryo. Upper part: dorsal view, lower part: ventral view. Nuclei are detected using His2Av-mRFP and suntag using scFv-GFP. Scale bar 100µm.

Movie S2: Confocal live imaging of an embryo containing *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+*. Translation foci are in red and mRNA molecules in green, scale bar 1µm.

Movie S3: Zoom on a confocal live imaging used for tracking in Figure 1D of an embryo containing *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_CRISPR/+*. Translation foci are in red and mRNA molecule in green, scale bar 1µm.

Movie S4: Reconstructed 5µm thick movie of a MuViSPIM live imaging of a *His2Av-mRFP/+;scFv-GFP-NLS/+>twi_suntag_CRISPR/+* embryo. Nuclei are detected using His2Av-mRFP and suntag using scFv-GFP. Scale bar 30µm.

Movie S5: Maximum intensity projection of confocal Z-stack live imaging of an embryo containing *UASp-mito-mCherry/+;nosGal4,scFv-sfGFP-NLS/+>twi_suntag_CRISPR/+*. Translation foci below nuclei are shown, scale bar 5µm.

Movie S6: Maximum intensity projection of confocal Z-stack live imaging of a *MCP-GFP,H2Av-RFP/+>twi_suntag_MS2_CRISPR/+* embryo. Transcription sites are visualized in green and nucleus in red, scale bars 1µm.

Movie S7: A Maximum intensity projection of confocal Z-stack live imaging of a *scFv-sfGFP-NLS/+>twi_suntag_CRISPR/+* embryo. Translation foci above nuclei are shown, scale bar 5µm. **B** Maximum intensity projection of confocal Z-stack live imaging of an embryo containing *scFv-sfGFP-NLS/+>twi_suntag_CRISPR/+*. Translation foci below nuclei are shown, scale bar 5µm.

Movie S8: Maximum intensity projection of confocal Z-stack live imaging of an embryo containing: scFv-sfGFP no NLS showing no formation of aggregates in **A**, scFv-sfGFP-NLS showing no formation of aggregates in **B**, containing scFv-TagRFPT-NLS showing nuclear formation of aggregates in **C** and scFv-mScarlet-NLS showing no formation of aggregates in **D**. Scale bars 10µm.

Movie S9: Maximum intensity projection of LightSheet Z-stacks live imaging of a *scFv-sfGFP-NLS/+ > twi_suntag/+* transgenic embryo.

Movie S10: Maximum intensity projection of confocal Z-stack live imaging of: *scFv-sfGFP-NLS/+>twi_suntag* transgenic embryo after HEPES injection in **A** *scFv-sfGFP-NLS/+>twi_suntag* transgenic embryo after puromycin injection in **B** and *scFv-sfGFP-NLS/+>twi_suntag_CRISPR/+* embryo after puromycin injection in **C**. Signal of scFv-GFP dots are false colored. Scale bars 50µm.

Movie S11: Confocal live imaging of an embryo containing *MCP-eGFP/+;scFv-mScarlet-NLS/+>twi_suntag_MS2_transgene*. Translation foci are in red and mRNA molecules in green, scale bar 1µm.

Movie S12: A, Maximum intensity projection of Light-Sheet Z-stack live imaging of *His2Av-mRFP/+;scFv-sfGFP-NLS/+>twi_suntag_MS2_CRISPR/+* embryo. Upper part: dorsal view, lower part: ventral view. **B**, Reconstructed sagittal view of 5µm thick movie of the same embryo. Nuclei are detected using His2Av-mRFP and suntag using scFv-GFP.

Movie S13: Maximum intensity projection of confocal Z-stack live imaging of an embryo containing *MCP-eGFP-His2Av-mRFP/+>twi_suntag_MS2_CRISPR/+*. mRNAs are shown as bright green dot. scale bar 5µm.

Movie S14: Maximum intensity projection of LightSheet Z-stack 10µm thick live imaging of *scFv-sfGFP-NLS, MCP-TagRFPT/+ > twi _suntag_MS2_CRISPR/+* embryo. Transcription sites are detected using MCP-TagRFPT and suntag using scFv-GFP.

Movie S15: Maximum intensity projection of confocal Z-stack live imaging of a *UASp-me31B-GFP/+;scFv-mScarlet-NLS,nosGal4/+>twi_suntag_CRISPR/+* embryo. Translation foci below nuclei are shown (red), scale bar 10µm.

Supplementary sequences

sequence of *twist suntag* inserted into *puc57* simple

gtttaaacaacataattatcattactctgaatgcactttttcaaaacttagaaactctgtcctatgaattcccgtcgatccaa
agatattctcaatccccttttgaatcaacaagtaaaatattcaaaaattgccgacaattcccctcgattccccgtcccgc
atcccaacacgcatactcccaggcattttcccaaatcgagagaaaaaccaagaataaccaagagaaaacagaa
aatccagagcgtcgagtcaaggctcttcaatttagctttgaattgtgtattttcgtttgcagccgccgctgccgctcg
agaaaaatcgaaaatccccgccgctgacgtcatacctgccgatgccgcagctccgccattgagtgaggagcggccg
cgggatggcaagacaagcagcagcagcgggacgacgatagagcggcggcagcgaatggccgtccatattggagc
agccgcaaaaatgtcaattgagcaatggccggaaggatcctgcgtcagttgcgttccgtaagtgcgtgcgagcagatc
gatccagcaaaaacgcccggccatggcgtgaagaatatctacggagtatacagttccgaaataagaatattgttaga
taccacaaattctaactgaagaagtgcgctaaaaagccaagcaagatcaccaaatgggtaccaagcttagcggca
gcccgcgaggagctgctgagcaagaactaccacctggagaacgaggtcgccagactgaagaagggttagcggctc
cggtaggagctgcttcaagaactaccattggagaatgaagtggcgcgattgaaaaggccagcgggtcagga
gaagaacttctcttaagaactatcacttagagaatgaagttagctcgattaaagaaggctctggctcaggtgaagag
ctctatcaaagaattaccatctgaaaatgaggtagcagacttaagaaggctccggatctggtgaagaattgctct
ctaagaactatcacctcgagaacgaagtcgctcgctcaagaagggttctgggagtgaggagaagaacttctgctaaa
aattaccacctagagaacgaggttcacgcctaaaaagggaagcgggaagcggtagagaattgctctcaagaat
tatcattggaaaatgaagtggcgcggtgaaaaagggaagcggcagcggtagggaactcctcagcaagaactatc
atctggaaaatgaagttagcagactgaaaaagggtcaggtagtgccgaagaattgctctcaaaaattaccactg
gagaatgaggttcacgctgaagaagggtctgggagtgaggagaagagttactgagtaagaactatcatttagaaaa
tgaggtcgcaaggctaaagaagggttctggctccggggaagaactcctgcttaagaactaccatcttgagaacgaag
ttgccagacttaaaaaagggttagcagcgggagaagagttgtatccaagaactaccatctcgaaaatgaagtggct
aggctgaaaaagggttctgggagtgaggagagctctcttaaaaactaccacctagaaaacgaggttagctcgg
ctaaaaagggtctggcagcggagaggagcttctcagtaagaattaccatctggagaacgaagtagcccggctga
aaaagggttagcgggtccgggtgaagaactattatctaaaaattaccatctcgagaatgaggtcgcaagactcaagaag
ggtagtggatctggagaagagttgctctcaaaaactaccatctgaaaacgaagttgcgcgacttaagaaggatc
aggcagtgaggagaagaactgttgagcaaaaactaccatctcgaaaatgaggtcgccaggctcaagaaggctcag
gtagcggtagaagagttgttctagaattaccatctcgagaatgaagtggcaagactcaaaaagggttccggctccg
gggaagagctactatcaaaaaactaccatctcgagaatgaagtgcacgactcaaaaagggttagctggatcagggg
aggagctttgagtaagaattatcacttgaaaacgaagttgcgagattgaagaagggttagctggctcaggggaggag
ctcctgagtaagaactatcacctagagaacgaggtggcagggctaaaaagggttccggtagtggtgaagaattact
atctaaaaactaccatctggaaaatgaagtggcagggctgaagaagggaagtggttagcggagaggagctattgag
taaaaattatcatttagagaatgaagtggcccgttaagaagggttaggtccggagaagagttgtgtcaaaaaa
ttatcatttagaaaacgaagtgccacgactaaagaagggtcaggatcaggagaagagctcctctcaagaattatc
acctggaaaatgaggtcgcccgactgaagaagggtcagggagcggcgaagaattgctcttaagaattatcactta
gaaaatgaggttcacgattaaaaaaaggtagtgaggagcggtagggaactactctcaaaaactatcatctcgaga
acgaggttagcacggctcaagaagggtctggaagcggtagggaattactttcaaaaactatcacctcgaaaacg
aagttgcgcgactcaaaaagggttaggtccggtagaagagctattatctagaactaccacctggaaaacgaggt
ggctcggctgaagaagggttagcgggagtggggaagagctgcttctaaaaattaccatttagagaatgaggtgcta
gataaaaaagggttagcggatcaggagaggaattgctaagtaagaactaccacttagaaaacgaagttgccgctt
aaaaagggttagcgaagtggtgaggaactgttgagcaagaattatcacctcgagaacgaagttgctcgcctcaag
aaaggctctggtagcggtaggaattactaagtaaaaaactatcatttagagaatgaagttagcaagattaaaaagg
cagcgatataccggtagatgagcgtcgcgtcgggtgcgccaaagtgtgctggacataagctacaagcccacac
tgccaacatcatggagctgcagaacaatgtgatcaagctgatacaggtggagcagcaggcctacatgcagtcgg
ctatcagctgcagcaccagcagcagcaccctccactcccaccagcaccaccagcagcaccaccagcagcagcatg
cccagtagcggccactgccctcggagtacggcctatggattaccgaactggaggacacagactacaacataacc
cagcaacgaggtcctgagcaccagcagcaaccagagtgcccagagcaccagctgaggctgaacaacaacaac

accagcagcaacaccaacagctccggcaacaatccgagtggttcgatggccaggccagcagcggatcctcttgg
aacgagcacggcaagagggccaggagcagtggtcgactacgattgccaaaccgggggatcactggatcagcagc
ccgagcacaagaagctaaccaccagcaacagcagcagcaacaacagcaccagcaacagatctatgtggattac
ttgccaccaccgtggacgaggtggctcggctcaatctgtcctggcgtccagagcacgtgcacctccccgcaatcc
cacttcgatcttcccgacgaggagctgcccagcacaaggcccagggtgttctgcccctctacaacaaccagcagca
gcagtcgcagcagctgcaacagcagcagccgcaccagcaaaagccacgcccagatgcacttccaaaacgcctac
agacaaagtttcgagggctacgagccggccaactcgtgaacggcagtgcttactccagctcggatcgggatgacat
ggagtacgcccgccacaacgcccctgagttcagttagcagatctcaacggaggagtcagtcgcccgcctgcttggcg
atgacggcagtgccggcagtttctggacggatccgatgccggcggaaggcctccgcaagccacgtcggcggt
gaagcgggaagcccagcaagacggaggagcggacgagttcagcaaccagcgggtcatggccaatgtgagggga
gcccagcgcacccagagcctcaacgacgcttcaagtccctgcagcagatcatccccacgctgcccagcgcaca
gctcagcaagatccagaccctcaaactggccacaaggtaataatacttttagtttacatcttcataaactagttagttgaaa
gcattcatagatagatagatagataatagatgaattgaatcgaactcgttaataatcgaactcgtttaaccaacttaca
gatacattgacttctgtgcccagatgctcagctcagtgatatacttctgtaaggccttggaggcccagggtatcgccct
cggcgtatggatcggccagctccctcctgagtgccgcccgaatggagccgaggcagatctgaagtgctcgcgca
ggccaacggagcaccattatcccggcggagaagctgagttatctgttctgggtgtggcgcatggaggggcgacgcg
cagcaccagaaggcatagcggcggatcaggacactctagatagatccgagaggggctccagcagcaactatcg
tgtgaattccagagccctggcgtctcactcttcaattctctgcacgcttccatataacggtcgcactagatgattctaa
atgtctaagcctaaaaacactactgttaatgcctatthaatgtcatagtctaaactaaattaattgaaaaagccaacaagc
caagaacaaagaaacgcatgaacaaaaccagagctagcggccggcc

sequence of pNosPE MCP-eGFP

cgccccctgacgagcatcacaanaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagata
ccaggcgtttccccctggaagctccctcgtgctcctctgttccgacctgcccgttaccggatacctgtccgcctttctc
cnttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcccaagctgggct
gtgtgcagcaacccccgtcagcccagcgtcgccttaccgtaactatcgtcttgagccaacccggtaagac
acgacttatcgccactggcagcagccactgtaacaggattagcagagcaggtatgtagggcgtgtacagagttc
ttgaagtgggtggcctaactacggctacactagaagaacagatatttggtatctgcgctcgtgaagccagttacctcgg
aaaaagagttgtagctcttgatccggcaacaaaccaccgctgtagcgggtggtttttgttgcaagcagcagatta
cgcgcaaaaaaaggatctcaagaagatcctttgatctttctacgggtctgacgctcagtggaacgaaaactcac
gtaagggttttggatgagattatcaaaaaggatcttcacctagatccttttaataaaaaatgaagtttaaatcaatc
taaagtataatgagtaaactggctgacagttaccaatgcttaatcagtgaggcacctatctcagcagatctgtctatctc
ttcatccatagttgctgactcccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaa
tgataccgcgagaccacgctcaccggctccagattatcagcaataaaccagccagccggaaggggccgagcgcga
gaagtggctcgaactttatccgctccatccagcttattaattgttggcgggaagctagagtaagtagttcggcagtta
atagtttgcgaacggtgttgcattgctacaggcatcgtggtgacgctcgtctgttggatggcttattcagctccggtt
ccaacgatcaaggcagttacatgatccccatgttgcgcaaaaaagcggtagctcctcggctcctccgatcgttgc
agaagtaagttggccgagtggtatcactcatggtatggcagcactgcataattcttactgtcatgccaatccgtaagat
gctttctgtgactggtgagtactcaaccaagtcattctgagaatagtgatgcggcgaccgagttgctcttggccggt
caatacgggataataccgcgccacatagcagaactttaaagtgtcatcattggaaaacggttctcggggcgaaaa
ctctcaaggatcttaccgctgttgagatccagttcagatgaaccactcgtgcaccaactgatcttcagcatctttactttc
accagcgtttctgggtgagcaaaaacaggaaggcaaatgcccgaanaagggaataagggcgacacggaaat
gttgaatactcactcttcttttcaatattatgaagcatttatcagggttattgtctcatgagcggatacatattgaatgta
tttagaaaaataaacaataggggttccgcgacatttccccgaaaagtccacctgagctaaagaaaccattattatc
atgacattaacctataaaaaataggcgtatcacgaggcccttctcgtctcaagaattcgtttaaaccggccggccagatcc
aggctgcagcggccgcttcgaccgttttaacctcgaaatgacatgtaaggacggatgtgagcgaacgcccagt
gatgaccgggatcagaggtaacctaccatggtggggattaggtgaccgttcgaggtagtttgatcggagcgaatgtt
cggggggtctggcgtcagaggctctaaactttatgtaattcctgcccgaacacgcacgtatcaagcagtcagctgtt

ctcttcgttcagcgcgcgcccgggtgtgcaaaacgagcgcctctcgcggcggtggctcgtgcgatagttcgtttgtcgg
aatccgatgtgcccgcgcatatcatgtgatgtgtcacagtgccgcaaatcgaatgggtgtgcagtgattgtgtgt
gacggcagtgggcgcgtgtgggtgcttagtttgggagatgtttcgtatttttgtgataactcaggcttgtgtgtgt
agtactatttccattgcccgggtgtccagcttttaattagtgccacatattcttagcaagtaaaaatttttgcatactattaa
atcttataaatttttctaaaattaagtttaccctttcaatttactaaaaatcgcataatatttattatcgctggaaaactacat
tattccacctctaagcaagaaccgtagttggcgcgtagctttaccacaaaattcctggaattgccgtacgcttcgcagtt
gtttcaagttgtctaagggacatacagattttttgcctctgcgtcacgattttaacccaaaagcgagtttagttacatgtacat
tattattagataaagaagatcgcgaatacttcagttgaataaactgtgcttggttttgggtgaggatttggtaaagtaga
gtgcccgcgataaccgtaactttcgacccggattttcgccctcgagatggcttctaactttactcagttcgttctcgtcgacaat
ggcggaaactggcgacgtgactgtgcctcgaagcaacttcgtaacgggatcgtgaatggatcagctctaactcgcg
ttcacaggcttacaagtaacctgtagcgttcgtcagagctcgtcgcagaatcgcataacacatcaaagtcgaggt
gcctaaaggcgcctggcgttcgtacttaaatatggaaactaacattccaattttcgccacgaattccgactgcgagcttat
tgtaaggcaatgcaaggtctcctaaaagatggaaacccgattccctcagcaatcgcagcaaacctccggcatctaca
agcttggcggatcaggctcgggatcatcggctagcatggtgagcaagggcgaggagctgttaccgggggtgtgccc
catctggtcgcagctggacggcgcgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccac
ctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgcccggcccaccctcgtgaccacc
tgacctacggcgtgcagtgcttcagccgtaccccaccacatgaagcagcagcacttctcaagtccgcatgcccg
aaggctacgtccaggagcgcacatcttctcaaggacgacggcaactacaagaccgcgcccagggtgaagttcg
agggcgacaccctggtgaaccgcatcgcagctgaagggcatcgcactcaaggaggacggcaacatcctgggacac
aagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaa
cttcaagatccgccacaacatcgaggacggcagcgtgcagctcgcgaccactaccagcagaacacccccatcg
gagcagggcccgtgctgctgcccgacaaccactacctgagcaccagctccgcccctgagcaaaagacccaacgag
aagcgcgatcacatggtcctgctggagttcgtgaccgcccggggatcactctcggcatggacgagctgtacaagta
aggatccgcgtcacgccactcaacgctcgcagtgaggcgtcattggtgggggtaaccgtcgaatcagtgtttac
gcttcaatcgcaaaaaaattcactgcaacactgaaaagcatacgaaaacgatgaagattgacgagaaacctat
aaagtattttatccacaaagacacgtatagcanaaaagccaagttaactcggcgataagttgtgtacacaagaataa
aatcggccagattcagtggttcagaaataagaaaacccactatgttttcttgccttttcttcccagcgatcattcatt
cgtggtgaaagaacggggcattgcacggagtttcgactgcccggaaagcagagctcggcttactctctataattag
cgctttctatttccccgattcgggcccgtgctgctgtttccgctgctgtttgtggcaagtgtagcagcaggctgtgcacg
cagtggtgcatgcacttggcttccaccgttggtatcgattctctgggacgatgagtcattccttccggggccacagcataa
tcgttccagctcaccgaaatggtgacttatttctaactgccgtcaagcatgcgattgtacatacataatattatata
gtacataatattgtagctatggttaggtcgataataatagcaatcaacgcaagcaaatgtgtcagctcctgcttacaggaacg
attctatttagtaatttctgtgtataaagtaattatgtatgtatgtaagccccataaatctgaacaattaggcaaaaccatg
cgaagctctctagagtcgacctgcatctacacaaggaacaaacactggatgtcactttcagttcaaatgtaacgtaaa
tactccgaacaggtcacaataaattacctaataaagtcataatattaaattagaataaataatagctgtgagggaaata
tatacaaatatattggagcaataaattgtacatacaaatatttactaatttctattgagacgaaatgaaccactcggga
accatttgagcgaaccgaatcgcgcggaactaacgcagctcctcaaggctcgtcgaacaaaaggatgtgtgtg
cggagagcgggtgggagacagcgaagagcaactacgaaacgtggtgtggtggaggatgaattatgaagagggc
gcccgatattgaaaagtagtatataaaaaatatacccgggtgtttatgtagcgataaacgagttttgatgtaaggtatgc
aggtgtgtaagcttttggtagaagacaaatccaaagtctactgtggggatgttcgaaggggaaataactgtattctata
ggtcatactgtttttattggcacaataataattacattagctttttgagggggcaataaacagtaaacacgatggtaata
atggtaaaaaaaaaaacaagcagttatttccgataatgtcggctactccttgcgtcgggcccgaagcttagagcca
gatatgcgagcaccgggaagctcacgatgagaatggccagaccacgtagtccagcggcagatcggcggcgggag
aagttaagcgtctccaggatgacctgcccgaactggggcacgtggtgtcgcagatgtgcagctaattcggccggct
ccacgtccgcccattggttaatcagcagaccctcgttggcgtaacggaaccatgagaggtacgacaaccattgaggt
atactggcaccgagcccagttcaagaanaagccgccaagagcaggaatggtatgataaccggcggaccac
agacagcgcctcgcaggtcgcaggtggtggcaggaatattagataccgaaggcgttgacacattggccaccag
agtgaccagcgcaggcagttgagaagtgacgactccggcccgcagtcgatcatcgataggcaatcgcggt
gaagaccagtggtcactgtgagaaaaagcggcaattcggcaatcgttttcccagaaagtagtgtcagcagcagataa
agtcgactcgggcccctcataaaaactggcagctcgtgaggtgaacacctaatacgaatcgattcattagaaagtta

gcggcagcctgaagctgagctgcgccgtgagcggctcagcctgaccgactacggcgtgaactgggtgcgccagg
ccccggccgcgccctggagtgatcggcgatctggggcgacggcatcaccgactacaacagcgccctgaagg
accgctcatcatcagcaaggacaacggcaagaacaccgtgtacctgcagatgagcaaggtgcgacgcgacgac
accgcccgtgactactgcgtgaccggcctgttcgactactggggccagggcaccctggtgaccgtgagcagctacc
atcagatgtccagattacgctggtggaggcggagggtctgggggaggaggtagtgccggtggtggttcaggaggcg
gcggaagcttgatccaggtggaggtggaagcggtagcaaaggagaagaactttcactggaggtgtccaattctg
ttgaattagatggtgatgtaatgggcacaaatctgtccgtggagaggggtgaaggtgatgtacaacggaaaact
caccctaaatatttgcactactggaaaaactacctgtccgtggccaacactgtcactactctgacctatggtgtcaat
gctttcccgttatccggatcacatgaaacggcatgacttttcaagagtccatgccgaagggtatgtacaggaacgc
actatatcttcaaagatgacgggacctacaagacgcgtgctgaagtaagttgaaggtgataccctgttaatcgtatc
gagttaaagggatgtgatttaagaagatggaacattctggacacaaactcgagtacaactttaactcacacaatgt
atacatcacggcagacaaaacaaagaatggaatcaaagctaactcaaaattcgccacaacggtgaagatggtcc
gttcaactagcagaccattatcaacaaaatactccaattggcgtatggccctgtcctttaccagacaaccattacctgtc
gacacaatctgtccttgcgaaagatcccaacgaaaagcgtgaccacatggtccttctgagtttgaactgctgctggga
ttacacatggcatggatgagctctacaaggtggaggtcggaccgaagagtacaagcttatcctgaacggtaaaacc
ctgaaaggtgaaaccaccaccgaagctgtgacgctgctaccgcggaaaaagtttcaaacagtacgctaacgaca
acggtgtgacggtgaatggacctacgacgacgctaccaaaccctcacggaaccgaataa

sequence of scFv-sfGFP-NLS

atgggccccgacatcgtgatgaccagagccccagcagcctgagcgcagcgtggggcgaccgctgaccatcac
ctgccgagcagcaccggcgccgtgaccaccagcaactacgccagctgggtgcaggagaagcccggcaagctgt
tcaagggcctgatcggcggcaccacaaccgccccggcgtgccagccgctcagcggcagcctgatcggcg
acaaggccaccctgaccatcagcagcctgcagcccaggacttcgccacctactctgcgccctgtggtacagcaac
cactgggtgttcggccagggcaccaaggtggagctgaagcgcgcgccggcggcagcggcggcggcggcagcg
gcgcgcgccgagcagcggcggcggcagcaggtgaagctgctggagagcggcggcggcctggtgcagcccg
gcgcgacctgaagctgagctgcgccgtgagcggctcagcctgaccgactacggcgtgaactgggtgcgccagg
ccccggccgcgccctggagtgatcggcgatctggggcgacggcatcaccgactacaacagcgccctgaagg
accgctcatcatcagcaaggacaacggcaagaacaccgtgtacctgcagatgagcaaggtgcgacgcgacgac
accgcccgtgactactgcgtgaccggcctgttcgactactggggccagggcaccctggtgaccgtgagcagctacc
atcagatgtccagattacgctggtggaggcggagggtctgggggaggaggtagtgccggtggtggttcaggaggcg
gcggaagcttgatccaggtggaggtggaagcggtagcaaaggagaagaactttcactggaggtgtccaattctg
ttgaattagatggtgatgtaatgggcacaaatctgtccgtggagaggggtgaaggtgatgtacaacggaaaact
caccctaaatatttgcactactggaaaaactacctgtccgtggccaacactgtcactactctgacctatggtgtcaat
gctttcccgttatccggatcacatgaaacggcatgacttttcaagagtccatgccgaagggtatgtacaggaacgc
actatatcttcaaagatgacgggacctacaagacgcgtgctgaagtaagttgaaggtgataccctgttaatcgtatc
gagttaaagggatgtgatttaagaagatggaacattctggacacaaactcgagtacaactttaactcacacaatgt
atacatcacggcagacaaaacaaagaatggaatcaaagctaactcaaaattcgccacaacggtgaagatggtcc
gttcaactagcagaccattatcaacaaaatactccaattggcgtatggccctgtcctttaccagacaaccattacctgtc
gacacaatctgtccttgcgaaagatcccaacgaaaagcgtgaccacatggtccttctgagtttgaactgctgctggga
ttacacatggcatggatgagctctacaaggtggaggtcggaccgaagagtacaagcttatcctgaacggtaaaacc
ctgaaaggtgaaaccaccaccgaagctgtgacgctgctaccgcggaaaaagtttcaaacagtacgctaacgaca
acggtgtgacggtgaatggacctacgacgacgctaccaaaccctcacggaaccgaaggtggtggttagcgggtgt
ggtactagtccaagaagaagcgaaggtgtaa

sequence of scFv-sfGFP-NLS- bcd3'UTR

atgggccccgacatcgtgatgacccagagccccagcagcctgagcgccagcgtgggcgaccgctgacccatcac
ctgccgagcagcaccggcgccgtgaccaccagcaactacgccagctgggtgcaggagaagccccggcaagctgt
tcaagggcctgatcggcggcaccaacaaccgcgccccggcgtgccagccgcttcagcggcagcctgatcggcg
acaaggccaccctgacccatcagcagcctgagcccaggacttcgccacctactctgcgccctgtggtacagcaac
cactgggtgttcggccagggcaccaaggtggagctgaagcgcggcggcggcgccagcggcggcgccgagcg
gcgccggcggcagcagcggcggcggcagcaggtgaagctgctggagagcggcggcggcctgggtgcagcccg
gcgccagcctgaagctgagctgcgccgtgagcggcttcagcctgaccgactacggcgtgaactgggtgcgccagg
ccccggccgcgccctggagtggatcggcgtgatctggggcgacggcatcaccgactacaacagcgcctgaagg
accgctcatcatcagcaaggacaacggcaagaacaccgtgtacctgcagatgagcaaggtgctgcagcagcagc
accgcccctgtactactgctgaccggcctgttcgactactggggccagggcaccctggtagccgtgagcagctacc
atacgatgtccagattacgctggaggcggagggtctggggaggaggtagtgccggtggtggttcaggaggcg
gcggaagctggatccagggtggagggtgaagcgtgagcaaggagaagaactttcactggagttgtccaattctg
ttgaattagatggtgatgttaatgggcacaaatctctcgtggagagggtgaaggtgatgctacaacggaaaact
caccctaaatatttgcactactggaactacctgtccgtggcaacactgtcactactgacctatggtgttcaat
gctttcccgttaccggatcacatgaaacggcatgactttcaagagtccatgccgaagggtatgtacaggaacgc
actatatcttcaaagatgacgggacctacaagcgcgtgctgaagtcaagttgaaggtgataccctgttaatcgtatc
gagttaaagggtatgtatgtaagaagatggaacattctggacacaaactcgagtacaactttaactcacacaatgt
atacatcacggcagacaaacaaaagaatggaatcaaagctaaactcaaaattcgccacaacgttgaagatggtcc
gttcaactagcagaccattatcaacaaaatactccaattggcgtatggccctgtcctttaccagacaaccattacctgtc
gacacaatctgtcctttcgaagatcccaacgaaaagcgtgaccacatggtccttctgagtttgaactgctgctggga
ttacacatggcatggatgagctctacaaggtggaggctggaccgaagagtacaagctatcctgaacggtaaaacc
ctgaaagggtgaaaccaccaccgaagctgtgacgctgctaccgcggaaaaagtttcaaacagtagcctaacgaca
acgggtgtgacgggtgaatggacctacgacgacgctacccaaacctcacggtaaccgaaggtggtgtagcgggtgt
ggtactagtcccaagaagaagcgcgaaggtgaacctggatgagaggcgtgttagagagttcattagcttaggttaac
cactgtgttctctgattgtacaaataccaagtgattgtagatctacgcgtagaaagttaggctagcttaagatccgtg
taaaggttccagggaagtttatgtactagcctagtcagcaggccgcacggattccagtgcatatcttagtgatactcc
agttactctatactttccctgcaatacgtattcgccttagatgtatctgggtggctgctccactaaagcccgggaatatg
caaccagttacattgaggccatttgggcttaagcgtattccatggaaagttatcgtcccacatttcggaattatattccg
agccagcaagaaaatcttctctgttacaatttgacatagctaaaaactgtactaatcaaatgaaaaatgtttctctggg
cgtaatctacataatgattaccctaaagatcgaacatttaacaataatattgatgatatttcaatttctatgctatgc
caaagtgctgacataatcaaacatttgcgacttcttgaccaagaatagtcagcaaatgtatttcaatcaatgcagac
cattgtttcagattcggagatttttgtgccaacggataactatcatagctcacatttctatcatcactaagaagag
cattgcaatctgttaggcctcaagtttaatttaaatgctgcacctttagtgtgtctttaaagcttgtattttaaattacgaaa
atatataagaactactctactcgggtaaatgtgactaa

sequence of scFv-mScarlet-NLS

atgggccccgacatcgtgatgacccagagccccagcagcctgagcgccagcgtgggcgaccgctgacccatcac
ctgccgagcagcaccggcgccgtgaccaccagcaactacgccagctgggtgcaggagaagccccggcaagctgt
tcaagggcctgatcggcggcaccaacaaccgcgccccggcgtgccagccgcttcagcggcagcctgatcggcg
acaaggccaccctgacccatcagcagcctgagcccaggacttcgccacctactctgcgccctgtggtacagcaac
cactgggtgttcggccagggcaccaaggtggagctgaagcgcggcggcggcgccagcggcggcgccgagcg
gcgccggcggcagcagcggcggcggcagcaggtgaagctgctggagagcggcggcggcctgggtgcagcccg
gcgccagcctgaagctgagctgcgccgtgagcggcttcagcctgaccgactacggcgtgaactgggtgcgccagg
ccccggccgcgccctggagtggatcggcgtgatctggggcgacggcatcaccgactacaacagcgcctgaagg
accgctcatcatcagcaaggacaacggcaagaacaccgtgtacctgcagatgagcaaggtgctgcagcagcagc
accgcccctgtactactgctgaccggcctgttcgactactggggccagggcaccctggtagccgtgagcagctacc
atacgatgtccagattacgctggaggcggagggtctggggaggaggtagtgccggtggtggttcaggaggcg
gcggaagctggatccagggtggagggtgaagcgtgtgagcaaggcggaggcagtgatcaaggagttcatgcggtt

caaggtgcacatggagggctccatgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacg
agggcaccagaccgccaagctgaaggtgaccaaggggtggccccctgcccttctcctgggacatcctgtcccctcag
ttcatgtacggctccagggcctcaccaagcaccgcccgcacatccccgactactataagcagtccttccccgagggc
ttcaagtgggagcgcgtgatgaacttcgaggacggcggcggcgtgaccgtgacccaggacacctccctggaggac
ggcaccctgatctacaaggtgaagctccgcgccaccaacttccctcctgacggccccgtaatgcagaagaagacia
tgggctgggaagcgtccaccgagcgggtgtacccccaggacggcgtgctgaagggcgacattaagatggccctgc
gcctgaaggacggcggccgctacctggcggactcaagaccctacaaggccaagaagccctgcagatgcc
ggcgctacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtgggtggaacagtacgaa
cgctccgagggccgcccactccaccggcggcatggacgagctgtacaaggggtggaggtcggaccgaagagtacaa
gcttatcctgaacggtaaaacctgaaaggtgaaaccaccaccgaagctgttgacgctgctaccgcggaaaaagt
tcaaacagtacgtaacgacaacgggtgtgacgggtgaatggacctacgacgacgctacaaaaccttcacggtaac
cgaaggtgggtgtagcgggtgggtactagctccaagaagaagcgcaaggtgtaa

sequence of *Twipe TwiPr*

aacataattatcattactctgaatgcacttttcaaaacttagaaactctgtcctatgaattcccgtcgatccaaagatattc
tcaatccccttttgaatcaacaagtaaaatattcaaaaaattgccgacaattcccctcgtattcccctcccgcacccaa
cacgcatactcccaggcattttcccaaactcgagagaaaacccaaagaataacccaagagaaaacagaaaaatcca
gagcgtcgagtcaaggctctctcaatttagcttgaattgtctgtatttctgtttgcagccgcccgtgccgctcgagaaaa
tcgaaatccccgcccgtgacgtcatacctgccgatgccgcagctccgccattgagtgaggagcggccgcccgggatg
gcaagacaagcgcgagcgggacgacgatagagcggcggcagcgaatggccgtccatattggagcagccgca
aaatgtcaattgagcaatggccggaaggatcctgcgtcagttgcgtccgtaagtgcgtgcgagcagatcgatccag
caaaacgcggg

sequence of *ilp4* transgene

gtttaaaccatggccaggtggctccgatgccgcaactattggcattggagccggaaccgatcgtgctaactcccagc
agcttctgtatgtcgcctcgtggagtagagtgtgtcaaggtggaggtcactaatgtgccagaagtagcctgtcaac
aaacaagtagaatcaagtaaatgtgttagttaatacccatagatatagtaaaagtgtgttttatttgtaagaaaagt
taatctatatcccagttttacacaacaaaattttatgtcctgagcaatttcgatgtatttccccttcgtaaagtaaggatcgag
atagactttgactttggttaagtccggcaattcctggccgggaaaggccatttcccttcgcccgggcatttcccgcgggct
ggtcgagcgcacaaaataagaaaaacctggtagttcaaatggaaatctcctgcagctgactgtttggttggtgactga
cctggcccgaatttaactttctacctggtcgcaatacgtgaagtcaaaaagtcaattagcagagcaacattttgagcgc
ggccaactccaaggatcagtatcatttggcatgccagcgcgatcggtttgccaagagcagcagagaagttcgagatagga
cccagagataccagagataaaggaggcatacctttatgcccgggtgagagcacggacggcggagtgaaagatcga
gcaggatgggtaccaagcttcccggggatatacgcctgattagactgggactggcgcgtgctgctcctgctggccaccg
tgtcgcagttactgcagccggtccagggacgcccgaagatgtgcggcggaggctctgatccaggcactggatgtgatt
gtgtaatggatttacgcgcctgtcaggcggagcagtggttaagttgggtactatgcatattcgattggctccatacatc
taacttctttcgacaagcgtctaaggatgctagagtgcgagacctatccgtaagctacagcagccggatgaggacat
tgaacaggaaacggaaacgggaagggttaaagcagaagcatacggatgcggatacggagaaggggtgtgccaccg
gccgtcggaaagtggacgaaagtgcgacgccatcggcgcgacgcatcggccagagtggtgcaaggaggggtgcacc
tacgacgataactggactactgcgcctgatgaccaggatggcaaaaacaaaacaaaataaaaaccagaaaccagat
cccaaaaaccaagtaccagatgaacacgacatggctgagattttgtgtggcggcaggggaaaacaccggacga
ccggcagggctatttgcaattcattttcctactacacttaaccctaaactataaacgtaatcgtatttccaaatatttcattgtaa
aatttctagtggaggcaataaagttactctccaagcagcagcaggctagcggccggcc

Supplementary table 1

Primer Assembly CRISPR:

TwicRISPRSuntag_Frag1_F	ctatagggcgaattgggtaccGGGTATTTTTGGTGGCCCC	
TwicRISPRSuntag_Frag1_R	cgtaacgatccgcggCCGCTCCCACTCAATGGC	
TwicRISPRSuntag_Frag2_F	gtgggagcggccgcggatcggttacgatcgctggccggccCGCCATTTGCATTCCATTC	
TwicRISPRSuntag_Frag2_R	agggaaacaaaagctggagctcAGCAATAGAAAGTGTTAAGTGG	

Primer gRNA CRISPR:

Guide1_Twi_F	gtcg GGGAATACGAGGGGAATTGT
Guide1_Twi_R	aaac ACAATCCCCTCGTATTCCC
Guide2_Twi_F	gtcg GGGTCGGTCCCCTCGAGATC
Guide2_Twi_R	aaac GATCTCGAGGGGACCGACCC

Primer Assembly all scFv-sfGFP:

HIFI_scfvGFPNls_F	ttcgaccggattttgcacctcgagATGGGCCCCGACATCGTG
HIFI_scfvGFPNls_bic_R	ctcatccaggTTACACCTTGCCTTCTTCTTGG
HIFI_Bic3UTR_F	caaggtgtaaCCTGGATGAGAGGCGTGTTAG
HIFI_Bic3UTR_R	tgtgtagatgcaggtcgactctagaGTAGTTAGTCACAATTTACCCGAGTAG
HIFI_scfv_noNLS_F	ttcgaccggattttgcacctcgagATGGGCCCCGACATCGTG
HIFI_scfv_noNLS_R	tgctaccgctTCCACCTCCACCTGGATC
HIFI_GFP_NLS_F	tggaggtggaAGCGGTAGCAAAGGAGAAG
HIFI_GFP_NLS_R	cgttgaagtggcgtgacgcggatccTTACACCTTGCCTTCTTC
scfvGFPnls_fwd	taactttcgaccggattttgcacctcgagATGGGCCCCGACATCGTG
scfvGFPnls_rev	atcgagcgttgaagtggcgtgacgcTTACACCTTGCCTTCTTCTTGG

Primer Assembly TwiPE TwiPr:

TwiPe_TwiPr_fwd	gccagatccaggtcgagcggcccaacataatttatcattactctgaatgcacttttc	
TwiPe_TwiPr_rev	cttggcaataagtaccgtaggatccccgcgttttgctggatc	

Primer Assembly tag-RFPt-NLS:

Frag1 F -	ttcgaccggattttgcacctcgagATGGCTTCTAACTTTACTCAGTTTCG
Frag1 R -	taccaccaccCTTGACAGCTCGTCCATG
Frag2 F -	gctgtacaagGGTGGTGGTAGCGGTGGTG

Frag2 R - cgttgaagtggcgtgacgcggatccTTACACCTTGCGCTTCTTCTGG

Primer PCR tag-RFPt:

RFP_F 5'	gctagc ATG GTGTCTAAGGGCGAAGA
RFP_R 5'	ggatcc TTA CTTGACAGCTCGTCCATGC

Primer Assembly scFv-mScarlet-NLS:

scFv_fwd	ttcgaccggattttcgcctcgagATGGGCCCCGACATCGTG
scFv_rev	ccttgctcacACCGTTCCACCTCCACC
mScarlet_fwd	tggaagcggGTGAGCAAGGGCGAGGCA
mScarlet_rev	gacctccaccCTTGACAGCTCGTCCATGCC
GB1nls_fwd	gctgtacaagGGTGGAGGTCGGACCGAAG
GB1nls_rev	acacaacttatcgccgagttaacTTGGCTTTTCTGCTATACGTGTC TTTG

Probes smFISH sequences:

<i>suntag</i> probes	
gagagaagttcttctctga	SunTag_32x_Q570_1
tcgagctacttcatttctcta	SunTag_32x_Q570_2
gataagagctcttcacctga	SunTag_32x_Q570_3
tcgctacctcattttcaaga	SunTag_32x_Q570_4
gagagcaattcttcaccaga	SunTag_32x_Q570_5
cttcgttctcgaggatag	SunTag_32x_Q570_6
agacagaagttcttccac	SunTag_32x_Q570_7
tcgttctctagtggaatt	SunTag_32x_Q570_8
ttggagagcaattcttcacc	SunTag_32x_Q570_9
gccacttctttccaaatg	SunTag_32x_Q570_10
ccagatgatagtcttgctg	SunTag_32x_Q570_11
ttttttagtctcgctactt	SunTag_32x_Q570_12
ttggagagcaattcttcgc	SunTag_32x_Q570_13
tactcagtaactcttcca	SunTag_32x_Q570_14
cttgcgacctcattttctaa	SunTag_32x_Q570_15
ggtagttcttagacaggagt	SunTag_32x_Q570_16
cactccctttttaagtctg	SunTag_32x_Q570_17
tcgagatggtagttcttgga	SunTag_32x_Q570_18
agaacctttttcagcctag	SunTag_32x_Q570_19
ggtagtttttagagagcagc	SunTag_32x_Q570_20
cgttctccagatggaattc	SunTag_32x_Q570_21
aatagttcttcaccggaacc	SunTag_32x_Q570_22
cctcattctcgagatggtaa	SunTag_32x_Q570_23
ttggagagcaactcttctc	SunTag_32x_Q570_24
ttcttaagtcgcgcaacttc	SunTag_32x_Q570_25
tttgctcaacagttcttctc	SunTag_32x_Q570_26

acctcattttcgagatggta	SunTag_32x_Q570_27
cacttcattctcgagatggg	SunTag_32x_Q570_28
gtgcaacttcattctcgaga	SunTag_32x_Q570_29
gcaacttcgttttccaagt	SunTag_32x_Q570_30
gttctctaggtgatagttct	SunTag_32x_Q570_31
agtaattcttcaccactacc	SunTag_32x_Q570_32
tttactcaatagctcctctc	SunTag_32x_Q570_33
ggccacttcattctctaaat	SunTag_32x_Q570_34
tgacaacaactcttctccgg	SunTag_32x_Q570_35
atcctgaacctttcttagt	SunTag_32x_Q570_36
ttctttgagaggagctcttc	SunTag_32x_Q570_37
ttagagagcaattcttcgcc	SunTag_32x_Q570_38
ccactaccttttttaatc	SunTag_32x_Q570_39
tttttgagagtagttcctc	SunTag_32x_Q570_40
aaagtaattctcaccgctt	SunTag_32x_Q570_41
caacttcgttttcgaggtga	SunTag_32x_Q570_42
agataatagctcttcaccgg	SunTag_32x_Q570_43
cgctacccttttttaatcta	SunTag_32x_Q570_44
gtggtagttcttacttagca	SunTag_32x_Q570_45
ttccagatccctttttaag	SunTag_32x_Q570_46
aacttcgttctcgaggtgat	SunTag_32x_Q570_47
atcttgctacttcattctct	SunTag_32x_Q570_48
<i>twist probes</i>	
acgcacttacggaacgcaac	twist_mRNA_1
cgttttgctggatcgatctg	twist_mRNA_2
agatattcttcacgcatgg	twist_mRNA_3
tttcggaactgtataactcc	twist_mRNA_4
cttcagttagaatttgggt	twist_mRNA_5
gatcttgctggcttttag	twist_mRNA_6
gagcgagcgctcatcattg	twist_mRNA_7
ttgtagcttatgtccagcag	twist_mRNA_8
catgatgttgggcagtggtg	twist_mRNA_9
cttgatcacattgttctgca	twist_mRNA_10
ccagttcggaataaccatag	twist_mRNA_11
tgggtatgtttagtctgtg	twist_mRNA_12
gtgctcgttccaagaggatc	twist_mRNA_13
cggtttggcaatcgtatcg	twist_mRNA_14
gattagcttctgtgctcgg	twist_mRNA_15
gtgggcaagtaatccacata	twist_mRNA_16
caggacaagattgagccgag	twist_mRNA_17
tcctcgtcgggaaaatcgaa	twist_mRNA_18
ttgtagaggggcaggaacac	twist_mRNA_19

cgtttgaagtgcatctgg	twist_mRNA_20
ccctcгааactttgtctgta	twist_mRNA_21
atccgagctggagtaagcac	twist_mRNA_22
taactgaactcagggcggtg	twist_mRNA_23
atgactcctccgttgagatc	twist_mRNA_24
atcggatccgtccagcaaac	twist_mRNA_25
gacttgaaggcgtcgttgag	twist_mRNA_26
tctggatcttgctgagcttg	twist_mRNA_27
tatcttgaggccagtttgag	twist_mRNA_28
catgcggcacaggaagtcaa	twist_mRNA_29
caaggccttcagaaagata	twist_mRNA_30
gataatgggtgctccgttgg	twist_mRNA_31
cagataactcagcttctcgg	twist_mRNA_32
ggactatatctagagtgtcc	twist_mRNA_33
acgatagttgctgctggaag	twist_mRNA_34
tgaagagtгagagcсcagg	twist_mRNA_35
tatggaaagcgtgacagaga	twist_mRNA_36
gaatcatctagtcgacacgt	twist_mRNA_37
ggtttttaggcttagacatc	twist_mRNA_38
ggcttgttggctttttaca	twist_mRNA_39
ttgttcatggcgtttctttg	twist_mRNA_40
<i>snail probes</i>	
tctcaacgagagctgaggtg	sna_mRNA_1
gagtatagagcgggtggttc	sna_mRNA_2
tgggtaaатcgggagatcgg	sna_mRNA_3
agttttagttccggtactg	sna_mRNA_4
ccattttgatgtgtgtgtg	sna_mRNA_5
ttagcgggcagctttttag	sna_mRNA_6
ctcctccacgaagacaatgg	sna_mRNA_7
caaactgtgagtccttggtc	sna_mRNA_8
cgtttcagggatagatcctg	sna_mRNA_9
tgctgataatcctgggtctc	sna_mRNA_10
acatagtcacgtttcggttc	sna_mRNA_11
ccggtgttttgaaagggttc	sna_mRNA_12
agttggagctagagctggag	sna_mRNA_13
tagtcacgcatatggattt	sna_mRNA_14
gattaatcgtggtgggggtg	sna_mRNA_15
atcacaaggcggactggaa	sna_mRNA_16
cagagatcggattgcaaccg	sna_mRNA_17
atctgctgtagctgtagac	sna_mRNA_18
aaccggtttcagatcggat	sna_mRNA_19

actgaaagatcctctggctc	sna_mRNA_20
cggcagtgggatgtcatttc	sna_mRNA_21
gcctcatcgaaaaggtggaa	sna_mRNA_22
tgtaggagtatcccgatgag	sna_mRNA_23
catgattggcggcaacactc	sna_mRNA_24
gcacttgaagcggtagtttt	sna_mRNA_25
atcgaggtggagtacatctt	sna_mRNA_26
aactgacgggtgcttgacag	sna_mRNA_27
ttcttctcctgattacactc	sna_mRNA_28
aatgggtggtgtacagctttc	sna_mRNA_29
gtgcggatgtgcatcttcag	sna_mRNA_30
caaatggggcacttgcaggg	sna_mRNA_31
agggtcgagagaaggccttg	sna_mRNA_32
aaaggcttctctccagtgtg	sna_mRNA_33
caaaggatcgtgggcagtcg	sna_mRNA_34
gatgagctcgcaggttcgag	sna_mRNA_35
tacttcttgacgtccacgtg	sna_mRNA_36
gaaagatttgggcacacct	sna_mRNA_37
tgcttggtcaggagcgacat	sna_mRNA_38
tagtgatggtgcagttggag	sna_mRNA_39
atatgtcgagaatcctacgc	sna_mRNA_40
taattgtgtcctgctaaggg	sna_mRNA_41
gcggaatgtgagtttgctta	sna_mRNA_42

attgtctgtttgttggtct	sna_mRNA_43
gcaccaaaccgaatcgact	sna_mRNA_44
atgctgcgtgtgacaatgag	sna_mRNA_45
acagttggcttaacagtact	sna_mRNA_46
ttcttctcttaagctagga	sna_mRNA_47

References and Notes

1. K. Becker, A. Bluhm, N. Casas-Vila, N. Dinges, M. Dejung, S. Sayols, C. Kreutz, J.-Y. Roignant, F. Butter, S. Legewie, Quantifying post-transcriptional regulation in the development of *Drosophila melanogaster*. *Nat. Commun.* **9**, 4970 (2018). [doi:10.1038/s41467-018-07455-9](https://doi.org/10.1038/s41467-018-07455-9) [Medline](#)
2. E. Lécuyer, H. Yoshida, N. Parthasarathy, C. Alm, T. Babak, T. Cerovina, T. R. Hughes, P. Tomancak, H. M. Krause, Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174–187 (2007). [doi:10.1016/j.cell.2007.08.003](https://doi.org/10.1016/j.cell.2007.08.003) [Medline](#)
3. A. Chin, E. Lécuyer, RNA localization: Making its way to the center stage. *Biochim. Biophys. Acta Gen. Subj.* **1861**, 2956–2970 (2017). [doi:10.1016/j.bbagen.2017.06.011](https://doi.org/10.1016/j.bbagen.2017.06.011) [Medline](#)
4. E. Bertrand, P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, R. M. Long, Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998). [doi:10.1016/S1097-2765\(00\)80143-4](https://doi.org/10.1016/S1097-2765(00)80143-4) [Medline](#)
5. X. Yan, T. A. Hoek, R. D. Vale, M. E. Tanenbaum, Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* **165**, 976–989 (2016). [doi:10.1016/j.cell.2016.04.034](https://doi.org/10.1016/j.cell.2016.04.034) [Medline](#)
6. T. Morisaki, K. Lyon, K. F. DeLuca, J. G. DeLuca, B. P. English, Z. Zhang, L. D. Lavis, J. B. Grimm, S. Viswanathan, L. L. Looger, T. Lionnet, T. J. Stasevich, Real-time quantification of single RNA translation dynamics in living cells. *Science* **352**, 1425–1429 (2016). [doi:10.1126/science.aaf0899](https://doi.org/10.1126/science.aaf0899) [Medline](#)
7. X. Pichon, A. Bastide, A. Safieddine, R. Chouaib, A. Samacoits, E. Basyuk, M. Peter, F. Mueller, E. Bertrand, Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *J. Cell Biol.* **214**, 769–781 (2016). [doi:10.1083/jcb.201605024](https://doi.org/10.1083/jcb.201605024) [Medline](#)
8. B. Wu, C. Eliscovich, Y. J. Yoon, R. H. Singer, Translation dynamics of single mRNAs in live cells and neurons. *Science* **352**, 1430–1435 (2016). [doi:10.1126/science.aaf1084](https://doi.org/10.1126/science.aaf1084) [Medline](#)
9. C. Wang, B. Han, R. Zhou, X. Zhuang, Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell* **165**, 990–1001 (2016). [doi:10.1016/j.cell.2016.04.040](https://doi.org/10.1016/j.cell.2016.04.040) [Medline](#)
10. M. E. Tanenbaum, L. A. Gilbert, L. S. Qi, J. S. Weissman, R. D. Vale, A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014). [doi:10.1016/j.cell.2014.09.039](https://doi.org/10.1016/j.cell.2014.09.039) [Medline](#)
11. T. Sandmann, C. Girardot, M. Brehme, W. Tongprasit, V. Stolc, E. E. M. Furlong, A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* **21**, 436–449 (2007). [doi:10.1101/gad.1509007](https://doi.org/10.1101/gad.1509007) [Medline](#)
12. K. Tantale, F. Mueller, A. Kozulic-Pirher, A. Lesne, J.-M. Victor, M.-C. Robert, S. Capozzi, R. Chouaib, V. Bäcker, J. Mateos-Langerak, X. Darzacq, C. Zimmer, E. Basyuk, E. Bertrand, A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* **7**, 12248 (2016). [doi:10.1038/ncomms12248](https://doi.org/10.1038/ncomms12248) [Medline](#)

13. J. E. Sandler, A. Stathopoulos, Quantitative Single-Embryo Profile of *Drosophila* Genome Activation and the Dorsal-Ventral Patterning Network. *Genetics* **202**, 1575–1584 (2016). [doi:10.1534/genetics.116.186783](https://doi.org/10.1534/genetics.116.186783) [Medline](#)
14. S. W. Eichhorn, A. O. Subtelny, I. Kronja, J. C. Kwasniewski, T. L. Orr-Weaver, D. P. Bartel, mRNA poly(A)-tail changes specified by deadenylation broadly reshape translation in *Drosophila* oocytes and early embryos. *eLife* **5**, e16955 (2016). [doi:10.7554/eLife.16955](https://doi.org/10.7554/eLife.16955) [Medline](#)
15. L. Chang, Y. Shav-Tal, T. Treck, R. H. Singer, R. D. Goldman, Assembling an intermediate filament network by dynamic cotranslation. *J. Cell Biol.* **172**, 747–758 (2006). [doi:10.1083/jcb.200511033](https://doi.org/10.1083/jcb.200511033) [Medline](#)
16. R. Chouaib, A. Safieddine, X. Pichon, A. Imbert, O. S. Kwon, A. Samacoits, A.-M. Traboulsi, M.-C. Robert, N. Tsanov, E. Coleno, I. Poser, C. Zimmer, A. Hyman, H. Le Hir, K. Zibara, M. Peter, F. Mueller, T. Walter, E. Bertrand, A Dual Protein-mRNA Localization Screen Reveals Compartmentalized Translation and Widespread Co-translational RNA Targeting. *Dev. Cell* **54**, 773–791.e5 (2020). [doi:10.1016/j.devcel.2020.07.010](https://doi.org/10.1016/j.devcel.2020.07.010) [Medline](#)
17. I. Davis, D. Ish-Horowicz, Apical localization of pair-rule transcripts requires 3' sequences and limits protein diffusion in the *Drosophila* blastoderm embryo. *Cell* **67**, 927–940 (1991). [doi:10.1016/0092-8674\(91\)90366-7](https://doi.org/10.1016/0092-8674(91)90366-7) [Medline](#)
18. S. L. Bullock, M. Stauber, A. Prell, J. R. Hughes, D. Ish-Horowicz, U. Schmidt-Ott, Differential cytoplasmic mRNA localisation adjusts pair-rule transcription factor activity to cytoarchitecture in dipteran evolution. *Development* **131**, 4251–4261 (2004). [doi:10.1242/dev.01289](https://doi.org/10.1242/dev.01289) [Medline](#)
19. I. Castanon, S. Von Stetina, J. Kass, M. K. Baylies, Dimerization partners determine the activity of the Twist bHLH protein during *Drosophila* mesoderm development. *Development* **128**, 3145–3159 (2001). [Medline](#)
20. J. Dufourt, A. Trullo, J. Hunter, C. Fernandez, J. Lazaro, M. Dejean, L. Morales, S. Nait-Amer, K. N. Schulz, M. M. Harrison, C. Favard, O. Radulescu, M. Lagha, Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun.* **9**, 5194 (2018). [doi:10.1038/s41467-018-07613-z](https://doi.org/10.1038/s41467-018-07613-z) [Medline](#)
21. T. Ferraro, E. Esposito, L. Mancini, S. Ng, T. Lucas, M. Coppey, N. Dostatni, A. M. Walczak, M. Levine, M. Lagha, Transcriptional Memory in the *Drosophila* Embryo. *Curr. Biol.* **26**, 212–218 (2016). [doi:10.1016/j.cub.2015.11.058](https://doi.org/10.1016/j.cub.2015.11.058) [Medline](#)
22. K. J. Venken, Y. He, R. A. Hoskins, H. J. Bellen, P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**, 1747–1751 (2006). [doi:10.1126/science.1134426](https://doi.org/10.1126/science.1134426) [Medline](#)
23. S. C. Little, M. Tikhonov, T. Gregor, Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789–800 (2013). [doi:10.1016/j.cell.2013.07.025](https://doi.org/10.1016/j.cell.2013.07.025) [Medline](#)
24. H. J. Motulsky, R. E. Brown, Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics* **7**, 123 (2006). [doi:10.1186/1471-2105-7-123](https://doi.org/10.1186/1471-2105-7-123) [Medline](#)

25. B. Rappaz, P. W. Wiseman, Image Correlation Spectroscopy for Measurements of Particle Densities and Colocalization. *Curr. Protoc. Cell Biol.* **59**, 4.27.1–4.27.15 (2013). [doi:10.1002/0471143030.cb0427s59](https://doi.org/10.1002/0471143030.cb0427s59)
26. N. O. Petersen, Scanning fluorescence correlation spectroscopy. I. Theory and simulation of aggregation measurements. *Biophys. J.* **49**, 809–815 (1986). [doi:10.1016/S0006-3495\(86\)83709-2](https://doi.org/10.1016/S0006-3495(86)83709-2) [Medline](#)
27. U. Krzic, S. Gunther, T. E. Saunders, S. J. Streichan, L. Hufnagel, Multiview light-sheet microscope for rapid in toto imaging. *Nat. Methods* **9**, 730–733 (2012). [doi:10.1038/nmeth.2064](https://doi.org/10.1038/nmeth.2064) [Medline](#)
28. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). [doi:10.1038/nmeth.2019](https://doi.org/10.1038/nmeth.2019) [Medline](#)
29. F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, J.-C. Olivo-Marin, Icy: An open bioimage informatics platform for extended reproducible research. *Nat. Methods* **9**, 690–696 (2012). [doi:10.1038/nmeth.2075](https://doi.org/10.1038/nmeth.2075) [Medline](#)
30. P. Müller, P. Schwille, T. Weidemann, PyCorrFit-generic data evaluation for fluorescence correlation spectroscopy. *Bioinformatics* **30**, 2532–2533 (2014). [doi:10.1093/bioinformatics/btu328](https://doi.org/10.1093/bioinformatics/btu328) [Medline](#)
31. J. Widengren, Ü. Mets, R. Rigler, Photodynamic properties of green fluorescent proteins investigated by fluorescence correlation spectroscopy. *Chem. Phys.* **250**, 171–186 (1999). [doi:10.1016/S0301-0104\(99\)00255-4](https://doi.org/10.1016/S0301-0104(99)00255-4)
32. Z. Liu, W. R. Legant, B.-C. Chen, L. Li, J. B. Grimm, L. D. Lavis, E. Betzig, R. Tjian, 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *eLife* **3**, e04236 (2014). [doi:10.7554/eLife.04236](https://doi.org/10.7554/eLife.04236) [Medline](#)
33. A. Sergé, N. Bertaux, H. Rigneault, D. Marguet, Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nat. Methods* **5**, 687–694 (2008). [doi:10.1038/nmeth.1233](https://doi.org/10.1038/nmeth.1233) [Medline](#)
34. D. Normanno, L. Boudarène, C. Dugast-Darzacq, J. Chen, C. Richter, F. Proux, O. Bénichou, R. Voituriez, X. Darzacq, M. Dahan, Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat. Commun.* **6**, 7357 (2015). [doi:10.1038/ncomms8357](https://doi.org/10.1038/ncomms8357) [Medline](#)
35. N. Tarantino, J.-Y. Tinevez, E. F. Crowell, B. Boisson, R. Henriques, M. Mhlanga, F. Agou, A. Israël, E. Laplantine, TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO-IKK supramolecular structures. *J. Cell Biol.* **204**, 231–245 (2014). [doi:10.1083/jcb.201307172](https://doi.org/10.1083/jcb.201307172) [Medline](#)
36. H. Qian, M. P. Sheetz, E. L. Elson, Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophys. J.* **60**, 910–921 (1991). [doi:10.1016/S0006-3495\(91\)82125-7](https://doi.org/10.1016/S0006-3495(91)82125-7) [Medline](#)
37. A. Stathopoulos, M. Van Drenth, A. Erives, M. Markstein, M. Levine, Whole-genome analysis of dorsal-ventral patterning in the *Drosophila* embryo. *Cell* **111**, 687–701 (2002). [doi:10.1016/S0092-8674\(02\)01087-5](https://doi.org/10.1016/S0092-8674(02)01087-5) [Medline](#)

Chapter 7. Discussion

New technological advances such as the SunTag method is a real breakthrough for our understanding of gene regulation. By implementing this technique in the *Drosophila* embryo, we are now able to monitor translation kinetics in a multicellular organism. I will discuss here which kind of questions can be addressed with this method combined with imaging transcription in *Drosophila* embryos but also try to elaborate what could be the applications in other systems.

The impact of mRNA localization on gene expression patterning

As described in the introduction, mRNA localization can be determinant for gene regulation. In the case of *twi*, we observed that its mRNA is located basally in the cell whereas *sna* is located apically. At this stage of *Drosophila* embryogenesis, cells are not completely cellularized as the membrane elongates during time in an apico-basal directionality. In 1991, I. Davis and D. Ish-Horowicz (Davis and Ish-Horowicz 1991) proposed that apico-basal partitioning of mRNA can constrain protein diffusion once translated. Interestingly, Twi protein pattern is distributed in gradient at the border of the pattern, contrary to Snail which has a very sharp border, well delimited by one row of cells. The hypothesis of these two researchers can now be revisited and addressed by looking at mRNA translation dynamics of these two genes and follow protein diffusion with live imaging using tags such as ALFA-tag (Götzke et al. 2019) or LlamaTag (Bothma et al. 2018). Ultimately, changing the localization of those mRNAs and see the consequences on translation and development of the embryo (mesoderm formation, gastrulation) would shed light on the importance of regulation of translation in space. To change localization, one could change 3'UTR sequences as these regulatory sequences are usually involved in mRNA localization (see introduction). This will give insights on how mRNA localization influences protein distribution and what is the impact on development.

The role of regulatory sequences on translation dynamics

Upstream Open Reading Frame (uORF) are regulatory sequences (with a start codon) present in 5'UTR of many eukaryotic mRNAs. They can play roles on regulating the level of translation of the main ORF by entering in competition and repress translation on the main ORF (reviewed in Orr et al. 2020), leading to a fine-tuning of gene expression. Those studies,

including in *Drosophila* embryos for the case of *msl-2* mRNA (Medenbach et al. 2011), has mostly been done on a bulk of cells. How uORF influences the precise dynamics of translation have been investigated in cultured cells (Boersma et al. 2019) but we can imagine that depending of the cell and the environmental context those uORF will be differentially used and this will create different protein output. As an example, *twist* has a uORF in its 5'UTR whereas *snail* don't. This difference might have a developmental significance in terms of translation efficiency in the early embryo.

Additionally, it would be interesting to test the effect of adding or removing uORF of development genes and/or testing the use of uORF under stress conditions.

Transcription-translation relationship

Another question that we can address with the type of SunTag and MS2/MCP technique is about the link between transcription and translation regulation.

In *Drosophila*, endogenous histone locus contains about one hundred gene repeats on each allele (McKay et al. 2015). Surprisingly, it has been shown that a transgene containing only twelve histone genes can rescue the entire deletion (Hur et al. 2020), even if its expression is way reduced at the transcription site and by FISH (Figure 4 of Hur et al. 2020). This suggests that a post-transcriptional mechanism might be at play to ensure a good level of histone protein expression or that only a low level of histone is necessary for early development. Is there a 'sensing' mechanism at play to adapt the protein output regarding the mRNA amount? Looking quantitatively at translation dynamics of histone locus would be a good approach to try to answer this question.

Some researchers proposed the concept of genetic compensation, more specifically transcriptional adaptation, in mouse and zebrafish (Rossi et al. 2015). This corresponds to the discovery that deleterious mutations of some genes (for instance the *egfl7* gene) lead to transcriptional output increase of genes that are similar in sequences. They have next shown that this genetic adaptation depends on mRNA degradation factors (El-Brolosy et al. 2019). However, this very recent concept is not fully understood in terms of mechanisms and on the temporal and spatial aspects. Combining with the fact that mRNA degradation proteins such as Ccr4-Not have been shown to shuttle between the nucleus and cytoplasm of the cells

(Collart 2016; Kerr et al. 2011), there might be a link between cytoplasmic mRNA, therefore its translation regulation, and transcription.

How long does it take for the cell to transcriptionally adapt? Are these mutated mRNAs transported to specific sub-cellular compartments? What are the consequences in terms of translation?

Tagging a mutated gene such as *egfl7* with MS2 and suntag would be very interesting in wild type condition as well as in mutant mRNA degradation.

Translation during regeneration

The capability of certain organs to regenerate is a field of particular interest as it could have many applications for treating several injuries or diseases. Recently, a study on regeneration of zebrafish scales, have shown that waves of Erk activity controls the growth of bones (De Simone et al. 2021). Interestingly, they have shown that this depends on translation using translation inhibitors (Extended Data Fig. 6b, c. in De Simone et al. 2021). Knowing that Erk controls translation is therefore not a surprise (Monick et al. 2006), but nothing is known about the dynamic in space and time of translation in that case of regeneration. Do targets of Erk translation occur in waves as well? Is Erk itself translated in waves?

Chapter 8. General conclusion

During my PhD I was fascinated by the fact that the development of an embryo can be so reproducible and robust. I learned from literature that this is possible only if gene expression is tightly regulated in space but also in time. Indeed, how from the pure stochastic biochemical process of transcription, development occurs with such a precision is clearly amazing. This gives hints on the existence of many layers of gene regulation. Development of new technologies of mRNA labeling such as MS2/MCP and single-molecule FISH, has significantly participated to the understanding of the gene expression precision at the transcriptional level. However, quite much remain to be answered about the complex mechanisms regulating transcription in space and time at the single molecule level. Now, through the development of the bipartite Tag/scFv or nanobody (such as SunTag) systems, a new layer of gene regulation, translation, has become accessible. Thus, improvement of live imaging technologies as well as optical engineering are really a breakthrough for the understanding of mechanisms underlying precision of gene regulation.

The work I contributed during my PhD participates to a better understanding of mechanisms behind developmental robustness, such as mitotic memory at the transcriptional level but also the translation efficiency of a fate determinant transcription factor.

With the results obtained during my PhD as well as from the literature, I learned new biological concepts. Among them, I came to realize: **i)** the impact of spatial localization on biological processes. Indeed, we learned that mRNA localization is important for the translatability of a given mRNA, which might differ regarding its position within the embryonic cell. Furthermore, the distribution of a given protein is not homogenous as seen in the case of GAF (but also seen for Zelda and Bicoid) where it forms speckles inside the nucleus. This might have consequences in terms of transcription, but also on biological events such as mitosis. In the case of GAF, its localization within large nuclear speckles at the centromeres is much probably involved in the correct separation of the two alleles during mitosis.

This leads to a second concept: **ii)** the importance of the timing of gene expression. Indeed, the existence of a transcriptional memory raises the question of how mRNA translation is regulated during the cell cycle. Indeed, if transcription can be a process with memory, translation could also be tightly regulated and compensate some lack of transcription occurring at a certain time point or in some cells. A study revealed that, in human cells,

translation is highly changing during mitosis although mRNA abundance does not change drastically (Tanenbaum et al. 2015). Live imaging approaches would be determinant to discover a precise timed control of translation during mitosis but also to investigate how translation evolved upon mitosis exit.

iii) the concept of community effect refers to the interaction among a group of nearby precursor cells, being necessary for them to maintain tissue-specific gene expression and differentiate in a coordinated manner. This is especially important during embryonic development. The study of the expression of genes in space and time takes on its full meaning in this context. But what dictates the community effect? How transcription, translation and/or signaling regulation impact this effect? I believe for the future that space and time controlled perturbations such as optogenetics will be key understand the gene regulation circuits, and particularly the timing of response of the system regarding a stimulus.

Finally, I was amazed by the reproducibility of development in terms of gene expression, and thus from a unique cell. Developmental genes such as *snail* and *twist* are expressed in a very defined pattern and position early *Drosophila* embryos as an ideal system to study which factors could perturb these pattern formations. I believe that it is important to understand the fundamental biology of embryogenesis in order to be able to understand and potentially cure diseases such as fetal malformations.

References

- Acar M, Becskei A, van Oudenaarden A. 2005. Enhancement of cellular memory by reducing stochastic transitions. *Nature* **435**: 228–232.
- Adkins NL, Hagerman TA, Georgel P. 2006. GAGA protein: A multi-faceted transcription factor. *Biochem Cell Biol* **84**: 559–567.
- Alexander JM, Guan J, Li B, Maliskova L, Song M, Shen Y, Huang B, Lomvardas S, Weiner OD. 2019. Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. *Elife* **8**.
- Allahyar A, Vermeulen C, Bouwman BAM, Krijger PHL, Verstegen MJAM, Geeven G, van Kranenburg M, Pieterse M, Straver R, Haarhuis JHI, et al. 2018. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat Genet* **50**: 1151–1160.
- Auer JMT, Stoddart JJ, Christodoulou I, Lima A, Skouloudaki K, Hall HN, Vukojević V, Papadopoulos DK. 2021. Of numbers and movement – Understanding transcription factor pathogenesis by advanced microscopy. *DMM Dis Model Mech* **13**.
- Axelrod D, Koppel DE, Schlessinger J, Elson E, Webb WW. 1976. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys J* **16**: 1055–1069. [http://dx.doi.org/10.1016/S0006-3495\(76\)85755-4](http://dx.doi.org/10.1016/S0006-3495(76)85755-4).
- Bagley JA, Yan Z, Zhang W, Wildonger J, Jan LY, Jan YN. 2014. Double-bromo and extraterminal (BET) domain proteins regulate dendrite morphology and mechanosensory function. *Genes Dev* **28**: 1940–1956.
- Balaskas N, Ribeiro A, Panovska J, Dessaud E, Sasai N, Page KM, Briscoe J, Ribes V. 2012. Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell* **148**: 273–284. <http://dx.doi.org/10.1016/j.cell.2011.10.047>.
- Bannister AJ, Schneider R, Kouzarides T. 2002. Histone methylation: dynamic or static? *Cell* **109**: 801–806.
- Bantignies F, Cavalli G. 2014. Topological Organization of Drosophila Hox Genes Using DNA Fluorescent In Situ Hybridization. *Methods Mol Biol* **1196**: 103–120.
- Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, Cavalli G. 2011. Polycomb-Dependent Regulatory Contacts between Distant Hox Loci

- in *Drosophila*. *Cell* **144**: 214–226.
<https://www.sciencedirect.com/science/article/pii/S0092867410014856>.
- Behera V, Stonestrom AJ, Hamagami N, Hsiung CC, Keller CA, Giardine B, Sidoli S, Yuan ZF, Bhanu N V., Werner MT, et al. 2019. Interrogating Histone Acetylation and BRD4 as Mitotic Bookmarks of Transcription. *Cell Rep* **27**: 400–415.e5.
<https://doi.org/10.1016/j.celrep.2019.03.057>.
- Bellec M, Radulescu O, Lagha M. 2018. Remembering the past: Mitotic bookmarking in a developing embryo. *Curr Opin Syst Biol* **11**: 41–49.
<https://doi.org/10.1016/j.coisb.2018.08.003>.
- Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, Grimes GR, Therizols P, Bickmore WA. 2019. Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Mol Cell* **76**: 473–484.e7.
- Bentovim L, Harden TT, DePace AH. 2017. Transcriptional precision and accuracy in development: From measurements to models and mechanisms. *Dev* **144**: 3855–3866.
- Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. 1998. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* **2**: 437–445.
- Bhat KM, Farkas G, Karch F, Gyurkovics H, Gausz J, Schedl P. 1996. The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division. *Development* **122**: 1113–1124.
- Bheda P, Aguilar-Gómez D, Becker NB, Becker J, Stavrou E, Kukhtevich I, Höfer T, Maerkl S, Charvin G, Marr C, et al. 2020. Single-Cell Tracing Dissects Regulation of Maintenance and Inheritance of Transcriptional Reinduction Memory. *Mol Cell* **78**: 915–925.e7.
- Bickmore WA. 2013. The spatial organization of the human genome. *Annu Rev Genomics Hum Genet* **14**: 67–84.
- Black KL, Petruk S, Fenstermaker TK, Hodgson JW, Caplan JL, Brock HW, Mazo A. 2016. Chromatin proteins and RNA are associated with DNA during all phases of mitosis. *Cell Discov* **2**. <http://dx.doi.org/10.1038/celldisc.2016.38>.
- Blair SS. 2008. Segmentation in animals. *Curr Biol* **18**: R991–5.
- Blythe SA, Wieschaus EF. 2016. Establishment and maintenance of heritable chromatin structure during early *drosophila* embryogenesis. *Elife* **5**: 1–21.
- Boersma S, Khuperkar D, Verhagen BMP, Sonneveld S, Grimm JB, Lavis LD, Tanenbaum

- ME. 2019. Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. *Cell* **178**: 458-472.e19.
- Boettiger AN, Levine M. 2009. Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science* **325**: 471-3.
- Bonifer C, Cockerill PN. 2017. Chromatin priming of genes in development: Concepts, mechanisms and consequences. *Exp Hematol* **49**: 1-8.
- Borun TW, Pearson D, Paik WK. 1972. Studies of Histone Methylation during the HeLa S-3 Cell Cycle. *J Biol Chem* **247**: 4288-4298.
<https://www.sciencedirect.com/science/article/pii/S0021925819450746>.
- Bothma JP, Norstad MR, Alamos S, Garcia HG. 2018. LlamaTags: A Versatile Tool to Image Transcription Factor Dynamics in Live Embryos. *Cell* **173**: 1810-1822.e16.
- Brickner DG, Cajigas I, Fondufe-Mittendorf Y, Ahmed S, Lee P-C, Widom J, Brickner JH. 2007. H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. *PLoS Biol* **5**: e81.
- Browne GJ, Proud CG. 2002. Regulation of peptide-chain elongation in mammalian cells. *Eur J Biochem* **269**: 5360-5368.
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**: 1213-1218.
<https://doi.org/10.1038/nmeth.2688>.
- Buxbaum AR, Haimovich G, Singer RH. 2015. In the right place at the right time: visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* **16**: 95-109.
- Byvoet P, Shepherd GR, Hardin JM, Noland BJ. 1972. The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch Biochem Biophys* **148**: 558-567.
- Callegari A, Sieben C, Benke A, Suter DM, Fierz B, Mazza D, Manley S. 2019. Single-molecule dynamics and genome-wide transcriptomics reveal that NF- κ B (p65)-DNA binding times can be decoupled from transcriptional activation. *PLoS Genet* **15**: e1007891.
- Cañestro C, Yokoi H, Postlethwait JH. 2007. Evolutionary developmental biology and genomics. *Nat Rev Genet* **8**: 932-942. <https://doi.org/10.1038/nrg2226>.

- Caravaca JM, Donahue G, Becker JS, He X, Vinson C, Zaret KS. 2013. Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev* **27**: 251–260.
- Cardozo Gizzi AM, Cattoni DI, Fiche J-B, Espinola SM, Gurgo J, Messina O, Houbron C, Ogiyama Y, Papadopoulos GL, Cavalli G, et al. 2019. Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. *Mol Cell* **74**: 212–222.e5.
- Caster SZ, Castillo K, Sachs MS, Bell-Pedersen D. 2016. Circadian clock regulation of mRNA translation through eukaryotic elongation factor eEF-2. *Proc Natl Acad Sci* **113**: 9605 LP – 9610. <http://www.pnas.org/content/113/34/9605.abstract>.
- Ceci M, Gaviraghi C, Gorrini C, Sala LA, Offenhäuser N, Marchisio PC, Biffo S. 2003. Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**: 579–584.
- Chan CS, Rastelli L, Pirrotta V. 1994. A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *EMBO J* **13**: 2553–2564.
- Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MGM, Qi J, Rahl PB, Sun HH, Yeda KT, Doench JG, et al. 2013. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* **24**: 777–790.
- Chen E, Sharma MR, Shi X, Agrawal RK, Joseph S. 2014a. Fragile X mental retardation protein regulates translation by binding directly to the ribosome. *Mol Cell* **54**: 407–417.
- Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, Gregor T. 2018. Dynamic interplay between enhancer-promoter topology and gene activity. *Nat Genet* **50**: 1296–1303.
- Chen J, Zhang Z, Li L, Chen B-C, Revyakin A, Hajj B, Legant W, Dahan M, Lionnet T, Betzig E, et al. 2014b. Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**: 1274–1285.
- Chen K, Johnston J, Shao W, Meier S, Staber C, Zeitlinger J. 2013. A global change in RNA polymerase II pausing during the *Drosophila* midblastula transition ed. R. Lehmann. *Elife* **2**: e00861. <https://doi.org/10.7554/eLife.00861>.
- Chetverina D, Erokhin M, Schedl P. 2021. GAGA factor: a multifunctional pioneering chromatin protein. *Cell Mol Life Sci* **78**: 4125–4141.
- Chiang C-M. 2009. Brd4 engagement from chromatin targeting to transcriptional

regulation: selective contact with acetylated histone H3 and H4. *F1000 Biol Rep* **1**: 98.

Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, Heckert A, Banala S, Lavis L, Darzacq X, et al. 2018. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**.

Chong S, Graham TGW, Dugast-Darzacq C, Dailey GM, Darzacq X, Tjian R. 2021. Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *bioRxiv* 2021.08.16.456551.

Chubb JR. 2016. Gene Regulation: Stable Noise. *Curr Biol* **26**: R61–R64.
<http://dx.doi.org/10.1016/j.cub.2015.12.002>.

Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell* **9**: 279–89. <http://www.ncbi.nlm.nih.gov/pubmed/11864602>.

Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, Clark KL, Burley SK, Zaret KS. 1998. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J* **17**: 244–254.

Cirillo LA, Zaret KS. 1999. An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. *Mol Cell* **4**: 961–969.

Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, Dugast-Darzacq C, Hajj B, Dahan M, Darzacq X. 2013. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* **341**: 664–667.

Clapier CR, Iwasa J, Cairns BR, Peterson CL. 2017. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**: 407–422.

Clauß K, Popp AP, Schulze L, Hettich J, Reisser M, Escoter Torres L, Uhlenhaut NH, Gebhardt JCM. 2017. DNA residence time is a regulatory factor of transcription repression. *Nucleic Acids Res* **45**: 11121–11130. <https://doi.org/10.1093/nar/gkx728>.

Coleman RT, Struhl G. 2017. Causal role for inheritance of H3K27me3 in maintaining the OFF state of a *Drosophila* HOX gene. *Science* **356**.

Collart MA. 2016. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip Rev RNA* **7**: 438–454.

Cramer P. 2019. Organization and regulation of gene transcription. *Nature* **573**: 45–54.

Dar RD, Razooky BS, Singh A, Trimeloni T V, McCollum JM, Cox CD, Simpson ML,

- Weinberger LS. 2012. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc Natl Acad Sci U S A* **109**: 17454–17459.
- Das S, Vera M, Gandin V, Singer RH, Tutucci E. 2021. Intracellular mRNA transport and localized translation. *Nat Rev Mol Cell Biol* **22**: 483–504.
<http://dx.doi.org/10.1038/s41580-021-00356-8>.
- Davidson EH. 1990. How embryos work: a comparative view of diverse modes of cell fate specification. *Development* **108**: 365–389.
- Davis I, Ish-Horowicz D. 1991. Apical localization of pair-rule transcripts requires 3' sequences and limits protein diffusion in the Drosophila blastoderm embryo. *Cell* **67**: 927–940. <https://www.sciencedirect.com/science/article/pii/0092867491903667>.
- de Castro IJ, Gokhan E, Vagnarelli P. 2016. Resetting a functional G1 nucleus after mitosis. *Chromosoma* **125**: 607–619.
- De Simone A, Evanitsky MN, Hayden L, Cox BD, Wang J, Tornini VA, Ou J, Chao A, Poss KD, Di Talia S. 2021. Control of osteoblast regeneration by a train of Erk activity waves. *Nature* **590**: 129–133.
- Deluz C, Friman ET, Strebinger D, Benke A, Raccaud M, Callegari A, Leleu M, Manley S, Suter DM. 2016. A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes Dev* **30**: 2538–2550.
- Deneke VE, Melbinger A, Vergassola M, Di Talia S. 2016. Waves of Cdk1 Activity in S Phase Synchronize the Cell Cycle in Drosophila Embryos. *Dev Cell* **38**: 399–412.
- Dertinger T, Loman A, Ewers B, Müller CB, Krämer B, Enderlein J. 2008. The optics and performance of dual-focus fluorescence correlation spectroscopy. *Opt Express* **16**: 14353.
- Devaiah BN, Case-Borden C, Gegonne A, Hsu CH, Chen Q, Meerzaman D, Dey A, Ozato K, Singer DS. 2016a. BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin. *Nat Struct Mol Biol* **23**: 540–548.
- Devaiah BN, Gegonne A, Singer DS. 2016b. Bromodomain 4: a cellular Swiss army knife. *J Leukoc Biol* **100**: 679–686. <http://www.jleukbio.org/cgi/doi/10.1189/jlb.2R10616-250R>.
- Dey A, Chitsaz F, Abbasi A, Misteli T, Ozato K. 2003. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A* **100**: 8758–8763.

- Dey A, Nishiyama A, Karpova T, McNally J, Ozato K. 2009. Brd4 Marks Select Genes on Mitotic Chromatin and Directs Postmitotic Transcription. *Mol Biol Cell* **20**: 4899–4909. <http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E09-05-0380>.
- Digan ME, Haynes SR, Mozer BA, Dawid IB, Forquignon F, Gans M. 1986. Genetic and molecular analysis of *fs(1)h*, a maternal effect homeotic gene in *Drosophila*. *Dev Biol* **114**: 161–169. <https://www.sciencedirect.com/science/article/pii/0012160686903921>.
- Ding Y, Fromm M, Avramova Z. 2012. Multiple exposures to drought “train” transcriptional responses in *Arabidopsis*. *Nat Commun* **3**: 740. <https://doi.org/10.1038/ncomms1732>.
- Ding Y, Liu N, Virilouvet L, Riethoven J-J, Fromm M, Avramova Z. 2013. Four distinct types of dehydration stress memory genes in *Arabidopsis thaliana*. *BMC Plant Biol* **13**: 229.
- Donovan BT, Huynh A, Ball DA, Patel HP, Poirier MG, Larson DR, Ferguson ML, Lenstra TL. 2019. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *EMBO J* **38**.
- Dorigi KM, Tamkun JW. 2013. The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in *Drosophila*. *Development* **140**: 4182–4192.
- Driever W, Siegel V, Nusslein-Volhard C. 1990. Autonomous determination of anterior structures in the early *Drosophila* embryo by the bicoid morphogen. *Development* **109**: 811–820. <https://doi.org/10.1242/dev.109.4.811>.
- Dubarry N, Pasta F, Lane D. 2006. ParABS systems of the four replicons of *Burkholderia cenocepacia*: new chromosome centromeres confer partition specificity. *J Bacteriol* **188**: 1489–1496.
- Dufourt J, Bellec M, Trullo A, Dejean M, De Rossi S, Favard C, Lagha M. 2021. Imaging translation dynamics in live embryos reveals spatial heterogeneities. *Science (80-)* **372**: 840–844.
- Dufourt J, Dennis C, Boivin A, Gueguen N, Théron E, Goriaux C, Pouchin P, Ronsseray S, Brassat E, Vaury C. 2014. Spatio-temporal requirements for transposable element piRNA-mediated silencing during *Drosophila* oogenesis. *Nucleic Acids Res* **42**: 2512–2524.
- Dufourt J, Trullo A, Hunter J, Fernandez C, Lazaro J, Dejean M, Morales L, Nait-Amer S, Schulz KN, Harrison MM, et al. 2018. Temporal control of gene expression by the

- pioneer factor Zelda through transient interactions in hubs. *Nat Commun* **9**: 1–13.
<http://dx.doi.org/10.1038/s41467-018-07613-z>.
- El-Brolosy MA, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, Kikhi K, Boezio GLM, Takacs CM, Lai S-L, et al. 2019. Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**: 193–197.
- Elf J, Li G-W, Xie XS. 2007. Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell. *Science (80-)* **316**: 1191 LP – 1194.
<http://science.sciencemag.org/content/316/5828/1191.abstract>.
- Elsherbiny A, Dobрева G. 2021. Epigenetic memory of cell fate commitment. *Curr Opin Cell Biol* **69**: 80–87. <https://doi.org/10.1016/j.ceb.2020.12.014>.
- ENCODE Project Consortium. 2011. A user’s guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol* **9**: e1001046.
- Engel KL, Arora A, Goering R, Lo H-YG, Taliaferro JM. 2020. Mechanisms and consequences of subcellular RNA localization across diverse cell types. *Traffic* **21**: 404–418.
- Escobar TM, Oksuz O, Saldaña-Meyer R, Descostes N, Bonasio R, Reinberg D. 2019. Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. *Cell* **179**: 953–963.e11.
- Escoffre JM, Hubert M, Teissié J, Rols MP, Favard C. 2014. Evidence for electro-induced membrane defects assessed by lateral mobility measurement of a gpi anchored protein. *Eur Biophys J* **43**: 277–286.
- Espinás ML, Jiménez-García E, Vaquero A, Canudas S, Bernués J, Azorín F. 1999. The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. *J Biol Chem* **274**: 16461–16469.
- Esposito E, Lim B, Guessous G, Falahati H, Levine M. 2016. Mitosis-associated repression in development. *Genes Dev* **30**: 1503–1508.
- Fabre B, Korona D, Groen A, Vowinckel J, Gatto L, Deery MJ, Ralser M, Russell S, Lilley KS. 2016. Analysis of *Drosophila melanogaster* proteome dynamics during embryonic development by a combination of label-free proteomics approaches. *Proteomics* **16**: 2068–2080.
- Falo-Sanjuan J, Bray SJ. 2021. Membrane architecture and adherens junctions contribute to strong Notch pathway activation. *bioRxiv* 2021.05.26.445755.

- <http://biorxiv.org/content/early/2021/05/26/2021.05.26.445755.abstract>.
- Farrell JA, O'Farrell PH. 2014a. From egg to gastrula: how the cell cycle is remodeled during the *Drosophila* mid-blastula transition. *Annu Rev Genet* **48**: 269–294.
- Farrell JA, O'Farrell PH. 2014b. From egg to gastrula: How the cell cycle is remodeled during the *drosophila* mid-blastula transition. *Annu Rev Genet* **48**: 269–294.
- Fasulo B, Koyama C, Yu KR, Homola EM, Hsieh TS, Campbell SD, Sullivan W. 2012. Chk1 and Wee1 kinases coordinate DNA replication, chromosome condensation, and anaphase entry. *Mol Biol Cell* **23**: 1047–1057.
- Fedorova E V, Dorogova N V, Bolobolova EU, Fedorova SA, Karagodin DA, Ogienko AA, Khruscheva AS, Baricheva EM. 2019. GAGA protein is required for multiple aspects of *Drosophila* oogenesis and female fertility. *Genesis* **57**: e23269.
- Fenouil R, Descostes N, Spinelli L, Kocj F, Maqbool M, Benoukraf T, Cauchy P, Innocenti C, Ferrier P, Andrau J. 2016. Pasha: a versatile R package for piling chromatin HTS data. *Bioinformatics* **15**: 2528–30.
- Ferraro T, Esposito E, Mancini L, Ng S, Lucas T, Coppey M, Dostatni N, Walczak AM, Levine M, Lagha M. 2016. Transcriptional Memory in the *Drosophila* Embryo. *Curr Biol* **26**: 212–218. <http://dx.doi.org/10.1016/j.cub.2015.11.058>.
- Festuccia N, Dubois A, Vandormael-Pournin S, Gallego Tejada E, Mouren A, Bessonard S, Mueller F, Proux C, Cohen-Tannoudji M, Navarro P. 2016. Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat Cell Biol* **18**: 1139–1148.
- Festuccia N, Gonzalez I, Owens N, Navarro P. 2017. Mitotic bookmarking in development and stem cells. *Development* **144**: 3633–3645.
<http://dev.biologists.org/lookup/doi/10.1242/dev.146522>.
- Festuccia N, Owens N, Papadopoulou T, Gonzalez I, Tachtsidi A, Vandoermel-Pournin S, Gallego E, Gutierrez N, Dubois A, Cohen-Tannoudji M, et al. 2019. Transcription factor activity and nucleosome organization in mitosis. *Genome Res* **29**: 250–260.
- Filipczyk A, Marr C, Hastreiter S, Feigelman J, Schwarzfischer M, Hoppe PS, Loeffler D, Kokkaliaris KD, Endeke M, Schauburger B, et al. 2015. Network plasticity of pluripotency transcription factors in embryonic stem cells. *Nat Cell Biol* **17**: 1235–1246.
- Florence BL, Faller D V. 2008. *Drosophila* female sterile (1) homeotic is a multifunctional transcriptional regulator that is modulated by Ras signaling. *Dev Dyn* **237**: 554–564.

- Foe VE, Alberts BM. 1983. Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. *J Cell Sci* **61**: 31–70.
- Foe VE, Odell GM, Edgar BA. 1993. Mitosis and morphogenesis in the *Drosophila* embryo: Point and counterpoint. In: The Development of *Drosophila melanogaster*. *Cold Spring Harb Lab Press* 149–300.
- Follmer NE, Wani AH, Francis NJ. 2012. A Polycomb Group Protein Is Retained at Specific Sites on Chromatin in Mitosis. *PLoS Genet* **8**.
- Formicola N, Heim M, Dufourt J, Lancelot A-S, Nakamura A, Lagha M, Besse F. 2021. Tyramine induces dynamic RNP granule remodeling and translation activation in the *Drosophila* brain. *Elife* **10**.
- Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**: 490–493. <https://doi.org/10.1038/nature09158>.
- Fu Z, Indrisiunaite G, Kaledhonkar S, Shah B, Sun M, Chen B, Grassucci RA, Ehrenberg M, Frank J. 2019. The structural basis for release-factor activation during translation termination revealed by time-resolved cryogenic electron microscopy. *Nat Commun* **10**: 2579.
- Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A, Lis JT. 2015. GAGA Factor Maintains Nucleosome-Free Regions and Has a Role in RNA Polymerase II Recruitment to Promoters. *PLoS Genet* **11**: 1–22.
- Garavís M, Méndez-Lago M, Gabelica V, Whitehead SL, González C, Villasante A. 2015. The structure of an endogenous *Drosophila* centromere reveals the prevalence of tandemly repeated sequences able to form i-motifs. *Sci Rep* **5**: 1–10.
- García-Muse T, Aguilera A. 2016. Transcription–replication conflicts: how they occur and how they are resolved. *Nat Rev Mol Cell Biol* **17**: 553–563. <https://doi.org/10.1038/nrm.2016.88>.
- Garvie CW, Wolberger C. 2001. Recognition of specific DNA sequences. *Mol Cell* **8**: 937–946.
- Gaskill MM, Gibson TJ., Larson ED., Harrison MM. 2021. GAF is essential for zygotic genome activation and chromatin accessibility in the early *Drosophila* embryo. *Elife* **0–43**.
- Gavis ER, Lehmann R. 1992. Localization of nanos RNA controls embryonic polarity. *Cell* **71**:

301–313.

- Gebhardt JCM, Suter DM, Roy R, Zhao ZW, Chapman AR, Basu S, Maniatis T, Xie XS. 2013. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat Methods* **10**: 421–426.
- Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EEM. 2014. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**: 96–100.
- Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, Shalek AK, Kelley DR, Shishkin AA, Issner R, et al. 2013. Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**: 1149–1163.
- Gilbert SF. 2000a. The Developmental Mechanics of Cell Specification. In *Developmental Biology*. 6th edition. Sunderland (MA).
- Gilbert SF. 2000b. The Origins of Anterior-Posterior Polarity. In *Developmental Biology*. 6th edition. Sunderland (MA).
- Gomez-Lamarca MJ, Falo-Sanjuan J, Stojnic R, Abdul Rehman S, Muresan L, Jones ML, Pillidge Z, Cerda-Moya G, Yuan Z, Baloul S, et al. 2018. Activation of the Notch Signaling Pathway In Vivo Elicits Changes in CSL Nuclear Dynamics. *Dev Cell* **44**: 611–623.e7.
- Götzke H, Kilisch M, Martínez-Carranza M, Sograte-Idrissi S, Rajavel A, Schlichthaerle T, Engels N, Jungmann R, Stenmark P, Opazo F, et al. 2019. The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nat Commun* **10**: 4403. <https://doi.org/10.1038/s41467-019-12301-7>.
- Granok H, Leibovitch BA, Shaffer CD, Elgin SC. 1995. Chromatin. Ga-ga over GAGA factor. *Curr Biol* **5**: 238–241.
- Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E, Blobel GA. 2007. Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol Cell Biol* **27**: 8466–8479.
- Gualdi R, Bossard P, Zheng M, Hamada Y, Coleman JR, Zaret KS. 1996. Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. *Genes Dev* **10**: 1670–1682.
- Gupta K, Sari-Ak D, Haffke M, Trowitzsch S, Berger I. 2016. Zooming in on Transcription

- Preinitiation. *J Mol Biol* **428**: 2581–2591.
- Gurdon JB, Bourillot PY. 2001. Morphogen gradient interpretation. *Nature* **413**: 797–803.
- Halstead JM, Lionnet T, Wilbertz JH, Wippich F, Ephrussi A, Singer RH, Chao JA. 2015. An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science (80-)* **347**: 1367 LP – 1671.
<http://science.sciencemag.org/content/347/6228/1367.abstract>.
- Hamperl S, Bocek MJ, Saldivar JC, Swigut T, Cimprich KA. 2017. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* **170**: 774-786.e19.
- Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* **6**.
- Harrison MM, Li XY, Kaplan T, Botchan MR, Eisen MB. 2011. Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS Genet* **7**.
- Harteis S, Schneider S. 2015. Making the Bend : DNA Tertiary Structure and Protein-DNA Interactions.
- Haynes SR, Mozer BA, Bhatia-Dey N, Dawid IB. 1989. The *Drosophila* fsh locus, a maternal effect homeotic gene, encodes apparent membrane proteins. *Dev Biol* **134**: 246–257.
<https://www.sciencedirect.com/science/article/pii/0012160689900948>.
- Hellen CU, Sarnow P. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**: 1593–1612.
- Hendrix DA, Hong JW, Zeitlinger J, Rokhsar DS, Levine MS. 2008. Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proc Natl Acad Sci U S A* **105**: 7762–7767.
- Henzel MJ, Wei Y, Mancini MA, Van Hooser A, Ranalli T, Brinkley BR, Bazett-Jones DP, Allis CD. 1997. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* **106**: 348–360.
- Hinde E, Pandžić E, Yang Z, Ng IHW, Jans DA, Bogoyevitch MA, Gratton E, Gaus K. 2016. Quantifying the dynamics of the oligomeric transcription factor STAT3 by pair correlation of molecular brightness. *Nat Commun* **7**: 11047.

- Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. 2017. A Phase Separation Model for Transcriptional Control. *Cell* **169**: 13–23.
<http://dx.doi.org/10.1016/j.cell.2017.02.007>.
- Holguera I, Desplan C. 2018. Neuronal specification in space and time. *Science* **362**: 176–180.
- Hong J-W, Hendrix DA, Levine MS. 2008. Shadow enhancers as a source of evolutionary novelty. *Science* **321**: 1314.
- Hoogenkamp M, Lichtinger M, Kryszyska H, Lancrin C, Clarke D, Williamson A, Mazzarella L, Ingram R, Jorgensen H, Fisher A, et al. 2009. Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. *Blood* **114**: 299–309.
- Hörmanseder E, Simeone A, Allen GE, Bradshaw CR, Figlmüller M, Gurdon J, Jullien J. 2017. H3K4 Methylation-Dependent Memory of Somatic Cell Identity Inhibits Reprogramming and Development of Nuclear Transfer Embryos. *Cell Stem Cell* **21**: 135-143.e6.
- Houston DW. 2013. Chapter Four - Regulation of Cell Polarity and RNA Localization in Vertebrate Oocytes. In *International Review of Cell and Molecular Biology* (ed. K.W. Jeon), Vol. 306 of *International Review of Cell and Molecular Biology*, pp. 127–185, Academic Press.
- Houzelstein D, Bullock SL, Lynch DE, Grigorieva EF, Wilson VA, Beddington RSP. 2002. Growth and Early Postimplantation Defects in Mice Deficient for the Bromodomain-Containing Protein Brd4. *Mol Cell Biol* **22**: 3794–3802.
- Hsiung CCS, Morrissey CS, Udugama M, Frank CL, Keller CA, Baek S, Giardine B, Crawford GE, Sung MH, Hardison RC, et al. 2015. Genome accessibility is widely preserved and locally modulated during mitosis. *Genome Res* **25**: 213–225.
- Huang A, Amourda C, Zhang S, Tolwinski NS, Saunders TE. 2017. Decoding temporal interpretation of the morphogen bicoid in the early drosophila embryo. *Elife* **6**.
- Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. 2017a. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169**: 216-228.e19. <http://dx.doi.org/10.1016/j.cell.2017.03.024>.
- Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. 2017b. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169**: 216-

228.e19.

- Hur W, Kemp JPJ, Tarzia M, Deneke VE, Marzluff WF, Duronio RJ, Di Talia S. 2020. CDK-Regulated Phase Separation Seeded by Histone Genes Ensures Precise Growth and Function of Histone Locus Bodies. *Dev Cell* **54**: 379-394.e6.
- Hüttelmaier S, Zenklusen D, Lederer M, Dichtenberg J, Lorenz M, Meng X, Bassell GJ, Condeelis J, Singer RH. 2005. Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* **438**: 512–515.
- Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218–223.
- Ip YT, Park RE, Kosman D, Yazdanbakhsh K, Levine M. 1992. dorsal-twist interactions establish snail expression in the presumptive mesoderm of the *Drosophila* embryo. *Genes Dev* **6**: 1518–1530.
- Iwafuchi-Doi M, Zaret KS. 2016. Cell fate control by pioneer transcription factors. *Development* **143**: 1833–1837. <http://dev.biologists.org/lookup/doi/10.1242/dev.133900>.
- Izeddin I, Récamier V, Bosanac L, Cissé II, Boudarene L, Dugast-Darzacq C, Proux F, Bénichou O, Voituriez R, Bensaude O, et al. 2014. Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *Elife* **3**.
- Jadhav U, Manieri E, Nalapareddy K, Madha S, Chakrabarti S, Wucherpennig K, Barefoot M, Shivdasani RA. 2020. Replicational Dilution of H₃K₂₇me₃ in Mammalian Cells and the Role of Poised Promoters. *Mol Cell* **78**: 141-151.e5.
- Javasky E, Shamir I, Gandhi S, Egri S, Sandler O, Rothbart SB, Kaplan N, Jaffe JD, Goren A, Simon I. 2018. Study of mitotic chromatin supports a model of bookmarking by histone modifications and reveals nucleosome deposition patterns. *Genome Res* **28**: 1455–1466.
- Judd J, Duarte FM, Lis JT. 2021. Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. *Genes Dev* **35**: 147–156.
- Kadauke S, Udugama MI, Pawlicki JM, Achtman JC, Jain DP, Cheng Y, Hardison RC, Blobel GA. 2012. Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell* **150**: 725–737. <http://dx.doi.org/10.1016/j.cell.2012.06.038>.
- Kamada R, Yang W, Zhang Y, Patel MC, Yang Y, Ouda R, Dey A, Wakabayashi Y, Sakaguchi

- K, Fujita T, et al. 2018. Interferon stimulation creates chromatin marks and establishes transcriptional memory. *Proc Natl Acad Sci* **115**: E9162 LP-E9171.
<http://www.pnas.org/content/115/39/E9162.abstract>.
- Kanodia JS, Rikhy R, Kim Y, Lund VK, DeLotto R, Lippincott-Schwartz J, Shvartsman SY. 2009. Dynamics of the Dorsal morphogen gradient. *Proc Natl Acad Sci* **106**: 21707 LP – 21712. <http://www.pnas.org/content/106/51/21707.abstract>.
- Kasinath BS, Mariappan MM, Sataranatarajan K, Lee MJ, Feliers D. 2006. mRNA Translation: Unexplored Territory in Renal Science. *J Am Soc Nephrol* **17**: 3281 LP – 3292. <http://jasn.asnjournals.org/content/17/12/3281.abstract>.
- Kellner WA, Van Bortle K, Li L, Ramos E, Takenaka N, Corces VG. 2013. Distinct isoforms of the *Drosophila* Brd4 homologue are present at enhancers, promoters and insulator sites. *Nucleic Acids Res* **41**: 9274–9283.
- Kellum R, Raff JW, Alberts BM. 1995. Heterochromatin protein 1 distribution during development and during the cell cycle in *Drosophila* embryos. *J Cell Sci* **108** (Pt 4): 1407–1418.
- Kent S, Brown K, Yang C, Alsaihati N, Tian C, Wang H, Ren X. 2020. Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. *Cell Rep* **33**: 108248.
<https://www.sciencedirect.com/science/article/pii/S2211124720312377>.
- Kerr SC, Azzouz N, Fuchs SM, Collart MA, Strahl BD, Corbett AH, Larabee RN. 2011. The Ccr4-Not complex interacts with the mRNA export machinery. *PLoS One* **6**: e18302.
- Kharchenko P V, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al. 2011. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* **471**: 480–485.
- Ko MS, Nakauchi H, Takahashi N. 1990. The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J* **9**: 2835–2842.
- Kockmann T, Gerstung M, Schlumpf T, Xhinzhou Z, Hess D, Beerenwinkel N, Beisel C, Paro R. 2013. The BET protein FSH functionally interacts with ASH1 to orchestrate global gene activity in *Drosophila*. *Genome Biol* **14**: R18.
- Koenecke N, Johnston J, Gaertner B, Natarajan M, Zeitlinger J. 2016a. Genome-wide

- identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol* **17**: 1–19. <http://dx.doi.org/10.1186/s13059-016-1057-2>.
- Koenecke N, Johnston J, Gaertner B, Natarajan M, Zeitlinger J. 2016b. Genome-wide identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol* **17**: 1–19.
- Koenecke N, Johnston J, He Q, Meier S, Zeitlinger J. 2017. *Drosophila* poised enhancers are generated during tissue patterning with the help of repression. *Genome Res* **27**: 64–74.
- Kogler AC, Kherdjemil Y, Bender K, Rabinowitz A, Marco-ferreres R, Furlong EEM. 2021. Extremely rapid and reversible optogenetic perturbation of nuclear proteins in living embryos. *Dev Cell* **2348–2363**.
- Kohwi M, Doe CQ. 2013. Temporal fate specification and neural progenitor competence during development. *Nat Rev Neurosci* **14**: 823–838.
- Komar AA, Hatzoglou M. 2005. Internal ribosome entry sites in cellular mRNAs: mystery of their existence. *J Biol Chem* **280**: 23425–23428.
- Kouskouti A, Talianidis I. 2005. Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *EMBO J* **24**: 347–357.
- Kouzarides T. 2007. Chromatin modifications and their function. *Cell* **128**: 693–705.
- Kozak M. 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15**: 1109–1123.
- Krebs JE, Fry CJ, Samuels ML, Peterson CL. 2000. Global Role for Chromatin Remodeling Enzymes in Mitotic Gene Expression. **102**: 587–598.
- Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, Dickson BJ, Stark A. 2014. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature* **512**: 91–95.
- Kwon SY, Jang B, Badenhorst P. 2021. The ISWI chromatin remodelling factor NURF is not required for mitotic male X chromosome organisation. *microPublication Biol* **2–7**.
- Lagha M, Bothma JP, Esposito E, Ng S, Stefanik L, Tsui C, Johnston J, Chen K, Gilmour DS, Zeitlinger J, et al. 2013. XPaused Pol II coordinates tissue morphogenesis in the *drosophila* embryo. *Cell* **153**: 976–987. <http://dx.doi.org/10.1016/j.cell.2013.04.045>.
- Lagha M, Bothma JP, Levine M. 2012. Mechanisms of transcriptional precision in animal development. *Trends Genet* **28**: 409–416. <http://dx.doi.org/10.1016/j.tig.2012.03.006>.

- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Laprell F, Finkl K, Müller J. 2017. Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science (80-)* **356**: 85 LP – 88.
<http://science.sciencemag.org/content/356/6333/85.abstract>.
- Larson DR, Fritzschn C, Sun L, Meng X, Lawrence DS, Singer RH. 2013. Direct observation of frequency modulated transcription in single cells using light activation. *Elife* **2**: e00750.
- Latchman DS. 1993. Transcription factors: an overview. *Int J Exp Pathol* **74**: 417–422.
- Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**: 174–187.
- Li J, Gilmour DS. 2013. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J* **32**: 1829–1841.
<http://dx.doi.org/10.1038/emboj.2013.111>.
- Li XY, Harrison MM, Villalta JE, Kaplan T, Eisen MB. 2014. Establishment of regions of genomic activity during the Drosophila maternal to zygotic transition. *Elife* **3**: 1–20.
- Liang HL, Nien CY, Liu HY, Metzstein MM, Kirov N, Rushlow C. 2008. The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* **456**: 400–403.
- Light WH, Freaney J, Sood V, Thompson A, D’Urso A, Horvath CM, Brickner JH. 2013. A conserved role for human Nup98 in altering chromatin structure and promoting epigenetic transcriptional memory. *PLoS Biol* **11**: e1001524.
- Liu Y, Pelham-Webb B, Di Giammartino DC, Li J, Kim D, Kita K, Saiz N, Garg V, Doane A, Giannakakou P, et al. 2017. Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells. *Cell Rep* **19**: 1283–1293.
- Liu Y, Zhao N, Kanemaki MT, Yamamoto Y, Sadamura Y, Ito Y, Tokunaga M, Stasevich TJ, Kimura H. 2021. Visualizing looping of two endogenous genomic loci using synthetic zinc-finger proteins with anti-FLAG and anti-HA frankenbodies in living cells. *Genes Cells*.
- Liu Z, Legant WR, Chen B-C, Li L, Grimm JB, Lavis LD, Betzig E, Tjian R. 2014. 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *Elife* **3**: e04236.

- Liu Z, Tjian R. 2018. Visualizing transcription factor dynamics in living cells. *J Cell Biol* **217**: 1181–1191.
- Loffreda A, Jacchetti E, Antunes S, Rainone P, Daniele T, Morisaki T, Bianchi ME, Tacchetti C, Mazza D. 2017. Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat Commun* **8**: 313. <https://doi.org/10.1038/s41467-017-00398-7>.
- Lomaev D, Mikhailova A, Erokhin M, Shaposhnikov A V., Moresco JJ, Blokhina T, Wolle D, Aoki T, Ryabykh V, Yates JR, et al. 2017. The GAGA factor regulatory network: Identification of GAGA factor associated proteins. *PLoS One* **12**: 1–20.
- Los G V, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, et al. 2008. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **3**: 373–382.
- Lott SE, Villalta JE, Schroth GP, Luo S, Tonkin LA, Eisen MB. 2011. Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-Seq. *PLoS Biol* **9**.
- Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA. 2013. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**: 320–334.
- Lyon K, Stasevich TJ. 2017. Imaging Translational and Post-Translational Gene Regulatory Dynamics in Living Cells with Antibody-Based Probes. *Trends Genet* **33**: 322–335.
- Mahmoudi T, Katsani KR, Verrijzer CP. 2002. GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J* **21**: 1775–1781.
- Martin KC, Ephrussi A. 2009. mRNA localization: gene expression in the spatial dimension. *Cell* **136**: 719–730.
- Martínez-Balbás MA, Dey A, Rabindran SK, Ozato K, Wu C. 1995. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**: 29–38. <https://www.sciencedirect.com/science/article/pii/0092867495902317>.
- Matharu NK, Yadav S, Kumar M, Mishra RK. 2021. Role of vertebrate GAGA associated factor (vGAF) in early development of zebrafish. *Cells Dev* **166**: 203682. <https://doi.org/10.1016/j.cdev.2021.203682>.
- Mazzocca M, Colombo E, Callegari A, Mazza D. 2021. Transcription factor binding kinetics

- and transcriptional bursting: What do we really know? *Curr Opin Struct Biol* **71**: 239–248. <https://www.sciencedirect.com/science/article/pii/S0959440X21001226>.
- McDaniel SL, Gibson TJ, Schulz KN, Fernandez Garcia M, Nevil M, Jain SU, Lewis PW, Zaret KS, Harrison MM. 2019. Continued Activity of the Pioneer Factor Zelda Is Required to Drive Zygotic Genome Activation. *Mol Cell* **74**: 185-195.e4.
- McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ. 1984. A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell* **37**: 403–408.
- McGinty RK, Tan S. 2015. Nucleosome structure and function. *Chem Rev* **115**: 2255–2273.
- McKay DJ, Klusza S, Penke TJR, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW, Tatomer DC, Lieb JD, et al. 2015. Interrogating the function of metazoan histones using engineered gene clusters. *Dev Cell* **32**: 373–386.
- Medenbach J, Seiler M, Hentze MW. 2011. Translational control via protein-regulated upstream open reading frames. *Cell* **145**: 902–913.
- Michelman-Ribeiro A, Mazza D, Rosales T, Stasevich TJ, Boukari H, Rishi V, Vinson C, Knutson JR, McNally JG. 2009. Direct measurement of association and dissociation rates of DNA binding in live cells by fluorescence correlation spectroscopy. *Biophys J* **97**: 337–346.
- Michelotti EF, Sanford S, Levens D. 1997. Marking of active genes on mitotic chromosomes. *Nature* **388**: 895–899. <https://doi.org/10.1038/42282>.
- Mir M, Reimer A, Haines JE, Li XY, Stadler M, Garcia H, Eisen MB, Darzacq X. 2017. Dense bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev* **31**: 1784–1794.
- Mir M, Stadler MR, Ortiz SA, Hannon CE, Harrison MM, Darzacq X, Eisen MB. 2018. Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos. *Elife* **7**: 1–27.
- Monick MM, Powers LS, Gross TJ, Flaherty DM, Barrett CW, Hunninghake GW. 2006. Active ERK Contributes to Protein Translation by Preventing JNK-Dependent Inhibition of Protein Phosphatase 1. *J Immunol* **177**: 1636 LP – 1645. <http://www.jimmunol.org/content/177/3/1636.abstract>.
- Moshe A, Kaplan T. 2017. Genome-wide search for Zelda-like chromatin signatures

- identifies GAF as a pioneer factor in early fly development. *Epigenetics and Chromatin* **10**: 1–14.
- Mueller F, Wach P, McNally JG. 2008. Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophys J* **94**: 3323–3339.
- Müller P, Schwille P, Weidemann T. 2014. PyCorrFit-generic data evaluation for fluorescence correlation spectroscopy. *Bioinformatics* **30**: 2532–2533.
- Muramoto T, Müller I, Thomas G, Melvin A, Chubb JR. 2010. Methylation of H3K4 Is Required for Inheritance of Active Transcriptional States. *Curr Biol* **20**: 397–406.
- Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, Feng X, Ahmad K, Russell S, White RAH, et al. 2010. A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* **6**.
- Neumann CJ, Cohen SM. 1997. Long-range action of Wingless organizes the dorsal-ventral axis of the *Drosophila* wing. *Development* **124**: 871–880.
- Nien C-Y, Liang H-L, Butcher S, Sun Y, Fu S, Gocha T, Kirov N, Manak JR, Rushlow C. 2011. Temporal coordination of gene networks by Zelda in the early *Drosophila* embryo. *PLoS Genet* **7**: e1002339.
- Niopek D, Wehler P, Roensch J, Eils R, Di Ventura B. 2016. Optogenetic control of nuclear protein export. *Nat Commun* **7**: 10624.
- Noonan JP, McCallion AS. 2010. Genomics of Long-Range Regulatory Elements. *Annu Rev Genomics Hum Genet* **11**: 1–23. <https://doi.org/10.1146/annurev-genom-082509-141651>.
- Normanno D, Boudarène L, Dugast-Darzacq C, Chen J, Richter C, Proux F, Bénichou O, Voituriez R, Darzacq X, Dahan M. 2015. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat Commun* **6**: 7357.
- Ogiyama Y, Schuettengruber B, Papadopoulos GL, Chang JM, Cavalli G. 2018. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during *Drosophila* Development. *Mol Cell* **71**: 73–88.e5. <https://doi.org/10.1016/j.molcel.2018.05.032>.
- Ohta S, Bukowski-Wills J-C, Sanchez-Pulido L, de Lima Alves F, Wood L, Chen ZA, Platani M, Fischer L, Hudson DF, Ponting CP, et al. 2010. The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell* **142**:

- 810–821. <https://www.sciencedirect.com/science/article/pii/S0092867410008950>.
- Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D. 1998. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92**: 105–116.
- Orr MW, Mao Y, Storz G, Qian S-B. 2020. Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Res* **48**: 1029–1042. <https://doi.org/10.1093/nar/gkz734>.
- Oudelaar AM, Davies JOJ, Hanssen LLP, Telenius JM, Schwessinger R, Liu Y, Brown JM, Downes DJ, Chiariello AM, Bianco S, et al. 2018. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat Genet* **50**: 1744–1751.
- Pallier C, Scaffidi P, Chopineau-Proust S, Agresti A, Nordmann P, Bianchi ME, Marechal V. 2003. Association of Chromatin Proteins High Mobility Group Box (HMGB) 1 and HMGB2 with Mitotic Chromosomes. *Mol Biol Cell* **14**: 2372–2384.
- Palozola KC, Donahue G, Liu H, Grant GR, Becker JS, Cote A, Yu H, Raj A, Zaret KS. 2017. Mitotic transcription and waves of gene reactivation during mitotic exit. *Science (80-)* **358**: 119–122.
- Parsons GG, Spencer CA. 1997. Mitotic repression of RNA polymerase II transcription is accompanied by release of transcription elongation complexes. *Mol Cell Biol* **17**: 5791–5802.
- Peng L, Li E-M, Xu L-Y. 2020. From start to end: Phase separation and transcriptional regulation. *Biochim Biophys Acta Gene Regul Mech* **1863**: 194641.
- Pengelly AR, Copur Ö, Jäckle H, Herzig A, Müller J. 2013. A Histone Mutant Reproduces the Phenotype Caused by Loss of Histone-Modifying Factor Polycomb. *Science (80-)* **339**: 698 LP – 699. <http://science.sciencemag.org/content/339/6120/698.abstract>.
- Perry MW, Boettiger AN, Bothma JP, Levine M. 2010. Shadow enhancers foster robustness of drosophila gastrulation. *Curr Biol* **20**: 1562–1567. <http://dx.doi.org/10.1016/j.cub.2010.07.043>.
- Phillips NE, Mandic A, Omid S, Naef F, Suter DM. 2019. Memory and relatedness of transcriptional activity in mammalian cell lineages. *Nat Commun* **10**: 1208.
- Pichon X, Bastide A, Safieddine A, Chouaib R, Samacoits A, Basyuk E, Peter M, Mueller F, Bertrand E. 2016. Visualization of single endogenous polysomes reveals the dynamics

- of translation in live human cells. *J Cell Biol* **214**: 769–781.
- Pichon X, Lagha M, Mueller F, Bertrand E. 2018. A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges. *Mol Cell* **71**: 468–480. <https://doi.org/10.1016/j.molcel.2018.07.022>.
- Pockwinse SM, Kota KP, Quaresma AJC, Imbalzano AN, Lian JB, van Wijnen AJ, Stein JL, Stein GS, Nickerson JA. 2011. Live cell imaging of the cancer-related transcription factor RUNX2 during mitotic progression. *J Cell Physiol* **226**: 1383–1389.
- Popp AP, Hettich J, Gebhardt JCM. 2021. Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. *Nucleic Acids Res* **49**: 6249–6266. <https://doi.org/10.1093/nar/gkab443>.
- Porcher A, Abu-Arish A, Huart S, Roelens B, Fradin C, Dostatni N. 2010. The time to measure positional information: Maternal Hunchback is required for the synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. *Development* **137**: 2795–2804.
- Prescott DM, Bender MA. 1962. Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Exp Cell Res* **26**: 260–268. <https://www.sciencedirect.com/science/article/pii/0014482762901763>.
- Price D, Rabinovitch S, O'Farrell PH, Campbell SD. 2000. Drosophila wee1 has an essential role in the nuclear divisions of early embryogenesis. *Genetics* **155**: 159–166.
- Rabinowitz M. 1941. Studies on the cytology and early embryology of the egg of *Drosophila melanogaster*. *J Morphol* **69**: 1–49. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmor.1050690102>.
- Raccaud M, Friman ET, Alber AB, Agarwal H, Deluz C, Kuhn T, Gebhardt JCM, Suter DM. 2019. Mitotic chromosome binding predicts transcription factor properties in interphase. *Nat Commun* **10**: 1–16. <http://dx.doi.org/10.1038/s41467-019-08417-5>.
- Raccaud M, Suter DM. 2017. Transcription factor retention on mitotic chromosomes: regulatory mechanisms and impact on cell fate decisions. *FEBS Lett* **592**: 878–887. <http://doi.wiley.com/10.1002/1873-3468.12828>.
- Raff JW, Kellum R, Alberts B. 1994. The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle. *EMBO J* **13**: 5977–5983.

- Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**: W160–W165.
- Reeves GT, Trisnadi N, Truong T V, Nahmad M, Katz S, Stathopoulos A. 2012. Dorsal-Ventral Gene Expression in the Drosophila Embryo Reflects the Dynamics and Precision of the Dorsal Nuclear Gradient. *Dev Cell* **22**: 544–557. <https://www.sciencedirect.com/science/article/pii/S1534580711005739>.
- Reiling JH, Hafen E. 2004. The hypoxia-induced paralogs Scylla and Charybdis inhibit growth by down-regulating S6K activity upstream of TSC in Drosophila. *Genes Dev* **18**: 2879–2892.
- Rieder LE, Koreski KP, Boltz KA, Kuzu G, Urban JA, Bowman SK, Zeidman A, Jordan WT, Tolstorukov MY, Marzluff WF, et al. 2017. Histone locus regulation by the Drosophila dosage compensation adaptor protein CLAMP. *Genes Dev* **31**: 1494–1508.
- Roberts SGE. 2000. Mechanisms of action of transcription activation and repression domains. *Cell Mol Life Sci C* **57**: 1149–1160. <https://doi.org/10.1007/PL00000755>.
- Rongo C, Gavis ER, Lehmann R. 1995. Localization of oskar RNA regulates oskar translation and requires Oskar protein. *Development* **121**: 2737–2746.
- Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, Stainier DYR. 2015. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**: 230–233.
- Royou A, McCusker D, Kellogg DR, Sullivan W. 2008. Grapes(Chk1) prevents nuclear CDK1 activation by delaying cyclin B nuclear accumulation. *J Cell Biol* **183**: 63–75.
- Rushlow CA, Shvartsman SY. 2012. Temporal dynamics, spatial range, and transcriptional interpretation of the Dorsal morphogen gradient. *Curr Opin Genet Dev* **22**: 542–546.
- Sabari BR, Dall’Agnese A, Young RA. 2020. Biomolecular Condensates in the Nucleus. *Trends Biochem Sci* **45**: 961–977.
- Sabari BR, Sabari BR, Agnese AD, Boija A, Klein IA, Coffey EL, Abraham BJ, Hannett NM, Zamudio A V, Manteiga JC, et al. 2018. Coactivator condensation at super-enhancers links phase separation and gene control. **3958**: 1–17.
- Sagner A, Zhang I, Watson T, Lazaro J, Melchionda M, Briscoe J. 2020. Temporal patterning of the central nervous system by a shared transcription factor code. *bioRxiv*

- 2020.11.10.376491.
<http://biorxiv.org/content/early/2020/11/10/2020.11.10.376491.abstract>.
- Saksouk N, Simboeck E, Déjardin J. 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**: 3.
- Samata M, Alexiadis A, Richard G, Georgiev P, Nuebler J, Kulkarni T, Renschler G, Basilicata MF, Zenk FL, Shvedunova M, et al. 2020. Intergenerationally Maintained Histone H4 Lysine 16 Acetylation Is Instructive for Future Gene Activation. *Cell* **182**: 127-144.e23.
<http://dx.doi.org/10.1016/j.cell.2020.05.026>.
- Sani E, Herzyk P, Perrella G, Colot V, Amtmann A. 2013. Hyperosmotic priming of Arabidopsis seedlings establishes a long-term somatic memory accompanied by specific changes of the epigenome. *Genome Biol* **14**: R59.
- Sarnataro S, Riba A, Molina N. 2020. Regulation of transcription reactivation dynamics exiting mitosis. *bioRxiv* 2020.04.15.042853.
<http://biorxiv.org/content/early/2020/04/16/2020.04.15.042853.abstract>.
- Schep AN, Adryan B. 2013. A comparative analysis of transcription factor expression during metazoan embryonic development. *PLoS One* **8**: e66826.
- Schmähling S, Meiler A, Lee Y, Mohammed A, Finkl K, Tauscher K, Israel L, Wirth M, Philippou-Massier J, Blum H, et al. 2018. Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. *Development* **145**.
- Schneider I, Amemiya C. 2016. Developmental-Genetic Toolkit for Evolutionary Developmental Biology. In *Encyclopedia of Evolutionary Biology* (ed. R.M. Kliman), pp. 404–408, Academic Press, Oxford.
- Schneider N, Wieland F-G, Kong D, Fischer AAM, Hörner M, Timmer J, Ye H, Weber W. 2021. Liquid-liquid phase separation of light-inducible transcription factors increases transcription activation in mammalian cells and mice. *Sci Adv* **7**: eabd3568.
- Schneuwly S, Klemenz R, Gehring WJ. 1987. Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia. *Nature* **325**: 816–818.
- Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. 2007. Genome regulation by polycomb and trithorax proteins. *Cell* **128**: 735–745.
- Schulz KN, Bondra ER, Moshe A, Villalta JE, Lieb JD, Kaplan T, McKay DJ, Harrison MM. 2015a. Zelda is differentially required for chromatin accessibility, transcription factor

- binding, and gene expression in the early *Drosophila* embryo. *Genome Res* **25**: 1715–1726.
- Schulz KN, Bondra ER, Moshe A, Villalta JE, Lieb JD, Kaplan T, McKay DJ, Harrison MM, Hill C, Hill C, et al. 2015b. Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early *Drosophila* embryo. 1715–1726.
- Schulz KN, Harrison MM. 2019. Mechanisms regulating zygotic genome activation. *Nat Rev Genet* **20**: 221–234. <http://dx.doi.org/10.1038/s41576-018-0087-x>.
- Schwarzer W, Spitz F. 2014. The architecture of gene expression: integrating dispersed cis-regulatory modules into coherent regulatory domains. *Curr Opin Genet Dev* **27**: 74–82.
- Scuderi A, Simin K, Kazuko SG, Metherall JE, Letsou A. 2006. scylla and charybde, homologues of the human apoptotic gene RTP801, are required for head involution in *Drosophila*. *Dev Biol* **291**: 110–122. <https://www.sciencedirect.com/science/article/pii/S0012160605008912>.
- Sekiya T, Muthurajan UM, Luger K, Tulin A V, Zaret KS. 2009. Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes Dev* **23**: 804–809.
- Senecal A, Munsky B, Proux F, Ly N, Braye FE, Zimmer C, Mueller F, Darzacq X. 2014. Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep* **8**: 75–83.
- Sérandour AA, Avner S, Percevault F, Demay F, Bizot M, Lucchetti-Miganeh C, Barloy-Hubler F, Brown M, Lupien M, Métivier R, et al. 2011. Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res* **21**: 555–565.
- Shearn A. 1989. The ash-1, ash-2 and trithorax genes of *Drosophila melanogaster* are functionally related. *Genetics* **121**: 517–525.
- Shi Z, Barna M. 2015. Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins. *Annu Rev Cell Dev Biol* **31**: 31–54.
- Shim EY, Woodcock C, Zaret KS. 1998. Nucleosome positioning by the winged helix transcription factor HNF3. *Genes Dev* **12**: 5–10.
- Shirokikh NE, Dutikova YS, Staroverova MA, Hannan RD, Preiss T. 2019. Migration of Small Ribosomal Subunits on the 5' Untranslated Regions of Capped Messenger RNA. *Int J Mol Sci* **20**.

- Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Liron Y, Rosenfeld N, Danon T, Perzov N, Alon U. 2006. Variability and memory of protein levels in human cells. *Nature* **444**: 643–646.
- Soeller WC, Oh CE, Kornberg TB. 1993. Isolation of cDNAs encoding the Drosophila GAGA transcription factor. *Mol Cell Biol* **13**: 7961–7970.
- Solnica-Krezel L, Sepich DS. 2012. Gastrulation: Making and Shaping Germ Layers. *Annu Rev Cell Dev Biol* **28**: 687–717. <https://doi.org/10.1146/annurev-cellbio-092910-154043>.
- Spitz F, Furlong EEM. 2012. Transcription factors: From enhancer binding to developmental control. *Nat Rev Genet* **13**: 613–626.
- Spradling A. 1993. Developmental genetics of oogenesis. *Cold Spring Harb Lab Press New York* 1–70.
- Sprague BL, Pego RL, Stavreva DA, McNally JG. 2004. Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching. *Biophys J* **86**: 3473–3495. <https://www.sciencedirect.com/science/article/pii/S0006349504743921>.
- Srivastava A, Kumar AS, Mishra RK. 2018. Vertebrate GAF/ThPOK: emerging functions in chromatin architecture and transcriptional regulation. *Cell Mol Life Sci Vol* **75**: 623–633.
- Stavreva DA, Garcia DA, Fettweis G, Gudla PR, Zaki GF, Soni V, McGowan A, Williams G, Huynh A, Palangat M, et al. 2019. Transcriptional Bursting and Co-bursting Regulation by Steroid Hormone Release Pattern and Transcription Factor Mobility. *Mol Cell* **75**: 1161-1177.e11.
- Steffen PA, Altmutter C, Dworschak E, Junttila S, Gyenesi A, Zhu X, Kockmann T, Ringrose L. 2021. The Trithorax group protein ASH1 requires a combination of BAH domain and AT hooks, but not the SET domain, for mitotic chromatin binding and survival. *Chromosoma*. <https://doi.org/10.1007/s00412-021-00762-z>.
- Steffen PA, Fonseca JP, Gänger C, Dworschak E, Kockmann T, Beisel C, Ringrose L. 2013. Quantitative in vivo analysis of chromatin binding of Polycomb and Trithorax group proteins reveals retention of ASH1 on mitotic chromatin. *Nucleic Acids Res* **41**: 5235–5250.
- Steffen PA, Ringrose L. 2014. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat Rev Mol Cell Biol* **15**: 340–356.
- Strom AR, Emelyanov A V, Mir M, Fyodorov D V, Darzacq X, Karpen GH. 2017. Phase

- separation drives heterochromatin domain formation. *Nature* **547**: 241–245.
- Stukenberg PT, Kirschner MW, Kingston RE. 1998. Mitotic inactivation of a human SWI / SNF chromatin remodeling complex. 2842–2851.
- Stumpff J, Duncan T, Homola E, Campbell SD, Su TT. 2004. Drosophila Wee1 kinase regulates Cdk1 and mitotic entry during embryogenesis. *Curr Biol* **14**: 2143–2148.
- Sun Y, Nien CY, Chen K, Liu HY, Johnston J, Zeitlinger J, Rushlow C. 2015. Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. *Genome Res* **25**: 1703–1714.
- Susor A, Jansova D, Cerna R, Danylevska A, Anger M, Toralova T, Malik R, Supolikova J, Cook MS, Oh JS, et al. 2015. Temporal and spatial regulation of translation in the mammalian oocyte via the mTOR-eIF4F pathway. *Nat Commun* **6**: 6078.
- Swinstead EE, Miranda TB, Paakinaho V, Baek S, Goldstein I, Hawkins M, Karpova TS, Ball D, Mazza D, Lavis LD, et al. 2016. Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. *Cell* **165**: 593–605.
- Tadros W, Lipshitz HD. 2009. The maternal-to-zygotic transition: a play in two acts. *Development* **136**: 3033–3042.
- Tadros W, Westwood JT, Lipshitz HD. 2007. The Mother-to-Child Transition. *Dev Cell* **12**: 847–849.
- Tanaka Y, Katagiri Z-I, Kawahashi K, Kioussis D, Kitajima S. 2007. Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* **397**: 161–168.
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**: 635–646.
- Tanenbaum ME, Stern-Ginossar N, Weissman JS, Vale RD. 2015. Regulation of mRNA translation during mitosis. *Elife* **4**.
- Tariq M, Wegrzyn R, Anwar S, Bukau B, Paro R. 2013. Drosophila GAGA factor polyglutamine domains exhibit prion-like behavior. *BMC Genomics* **14**: 374.
- Teves SS, An L, Bhargava-Shah A, Xie L, Darzacq X, Tjian R. 2018. A stable mode of bookmarking by TBP recruits RNA polymerase II to mitotic chromosomes. *Elife* **7**: 1–22. <https://elifesciences.org/articles/35621>.
- Teves SS, An L, Hansen AS, Xie L, Darzacq X, Tjian R. 2016. A dynamic mode of mitotic

- bookmarking by transcription factors. *Elife* **5**.
- Thoma C, Bergamini G, Galy B, Hundsdoerfer P, Hentze MW. 2004. Enhancement of IRES-mediated translation of the c-myc and BiP mRNAs by the poly(A) tail is independent of intact eIF4G and PABP. *Mol Cell* **15**: 925–935.
- Thomsen S, Anders S, Janga SC, Huber W, Alonso CR. 2010. Genome-wide analysis of mRNA decay patterns during early Drosophila development. *Genome Biol* **11**: R93.
- Toyama R, Rebbert ML, Dey A, Ozato K, Dawid IB. 2008. Brd4 associates with mitotic chromosomes throughout early zebrafish embryogenesis. *Dev Dyn* **237**: 1636–1644.
- Treen N, Shimobayashi SF, Eeftens J, Brangwynne CP, Levine M. 2021. Properties of repression condensates in living Ciona embryos. *Nat Commun* **12**: 1561.
<https://doi.org/10.1038/s41467-021-21606-5>.
- Tripoulas NA, Hersperger E, La Jeunesse D, Shearn A. 1994. Molecular genetic analysis of the Drosophila melanogaster gene absent, small or homeotic discs1 (ash1). *Genetics* **137**: 1027–1038.
- Trojanowski J, Frank L, Rademacher A, Grigaitis P, Rippe K. 2021. Transcription activation is enhanced by multivalent interactions independent of liquid-liquid phase separation. *bioRxiv* 2021.01.27.428421.
- Trullo A, Dufourt J, Lagha M. 2020. MitoTrack, a user-friendly semi-automatic software for lineage tracking in living embryos. *Bioinformatics* **36**: 1300–1302.
- Tsai A, Muthusamy AK, Alves MR, Lavis LD, Singer RH, Stern DL, Crocker J. 2017. Nuclear microenvironments modulate transcription from low-affinity enhancers. *Elife* **6**.
- Tsai S-Y, Chang Y-L, Swamy KBS, Chiang R-L, Huang D-H. 2016. GAGA factor, a positive regulator of global gene expression, modulates transcriptional pausing and organization of upstream nucleosomes. *Epigenetics Chromatin* **9**: 32.
- Vallot A, Tachibana K. 2020. The emergence of genome architecture and zygotic genome activation. *Curr Opin Cell Biol* **64**: 50–57.
- Valls E, Sánchez-Molina S, Martínez-Balbás MA. 2005. Role of histone modifications in marking and activating genes through mitosis. *J Biol Chem* **280**: 42592–42600.
- Van Steensel B, Delrow J, Bussemaker HJ. 2003. Genomewide analysis of Drosophila GAGA factor target genes reveals context-dependent DNA binding. *Proc Natl Acad Sci U S A* **100**: 2580–2585.

- Vastenhouw NL, Cao WX, Lipshitz HD. 2019. The maternal-to-zygotic transition revisited. *Development* **146**.
- Venken KJT, He Y, Hoskins RA, Bellen HJ. 2006. P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science (80-)* **314**: 1747–1751.
- WADDINGTON CH. 1942. CANALIZATION OF DEVELOPMENT AND THE INHERITANCE OF ACQUIRED CHARACTERS. *Nature* **150**: 563–565. <https://doi.org/10.1038/150563ao>.
- Wang C, Han B, Zhou R, Zhuang X. 2016. Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell* **165**: 990–1001.
- Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, Chen C, Kou X, Zhao Y, Chen J, et al. 2018. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat Cell Biol* **20**: 620–631.
- Wang F, Higgins JMG. 2013. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol* **23**: 175–184.
- Wei M-T, Chang Y-C, Shimobayashi SF, Shin Y, Strom AR, Brangwynne CP. 2020. Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol* **22**: 1187–1196.
- Wieschaus E, Nüsslein-Volhard C. 2016. The Heidelberg Screen for Pattern Mutants of *Drosophila*: A Personal Account. *Annu Rev Cell Dev Biol* **32**: 1–46. <https://doi.org/10.1146/annurev-cellbio-113015-023138>.
- Wilk R, Hu J, Blotsky D, Krause HM. 2016. Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. *Genes Dev* **30**: 594–609.
- Wilkins RC, Lis JT. 1999. DNA distortion and multimerization: novel functions of the glutamine-rich domain of GAGA factor. *J Mol Biol* **285**: 515–525.
- Wojenski LA, Wainman LM, Villafano G, Kuhlberg C, Nanda P, Core L. 2021. Multiple modes of regulation control dynamic transcription patterns during the mitosis-G1 transition. *bioRxiv* 2021.06.22.449286. <http://biorxiv.org/content/early/2021/06/22/2021.06.22.449286.abstract>.
- Wolffe A, Wong J, Pruss D. 1997. Activators and repressors: making use of chromatin to regulate transcription. *Genes to Cells* **2**.
- Wu B, Eliscovich C, Yoon YJ, Singer RH. 2016. Translation dynamics of single mRNAs in live

- cells and neurons. *Science* **352**: 1430–1435.
- Wu T, Kamikawa YF, Donohoe ME. 2018. Brd4's Bromodomains Mediate Histone H3 Acetylation and Chromatin Remodeling in Pluripotent Cells through P300 and Brg1. *Cell Rep* **25**: 1756–1771.
- Xu J, Pope SD, Jazirehi AR, Attema JL, Papathanasiou P, Watts JA, Zaret KS, Weissman IL, Smale ST. 2007. Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. *Proc Natl Acad Sci* **104**: 12377 LP – 12382. <http://www.pnas.org/content/104/30/12377.abstract>.
- Yan X, Hoek TA, Vale RD, Tanenbaum ME. 2016. Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* **165**: 976–989.
- Yang Z, Yik JHN, Chen R, He N, Jang MK, Ozato K, Zhou Q. 2005. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19**: 535–545.
- Yguerabide J, Schmidt JA, Yguerabide EE. 1982. Lateral mobility in membranes as detected by fluorescence recovery after photobleaching. *Biophys J* **40**: 69–75. [http://dx.doi.org/10.1016/S0006-3495\(82\)84459-7](http://dx.doi.org/10.1016/S0006-3495(82)84459-7).
- Yuan K, O'Farrell PH. 2016. TALE-light imaging reveals maternally guided, H3K9me2/3-independent emergence of functional heterochromatin in Drosophila embryos. *Genes Dev* **30**: 579–593.
- Zaessinger S, Busseau I, Simonelig M. 2006. Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. *Development* **133**: 4573–4583.
- Zaret KS. 2020a. Pioneer Transcription Factors Initiating Gene Network Changes. *Annu Rev Genet* **54**: 367–385.
- Zaret KS. 2020b. Pioneer Transcription Factors Initiating Gene Network Changes. *Annu Rev Genet* **54**: 367–385.
- Zaret KS. 2018. Pioneering the chromatin landscape. *Nat Genet* **50**: 167–169.
- Zaret KS, Mango SE. 2016. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr Opin Genet Dev* **37**: 76–81. <http://linkinghub.elsevier.com/retrieve/pii/S0959437X15001446>.
- Zenk F, Loeser E, Schiavo R, Kilpert F, Bogdanovic O, Iovino N. 2017. Germ line–inherited

- H3K27me3 restricts enhancer function during maternal-to-zygotic transition. *Science* (80-) **357**: 212–216.
- Zhang H, Emerson DJ, Gilgenast TG, Titus KR, Lan Y, Huang P, Zhang D, Wang H, Keller CA, Giardine B, et al. 2019a. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**: 158–162. <http://dx.doi.org/10.1038/s41586-019-1778-y>.
- Zhang H, Emerson DJ, Gilgenast TG, Titus KR, Lan Y, Huang P, Zhang D, Wang H, Keller CA, Giardine B, et al. 2019b. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**: 158–162.
- Zhang H, Lam J, Zhang D, Lan Y, Vermunt MW, Cheryl A. 2021. CTCF and transcription influence chromatin structure re-configuration after mitosis. *bioRxiv*.
- Zhang T, Cooper S, Brockdorff N. 2015. The interplay of histone modifications - writers that read. *EMBO Rep* **16**: 1467–1481.
- Zhao R, Nakamura T, Fu Y, Lazar Z, Spector DL. 2011. Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nat Cell Biol* **13**: 1295–1304.
- Zheng Y, Wildonger J, Ye B, Zhang Y, Kita A, Younger SH, Zimmerman S, Jan LY, Jan YN. 2008. Dynein is required for polarized dendritic transport and uniform microtubule orientation in axons. *Nat Cell Biol* **10**: 1172–1180.
- Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, Dodonova SO, Nitta KR, Morgunova E, Taipale M, et al. 2018. The interaction landscape between transcription factors and the nucleosome. *Nature* **562**: 76–81.

Annexes

Annexes comprise unpublished supplementary results and publications in which I participated during my thesis. The supplementary results belong to the project on mitotic bookmarking and more specifically on our efforts to characterize GAF functions as a mitotic bookmarker.

1. Transcription activation of *doc* transgene

It seems that for a given gene, experiencing transcription at a given cycle does not necessarily lead to a rapid post-mitotic reactivation in the following cycle. It is the case for the transgene with *doc enhancer* element combined with the minimal *snail* promoter (Dufourt et al. 2018). In that case, nuclei derived from active nuclei do not activate faster than the ones derived from inactive nuclei (Supplementary Fig. 2e–h of Dufourt et al. 2018). As this *doc enhancer* is not bound by GAF, we thought to add GAF binding sites to investigate if GAF would be sufficient to trigger memory in those conditions. We added a stretch of GAGAG motifs originally from the *snail primary enhancer* (Figure Annexe 1 A). Surprisingly, this GAGAG addition does not drastically change the post-mitotic activation synchrony of nuclei compared to the transgene without GAGAG sites (Figure Annexe 1 B). We next looked at the transcriptional memory bias by quantifying the timing of nuclei derived from active nuclei, and nuclei derived from inactive nuclei (Figure Annexe 1 C). We confirmed that *doc* transgene without GAGAG sites do not harbor a transcriptional memory (Figure Annexe 1 D) but adding GAGAG sites does not seem to have a significant impact on transcriptional memory (Figure Annexe 1 D). Taking together, those results demonstrate that GAF is not sufficient to trigger a transcriptional memory in the case of the *doc* transgene. It could be explained by the requirement of other factors that GAF is not able to recruit by itself, or that the activation timing is already too fast to allow us to see a difference between the from active and the from inactive. This fast activation can be due to Zelda, as Zelda is known to regulate *doc* gene and binds to this *doc* enhancer (Harrison et al. 2011), and is known to accelerate transcription (Dufourt et al. 2018).

Nonetheless an important control here would be to verify the binding of GAF by CHIP-qPCR on these transgenes, to see if GAF binds to the *snail* promoter and if we see a real increase of binding by adding the GAGA stretch.

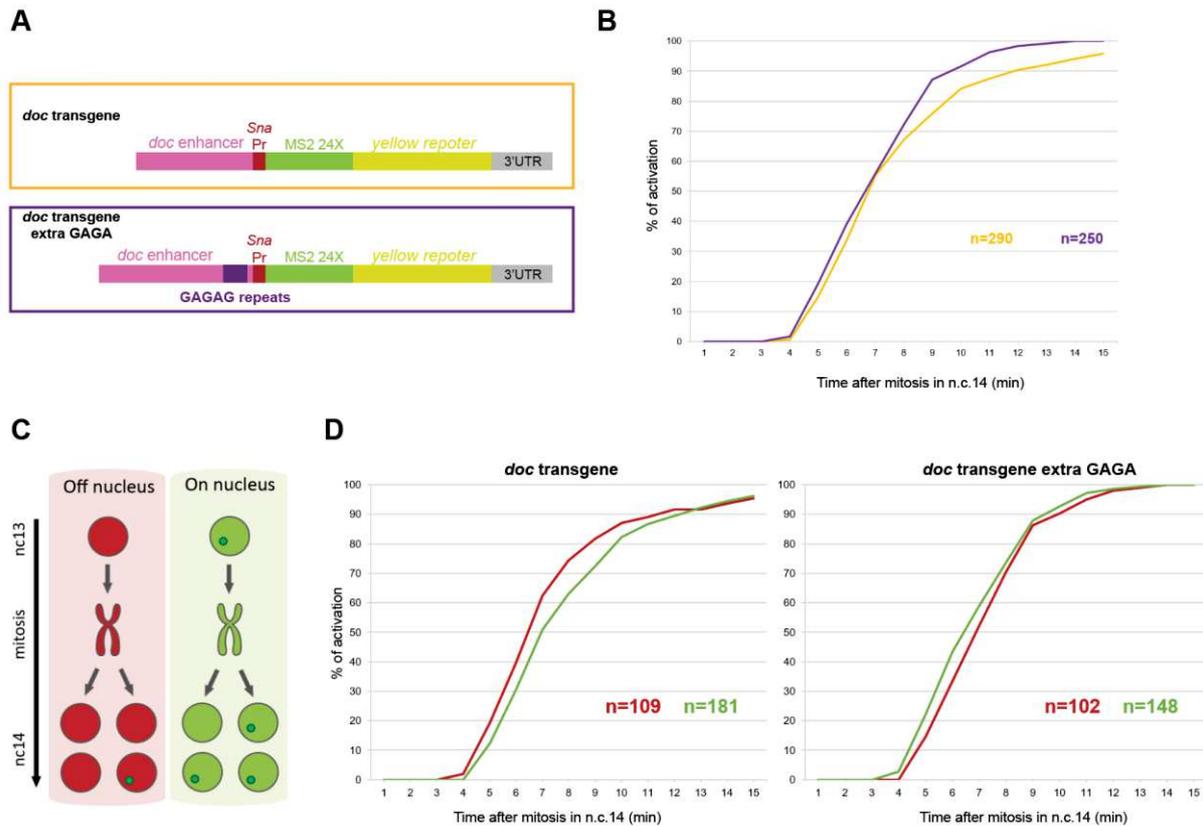


Figure Annexe 1: GAF is not sufficient to trigger memory on the *doc* transgene.

(A) Schematic of the transgenic constructs used. Those constructs contain 24 MS2 loops for transcription visualization. (B) Quantification of transcriptional synchrony of *doc* transgene (yellow) and *doc* transgene extra GAGA (purple) embryos after mitosis. *n*=number of nuclei analyzed from 3 movies of 3 independent embryos for each transgene. (C) Schematic of the two populations of nuclei analyzed: nuclei derived from inactive nuclei (red) and nuclei derived from active nuclei (green). Green dots inside nuclei represent the detected transcription activation of the MS2 transgene. (D) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (red) of *doc* transgene (left) and *doc* transgene extra GAGA (right). *n*=number of analyzed nuclei from 3 movies of 3 independent embryos for each transgene.

2. Transcriptional memory of the *Tom* gene

To examine transcriptional dynamics of GAF-bound targets, we selected eight genes (*bacchus*, *Tom*, *scylla*, *lola*, *frizzled*, *disco*, *pipesqueak* and *tonalli*) to tag with 24 MS2 loops by CRISPR editing. These genes were selected after validating their expression in the early *Drosophila* embryo with the literature and we took the genes with the highest peak of GAF during mitosis. Among these candidates, we obtained transformants for five genes.

For three of the transformants (*bacchus*, *frizzled*, *disco*) the MS2/MCP signal was too weak to be detectable by confocal live imaging (only by RNA-FISH on fixed embryos). So I focused my work on imaging *scylla* dynamics (results presented in Chapter 2) and *Tom*. Here I will describe the results obtained with the gene *Tom* (Twin of m_4).

Tom gene is highly bookmarked by GAF. Indeed, GAF is present at the promoter of *Tom* in interphase and mitosis (Figure Annexe 2 A). We validated the expression of *Tom* in the embryo conform to what has been observed (<http://fly-fish.cabr.utoronto.ca/>). *Tom* is expressed in the dorsal side of the embryo, antagonist to the snail pattern (Figure Annexe 2 B). We then sought to measure the potential transcriptional memory bias. We observed a small difference of timing between the nuclei derived from active nuclei compare to the nuclei derived from inactive nuclei (Figure Annexe 2 C). This was confirmed by the mean timing of activation after mitosis (Figure Annexe 2 D). This shows that this gene may have a transcriptional memory but also that this memory would be gene dependent. One would need to look at the memory bias in a GAF depleted embryos.

Interestingly, Zelda binding at the *Tom* gene seems more potent than on the *scylla* gene (ChIP-seq from Harrison et al. 2011) (Figure Annexe 2 A). Knowing that Zelda has been shown to accelerate transcription on both from active and from inactive nuclei and that can lead to bypass memory (Dufourt et al. 2018), the small memory bias in the case of *Tom* might be due to Zelda effect on transcription activation. We hypothesize that in some cases, GAF driven memory can be masked by an overall acceleration of the timing of transcriptional activation by Zelda.

To test this hypothesis, we would need to look at the memory bias of *Tom* Zelda depleted embryos.

The GAF bookmarked gene *Tom* might therefore be subject to transcriptional memory, but the regulation by a combination of different factors would lead to a different output than for other genes such as *scylla*.

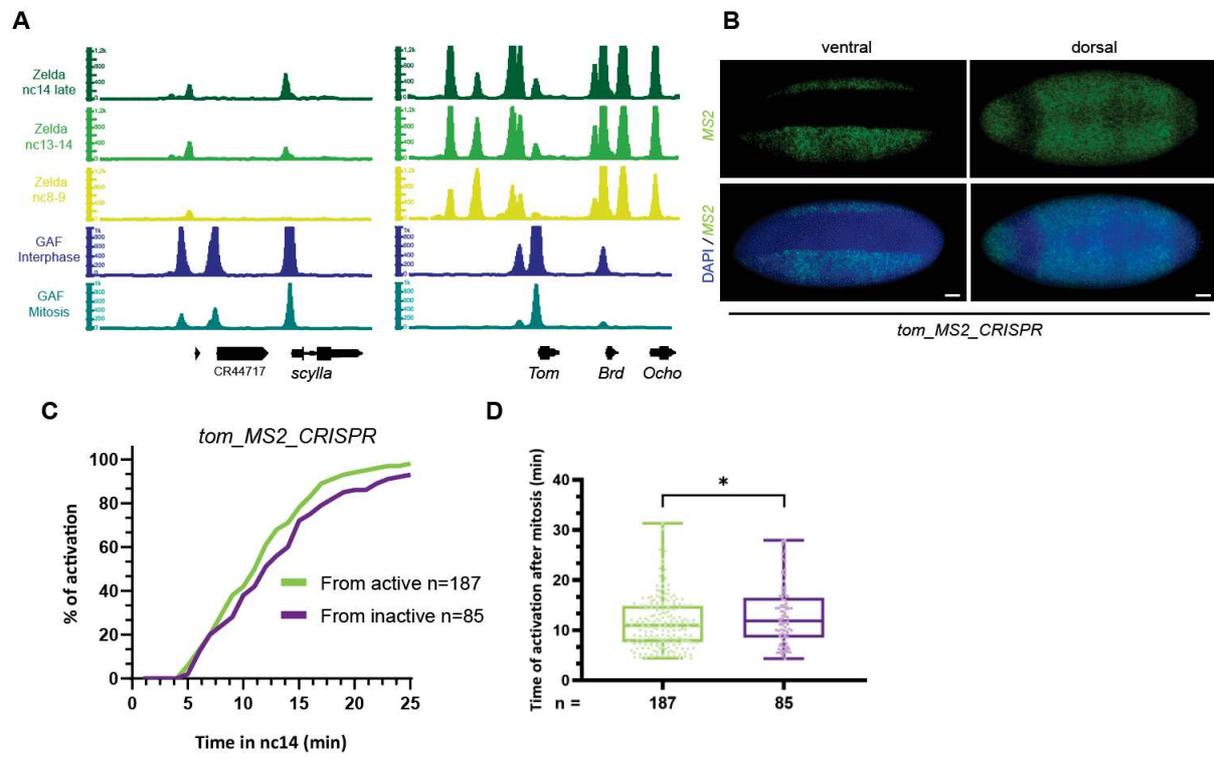


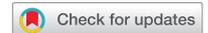
Figure Annexe 2: *Tom* gene transcription activation.

(A) Genome browser image of Zelda (yellow and greens) and interphase and mitotic GAF (dark blue and turquoise) ChIP-seq signal at the *scyl* and *Tom* locus. (B) Maximum intensity projected Z-planes of confocal images from smiFISH with MS2 probes (green) counterstained with DAPI (blue) of *Tom*_{24X}-MS2_CRISPR/+ embryos in ventral and dorsal view. Scale bars are 30µm. (C) Cumulative activation of the first activated nuclei coming from active nuclei (green) and from inactive (purple) of *Tom*_{24X}-MS2_CRISPR. n=number of analyzed nuclei from 3 movies of 3 independent embryos. (D) Box plot representing the mean time of the first activation after mitosis of nuclei derived from active (green) and inactive (purple) nuclei in *Tom*_{24X}-MS2_CRISPR/+ embryos. Centered horizontal line represents the median. Two tailed Welch's t-test * $p < 0.05$.

3. Publication: Cis-regulatory chromatin loops during early *Drosophila* development

This manuscript was recently published (2021). This was a collaborative project in which I participated mostly by generating a CRISPR fly strain (together with Matthieu Dejean) and performing RNA-FISH and generating *RNAi-Zelda* embryos. As a strategy to obtain a CRISPR fly strain deleted for a *doc* enhancer, M Dejean introduced FRT sequences from either side of the enhancer, and a dsRed marker, flanked with loxP sequences, to follow the insertion once flies were transformed. Then, I had to cross with CRE expressing strain to remove the dsRed because of the potential effect of the dsRed sequence for further analysis (here proximity analysis). After the screen for effective dsRed removal, I crossed with a Flippase expressing fly strain to remove the *doc* enhancer. After PCR screening of tens of strains, I could genotype and select one positive for further analysis. The strain was crossed with balancer chromosomes and did not show any phenotypic defects when the deletion was homozygous. As this enhancer potentially regulates three genes (*doc1*, *doc2*, and *doc3*), I performed RNA-FISH to study the expression pattern of these genes. This deletion did not seem to affect their expression patterns.

This work aimed at exploring the relationship between chromatin structure and gene expression in the *Drosophila* embryo, by employing a recently developed technique to look at *cis*-regulatory regions proximity at the single cell level (Cardozo Gizzi et al. 2019). This work was briefly introduced in Chapter 1, section B1.3. and D.3.2.



Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early *Drosophila* development

Sergio Martin Espinola^{1,5}, Markus Götz^{1,5}, Maelle Bellec², Olivier Messina^{1,2}, Jean-Bernard Fiche¹, Christophe Houbbron¹, Matthieu Dejean², Ingolf Reim³, Andrés M. Cardozo Gizzi⁴, Mounia Lagha²✉ and Marcelo Nollmann¹✉

Acquisition of cell fate is thought to rely on the specific interaction of remote cis-regulatory modules (CRMs), for example, enhancers and target promoters. However, the precise interplay between chromatin structure and gene expression is still unclear, particularly within multicellular developing organisms. In the present study, we employ Hi-M, a single-cell spatial genomics approach, to detect CRM-promoter looping interactions within topologically associating domains (TADs) during early *Drosophila* development. By comparing cis-regulatory loops in alternate cell types, we show that physical proximity does not necessarily instruct transcriptional states. Moreover, multi-way analyses reveal that multiple CRMs spatially coalesce to form hubs. Loops and CRM hubs are established early during development, before the emergence of TADs. Moreover, CRM hubs are formed, in part, via the action of the pioneer transcription factor Zelda and precede transcriptional activation. Our approach provides insight into the role of CRM-promoter interactions in defining transcriptional states, as well as distinct cell types.

Chromosomes are organized at different levels—nucleosomes, chromatin loops, TADs and chromosome territories—and each of these layers contributes to the regulation of transcription^{1,2}. In particular, loops between enhancers and promoters are critical for the precise regulation of transcriptional activation^{3–7}. In addition, organization of chromosomes into TADs plays a role in transcriptional regulation⁸, primarily by facilitating communication between enhancers and promoters through enhancer–promoter (E–P) loops within a TAD and restricting contacts from enhancers of neighboring TADs^{5,9–13}. However, the interplay of formation of E–P loops, emergence of TADs and transcriptional output is still poorly understood¹⁴.

Tissue-specific enhancers have been shown to be in proximity⁸ to their cognate promoters, indicating that E–P contacts are needed for precise gene regulation^{15–18}. Indeed, introduction of ectopically enforced E–P contacts can lead to transcriptional activation of a reporter gene during *Drosophila* development¹⁹. In some cases, enhancers can increase transcriptional output by modulating transcriptional bursting^{6,20–23}. However, in other cases, E–P contacts seem to be dissociated from gene activation^{24,25}, suggesting that an enhancer may not necessarily need to be in continuous physical contact with a promoter to influence transcription. The mechanisms by which E–P contacts may regulate transcription are currently under intense debate^{14,26–28}.

Promoters can contact several distant enhancers^{15–17}, raising the possibility that more than one enhancer may contact a promoter at any given time. More recently, use of multi-way 3C and 4C methods showed that, indeed, enhancers can cluster together to form enhancer hubs^{29–32}. This is supported by evidence of nuclear microenvironments containing multiple enhancers and clusters of

transcription factors (TFs)^{33–39}. This model is consistent with multi-way interactions between distal enhancers to regulate promoter activity of single or multiple genes by sharing resources. Whether and how formation of multi-way interactions may be related to the emergence of TADs during development^{40,41} are still open questions.

To shed light on to these questions, we investigated the interplay of transcriptional state and physical proximity between promoters and large sets of CRMs (for example, enhancers and silencers) during the awakening of the zygotic genome in early *Drosophila* embryos. During the first hours of development, *Drosophila* embryos offer an ideal biological context to decipher how CRMs are employed to establish precise spatiotemporal patterns of gene expression. Decades of genetic and genomic studies have characterized CRMs on a large scale and their usage to interpret morphogen gradients^{42–44}. In particular, the pioneering activity of factors such as Zelda (Zld) establishes early accessibility of CRMs (reviewed in ref. ⁴⁵) and is involved in the emergence of TAD organization^{40,46}.

In the present study, we used Hi-M, an imaging-based technology enabling the detection of chromatin organization and transcriptional status in intact embryos^{47,48}. This technology allowed us to visualize where and when interactions between CRMs occur and investigate their impact on transcriptional states. We first used Hi-M to detect intra-TAD chromatin loops in *Drosophila* embryos. We show that most of these loops involve CRMs. In fact, we identified not only E–P loops but multiple CRM contacts (E–P, P–P and E–E) co-interacting locally in single nuclei and referred to as CRM hubs. Unexpectedly, these contacts were not found to be specific to transcriptionally active nuclei. Hence, tissues with different cell fates exhibit similar CRMs contacts and E–P loops. Moreover, networks

¹Centre de Biologie Structurale, CNRS UMR 5048, INSERM U1054, Univ Montpellier, Montpellier, France. ²IGMM, CNRS, Univ Montpellier, Montpellier, France. ³Department of Biology, Friedrich-Alexander University of Erlangen-Nürnberg, Erlangen, Germany. ⁴Centro de Investigación en Medicina Traslacional Severo Amuchastegui, Instituto Universitario de Ciencias Biomédicas de Córdoba, Consejo Nacional de Investigaciones Científicas y Técnicas, Córdoba, Argentina. ⁵These authors contributed equally: Sergio Martin Espinola, Markus Götz. ✉e-mail: mounia.lagha@igmm.cnrs.fr; marcelo.nollmann@cbs.cnrs.fr

of CRM loops are established at early stages, before both the emergence of TADs and transcriptional activation. Finally, the pioneer factor Zld is required for the establishment of subsets of CRM hubs.

Results

High-resolution Hi-M reveals preferential interactions between *cis*-regulatory modules. Functional characterization of specific chromatin loops between CRMs within TADs (Fig. 1a) requires the development of technologies adapted for the simultaneous detection of such looping interactions and of transcriptional output. Recently, we and others established a new family of imaging-based methods able to retrieve chromatin architecture and transcriptional status simultaneously in single cells (Hi-M and ORCA)^{47–49}. Hi-M relies on the labeling and imaging of the expression pattern of genes by direct detection of transcripts via RNA–FISH, followed by the sequential imaging of tens of distinct DNA loci by Oligopaint–FISH⁵⁰ in intact *Drosophila* embryos^{47,48}. First, we tested whether conventional Hi-M could detect intra-TAD chromatin loops in two genomic regions harboring early developmental genes expressed at different timings and regions of the embryo (*dorsocross* (*doc*)- and *snail* (*sna*)-TADs).

The *doc*-TAD contains a family of three genes, the *dorsocross* genes *doc1*, *doc2* and *doc3* encoding functionally redundant T-box TFs essential for the development of the amnioserosa and cardiogenesis⁵¹. These genes display similar expression patterns, particularly during the early stages of embryogenesis, in the blastoderm embryo (nuclear cycle (nc) 11–14), which is the focus of the present study (Extended Data Fig. 1a,b). In early embryos, the *doc*-TAD is flanked by insulator-binding sites (for example, CP190) and displays extensive H3K27me3 marks as well as several prominent Zld peaks (Fig. 1b)^{52–55}. At nc14, the Hi-M contact probability map of this genomic region displays two clear TADs, similar to those detected by Hi-C (TAD1 and *doc*-TAD; Fig. 1b and Extended Data Fig. 1a)⁴⁶. Inspection of assay for transposase-accessible chromatin using sequencing (ATAC-seq)⁵⁶, H3K4me3, H3K4me1 and H3K27ac profiles⁵⁷, as well as of enhancer databases⁵⁸, revealed that the *doc*-TAD contains several putative CRMs, including four potential enhancers (CRM_a, CRM_b, CRM_c and CRM_d) for the three *doc* promoters (Fig. 1c and Supplementary Table 1)^{52,53,58,59}. We note that only CRM_a displayed exhaustive binding by several chromatin insulators (Extended Data Fig. 1b). Conventional Hi-M/Hi-C did not exhibit clear specific looping interactions within the *doc*-TAD, most probably due to insufficient genomic resolution and coverage (Fig. 1b).

To overcome these limitations and probe communications between CRMs and promoters within TADs in an unbiased manner, we improved the genomic resolution and coverage of Hi-M by threefold (from ~8–10 kb to ~3 kb) and painted the entire *doc*-TAD with contiguous barcodes (Extended Data Fig. 1a,c–i), particularly targeting promoters and predicted CRMs (Fig. 1c and Extended Data Fig. 1a). We first focused on enhancers already validated by transgenic assays (CRM_{b–d}; Supplementary Table 1). The three *doc* genes within the *doc*-TAD exhibit a shared spatiotemporal profile of expression in late nc14 (Extended Data Fig. 2a). The frequency of activation, estimated by the number of alleles transcribing per nucleus^{20,60}, was elevated for both *doc1* and *doc2* (~90%; Extended Data Fig. 2b). In addition, *doc* genes displayed a high degree of co-activation (~70%; Extended Data Fig. 2c). Thus, we hypothesized that multiple putative CRMs are likely to contact *doc* promoters to regulate their common expression pattern.

To test this hypothesis, we selected nuclei displaying at least one nascent messenger RNA–FISH *doc1* spot and built a high-resolution Hi-M contact map containing only these nuclei (Fig. 1d–f). Remarkably, the improvement in genomic coverage in Hi-M now enabled the detection of specific looping interactions between genetic elements within the *doc*-TAD in intact embryos (Fig. 1f and Extended Data Fig. 1g). The strongest contacts represented, in

all cases, interactions between CRMs (Fig. 1f, yellow arrows), but there was a considerable internuclear variation (see single-nucleus snapshots in Fig. 1f). Contact frequencies did not vary considerably when only nuclei displaying the strongest *doc1* RNA–FISH signals were used to construct the matrix (Extended Data Fig. 2d), suggesting that stronger transcriptional activity did not involve different interactions. To quantify the strength of looping interactions, we calculated the intensity of the Hi-M map across an anchor to generate virtual interaction profiles (hereafter referred to as 4M plots). We observed that CRM_c predominantly interacts with CRM_a and CRM_b, with similar probabilities (Extended Data Fig. 2e,vii). In contrast, we did not observe specific loops between CRMs and barcodes not containing early CRMs (for example, ctrl barcode; Fig. 1c and Extended Data Fig. 2e,iv). Interactions between CRMs and promoter regions (for example, P1, P2 and P3) were present but displayed lower frequencies than interactions between CRMs (red arrow, Fig. 1f). Interactions between CRMs and the *doc1* promoter did not depend on the activation level of *doc1* (Extended Data Fig. 2f).

Next, we investigated whether all putative CRMs displayed chromatin loops. It is interesting that a CRM predicted from epigenetic profiling but not present in enhancer databases (for example, CRM_e; see complete list of reported enhancers in Supplementary Table 1) displayed extensive interactions with reported enhancers (for example, CRM_b, CRM_c and CRM_d), as well as with the promoters of *doc* genes (Extended Data Fig. 2e,v). In contrast, a subset of barcodes harboring previously described enhancers or displaying enhancer marks (see barcodes 2, 12, 13 and 15 in Fig. 1c) failed to exhibit looping interactions with other CRMs (for example, black circle in Fig. 1f). We observed similar results at the *sna*-TAD (Supplementary Note 1 and Extended Data Fig. 2g,h). Thus, high-resolution Hi-M reveals unforeseen interactions between CRMs and other regulatory regions within the *doc*-TAD, and permits the quantification of the frequencies with which putative enhancers actively contact cognate target promoters in a specific tissue and developmental timing. Collectively, these data suggest that promoters interact with a panoply of enhancers that can be shared between different genes within a TAD.

Shared enhancers, promoter competition and CRM hubs. The existence of multiple pairwise interactions between CRMs within the *doc*-TAD and the naturally occurring overlapping expression patterns of *doc* genes (Extended Data Fig. 2a) suggest that multiple CRMs may compete or cooperate for gene activation in single cells⁶¹. To discriminate between these two hypotheses, we tested whether multi-way interactions are formed by excluding an anchor of interest and plotting the frequencies with which two barcodes interact with this given anchor³⁰. First, we selected promoters as anchors. We observed that three-way interactions with multiple promoters were infrequent (green arrows, Fig. 1g,i,ii,iii), consistent with our previous observations from Hi-M contact maps (Fig. 1f and Extended Data Fig. 2e). Instead, the three *doc* promoters preferentially looped to multiple CRMs in single nuclei (yellow arrows, Fig. 1g,i,ii,iii). A control locus with no promoter marks failed to display specific looping interactions (Fig. 1g,iv).

Genomic methods revealed the spatial clustering of multiple enhancers in cultured mammalian cells^{29,30}. To test whether spatial clustering of multiple enhancers could be directly visualized in intact embryos, we mapped three-way interactions using CRMs as anchors. It is interesting that we observed CRM_{a–d} displaying high frequencies of multi-way interactions (Fig. 1h, see examples labeled by yellow arrows). By analogy with previous studies^{29,30}, we termed these CRM interaction networks ‘CRM hubs’. CRM hubs can contain promoters (green arrows, Fig. 1h), but most often contained known or putative enhancers. Analysis of the *sna*-TAD reveals a similar scenario, where CRMs are involved in most three-way interactions (Extended Data Fig. 2i,iv,y,vi).

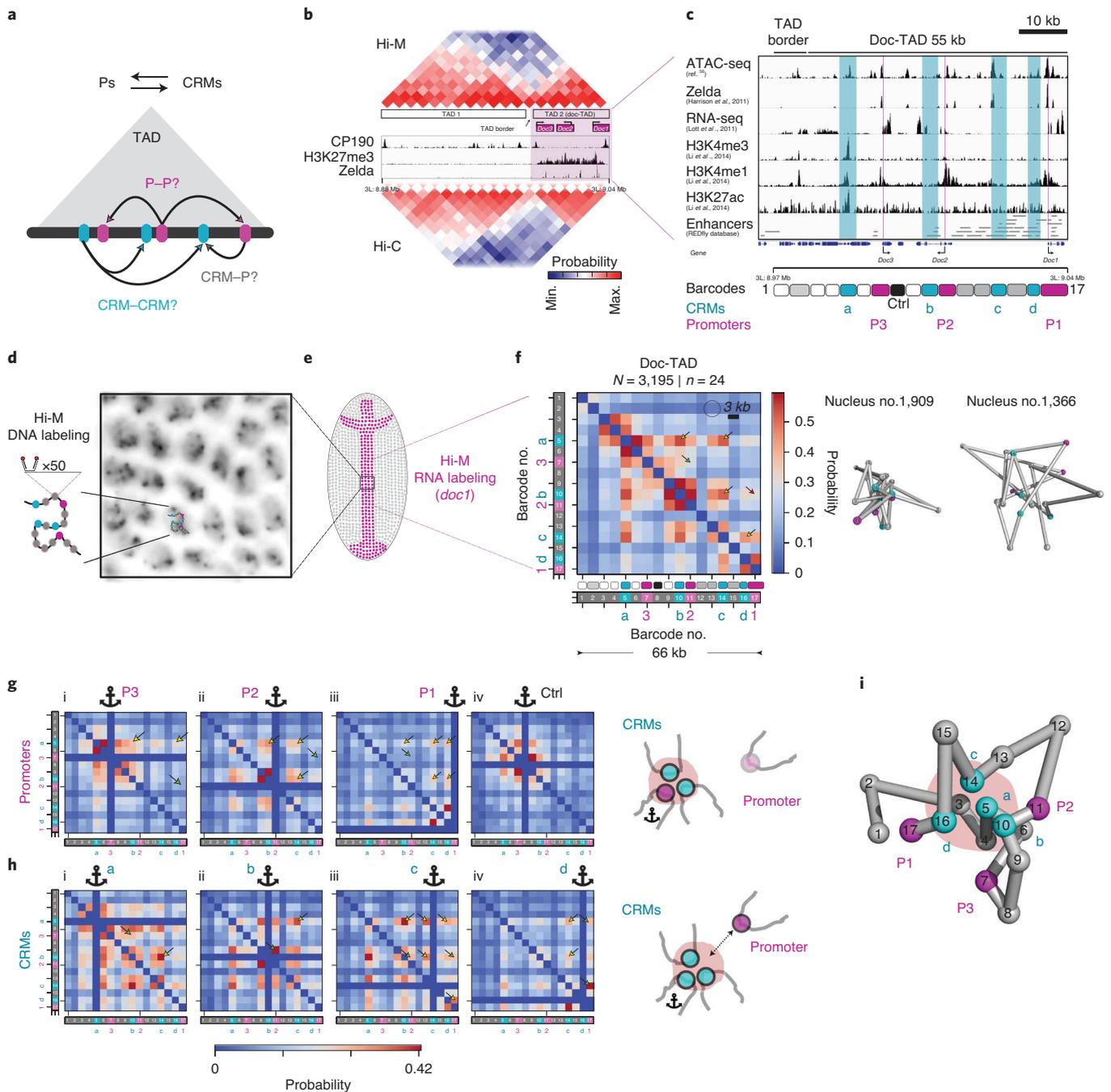


Fig. 1 | Hi-M reveals widespread *cis*-regulatory chromatin loops and hubs within TADs. **a**, Schematic of the networks of contacts between CRMs and promoters within TADs. **b**, The *doc* locus (Chr3L:8.88..9.03Mb) in *Drosophila melanogaster*. Low-resolution Hi-M and Hi-C⁴⁶ contact probability maps are shown at the top and bottom, respectively. Blue and red indicate low and high contact probabilities, respectively. **c**, Tracks for chromatin accessibility (ATAC-seq), Zld binding, transcriptional activity (RNA-seq), chromatin marks for active promoters (H3K4me3) and active enhancers (H3K4me1, H3K27ac) from nc14 embryos, and RedFly enhancers are shown. Barcodes used for high-resolution Hi-M included: regions with enhancer marks (CRM_{a-d}), *doc* promoters (P1–P3), intervening regions with no mark (for example, ‘Ctrl’) and other regions also documented as enhancers (gray). See Supplementary Table 1 for assignment of CRM_{b-d}. Accession codes for the datasets associated with the cited papers are listed in Supplementary Table 9. **d**, Schematic diagram of the labeling strategy. **e**, Schematic representation of a dorsally oriented *Drosophila* embryo. Segmentation of actively transcribing nuclei (magenta) is based on nascent RNA-FISH labeling. **f**, The high-resolution Hi-M contact probability map of *doc*-TAD in nuclei displaying *doc1* expression in nc14 embryos. Boxes with irregular sizes above barcodes represent the relative genomic lengths of each barcode. Arrows: strong looping interactions of CRMs (yellow), expected P–P interactions (green) and CRM–P interactions (red). Single-allele example reconstructions of spatially clustered and open CRM conformations. Statistics (**f–h**), $N = 3,195$ (number of nuclei with *doc1* expression), $n = 24$ (number of embryos with *doc1* expression), $N_T = 37,129$ (total number of nuclei), $n_T = 29$ (total number of embryos). **g**, Multi-way interactions between promoter regions. Anchoring barcodes are highlighted by a pictogram. A control barcode is depicted in iv. Prominent peaks (yellow) comprise one promoter and two CRMs but not multiple promoters (green). The scheme illustrates the spatial arrangement of CRMs and promoter regions when the anchor is placed at a promoter. **h**, Multi-way interactions between CRMs. Yellow arrows highlight prominent peaks involving three CRMs. The scheme illustrates the spatial arrangement of CRMs and promoter regions when the anchor is placed at a CRM. **i**, The 3D topological representation of the *doc*-TAD.

Finally, we used ShRec3D⁶² to obtain an ensemble topological reconstruction from the Hi-M matrix. In this reconstruction CRMs clustered at the center of the TAD, whereas promoter elements tended to be at the periphery (Fig. 1i). Similarly, we observed that CRMs within *sna*-TAD also tended to cluster together at the center of the TAD (Extended Data Fig. 2j). Collectively, three-way and topological analyses suggest that multiple enhancers physically interact to form CRM hubs. It is of interest that CRM hubs can but do not tend to contain multiple promoters (Supplementary Note 2).

Networks of CRM contacts are indistinguishable between cells of different cell fates. Next, we examined whether chromatin structure in this locus depended on transcriptional status (repression/activation). For this, we used Hi-M in three populations of nuclei established along the dorsoventral axis during nc14: mesoderm (M), neuroectoderm (NE) and dorsal ectoderm (DE)⁶³. To distinguish between these cell fates, we employed RNA-FISH labeling before Hi-M (with *sna* and *doc* probes directly labeling M and DE cells, respectively, Fig. 2a). Nuclei were classified as: (1) DE nuclei when an active *doc1* transcription hotspot could be visualized (Extended Data Fig. 2a); (2) M nuclei when located within the *sna* expression pattern (Fig. 2a); and (3) NE nuclei when located between the pattern of *sna* and the edge of the *doc1* pattern (Fig. 2a).

Unexpectedly, Hi-M interaction matrices for DE, NE and M displayed only minor differences (Fig. 2b–d), indicating that the same network of chromatin loops is present in nuclei that are actively transcribing and where *doc* gene expression is silent. In addition, 4M profiles were almost identical in nuclei with different cell fates and activation status, independently of whether promoters or CRMs were used as anchors (Fig. 2e and Extended Data Fig. 3a,b). For example, the *doc1* promoter (P1) showed identical interactions with the four CRMs (CRM_{a–d}) in the DE, the NE and the M (Fig. 2e,iii,vi). Likewise, CRM_a and CRM_c displayed patterns of interactions with other CRMs that were indistinguishable between tissues (Fig. 2e,i,ii,iv,v). Finally, to detect whether CRM hubs existed in tissues where *doc* genes are repressed, we performed single-nucleus three-way analyses. Indeed, comparison of three-way interaction matrices of NE and M with those of DE revealed the persistence of CRM hubs in nuclei where transcription is repressed (Fig. 2f), suggesting that CRM hubs also exist in these cell types. We observed similar results in the *sna* locus (Supplementary Note 3 and Extended Data Fig. 4a–d).

To search for a possible explanation of these results, we explored the TF-binding profiles of known activators and repressors in the *doc* locus^{53,64,65} (Supplementary Note 4). CRM_{a–g} are bound by activators in the DE, and tend to be occupied by transcriptional repressors in the M and NE (Fig. 2g,h). Thus, contacts between *doc* promoters and CRM hubs in the DE may promote activation and those in the M/NE may instead facilitate repression (Fig. 2h).

Cis-regulatory networks emerge before TADs and gene expression. Previous genome-wide and Hi-M studies have established that most *Drosophila* TADs emerge at nc14 during the major wave of zygotic gene activation^{40,46,47}. To explore whether the *doc*-TAD also emerges at this nuclear cycle, we performed low-resolution Hi-M experiments in nc11–nc12 and nc14 embryos (Fig. 3a). We used density of nuclei to unequivocally score developmental timing (Fig. 3b, insets). Hi-M contact maps revealed that the *doc*-TAD can be detected at nc14 but not at earlier stages (Fig. 3a), so emergence of this TAD coincides with the onset of *doc* expression (Fig. 3c).

To determine whether specific looping interactions between CRMs appear before the emergence of TADs, we performed high-resolution Hi-M between nc11 and nc14. As our previous data showed that Hi-M maps are similar in different presumptive tissues (Fig. 2), we built Hi-M maps for the different nuclear cycles using all detectable nuclei independently of their location in the embryo. Surprisingly, chromatin loops between CRMs were

observed very early in development (nc11) and remained almost unaffected at least until nc14 (Fig. 3b). For example, loops between CRM_c and CRM_a, and CRM_b and CRM_d were readily detected as early as nc11, and assumed their final contact frequencies at nc12 (Fig. 3d). Similar behaviors were observed when using other CRMs as anchors (Extended Data Fig. 5a–c).

These results are consistent with three-way analysis, where we observed that three-way interactions are almost indistinguishable from nc12–nc14 (Fig. 3e and Extended Data Fig. 5d–f). To gather further evidence for the formation of CRM hubs during early development, we obtained ShRec3D structures for each nuclear cycle. Notably, these structures show that CRMs cluster at the center of the TAD as early as nc11, with clusters becoming tighter as development progresses (Fig. 3f). We reached similar conclusions when analyzing the *sna*-TAD, which also emerges at nc14 (ref. 47) (Supplementary Note 5 and Extended Data Fig. 6). Overall, these data indicate that pairwise looping interactions between CRMs in *doc*- and *sna*-TADs are established from nc11 (or before), whereas three-way interactions are progressively acquired during development. Importantly, both pairwise and multi-way looping interactions are formed before the emergence of TADs.

To investigate whether specific loops between *doc* promoters and CRMs displayed quantitative changes before the onset of *doc* gene expression (*doc1* is the first to be activated, followed by *doc3* and *doc2*; Fig. 3c), we plotted virtual 4M profiles with promoters as viewpoints. Notably, we observed that promoters contact CRMs as early as nc11, and that frequencies of interactions ceased to change after nc12 (see P1 in Fig. 3g, and P2–P3 in Extended Data Fig. 5a–c). Three-way interactions involving promoters could also be already detected at nc11, and became more frequent at later nuclear cycles (Fig. 3h and Extended Data Fig. 5d–f). Thus, our results indicate that loops involving promoters and one or several CRMs precede TAD formation and gene expression, and are equally frequent in pluripotent cells, which do not express *doc* genes.

Redundancy of CRM usage at the *doc*-TAD. To test whether enhancers were redundant in this locus, we searched the literature^{36,66,67} and performed enhancer reporter assays to identify the activation pattern of CRMs located around the *doc* genes (Methods). Several enhancers displayed partially overlapping patterns of activation at this stage of development, particularly in the region around CRM_c (Fig. 4a,i,ii). This finding is consistent with multiple enhancers being able to activate the transcription of *doc* genes at this developmental stage.

To further test this hypothesis, we deleted CRM_c, which is located midway between two co-regulated promoters (*doc1* and *doc2*) (Methods, Fig. 4b and Extended Data Fig. 7a). Notably, deletion of CRM_c did not lead to detectable changes in *doc1* or *doc2* expression or on their co-activation frequency (Extended Data Fig. 7b), but produced large changes in the organization of the *doc*-TAD (Fig. 4c,d). Interactions involving the barcode originally containing CRM_c were considerably diminished, consistent with the binding of factors to this region being responsible for the formation of long-range CRM and promoter contacts (Fig. 4c,d, yellow arrows). Remarkably, we still observed interactions between the *doc1* promoter and the other enhancers (CRM_a, CRM_b and CRM_d), indicating that removal of CRM_c did not affect the ability of the *doc1* promoter to be frequently in proximity to the other CRMs within the *doc*-TAD (Fig. 4c,d, red arrows). Finally, three-way interactions between P1 and other CRMs (CRM_b, CRM_a and CRM_d), as well as with P2, persisted despite deletion of CRM_c (Fig. 4e, red arrows). Thus, CRM–CRM and CRM–P interactions still occur in the absence of CRM_c.

Formation of CRM hubs requires the pioneer factor Zld. Having shown that interactions between multiple CRMs do not depend on transcriptional state or developmental timing, we searched for

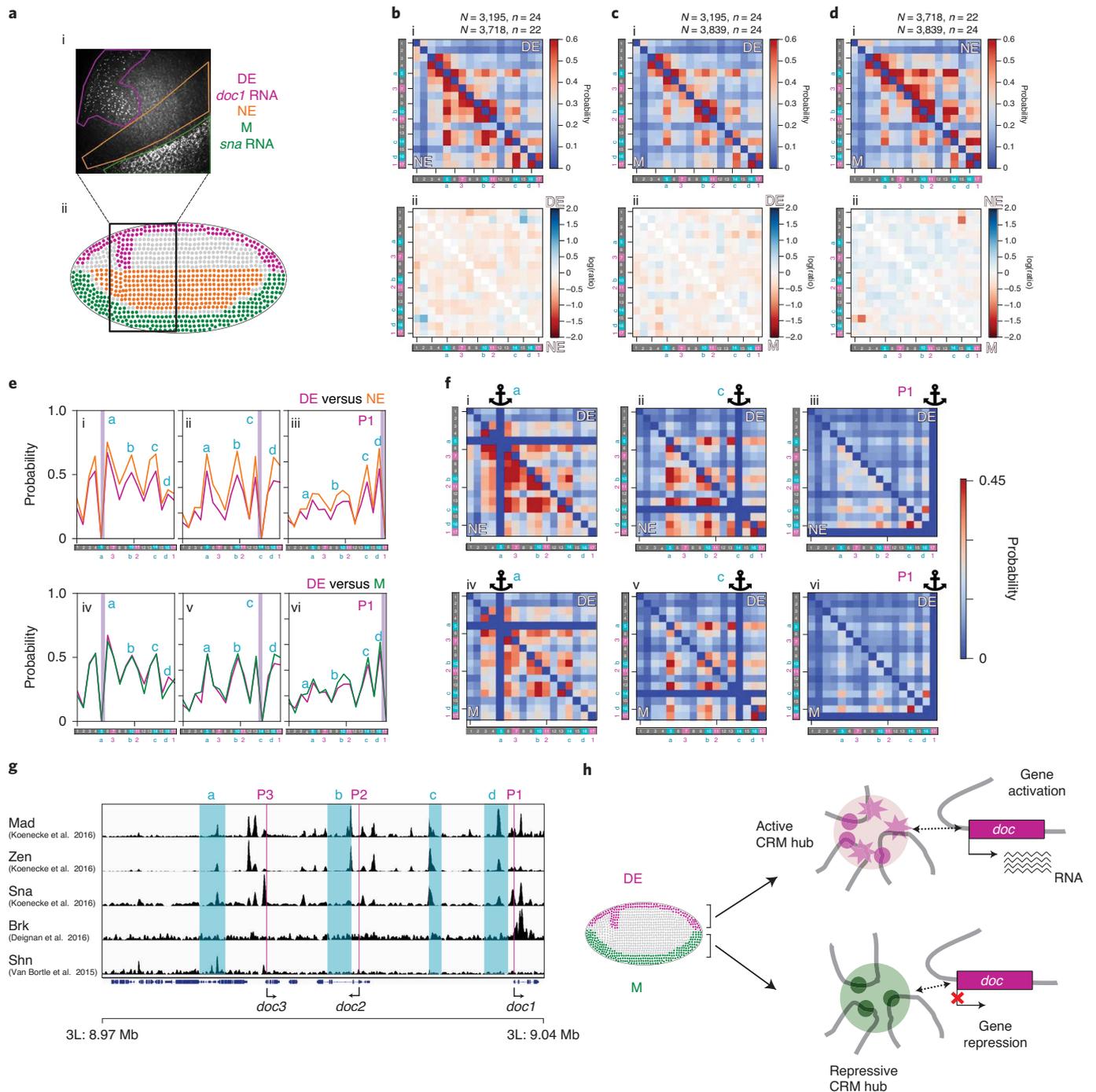


Fig. 2 | CRM-CRM and CRM-P loop frequencies are similar between cell types. **a**, Scheme indicating the three presumptive tissues and their segmentation (ii) based on RNA-FISH labeling (i). **b**, Contact probability maps for DE (upper-right half) and NE (lower-left half) (i). ii, Map of the natural log of the ratio between contact probabilities of DE and NE. Blue indicates a larger contact probability in DE and red in NE. **c**, Similar to **b**, but for DE and M. **d**, Similar to **b**, but for NE and M. **e**, The 4M profiles derived from Hi-M maps for a selected number of anchors. Anchors were placed at CRM_a (i and iv), CRM_c (ii and v) and P1 (iii and vi). **f**, Comparison of three-way contacts for the same tissues and anchors as in **e**. Number of examined nuclei and embryos as in **b-d**, respectively. **g**, ChIP profiles of key transcriptional regulators in the *doc*-TAD. Accession codes for the datasets associated with the cited papers are listed in Supplementary Table 9. Datasets are from whole embryos at nc14. **h**, Illustration of the double role of CRMs in the *doc*-TAD.

factors that may be required for the formation of CRM hubs. The pioneer factor Zld has the unusual ability to overcome nucleosome barriers at specific regulatory elements, making them accessible for binding by other classic TFs before activation, as early as nc8–nc11 (refs. 54–56,59,68,69). The *doc*-TAD is enriched in Zld binding, particularly CRM_{a-d} (Fig. 4f), and Zld was required to ensure proper

expression of *doc* genes, as well as ensuring Pol 2 binding and chromatin accessibility at the *doc* locus (Supplementary Note 6 and Extended Data Fig. 7c–e). To explore whether Zld depletion led to changes in the *doc*-TAD structure, we performed Hi-M experiments on Zld maternally depleted embryos by using RNA interference (RNAi)⁶⁸. Given the widespread developmental defects

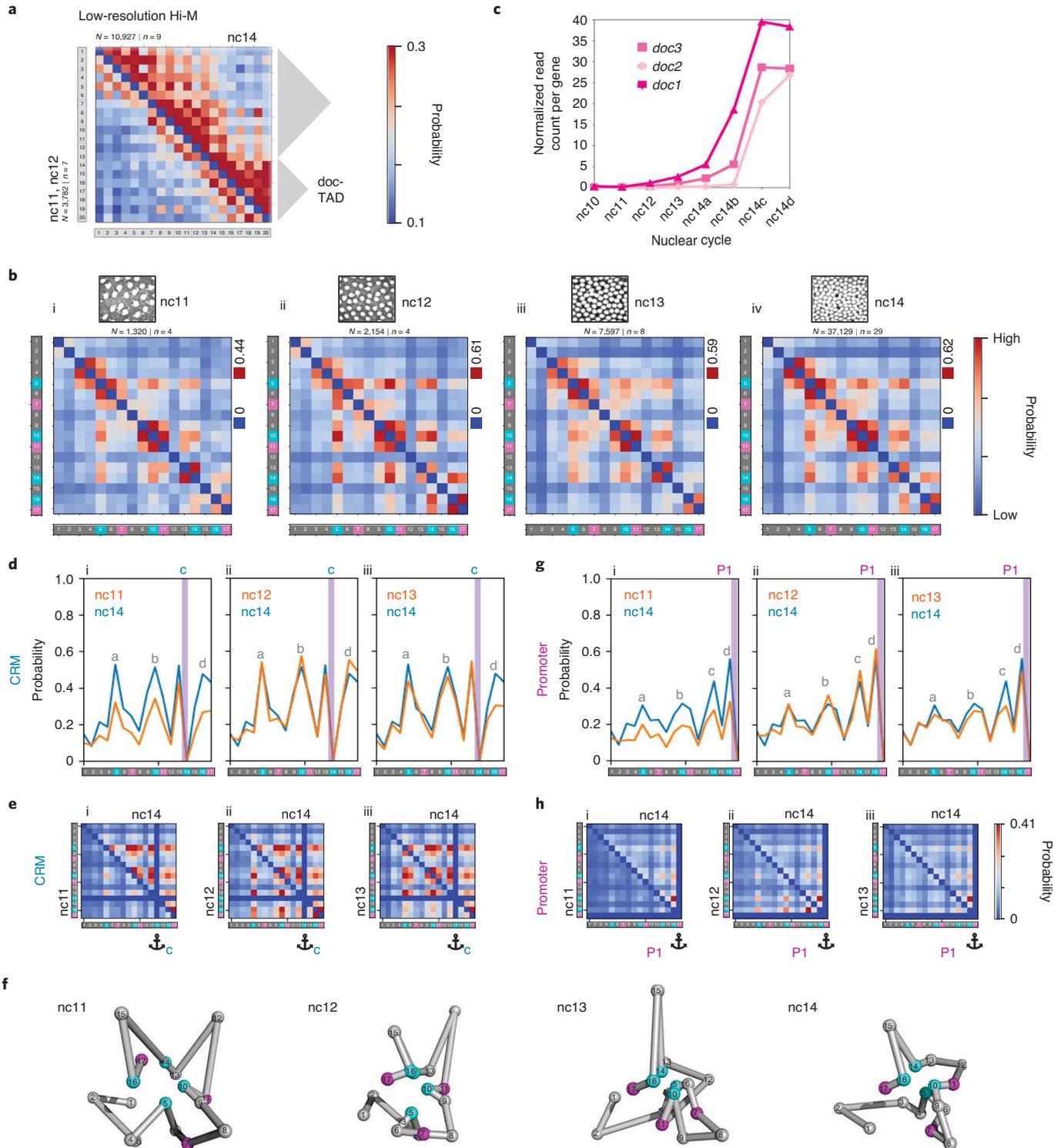


Fig. 3 | CRM loops and hubs precede TAD formation and gene expression. a, Low-resolution Hi-M contact probability map of an extended genomic region around the doc-TAD. N , number of nuclei; n , number of embryos. **b**, Representative images of DAPI-stained nuclei for embryos in nc11–nc14 (upper panel). High-resolution Hi-M contact probability maps of the doc-TAD for embryos in nc11–nc14. The minimum and maximum values of the linear color scale are indicated for each matrix contact using blue and red boxes. **c**, Expression profile of *doc1*, *doc2* and *doc3* during nc10–nc14. The nc14 was divided into four time points according to the extent of cellularization (a, earliest; d, last). Accession codes for the datasets associated with the cited papers are listed in Supplementary Table 9. **d**, Comparison of 4M profiles derived from Hi-M maps at different nuclear cycles. The position of the anchor (CRM_c) is indicated by a vertical purple line. Profiles for nc11, nc12 and nc13 (orange lines) are compared with nc14 (blue lines) in i–iii, respectively. **e**, Comparison of three-way contacts between nc14 and other nuclear cycles, using CRM_c as anchor. Upper-right half of the matrix always depicts nc14, and the bottom-left half shows nc11 (i), nc12 (ii) and nc13 (iii). Number of examined nuclei and embryos are as indicated for the respective nuclear cycle in **c**. **f**, Topological reconstructions of the doc-TAD for nc11–nc14. CRMs and promoter regions are indicated as cyan and magenta spheres, respectively. **g**, Similar to **c**, but anchor: *doc1* promoter (P1). **h**, Similar to **d**, but anchor: *doc1* promoter (P1).

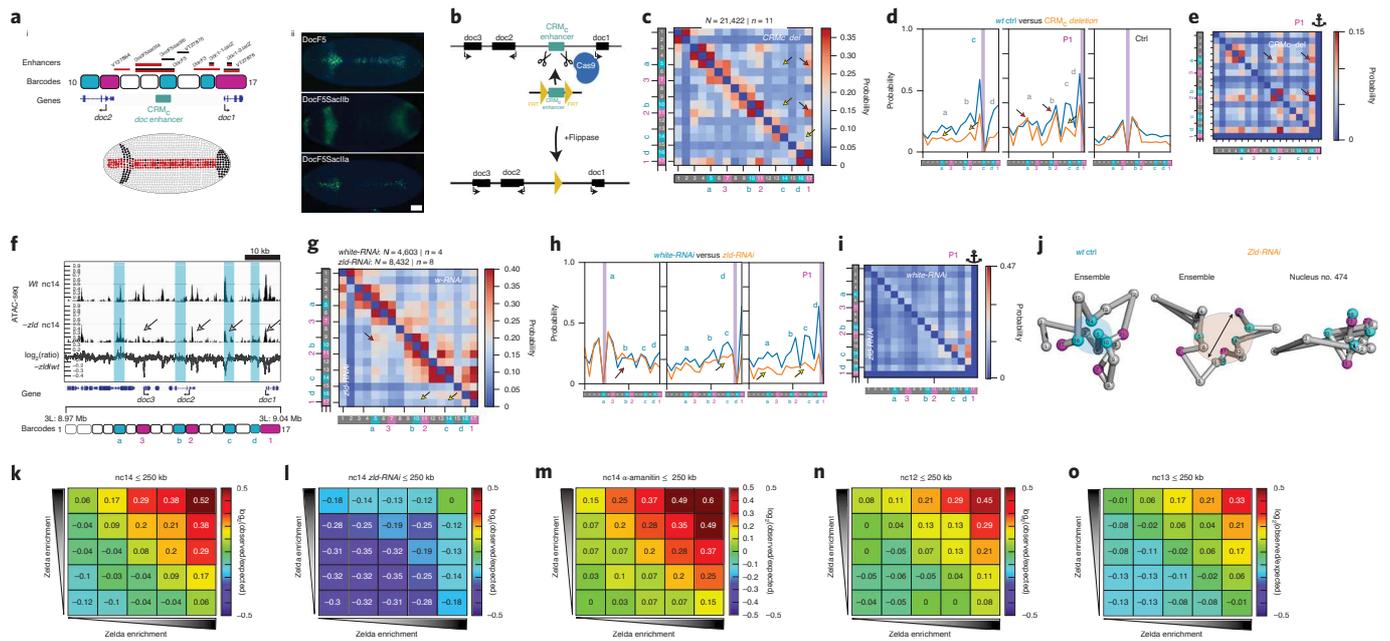


Fig. 4 | Formation of CRM loops and hubs in the *doc*-TAD requires the pioneer factor Zld. **a**, (i) Active enhancers in the *doc*-TAD and expression patterns: red represents pattern of activation of enhancers in the anteroposterior axis of the DE and black the second sub-pattern of activation observed for *doc* enhancers. (ii) Representative images, as detected by lacZ RNA in situ hybridization, for the activation patterns of DocF5, DocF5SaclIb and DocF5SaclIa. Scale bar, 50 μ m. **b**, Simplified scheme for the conditional deletion of CRM_c. **c**, Hi-M contact map for the CRM_c deletion mutant. Number of nuclei (N) and embryos (n) are the same for **c–e**. **d**, The 4M profiles for wild-type (*yw*) and CRM_c-deleted embryos. Yellow arrows indicate missing interactions between CRM_c and other CRMs or promoters. Red arrows represent interactions between the *doc1* promoter and other CRMs. **e**, Multi-way interactions in CRM_c-deleted embryos. Red arrows represent interactions between the *doc1* promoter and other CRMs/promoters. **f**, ATAC-seq profiles of wild-type and *Zld* mutant, and log₂(ratio) between *zld* and wt nc14 embryos. Barcodes are annotated below. Arrows: peaks showing a decrease in chromatin accessibility upon *Zld* depletion. **g**, High-resolution Hi-M contact probability maps for *white-RNAi* (upper-right matrix) and *zld-RNAi* (bottom-left matrix). Yellow arrows indicate missing interactions between CRMs and other CRMs/promoters. Red arrows represent the conserved interactions between CRMs. Number of nuclei (N) and embryos (n) are the same for **g–i**. **h**, The 4M profiles for *white-RNAi* and *zld-RNAi* embryos. **i**, Multi-way interactions in *white-RNAi* (upper-right map) and *zld-RNAi* (bottom-left map) embryos. **j**, Topological reconstructions of *doc*-TAD for wt and *zld-RNAi* embryos (left, center). The arrow indicates a separation of CRMs in the absence of Zld. Single-allele structure from a *zld-RNAi* embryo (right). **k–o**, The log₂(observed/expected) average contact frequencies between Zld-bound regions at short-range distances (≤ 250 kb) ranked by increasing Zld enrichment in nc14 (**k**), nc14 *zld-RNAi* (**l**), nc14 α -amanitin-treated (**m**), nc12 (**n**) and nc13 (**o**) embryos.

exhibited by *Zld* RNAi embryos at stage 5 (ref. ⁷⁰), we restricted our analyses to early nc14 *Zld* RNAi embryos.

Depletion of Zld did not affect TAD borders at the *doc* locus (Extended Data Fig. 7f) or change the overall compaction of the *doc*-TAD (Extended Data Figs. 1i and 7g). However, we observed large changes in networks of CRM–CRM and CRM–P interactions (Fig. 4g). Contacts between CRM_a and other CRMs or promoters were very similar in *Zld*-depleted and RNAi *white* control embryos (Fig. 4g,h, red arrows, and Supplementary Note 7). Remarkably, the main differences in contacts occurring in *Zld*-bound genomic regions that lose accessibility upon *Zld* depletion (CRM_c, CRM_d, CRM_b, P1, P2 and P3; Fig. 4g, yellow arrows, and Fig. 4h, middle and right). For example, on *Zld* depletion, CRM_c showed the largest drop in ATAC-seq signal among CRMs (Fig. 4f), and a dramatic drop in its interactions with other CRMs (Fig. 4g,h, yellow arrows). Finally, in *Zld* RNAi embryos, formation of CRM hubs was also considerably impacted (Fig. 4i) and topological reconstructions showed a loss of CRM clustering (Fig. 4j). Altogether, these results suggest a model whereby the pioneering activity of Zld plays a key role in the activation of *doc* genes and participates in the formation of CRM–CRM and CRM–P loops during early embryogenesis, possibly through its ability to open chromatin at specific CRMs.

To shed further light on the role of Zld in the formation of preferential interaction networks, we selected 5,038 genomic regions displaying Zld binding and calculated their pairwise, intra-arm

interaction frequencies⁷¹ using publicly available datasets⁴⁰. In nc14 embryos, Zld-bound regions interacted more frequently with each other than with control regions (Fig. 4k). This bias increased with the level of Zld binding and was present for short- (<250 kb) and long-range (>250 kb) genomic distances (Fig. 4k and Extended Data Fig. 7h). Zld depletion led to a considerable decrease in interactions between Zld-bound genomic regions (Fig. 4l and Extended Data Fig. 7i), consistent with *Zld* depletion results in the *doc*-TAD. However, this decrease in interactions was not observed upon transcriptional inhibition (Fig. 4m and Extended Data Fig. 7j), in agreement with other analyses⁴⁰. Importantly, interactions between Zld-bound genomic regions were already present in nc12–nc13 embryos (Fig. 4n,o), and were also specific to Zld-bound pre-midblastula-transition (MBT) enhancers (Supplementary Note 8 and Extended Data Fig. 7k). Overall, these results are consistent with Zld being needed for the formation of a subset of CRM–CRM interactions during early embryogenesis.

Discussion

In the present study, we use a high-resolution, imaging-based, single-cell spatial genomics approach (Hi-M) to link chromosome topology and transcriptional regulation during early *Drosophila* development. This approach has notable advantages, such as the detection of multi-way interactions and transcriptional output with spatial resolution (Supplementary Note 9). We reveal extensive

interaction networks within developmental TADs primarily involving CRMs. Critically, these networks arise thanks to the spatial clustering of multiple enhancers (CRM hubs) and are mostly invariant during cell fate specification and gene activation. Networks of pairwise CRM contacts and CRM hubs arise during early development, before the onset of gene expression and before the emergence of TADs, and require the pioneering activity of the transcription factor Zld.

One of the important results of the present study is that physical proximity between multiple CRMs and promoters is observed with very similar frequencies in cells with distinct fates and appeared during early embryogenesis. These results are consistent with those obtained at later stages of *Drosophila* embryogenesis, showing that enhancers located at considerably larger distances (~100 kb) can also form binary loops that are present in cells from different tissues¹⁷. Similarly, E–P interactions at the mouse *Hoxd* locus were detected in tissues where target genes were not expressed¹⁸. Our results are further supported by a companion paper⁷² that applied Hi-C and micro-C to study tissue-specific *Drosophila* chromosome organization at similar stages of development (Supplementary Note 10). From a developmental perspective, the formation of loops between promoters and distal regulatory elements in cells where genes need to be repressed can be seen as a ‘dangerous liaison’. Indeed, once a loop has been established, transcriptional activation could rapidly occur in cells where that specific promoter should be kept inactive.

This apparent dichotomy can, however, be rationalized in terms of the spatiotemporal patterning of the *cis*-regulatory logic of TFs during embryogenesis. For instance, in M cells, most *doc* CRMs are bound by the spatially localized transcriptional repressor Sna⁴⁴, which acts as a silencer in the M. In this case, communication between promoters and distal CRMs may reinforce transcriptional repression. This interpretation is in agreement with the finding that many enhancers can act as silencers in alternate cell types during *Drosophila* development⁷³; however, other silencing mechanisms may also be at play⁷⁴. Thus, we hypothesize that the optimal mechanism to ensure rapid and efficient activation or repression during development may involve two steps: the rapid priming of key CRMs via ubiquitously maternally deposited pioneer factors (for example, Zld), followed by regulation of transcriptional output by spatially and temporally localized transcriptional activators and repressors. In this model, three-dimensional (3D) chromatin architecture plays a double role because 3D contacts could serve to reinforce both activation and repression at a particular developmental stage while allowing for flexibility at later stages. For example, a repressive CRM loop in a tissue at an early developmental stage may switch to a CRM loop with activation capacities at later stages by changing TF occupancy. Future experiments testing whether CRM loops and hubs display more differences in active and repressed tissues at later stages of development will be important to test these hypotheses.

Previous studies suggested that invariant E–P loops may be pre-established and stable^{14,17,75,76}. In agreement with these results, our data indicate that E–P loops can form early, well before the onset of gene expression. However, in all cases, we measured low frequencies of looping interactions between functional elements. These results are consistent with previous measurements of absolute contact frequencies within TADs and between E–Ps^{49,77–79}. Thus, these results indicate that different sets of multi-way E–E and E–P contacts are present in different cells, and that these contacts may be highly dynamic.

Recent studies reported the existence of enhancer hubs: spatially localized clusters containing multiple enhancers^{29,30,33} that may facilitate transcriptional activation by creating a local microenvironment whereby transcriptional resources are shared, akin to early models of ‘transcription hubs’⁸⁰. Formation of enhancer hubs may require interactions between components of the transcriptional machinery, which could contribute to, or result from, the assembly

of phase-separated condensates^{32,37,38,81–83}. In this model, enhancers need not directly touch their target promoters but merely come into close proximity (~200–300 nm)^{25,84}. Overall, these findings and models are consistent with our observation that multiple endogenous CRMs within a TAD come together in space to form hubs in single, actively transcribing nuclei. We also observed the formation of similar hubs in inactive nuclei, suggesting that repressive elements may also form spatially localized clusters of transcriptional repressors to share resources and reinforce their silencing activities. CRM hubs are formed at early stages of development in pluripotent cells. Thus, we favor a model in which preferential CRM interaction networks are pre-formed at early stages and are subsequently specified (into activation or repression hubs) during nc14 or later.

In *Drosophila*, TADs emerge concomitant with the major wave of zygotic gene activation^{40,46,47}. Previous studies reported the existence of chromatin loops typically at considerably large genomic distances spanning two or more TADs^{17,46} or concerning Polycomb-binding sites^{46,85}. In the present study, we observed that chromatin loops between CRMs within *Drosophila* TADs are widespread, mimicking the common CTCF-mediated chromatin loops present within mammalian TADs^{15,41}. In addition, we found that multiple CRMs can cluster together to form *cis*-regulatory hubs located within TADs, suggesting a mechanism to sequester enhancers in space to reduce the activation of genes in neighboring TADs. Importantly, formation of CRM hubs precedes the emergence of TADs, consistent with the finding in mammalian cells that subsets of E–P contacts arise rapidly after mitosis before TADs are re-formed⁸⁶. Thus, our results suggest that CRM hubs and TADs probably form by different mechanisms. Overall, we hypothesize that CRM hubs represent an additional functional level of genome organization, independent of TADs. This additional layer can also be regulated by priming of enhancers and promoters by paused polymerases^{87–89} or pioneer factors^{54,55}, as well as by chromatin marks^{90,91}. As interactions between Zld CRMs appear before TADs, it is unclear how specificity of CRM interactions may be regulated to favor intra-TAD contacts (Supplementary Note 11).

It is interesting that we observed that interactions between Zld-bound CRMs, as well as interactions between CRMs and cognate promoters, are established very early in pluripotent nuclei, before cell fate commitments. These long-range interactions occur between related CRMs (within *doc*- and *sna*-TADs) as well as between unrelated but Zld-bound CRMs (Fig. 4k and Extended Data Fig. 7k), suggesting that a common link could be their regulation by broad factors such as Zld. Critically, preferential contacts involving Zld-bound CRMs were considerably attenuated upon Zld depletion. We and others have recently shown that Zld forms nuclear hubs in early *Drosophila* embryos^{35,36}, and that Zld hubs are re-established by the end of mitosis, before transcriptional activation. Taken together, our results suggest that Zld fosters the formation of CRM hubs by rendering chromatin accessible during early development, as a first step of cell specification to ensure maximum plasticity. Future work involving the detection of a larger number of CRMs will be needed to elucidate the factors and mechanisms involved in spatial clustering of developmental CRMs into nuclear microenvironments.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-021-00816-z>.

Received: 6 July 2020; Accepted: 9 February 2021;
Published online: 1 April 2021

References

- Bickmore, W. A. The spatial organization of the human genome. *Annu. Rev. Genomics Hum. Genet.* **14**, 67–84 (2013).
- Cavalli, G. & Misteli, T. Functional implications of genome topology. *Nat. Struct. Mol. Biol.* **20**, 290–299 (2013).
- Bulger, M. & Groudine, M. Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144**, 327–339 (2011).
- Alberts, B. et al. *Molecular Biology of the Cell* 6th edn (W. W. Norton & Company, 2017).
- de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
- Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2016).
- Bartman, C. R., Hsu, S. C., Hsiung, C. C.-S., Raj, A. & Blobel, G. A. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol. Cell* **62**, 237–247 (2016).
- Schwarzer, W. & Spitz, F. The architecture of gene expression: integrating dispersed *cis*-regulatory modules into coherent regulatory domains. *Curr. Opin. Genet. Dev.* **27**, 74–82 (2014).
- Symmons, O. et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).
- Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene–enhancer interactions. *Cell* **161**, 1012–1025 (2015).
- Ji, X. et al. 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* **18**, 262–275 (2016).
- Down, J. M. et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
- Ron, G., Globerson, Y., Moran, D. & Kaplan, T. Promoter–enhancer interactions identified from Hi-C data using probabilistic models and hierarchical topological domains. *Nat. Commun.* **8**, 2237 (2017).
- Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
- Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **162**, 687–688 (2015).
- Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–113 (2012).
- Ghavi-Helm, Y. et al. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96–100 (2014).
- Montavon, T. et al. A regulatory archipelago controls *Hox* genes transcription in digits. *Cell* **147**, 1132–1145 (2011).
- Chen, H. et al. Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.* **50**, 1296–1303 (2018).
- Bartman, C. R. et al. Transcriptional burst initiation and polymerase pause release are key control points of transcriptional regulation. *Mol. Cell* **73**, 519–532.e4 (2019).
- Morgan, S. L. et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.* **8**, 15993 (2017).
- Deng, W. et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
- Larsson, A. J. M. et al. Genomic encoding of transcriptional burst kinetics. *Nature* **565**, 251–254 (2019).
- Alexander, J. M. et al. Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. *eLife* **8**, e41769 (2019).
- Benabdallah, N. S. et al. Decreased enhancer–promoter proximity accompanying enhancer activation. *Mol. Cell* **76**, 473–484.e7 (2019).
- McCord, R. P., Kaplan, N. & Giorgetti, L. Chromosome conformation capture and beyond: toward an integrative view of chromosome structure and function. *Mol. Cell* **77**, 688–708 (2020).
- van Steensel, B. & Furlong, E. E. M. The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.* **20**, 327–337 (2019).
- Lim, B. & Levine, M. S. Enhancer–promoter communication: hubs or loops? *Curr. Opin. Genet. Dev.* **67**, 5–9 (2021).
- Allahyar, A. et al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat. Genet.* **50**, 1151–1160 (2018).
- Oudelaar, A. M. et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat. Genet.* **50**, 1744–1751 (2018).
- Oudelaar, A. M. et al. A revised model for promoter competition based on multi-way chromatin interactions at the α -globin locus. *Nat. Commun.* **10**, 5412 (2019).
- Hsieh, T.-H. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553.e8 (2020).
- Tsai, A., Alves, M. R. & Crocker, J. Multi-enhancer transcriptional hubs confer phenotypic robustness. *eLife* **8**, e45325 (2019).
- Baudement, M.-O. et al. High-salt-recovered sequences are associated with the active chromosomal compartment and with large ribonucleoprotein complexes including nuclear bodies. *Genome Res.* **28**, 1733–1746 (2018).
- Mir, M. et al. Dynamic multifactor hubs interact transiently with sites of active transcription in embryos. *eLife* **7**, e40497 (2018).
- Dufourt, J. et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun.* **9**, 5194 (2018).
- Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
- Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e16 (2018).
- Tsai, A. et al. Nuclear microenvironments modulate transcription from low-affinity enhancers. *eLife* **6**, e28975 (2017).
- 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).
- Vallot, A. & Tachibana, K. The emergence of genome architecture and zygotic genome activation. *Curr. Opin. Cell Biol.* **64**, 50–57 (2020).
- Porcher, A. & Dostatni, N. The bicoid morphogen system. *Curr. Biol.* **20**, R249–R254 (2010).
- Bonn, S. & Furlong, E. E. M. *cis*-Regulatory networks during development: a view of *Drosophila*. *Curr. Opin. Genet. Dev.* **18**, 513–520 (2008).
- Stathopoulos, A. & Levine, M. Localized repressors delineate the neurogenic ectoderm in the early *Drosophila* embryo. *Dev. Biol.* **280**, 482–493 (2005).
- Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat. Rev. Genet.* **20**, 221–234 (2019).
- 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).
- Cardozo Gizzi, A. M. et al. Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* <https://doi.org/10.1016/j.molcel.2019.01.011> (2019).
- 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).
- Mateo, L. J. et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
- Beliveau, B. J. et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl Acad. Sci. USA* **109**, 21301–21306 (2012).
- Reim, I. & Frasch, M. The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in *Drosophila*. *Development* **132**, 4911–4925 (2005).
- Hamm, D. C. et al. A conserved maternal-specific repressive domain in Zelda revealed by Cas9-mediated mutagenesis in *Drosophila melanogaster*. *PLoS Genet.* **13**, e1007120 (2017).
- Koenecke, N., Johnston, J., Gaertner, B., Natarajan, M. & Zeitlinger, J. Genome-wide identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol.* **17**, 196 (2016).
- Harrison, M. M., Li, X.-Y., Kaplan, T., Botchan, M. R. & Eisen, M. B. Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS Genet.* **7**, e1002266 (2011).
- Nien, C.-Y. et al. Temporal coordination of gene networks by Zelda in the early *Drosophila* embryo. *PLoS Genet.* **7**, e1002339 (2011).
- Hannon, C. E., Blythe, S. A. & Wieschaus, E. F. Concentration dependent chromatin states induced by the bicoid morphogen gradient. *eLife* **6**, e28275 (2017).
- Li, X.-Y., Harrison, M. M., Villalta, J. E., Kaplan, T. & Eisen, M. B. Establishment of regions of genomic activity during the *Drosophila* maternal to zygotic transition. *eLife* **3**, 127–144.e23 (2014).
- Rivera, J., Keränen, S. V. E., Gallo, S. M. & Halfon, M. S. REDfly: the transcriptional regulatory element database for *Drosophila*. *Nucleic Acids Res.* **47**, D828–D834 (2019).
- Blythe, S. A. & Wieschaus, E. F. Establishment and maintenance of heritable chromatin structure during early embryogenesis. *eLife* **5**, e20148 (2016).
- Senecal, A. et al. Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep.* **8**, 75–83 (2014).
- Hao, N., Shearwin, K. E. & Dodd, I. B. Positive and negative control of enhancer–promoter interactions by other DNA loops generates specificity and tunability. *Cell Rep.* **26**, 2419–2433.e3 (2019).
- Lesne, A., Riposo, J., Roger, P., Cournac, A. & Mozziconacci, J. 3D genome reconstruction from chromosomal contacts. *Nat. Methods* **11**, 1141–1143 (2014).
- Stein, D. S. & Stevens, L. M. Maternal control of the *Drosophila* dorsal–ventral body axis. *Wiley Interdiscip. Rev. Dev. Biol.* **3**, 301–330 (2014).
- Deignan, L. et al. Regulation of the BMP signaling-responsive transcriptional network in the *Drosophila* embryo. *PLoS Genet.* **12**, e1006164 (2016).
- Van Bortle, K., Peterson, A. J., Takenaka, N., O’Connor, M. B. & Corces, V. G. CTCF-dependent co-localization of canonical Smad signaling factors at architectural protein binding sites in *D. melanogaster*. *Cell Cycle* **14**, 2677–2687 (2015).

66. Kvon, E. Z. et al. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature* **512**, 91–95 (2014).
67. Weiss, A. et al. A conserved activation element in BMP signaling during *Drosophila* development. *Nat. Struct. Mol. Biol.* **17**, 69–76 (2010).
68. Sun, Y. et al. Zelda overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. *Genome Res.* **25**, 1703–1714 (2015).
69. Schulz, K. N. et al. Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early *Drosophila* embryo. *Genome Res.* **25**, 1715–1726 (2015).
70. Liang, H.-L. et al. The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. *Nature* **456**, 400–403 (2008).
71. 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).
72. Ing-Simmons, E. et al. Independence of chromatin conformation and gene regulation during *Drosophila* dorsoventral patterning. *Nat. Genet.* <https://doi.org/10.1038/s41588-021-00799-x> (2021).
73. Gisselbrecht, S. S. et al. Transcriptional silencers in *Drosophila* serve a dual role as transcriptional enhancers in alternate cellular contexts. *Mol. Cell* **77**, 324–337.e8 (2020).
74. Chopra, V. S., Kong, N. & Levine, M. Transcriptional repression via antilooping in the *Drosophila* embryo. *Proc. Natl Acad. Sci. USA* **109**, 9460–9464 (2012).
75. Rubin, A. J. et al. Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. *Nat. Genet.* **49**, 1522–1528 (2017).
76. Paliou, C. et al. Preformed chromatin topology assists transcriptional robustness of during limb development. *Proc. Natl Acad. Sci. USA* **116**, 12390–12399 (2019).
77. Cattoni, D. I. et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat. Commun.* **8**, 1753 (2017).
78. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019).
79. Gu, B. et al. Transcription-coupled changes in nuclear mobility of mammalian *cis*-regulatory elements. *Science* **359**, 1050–1055 (2018).
80. Jackson, D. A., Iborra, F. J., Manders, E. M. & Cook, P. R. Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. *Mol. Biol. Cell* **9**, 1523–1536 (1998).
81. Cho, W.-K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
82. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
83. Lesne, A., Baudement, M.-O., Rebouissou, C. & Forné, T. Exploring mammalian genome within phase-separated nuclear bodies: experimental methods and implications for gene expression. *Genes* **10**, 1049 (2019).
84. Lim, B., Heist, T., Levine, M. & Fukaya, T. Visualization of transvection in living *Drosophila* embryos. *Mol. Cell* **70**, 287–296.e6 (2018).
85. 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. USA* <https://doi.org/10.1073/pnas.1701291114> (2017).
86. Zhang, H. et al. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
87. Ghavi-Helm, Y. et al. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* **51**, 1272–1282 (2019).
88. Lagha, M. et al. Paused Pol II coordinates tissue morphogenesis in the *Drosophila* embryo. *Cell* **153**, 976–987 (2013).
89. Saunders, A., Core, L. J., Sutcliffe, C., Lis, J. T. & Ashe, H. L. Extensive polymerase pausing during *Drosophila* axis patterning enables high-level and pliable transcription. *Genes Dev.* **27**, 1146–1158 (2013).
90. Negre, N. et al. A *cis*-regulatory map of the *Drosophila* genome. *Nature* **471**, 527–531 (2011).
91. Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

Drosophila stocks and embryo collection. Fly stocks were maintained at room temperature (RT) with a natural light–dark cycle and raised in standard cornmeal yeast medium. The *yw* or UASp-shRNA-w (BL no. 35573) stocks were used as a control. Flippase stock (BL no. 26902) and CRE stock (BL no. 851) were used for the generation of the CRM_c deletion strain. Zld-depleted embryos were obtained from females from the cross between *nos-Gal4:VP16* (BL no. 4937) and UASp-shRNA-zld⁸⁸, a method with a Zld depletion efficiency of ~90% (ref. 36). After a pre-laying step, flies were allowed to lay eggs for 1.5 h on new yeast 0.4% acetic acid plates. Embryos were then incubated at 25 °C until they reached the desired developmental stage. Embryos were collected and fixed as previously described⁴⁸. Briefly, embryos were dechorionated with 2.6% bleach, rinsed and fixed with a 1:1 mixture of 4% methanol-free formaldehyde in phosphate-buffered saline (PBS) and heptane. Embryos were stored in methanol at –20 °C until further use.

Generation of the CRM_c deletion and reporter lines. A CRISPR (clustered regularly interspaced short palindromic repeats)–Cas9-based strategy was employed to specifically delete CRM_c in a conditional manner. Two flippase recognition target (FRT) sites were inserted flanking an ~860-bp region (3L:9021947-9022805) surrounding the most prominent Zld peaks in CRM_c (impacting barcodes 13–14). Recombination template was cloned into pH-DsRed plasmid (Addgene, catalog no. 1434). For this, a 5′-homology arm (PCR amplified from genomic DNA) was inserted into a vector previously digested with XmaI/NheI. Then, the PCR-amplified 3′-homology arm was inserted after digestion with SpeI/AscI. Finally, the PCR-amplified CRM_c flanked by FRT sequences was cloned after digestion by NotI. Guide RNAs were cloned into pCFD3-dU6:3gRNA (Addgene, catalog no. 49410) and digested by BbsI using annealed oligonucleotides. Recombination template and guide RNAs were injected by BestGene Inc. After obtaining the CRISPR-edited stocks, males were crossed with *CRE/CRE;D⁺/TM3,Sb* virgin females to remove the dsRed marker by the action of a Cre recombinase. Then, *dsRED^{-/-}* males were crossed with *Dp/TM3,Sb* virgin females. Males from this cross were then crossed with *hs-FLP/hs-FLP;Dr/TM3,Sb* virgin females. Larvae from these crosses were heat shocked at 37 °C for 30 min in a water bath for the flippase to be expressed and CRM_c to be deleted. Adult males were then PCR genotyped. Oligonucleotide sequences used for cloning and genotyping are listed in Supplementary Table 2. The deletion removes 287 bp of barcode 13 (5 oligonucleotides out of 50) and 562 bp (10 oligonucleotides out of 50) of barcode 14.

DocFX-lacZ reporter lines were generated by promoter transgenesis of pH-Pelican vectors with CRM fragments in between KpnI and NotI sites analogous to the procedure described in Kahn et al.⁹². Subfragments F5SacIIa and F5SacIIb were generated by removing KpnI–SacII or SacII–NotI fragments, respectively, with subsequent blunt end religation.

Hi-M libraries. Oligopaint libraries, consisting of unique ~35/41-mer sequences with genome homology, were obtained from the Oligopaint public database (<http://genetics.med.harvard.edu/oligopaints>). We selected 20 barcodes in the *doc* locus (3L:8882600..9039000 *Drosophila* release 5 reference genome in all cases) for the low-resolution Hi-M library, 17 barcodes encompassing the *doc*-TAD (3L:8974562..9038920) for the high-resolution Hi-M library and 65 barcodes (2L:15244500..15630000) for the high-resolution *sna* locus library. For each barcode, we used 45–50 probes, covering ~3 kb. An additional fiducial barcode located at least ~1 Mb away was used for drift correction (see below). The coordinates of the targeted genomic regions are listed in Supplementary Table 3.

Each oligonucleotide in the pool consisted of five regions: (1) a 21-mer forward priming region; (2) a 32-mer (low-res library) or two 20-mers separated by an AT sequence (high-res libraries) readout region unique for each barcode; (3) a 35/41-mer genome homology region; (4) a 32-mer (low-res library) or 20-mer (high-res libraries) readout region; and (5) a 21-mer reverse priming region. The designed template oligonucleotide pools were ordered from CustomArray. The procedure to amplify oligonucleotide pools to obtain the primary libraries was as previously described⁴⁸. It involved a five-step procedure consisting of (1) limited-cycle PCR, (2) amplification via T7 in vitro transcription, (3) reverse transcription, (4) alkaline hydrolysis and (5) purification. The sequences of the primers used for amplification of the libraries are listed in Supplementary Table 4.

For the low-resolution library, we employed 20 unique Alexa Fluor-647-labeled sequence oligonucleotides (imaging oligonucleotides), complementary to the readout region present in the primary oligonucleotide. The fluorophore was attached via a disulfide linkage cleavable by the mild reducing agent tris(2-carboxyethyl)phosphine (TCEP), using a previously described strategy⁴⁸. Alternatively, for the high-resolution libraries, we used ‘adapter’ oligonucleotides, consisting of a 20-mer region complementary to the readout sequence that can recognize the barcode being targeted, a 10-mer spacer sequence and a 32-mer region able to bind to a unique Alexa Fluor-647-labeled oligonucleotide (containing a disulfide linkage). In this approach, a single fluorescent oligonucleotide is required⁴⁹. For fiducial barcodes, a noncleavable, rhodamine-labeled oligonucleotide was used. The sequences of the imaging and adapter oligonucleotides are listed in Supplementary Table 5. PCR and reverse transcription primers used in

probe synthesis, as well as adapter oligonucleotides and fluorescently labeled oligonucleotides, were purchased from Integrated DNA Technology. The whole set of Oligopaints used can be found in Supplementary Table 6.

RNA-FISH probes. RNA probes were obtained by in vitro transcription from a vector containing the sequences targeting *sna* (previously described in ref. 47), *doc1*, *doc2* or *doc3* genes in the presence of digoxigenin (DIG) or biotin (BIO) haptens. Vector was linearized before the in vitro transcription with a specific restriction enzyme. RNA probes produced in this manner were then treated with carbonate buffer at 65 °C for 5 min (*sna* probe) or 2 min (*doc1*, *doc2* and *doc3* probes). The information on each probe, including the primers used to clone the target sequences by amplification of genomic DNA, is listed in Supplementary Table 7.

RNA-FISH. In situ hybridization was as described previously⁴⁸, with modifications to allow for the detection of two different species of RNA. The reader is invited to read our detailed protocol in the aforementioned reference. Briefly, fixed embryos were passed through a 1:1 mixture of methanol:ethanol and then pure ethanol. Embryos were then post-fixed with 5% formaldehyde in PBS-T (PBS-T = 0.1% Tween-20 + PBS) for 25 min. Then, embryos were incubated four times with PBS-T for 15 min and permeabilized for 1 h with 0.3% Triton X-100 in PBS. Embryos were rinsed with PBS-T and incubated for 2 h with RHS at 55 °C (where RHS = 50% formamide, 2× saline sodium citrate (SSC), 0.1% Tween-20, 0.05 mg ml⁻¹ of heparin and 0.1 mg ml⁻¹ of salmon sperm). In the meantime, RNA probes were heated at 85 °C for 2 min, transferred to ice for 2 min and then incubated with the embryos in RHS for 16–20 h at 55 °C for RNA hybridization. The next day, embryos were washed four times with RHS at 55 °C and three times with PBS-T at RT. Then, a saturation step was performed with blocking solution (blocking reagent (Sigma-Aldrich, catalog no. 11096176001): 100 mM maleic acid and 150 mM NaCl, pH 7.5) for 45 min.

Then the protocol depends on whether embryos were used for Hi-M (*sna/doc1* double labeling) or to reveal *doc1*, *doc2* and *doc3* expression patterns (Extended Data Fig. 1j). To reveal the expression patterns of *doc* genes, the combination *doc1*-DIG/*doc2*-BIO or *doc2*-DIG/*doc3*-BIO was used. After the saturation step, embryos were incubated with primary antibodies at 1:375 dilution (sheep anti-DIG (Roche, catalog no. 1133089001) and mouse anti-BIO (Life Technologies, catalog no. 03-3700)) overnight at 4 °C. The next day embryos were washed six times in PBS-T for 10 min. Embryos were incubated for 1 h in blocking solution, then for 2 h with secondary antibodies at 1:500 dilution (anti-mouse Alexa Fluor-488 conjugated (Life Technologies, catalog no. A21202)) and anti-sheep Alexa Fluor-555 conjugated (Life Technologies, catalog no. A21436) and washed six times in PBS-T. Finally, embryos were incubated for 10 min with 0.5 mg ml⁻¹ of DAPI solution, washed with PBS-T and mounted in ProLong Diamond Antifade.

For Hi-M, both *sna* and *doc1* probes were DIG labeled. By taking advantage of the differential spatial expression pattern, we labeled both RNAs simultaneously by the combination of both probes during incubation, and the use of a single anti-DIG antibody and a tyramide signal amplification reaction. After RNA hybridization and the saturation step, the activity of endogenous peroxidases was eliminated by incubating with 1% H₂O₂ in PBS-T for 30 min. After rinsing with PBS-T, embryos were incubated overnight at 4 °C with sheep anti-DIG conjugated with POD (Sigma-Aldrich, catalog no. 11207733910) with 1:500 working dilution in PBS-T. The next day, embryos were washed with PBS-T and incubated for 30 min with tyramide-coupled Alexa Fluor-488. Next, H₂O₂ was added to a final concentration of 0.012% over another 30 min. Embryos were washed with PBS-T and stored at 4 °C until further use.

Hybridization of Hi-M primary library. Hybridization followed a previously described protocol⁴⁸. Briefly, embryos were RNase treated for 2 h, permeabilized for 1 h with 0.5% Triton X-100 in PBS and rinsed with sequential dilutions of Triton X-100/pHM buffer to 100% pHM (2× SSC, 0.1 M NaH₂PO₄, pH 7, 0.1% Tween-20 and 50% formamide (v:v)). Embryos in pHM were preheated at 80 °C, the supernatant was aspirated and 30 μl of FHB (50% formamide, 10% dextran sulfate, 2× SSC and 0.5 mg ml⁻¹ of salmon sperm DNA) containing 225 pmol of the primary library was pipetted directly on to the embryos. Mineral oil was added on top and the tube was incubated overnight at 37 °C. The next day, oil was carefully removed and embryos were washed twice for 20 min at 37 °C with 50% formamide, 2× SSC and 0.3% 3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate (CHAPS). Next, embryos were sequentially washed for 20 min at 37 °C with serial dilutions of formamide/PBS-T to 100% PBS-T. Embryos were rinsed with PBS-T and stored in PBS-T at 4 °C until the imaging step.

Imaging system. All experiments were performed on a homemade, wide-field epifluorescence microscope built on a RAMM modular microscope system (Applied Scientific Instrumentation) coupled to a microfluidic device, as described previously^{47,48}. Samples were imaged using an ×60 Plan-Achromat water-immersion objective (numerical aperture = 1.2; Nikon). The objective lens was mounted on a closed-loop piezoelectric stage (Nano-F100, Mad City Labs Inc.). Illumination was provided by four lasers (OBIS-405/488/640 nm and

Sapphire-LP-561 nm, Coherent). Images were acquired using an sCMOS camera (ORCA Flash 4.0V3), with a final pixel size calibrated to 106 nm. A homemade autofocus system was used to correct for axial drift in real time and to maintain the sample in focus as previously described⁴⁷.

A fluidic system was used for automated sequential hybridizations, by computer-controlling a combination of three eight-way valves (HVXM 8-5, Hamilton) and a negative pressure pump (MFCS-EZ, Fluigent) to deliver buffers and secondary readout probes on to a FCS2 flow chamber (Biopetechs). Software-controlled microscope components, including camera, stages, lasers, pump and valves, were run using a homemade software package developed in LabView 2015 (National Instruments).

Acquisition of Hi-M datasets. Embryos were attached to a poly(L-lysine)-coated coverslip and mounted into the FCS2 flow chamber. Fiducial readout probe (25 nM rhodamine-labeled probe, 2× SSC, 40% v:v formamide) was flowed on to the sample and hybridized for 15 min, washed for 10 min with readout washing buffer (2× SSC, 40% v:v formamide) and for 5 min with 2× SSC before injecting 0.5 mg ml⁻¹ of DAPI in PBS to stain nuclei. The imaging buffer (1× PBS, 5% w:v glucose, 0.5 mg ml⁻¹ of glucose oxidase and 0.05 mg ml⁻¹ of catalase) was injected. Subsequently, 10–15 embryos were selected according to developmental stage and orientation, and segmented into a mosaic of multiple fields of view (200 × 200 μm²). After brightfield image recording, z-stacks were taken with 405-, 488- and 561-nm laser illuminations. The z-stacks had a step size of 250 nm with a total range of 15 μm.

Next, the sample was sequentially hybridized with different secondary readout probes, imaged in the rhodamine and the Alexa Fluor-647 channels, and photobleached. For each round of secondary hybridization, the sample was treated with secondary hybridization buffer (25–50 nM imaging oligonucleotide, 2× SSC, 40% v:v formamide, which also included 50 nM of adapter oligonucleotide in the case of the high-res libraries; see Hi-M libraries) for 15 min, then washed with readout washing buffer and with 2× SSC before injecting imaging buffer. After imaging, the fluorescence of the readout probes was extinguished using a chemical bleaching buffer (2× SSC, 50 mM TCEP hydrochloride) for 10 min, and then the sample was washed with 2× SSC for 5 min before a new hybridization cycle started. All buffers were freshly prepared and filtered for each experiment. The imaging buffer used for a single experiment was stored under a layer of mineral oil and renewed every 12–15 h. Further details can be found on our previously published protocol⁴⁸.

Image processing. Our homemade Hi-M microscope produced z-stacks in DCIMG format, which were converted to TIFF using proprietary software from Hamamatsu. TIFF images were then deconvolved using Huygens Professional v.20.04 (Scientific Volume Imaging, <https://svi.nl>). Further analysis steps were performed using a homemade analysis software that implemented the steps described previously⁴⁸. Briefly, images were first z-projected using either sum (DAPI channel) or maximum intensity projection (barcodes, fiducials) and image-based cross-correlation was used to align the fiducial channels. These corrections were then used to align DAPI and barcode images. Next, the positions of the centers of barcodes were detected with subpixel resolution using local maximum-fitting functions from the ASTROPY package⁹³. Nuclei were segmented from projected DAPI images by adaptive local thresholding and watershed filtering⁴⁸. RNA images were segmented by manually drawing polygons over the nuclei displaying a pattern of active transcription. Barcodes and RNA status were then attributed to each single nucleus by using the XY coordinates of the barcodes, the projected DAPI masks of nuclei and the transcriptional status from manual masking. Finally, pairwise distance matrices were calculated for each single nucleus. From the list of pairwise distances obtained from any two barcodes, we calculated the contact probability as the number of nuclei in which pairwise distances were within 250 nm normalized by the number of nuclei containing both barcodes. This definition was similar to that used in other studies^{77,78}, and avoided biases due to uneven barcode detections. Contact frequencies obtained using this pipeline and those using previous pipelines⁴⁸ produced highly correlated results. Image processing was carried out from Linux terminals connected to a server running Linux PopOS 19.10, with four GeForce GTX 1080Ti GPU cards (SCAN computers). Assessment of Hi-M dataset size was done using a bootstrapping approach (Extended Data Fig. 1h).

4M profiles and multi-way interactions. The 4M profiles were obtained by slicing the corresponding Hi-M contact map across a given anchor. Multi-way interactions were obtained by selecting an anchoring barcode and calculating the single-nucleus pairwise distances to all possible pairs of barcodes. If both barcode-anchor distances for a given barcode pair in a single nucleus are below the contact threshold (250 nm), this nucleus is considered to have a three-way interaction for this anchor and barcode pair. The three-way contact frequency is then obtained by dividing the number of nuclei that show a three-way interaction by the number of nuclei where the three barcodes involved in the three-way interaction have been detected. The calculation of three-way interactions does not require the detection of contiguous barcodes, and therefore the calculation of three-way frequencies is not restricted to nuclei displaying all the barcodes.

ShRec3D structures. The 3D topological representations were obtained from Hi-M pairwise distance maps using our own Python implementation of the approaches described by Lesne et al. and Morlot et al. for ShRec3D^{63,94}. Starting from the single-cell pairwise distance matrix, an ensemble pairwise distance matrix was calculated using the first maximum of the kernel density estimation. These pairwise distances were converted into 3D coordinates for each barcode using nonclassical metric multidimensional scaling. When necessary, structures were mirrored and a ball-and-stick representation was rendered with PyMOL (PyMOL Molecular Graphics System, v.2.3 Schrödinger, LLC.).

Genome-wide analysis of Zld-mediated interactions. First, we extracted lists of Zld peaks genome wide. For the Zld peaks used in Fig. 4k–o and Extended Data Fig. 7i–k: datasets from Harrison et al.⁵⁴ (accession no. [GSM763062](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM763062)) were used to extract the autosomal regions bound by Zld at 3 hours post fertilization using their corresponding chromatin immunoprecipitation (ChIP)-sequencing intensity⁵⁴. For the Zld peaks used in Extended Data Fig. 7l: a list of putative enhancers of pre-MBT genes ($N=62$) was obtained by selecting the Zld peaks nearest to the transcription start site of pre-MBT genes⁹⁵ (Supplementary Table 8). BED coordinates were remapped from dm3 (BDGP R5) to dm6 (BDGP R6) using FlyBase's sequence coordinates converter (FB2020_05, released 14 October 2020).

Second, we characterized interactions between Zld-bound regions using the 5-kb resolution Hi-C dataset from Hug et al.⁴⁰. Peaks were sorted and classified into different categories based on protein occupancy. If multiple Zld peaks were contained within the same 5-kb bin, only the one with the highest intensity was considered. After filtering, we analyzed 5,038 bins occupied by Zld in the different autosomal chromosomes. For each biological condition, intra-arm chromosomal contacts were distance normalized by computing the $\log_2(\text{observed/expected})$. The average interaction frequencies at long (>250 kb) or short (≤250 kb) ranges were then ranked in five groups by increasing the Zld ChIP signal. From low to high Zld peak intensity, each group contains, respectively, 3,124, 706, 480, 248 and 480 peaks.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The Oligopaint public database (<http://genetics.med.harvard.edu>) was used to select Oligopaints. Publicly available datasets used in the present study (accession nos. [GSE86966](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE86966), [GSE25180](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE25180), [E-MTAB-4918](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=E-MTAB-4918), [GSM763062](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM763062), [GSE58935](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58935), [GSE16245](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE16245), [GSE68983](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68983), [GSE68654](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68654), [E-MTAB-1673](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=E-MTAB-1673), [GSE62904](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62904) and [GSE65441](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65441)) are detailed in Supplementary Table 9. Data for matrices in Figs. 1–4 and in Extended Data Figs. are publicly available at <https://github.com/NollmannLab/Espinola-Goetz-2021>. Source data are provided with this paper.

Code availability

Code used in this manuscript is available at <https://github.com/NollmannLab/Espinola-Goetz-2021>.

References

- Kahn, T. G. et al. Interdependence of PRC1 and PRC2 for recruitment to polycomb response elements. *Nucleic Acids Res.* **44**, 10132–10149 (2016).
- The Astropy Collaboration et al. The Astropy Project: building an open-science project and status of the v2.0 core package. *Astronom. J.* **156**, 123–1412 (2018).
- Morlot, J.-B., Mozziconacci, J. & Lesne, A. Network concepts for analyzing 3D genome structure from chromosomal contact maps. *EPJ Nonlinear Biomed. Physics* **4**, 2 (2016).
- Chen, K. et al. A global change in RNA polymerase II pausing during the *Drosophila* midblastula transition. *eLife* **2**, e00861 (2013).
- Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).

Acknowledgements

We thank N. Benabdallah, G. Cavalli, T. Forne, T. Robert, J. Bonnet and members of the Lagha and Nollmann laboratories for their critical reading of the manuscript. We thank A. Makrini and D. Cattoni for help with bioinformatic analysis. This project was funded by an ERC Consolidator Grant from the European Union's Horizon 2020 Research and Innovation Program (grant no. 724429 to M.N.). We thank the Bettencourt-Schueller Foundation for their prize Coup d'élan pour la recherche Française, the France-BioImaging infrastructure supported by the French National Research Agency (grant no. ANR-10-INBS-04, Investments for the Future), the Labex EpiGenMed (ANR-10-LABX-12-01) and the *Drosophila* facility (BioCampus Montpellier, CNRS, INSERM, Université de Montpellier, France). M.G. was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation; project no. 431471305). M.L.'s laboratory is supported by an ERC Starting Grant (SyncDev, grant no. 679792) and CNRS. M.B. and O.M. are supported by an FRM PhD fellowship.

Author contributions

A.M.C.G., M.L. and M.N. conceived the study and the design. S.M.E., C.H., M.B. and I.R. acquired the data. M.G., S.M.E., M.B., O.M., I.R. and M.N. analyzed the data. M.G., M.N., J.B.F. and O.M. provided the software. S.M.E., M.G., O.M., I.R., M.B., M.N. and M.L. interpreted the data. M.L. and M.N. wrote the manuscript. J.-B.F., C.H. and I.R. provided the reagents. S.M.E., M.G. and M.N. did the visualization of the study. M.N. and M.L. supervised the study. M.N. and M.L. acquired funds.

Competing interests

The authors declare no competing interests.

Additional information

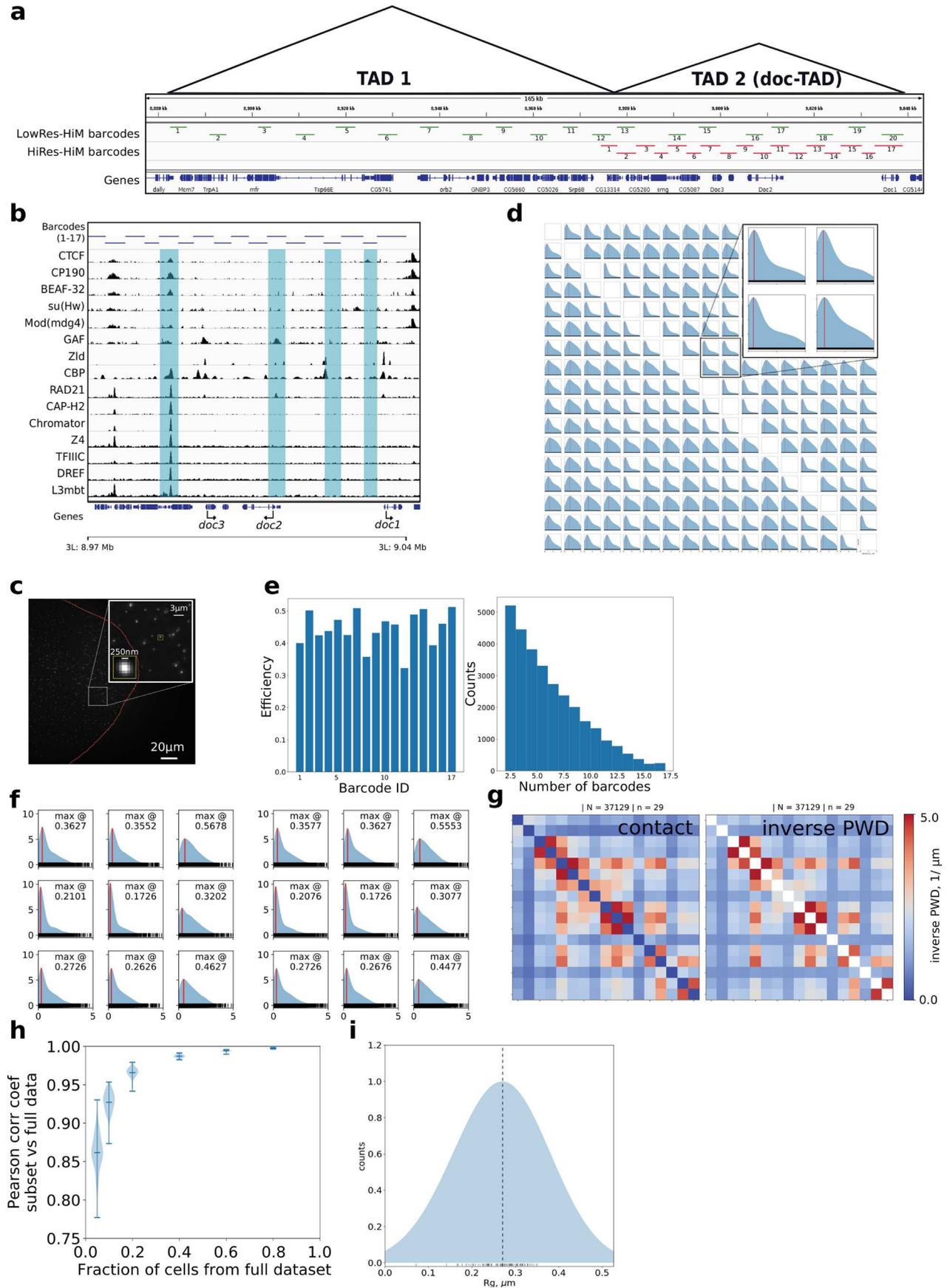
Extended data is available for this paper at <https://doi.org/10.1038/s41588-021-00816-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41588-021-00816-z>.

Correspondence and requests for materials should be addressed to M.L. or M.N.

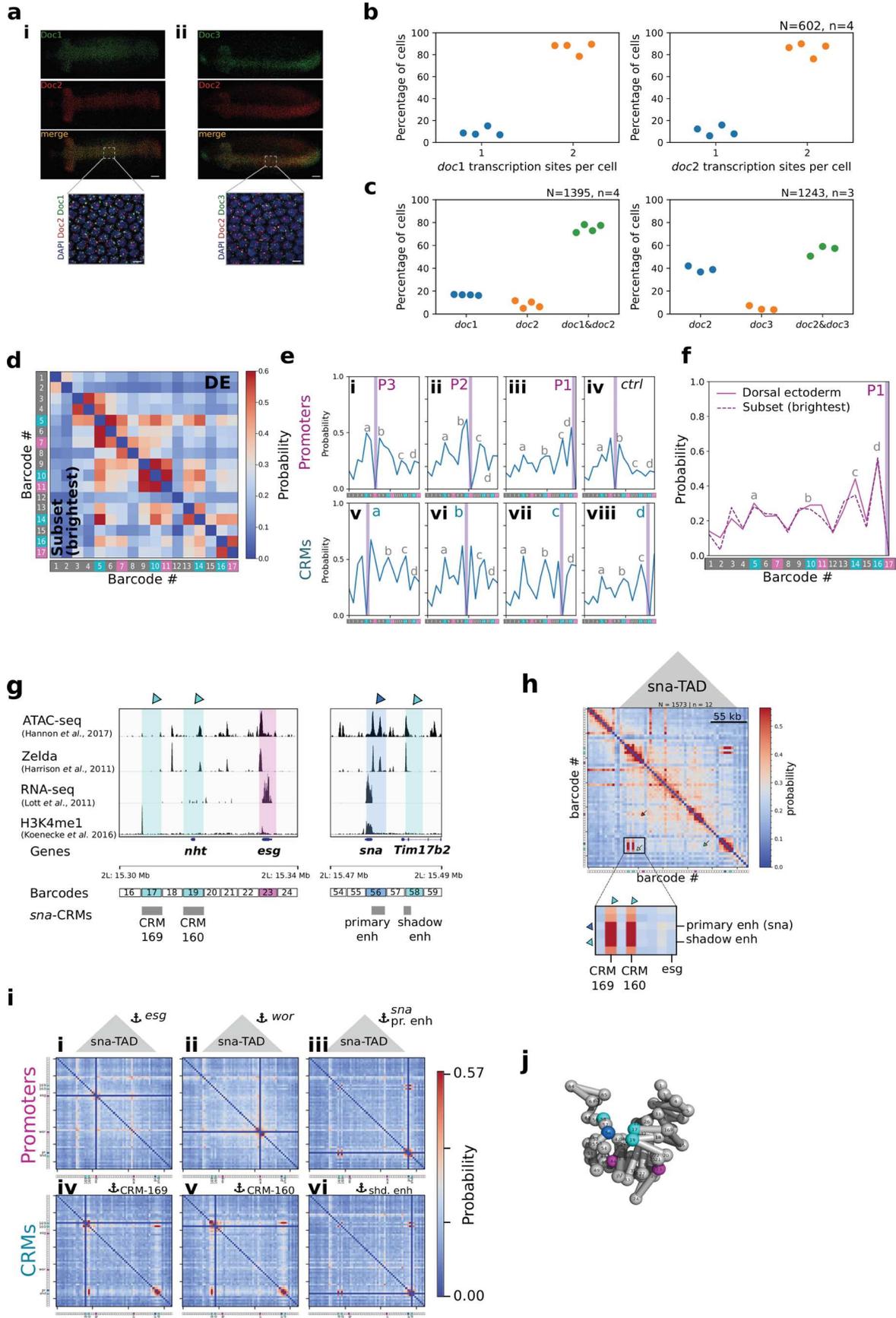
Peer review information *Nature Genetics* thanks Justin Crocker and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.



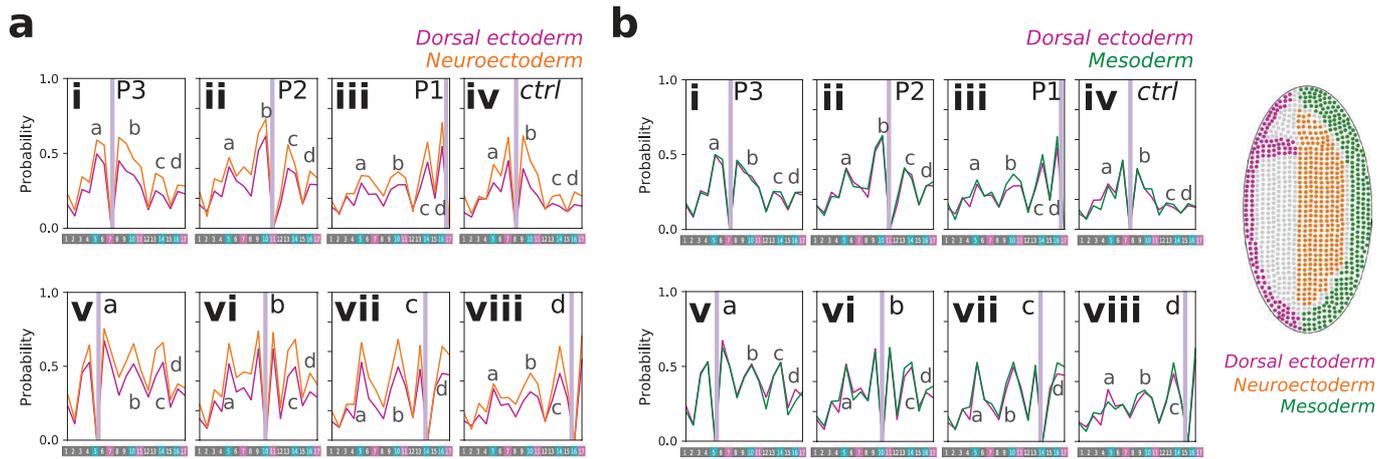
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Hi-M allows high-resolution chromatin tracing in the doc-TAD. **a**, Schematic representation of the genomic positions of barcodes for the low (green) and high (red) resolution *doc* Hi-M libraries. Triangles demarcate the two TADs registered in this genomic region⁹⁰. **b**, Chip-seq profiles for architectural proteins in the doc-TAD^{54,90}. Cis-regulatory modules (CRM_{a-d}) from Fig. 1c are highlighted by blue bars. **c**, Typical maximum intensity projection displaying the fluorescence emission signal from a single barcode in a section of an embryo (outline in red). Emissions from individual barcodes appear as diffraction-limited spots. **d**, Map of pairwise distance distributions for all barcode combinations. The order of the distributions follow that in the Hi-M matrix (Fig. 1f). Blue shade represents a kernel density estimation with a bandwidth of 0.2 μm , red line represents the maximum of the distribution, and black vertical lines on the x-axis represent individual data points. **e**, Efficiency of barcode detection and distribution of number of barcodes detected per cell. **f**, To verify that uneven barcode efficiencies did not affect our results, we plotted the pairwise distance distributions for the full dataset (right) and half the data (left, here cells were randomly chosen). Map of pairwise distributions is centered at the barcode bin (4,13). **g**, Hi-M contact probability map (left) and inverse pairwise distance map (right) for the same experiment (doc-TAD, all cells). $N = 37129$, $n = 29$. **h**, Pearson correlation coefficient of the contact probability of the full doc-TAD Hi-M dataset (nc14, dorsal cells displaying *doc1* expression) against subsets with a fraction of cells. One hundred random subsets were generated for each tested subset size. The central bar indicates the mean and the error bars indicate the extreme values of the distribution. **i**, Distribution of radii of gyration for the doc-TAD calculated from single cells. Blue shade represents a kernel density estimation with a bandwidth of 0.1 μm , black vertical lines on the x-axis represent individual data points. Dashed line represents the maximum of the distribution. The size of the doc-TAD, as estimated from its radius of gyration ($0.27 \pm 0.1 \mu\text{m}$), was comparable with that of TADs of similar genomic sizes⁹⁶.

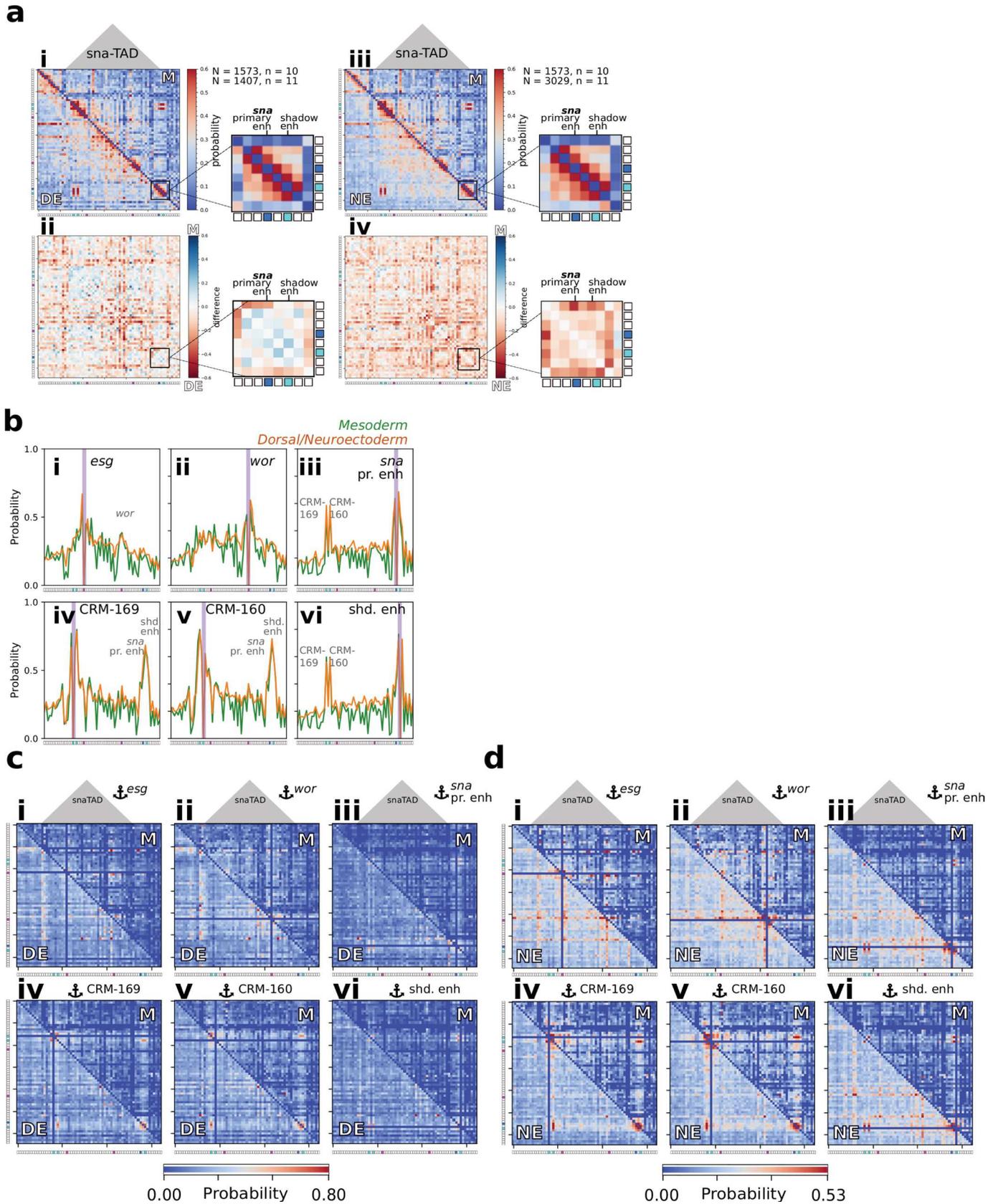


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | *doc* genes are highly co-expressed and *doc* CRMs spatially cluster, as do CRMs in the *sna* locus. **a, RNA-FISH staining for *doc1*, *doc2* and *doc3* in late nc14 embryos. Scale bars: 50 μm / 5 μm (inset). **b**, Percentage of one or two active transcription sites/nuclei for *doc1* and *doc2*. **c**, Percentage of cells displaying active transcription spots from 2-color RNA-FISH imaging of *doc1-doc2* and *doc2-doc3*. Most nuclei (>70%) displayed co-activation of *doc1+doc2* and of *doc2+doc3*. For this latter, a larger percentage of nuclei expressed only *doc2* (~40%), because of the low efficiency of labeling of *doc3* nascent mRNA (small intronic size). **d**, Comparison of contact maps from nuclei displaying at least one active *doc1* RNA-FISH spot (top right matrix) and from a subset (33%) of nuclei displaying the strongest *doc1* RNA-FISH signals. **e**, 4 M profiles derived from Hi-M maps of dorsal ectoderm cells in nc14. **f**, 4 M virtual profile for nuclei displaying at least one active *doc1* RNA-FISH spot (solid pink line) and from a subset of nuclei (33%) with the highest *doc1* signals (dashed dark pink). **g**, Epigenetic profile of selected regions around the *esg* and *sna* genes within the *sna*-TAD. Accessibility (ATAC-seq), pioneer factor binding (Zelda), transcriptional activity (RNA-seq), chromatin marks for active enhancers marks (H3K4me1), and for the transcriptional activators Dorsal, Zen and Mad are shown. A subset of barcodes were annotated as cis-regulatory modules (shown in cyan): CRM₁₆₉ harbors the canonical H3K4me1 active enhancer mark; CRM_{160r} and shadow *sna* enhancer were described in the RedFly database. Magenta barcode harbours the *esg* promoter and the blue barcode contains the *sna* promoter and its primary enhancer. See Supplementary Table 1 for more details. **h**, Hi-M contact probability map of the *sna* locus. Yellow arrow shows interactions between CRMs, red arrow between CRMs and promoters, and green arrow between promoters. **i**, Multi-way interactions between promoters (panels i-iii) and CRMs (panels iv-vi). Number of nuclei and embryos examined as in panel r. **j**, 3D topological representation of the *sna*-TAD. Bead colors are as in panel d. Barcode 44 contains the *wor* promoter.**

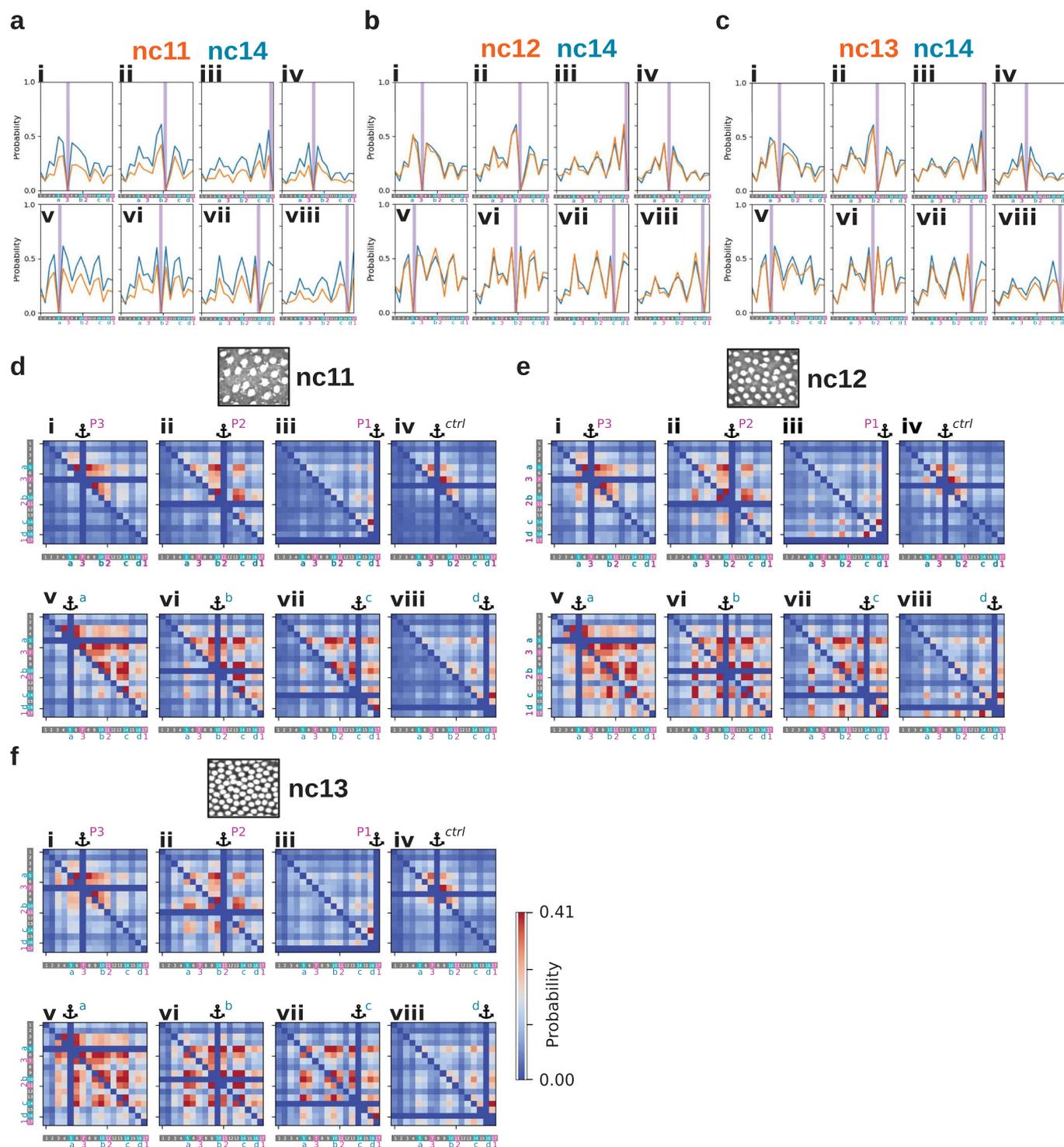


Extended Data Fig. 3 | doc CRM loops are similar between three presumptive tissues. a, Comparison of 4 M profiles between DE (magenta) and NE (orange) for different anchors within the doc-TAD (Panels i-iii: promoters. Panel iv: control. Panels v-viii: CRMs). Anchors are indicated by vertical purple lines. Peaks in the profiles are annotated with the corresponding CRMs (a-d) **b**, Comparison of 4 M profiles between DE (magenta) and M (green) for different anchors within the doc-TAD (Panels i-iii: promoters. Panel iv: control. Panels v-viii: CRMs). Anchors are indicated by vertical purple lines. Peaks in the profiles are annotated with the corresponding CRMs (a-d). Right panel: scheme indicating the three presumptive tissues.

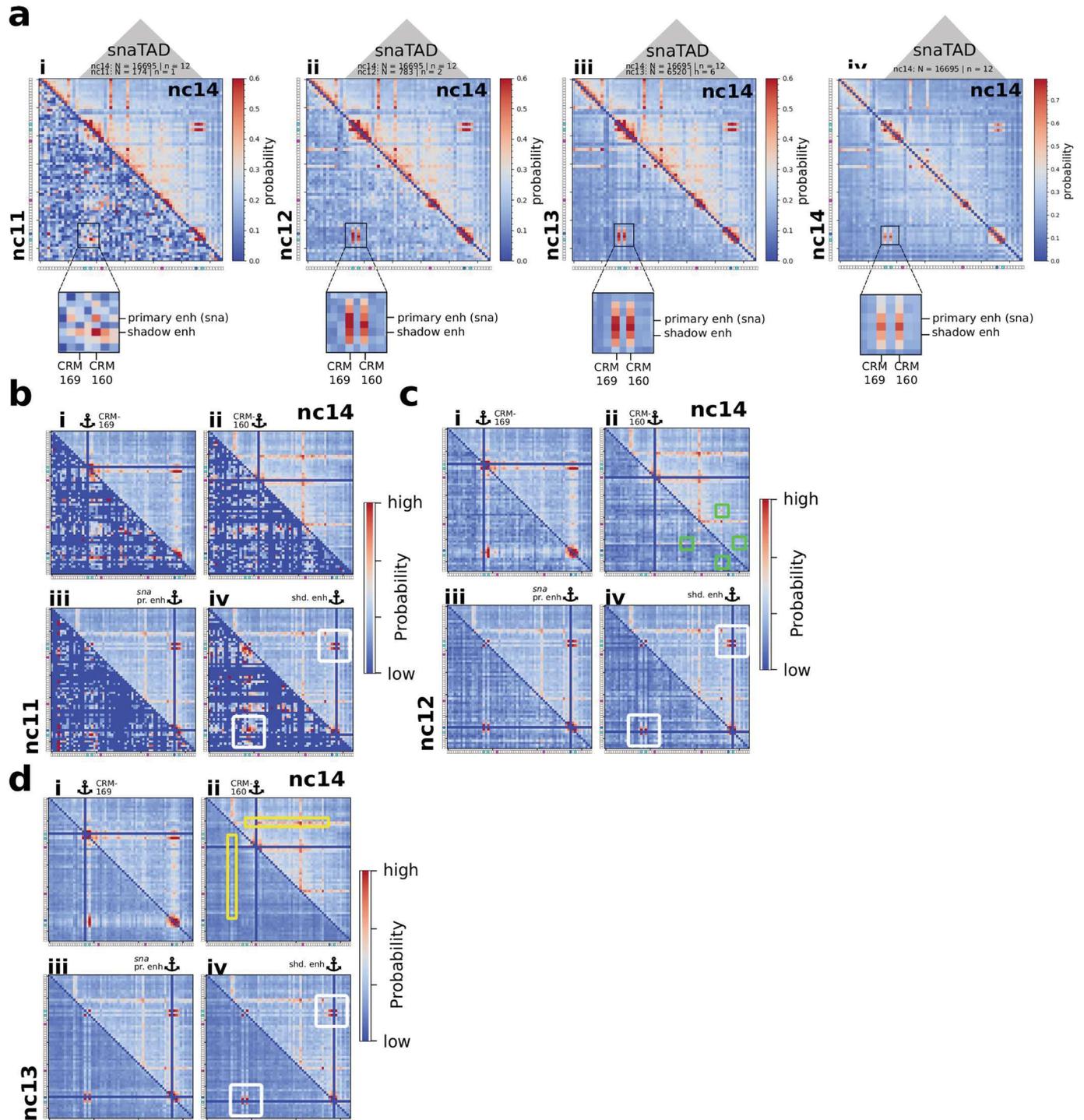


Extended Data Fig. 4 | See next page for caption.

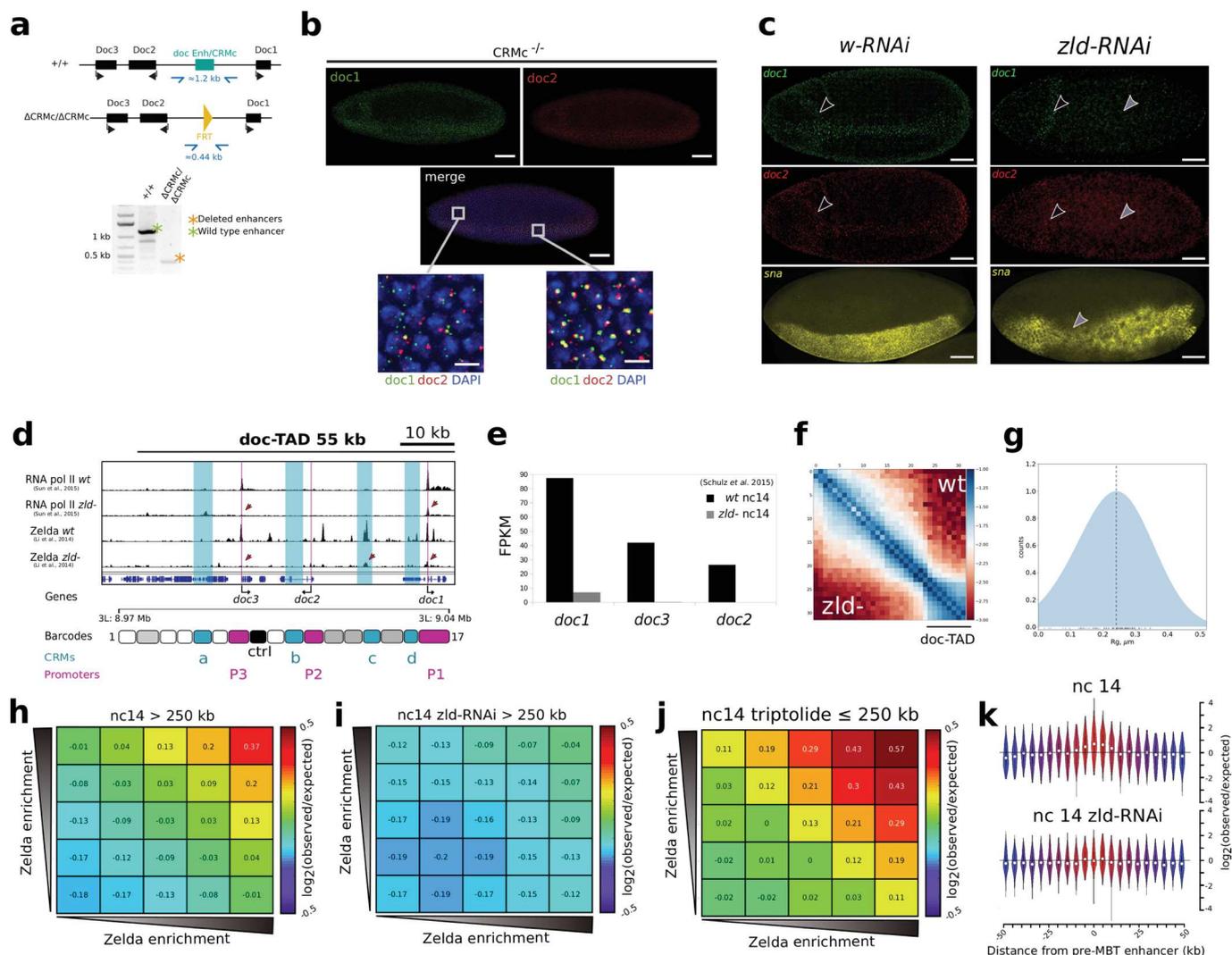
Extended Data Fig. 4 | *sna* CRM loops are similar between three presumptive tissues. a, Panel i: Hi-M contact probability map of the *sna* locus for M (upper-right map) versus DE (lower-left map). Inset show a magnification of the region around *sna*. Panel ii: Same but for the difference between M and DE Hi-M maps. Blue indicates larger contact probabilities in M whereas red indicates larger contact probabilities in DE. Panel iii: Similar to panel i, but for M (upper-right map) versus NE (lower-left map). Panel iv: Similar to panel ii, but for M versus NE. N: number of nuclei. n: number of embryos. **b**, Comparison of 4 M profiles between M (green) and DE/NE (orange). Anchors within the *sna*-TAD are indicated in each panel by vertical purple lines. A subset of peaks is annotated using the nomenclature from Fig. 1d. **c**, Comparisons of multi-way maps for M (upper-right map) versus DE (lower-left map) in the *sna* locus using the anchors indicated in each panel by pictograms and dark blue crosses. Maps are color-coded according to the scale bar on the right. Number of embryos and nuclei as in panel c. **d**, Similar to panel e, but for M (upper-right map) versus NE (lower-left map). Number of embryos and nuclei as in panel c.



Extended Data Fig. 5 | doc CRM loops are established early in development. **a**, Comparison of 4 M profiles between embryos in nc14 (blue lines) and nc11 (orange) for different anchors within the doc-TAD (Panels i-iii: promoters. Panel iv: control. Panels v-viii: CRMs). The position of the anchor is indicated by a vertical purple line. Peaks in the profiles are annotated with the corresponding CRMs (a-d). **b**, Similar to panel A, but comparing 4 M profiles between embryos in nc14 (blue lines) and nc12 (orange). **c**, Similar to panel A, but comparing 4 M profiles between embryos in nc14 (blue lines) and nc13 (orange). **d**, Comparison of multi-way interaction matrices of nc14 (upper-right map) and nc11 (lower-left map). Anchors (dark blue crosses) are as follows: *doc3*, *doc3*, *doc1* promoters (panels i-iii), control region (panel iv), CRM_{a-d} (panels v to viii). Representative image of DAPI-stained nuclei for nc11 is shown on top. Barcodes are shown on the left and bottom of multi-way maps. Number of nuclei (nc11): N = 37129, number of embryos (nc11); n = 4. Number of nuclei (nc14): N = 2154, number of embryos (nc14); n = 29. **e**, Similar to panel d, but for nc14 (upper-right map) and nc12 (lower-left map). Representative image of DAPI-stained nuclei for nc12 is shown on top. Number of nuclei (nc12): N = 2154, number of embryos (nc12); n = 4. **f**, Similar to panel d, but for nc14 (upper-right map) and nc13 (lower-left map). Representative image of DAPI-stained nuclei for nc13 is shown on top. Number of nuclei (nc13): N = 7597, number of embryos (nc13); n = 8.



Extended Data Fig. 6 | *sna* CRM loops are established early in development. **a**, Comparison of Hi-M contact probability maps in the *sna* locus for nc14 (upper-right map) and nc11 (panel i), nc12 (panel ii), nc13 (panel iii) and 14 (panel iv) (lower-left maps). Maps are color-coded according to the scale bar on the right. Inset on the bottom of each map shows a magnification of the region around *esg* and *sna* CRMs (see Extended Data Fig. 2 and Supplementary Table 1). **b**, Comparison of multi-way contact maps between nc14 (upper-right maps) and nc11 (lower-left maps). Maps are color-coded according to the scale bar on the right. The position of anchors are indicated by dark blue crosses. White boxes indicate contacts already present at nc11 that persist through nc14. Number of nuclei and embryos examined as indicated in panel a. **c**, Similar to panel b, but comparing nc14 to nc12. Green boxes indicate contacts that emerge at nc12 and persist at nc14. Number of nuclei and embryos examined as indicated in panel a. **d**, Similar to panel b, but comparing nc14 to nc13. Yellow boxes indicate interactions that appear at nc13 (at the TAD border). Number of nuclei and embryos examined as indicated in panel a.



Extended Data Fig. 7 | Perturbation of gene expression and CRM loops by enhancer deletion and Zld depletion. **a**, Scheme of the wild type *doc* locus (+/+) and the *doc* locus after CRISPR/Cas9 genome editing (Δ CRMc/ Δ CRMc). *doc* enhancer/CRM_c, FRT sequence and primers used for genotyping are in teal, yellow and blue, respectively. Genotyping PCR products on agarose gel electrophoresis are shown in the lower panel. Orange and green stars correspond to the bands of the expected sizes after amplification using primers flanking the *doc* enhancer/CRM_c sequence. See Methods for further details. **b**, RNA-FISH imaging of *doc1* and *doc2* in the CRM_c-deletion mutant. Scale bars: 50 μ m/ 5 μ m (inset). **c**, RNA-FISH imaging of *doc1*, *doc2* and *sna* in control (RNAi *white*) and RNAi *Zld* embryos. Black arrows show the *doc1* and *doc2* expression patterns in the anterior part of the embryo. Grey arrows indicate the absence (*doc1*, *doc2*) or perturbation (*sna*) of gene expression patterns in RNAi-*Zld* embryos. Scale bar: 50 μ m. **d**, Tracks for pioneer factor binding (Zelda) and RNA Pol2 binding in the *doc*-TAD. See Supplementary Table 1 for assignment of CRM_{b-d}. **e**, Transcription levels (RNAseq) of *doc1*, *doc2* and *doc3* in wild-type versus *zld*- embryos⁶⁹. **f**, Hi-C matrix for a genomic region containing *doc*-TAD in wild-type and *Zld*-depleted embryos. Data from Hug et al. (2017)⁴⁰. **g**, Distribution of radius of gyration for the *doc*-TAD in *Zld*-depleted embryos (see Extended Fig. 1i for wild-type). **h-j**, Log₂(observed/expected) average contact frequencies between Zelda bound regions at long-range distances (> 250 kb) ranked by increasing Zelda enrichment in nc14 (panel g), nc14 *zld*-RNAi (panel h) and at short-range (\leq 250 kb) in nc14 triptolide-treated embryos (panel i). **k**, Violin plot of intragroup Log₂(observed/expected) distribution between 62 selected pre-MBT enhancers and neighbouring sequences (\pm 5 kb) in nc14 (upper panel) and nc14 *zld*-RNAi (lower panel). The central white marker indicates the median and the vertical black lines indicate the extreme values of the distribution. The coordinates of enhancers and closest pre-MBT genes are listed in Supplementary Table 8.

Supplementary information

***Cis*-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early *Drosophila* development**

In the format provided by the authors and unedited

SUPPLEMENTARY INFORMATION RELATED TO

Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development

Sergio Martin Espinola^{1*}, Markus Götz^{1*}, Maelle Bellec², Olivier Messina^{1,2}, Jean-Bernard Fiche¹, Christophe Houbron¹, Matthieu Dejean², Ingolf Reim³, Andrés M. Cardozo Gizzi⁴, Mounia Lagha^{2#}, Marcelo Nollmann^{1#}

* *Co-first authors*

Correspondence: marcelo.nollmann@cbs.cnrs.fr, mounia.lagha@igmm.cnrs.fr

SUPPLEMENTARY NOTE 1.

Encouraged by these results, we applied a similar procedure to the *sna*-TAD, containing a family of paralogous genes encoding the zinc finger transcription factor *snail* (*sna*), in addition to *worniu* (*wor*) and *escargot* (*esg*) genes, as well as multiple CRMs (Extended Data Fig. 2g). Remarkably, loops between CRMs were also highly common in this locus, particularly between known enhancers of *sna* and of *esg* (Extended Data Fig. 2g,h, yellow arrow, Supplementary Table 1), but also between enhancer/promoters and promoter/promoter (Extended Data Fig. 2g,h, red and green arrows, respectively).

SUPPLEMENTARY NOTE 2.

Critically, CRM hubs can but do not tend to contain multiple promoters. For example, the putative shared enhancer CRM_c, located at 10 kb and 11 kb from *doc1* and *doc2* TSS respectively, is contacted in most nuclei by other CRMs (yellow arrows, Fig. 1h iii) and is contacted by multiple simultaneous promoters less often (green arrows, Fig. 1h iii). These observations are consistent with our previous analyses showing that formation of promoter-promoter loops and promoter clusters is uncommon (Fig. 1f,g and Extended Data 2e). Taken together, our results are inconsistent with: (a) promoters forming mutually-exclusive interactions with enhancers; and (2) multiple promoters coming together in space to share a common enhancer. Instead, our data suggest a model in which promoters contact CRM hubs containing multiple enhancers, as observed recently in bacteria and in mouse cells ^{1,2}.

SUPPLEMENTARY NOTE 3.

To test whether this surprising similarity between CRM loops in alternative cell fates could be observed in other genomic regions, we explored chromatin organization in the *sna* locus. We obtained similar networks of 2-way and 3-way interactions in mesoderm nuclei versus nuclei in other tissues (neuro and dorsal- ectoderms, Extended Data Fig. 4a-d). The major differences in Hi-M contact maps between tissues in the *sna*-TAD arose from a depletion of interactions in the mesoderm with respect to those in the dorsal- and neuro-ectoderms (see red stripes, Extended Data Fig. 4a). However, local chromatin structure (e.g. insets of Extended Data Fig. 4a), long-range E-E and E-P interactions (e.g. between *sna* and *esg* enhancers, Extended Data Fig. 4a) and 3-way networks (Extended Data Fig. 4c) appear mostly unchanged between tissues. Overall, these results indicate that very similar networks of preferential interactions between CRMs within a TAD are present in nuclei where transcription is either active or silent.

SUPPLEMENTARY NOTE 4.

Notably, the four CRMs for which we observed looping interactions (CRM_{a-d}) displayed strong binding of 'Mothers against dpp' (Mad) and Zerknullt (Zen), two transcriptional activators of *doc* genes that tend to localize specifically to the dorsal ectoderm at nc14 ^{3,4} (Fig. 2g). Therefore, contacts between *doc* promoters and CRM hubs in the dorsal ectoderm would presumably facilitate transcriptional activation (Fig. 2h). In contrast, in the mesoderm and neuroectoderm, CRM_{a-d} tend to be occupied by spatially-localized transcriptional repressors: Sna in the mesoderm, and Brinker (Brk)/Schnurri (Shn) in the neuroectoderm ⁵ (Fig. 2g) ⁶⁻⁸.

SUPPLEMENTARY NOTE 5.

We also analysed whether CRM interactions arose during early development for *sna*-TAD, which also emerges at nc14 ⁹. Remarkably, pairwise contacts between *esg* and *sna* enhancers/promoters appeared at nc11 and reached their maximum intensities between nc 12-13 (Extended Data Fig. 6a). As for the *doc*-TAD, we observed that the strongest 3-way contacts visible at nc14 are already present at nc11 (e.g. *sna* enhancers and the *nht* locus,

white boxes, Extended Data Fig. 6b), while other 3-way interactions were established at later cycles and persisted until nc14 (e.g. green boxes, Extended Data Fig. 6c). In particular, 3-way interactions involving the TAD border were the last to be acquired (nc13-nc14, e.g. yellow box, Extended Data Fig. 6d). All in all, these data indicate that pairwise looping interactions between CRMs in *doc* and *sna*-TADs are established from nc11 (or before) while 3-way interactions are progressively acquired. Importantly, both pairwise and multi-way looping interactions are formed before the emergence of TADs.

SUPPLEMENTARY NOTE 6.

Importantly, the pattern of expression of the *doc* genes and of *sna* were perturbed in Zld depleted embryos (Extended Data Fig. 7c). In addition, removal of Zld led to a decrease in Pol 2 binding (Extended Data Fig. 7d), to a decrease in *doc1-3* transcription (Extended Data Fig. 7e), and to an overall reduction in chromatin accessibility, with CRMs and promoters being particularly impacted by Zld depletion (e.g. CRM_{c-d}, gray arrows in Fig. 4f)^{10,11}. These experiments are consistent with CRM redundancy in the *doc* region, and suggest that a subset of Zld-independent regulatory elements can activate *doc* expression but cannot reproduce the whole wild-type pattern. Overall, these results indicate that depletion of Zld leads to important changes in the levels and patterns of *doc* expression as well as in the binding of Pol 2.

SUPPLEMENTARY NOTE 7.

CRM_a exhibited a mild increase in ATAC-seq signal upon Zld depletion (Fig. 4f) and displayed a very similar pattern of interactions in Zld depleted and control embryos (Fig. 4h, left panel).

SUPPLEMENTARY NOTE 8.

To provide further support for the role of Zld in the formation of Zld-bound CRM-CRM interactions, we calculated the intra-arm interactions between a subset of 62 Zld-bound pre-MBT enhancers¹². Zld-bound CRMs interacted preferentially, and their interaction gradually decreased as anchors were shifted from the position of the Zld peak (Extended Data Fig. 7k, upper panel). Preferential interactions centered at Zld peaks vanished in RNAi-Zld embryos (Extended Data Fig. 7k, lower panel).

SUPPLEMENTARY NOTE 9.

This study provides clear evidence of the advantages of imaging-based technologies to detect multi-way chromatin loops and transcriptional output with spatial resolution. First, Hi-M enables the detection of interactions in cells within different tissues and in distinct transcriptional states in a single experiment. Second, this tool may serve to discover novel *cis*-regulatory modules, and to assess whether predicted CRMs actually contact their target promoters and measure its activity in a given tissue and at a specific developmental time, without the burden of genetic manipulation. A current limitation of these methods is the genomic resolution, defined as the size of the genomic region covered by a barcode (~2kb¹³). This can be a limitation for compact genomes (e.g. *Drosophila*) where promoter and proximal regulatory regions often fall within the same barcode. Third, the ability to detect multi-way interactions will be important to further dissect the mechanisms of transcriptional control by distal CRMs, as well as mechanisms of transcriptional co-regulation. Finally, combination of these approaches with opto-genetic manipulations may open exciting avenues to refine our understanding of spatio-temporal control of gene expression during development.

SUPPLEMENTARY NOTE 10.

For this, the authors perturbed the dorsoventral patterning network to obtain embryos where most nuclei behaved either as mesoderm, neuroectoderm or dorsal ectoderm. It is important to note, however, that the principles of sequencing-based and microscopy-based techniques are very different: the former isolates protein-mediated chromatin interactions at a defined length scale while the latter measures spatial proximity. Thus, care must be taken when performing direct comparisons between these distinct technologies. Despite these differences, Hi-C and micro-C also displayed chromatin organization patterns that were strikingly similar amongst tissues.

SUPPLEMENTARY NOTE 11.

Our results suggest that Zld participates in the regulation of CRM interactions (Fig. 4k-o). In addition, we note that the binding of developmental transcriptional activators (e.g. Mad, Zen) is often highly inhomogeneous, and tends to localize within TADs and excluded from neighboring TADs. Finally, in *Drosophila* architectural proteins (e.g. BEAF, GAF) bind preferentially to TAD borders^{14,15}, and TADs are highly correlated to epigenetic domains¹⁴. Thus, it is likely that a combination of these factors may be involved in providing specificity to CRM interaction networks.

SUPPLEMENTARY REFERENCES

1. Hao, N., Shearwin, K. E. & Dodd, I. B. Positive and Negative Control of Enhancer-Promoter Interactions by Other DNA Loops Generates Specificity and Tunability. *Cell Reports* vol. 26 2419–2433.e3 (2019).
2. Oudelaar, A. M. *et al.* A revised model for promoter competition based on multi-way chromatin interactions at the α -globin locus. *Nat. Commun.* **10**, 5412 (2019).
3. Stein, D. S. & Stevens, L. M. Maternal control of the *Drosophila* dorsal-ventral body axis. *Wiley Interdiscip. Rev. Dev. Biol.* **3**, 301–330 (2014).
4. Hammonds, A. S. *et al.* Spatial expression of transcription factors in *Drosophila* embryonic organ development. *Genome Biol.* **14**, R140 (2013).
5. Gray, S., Cai, H., Barolo, S. & Levine, M. Transcriptional repression in the *Drosophila* embryo. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **349**, 257–262 (1995).
6. Koenecke, N., Johnston, J., Gaertner, B., Natarajan, M. & Zeitlinger, J. Genome-wide identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol.* **17**, 196 (2016).
7. Van Bortle, K., Peterson, A. J., Takenaka, N., O'Connor, M. B. & Corces, V. G.

- CTCF-dependent co-localization of canonical Smad signaling factors at architectural protein binding sites in *D. melanogaster*. *Cell Cycle* **14**, 2677–2687 (2015).
8. Deignan, L. *et al.* Regulation of the BMP Signaling-Responsive Transcriptional Network in the *Drosophila* Embryo. *PLoS Genet.* **12**, e1006164 (2016).
 9. Cardozo Gizzi, A. M. *et al.* Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. *Mol. Cell* (2019) doi:10.1016/j.molcel.2019.01.011.
 10. Hannon, C. E., Blythe, S. A. & Wieschaus, E. F. Concentration dependent chromatin states induced by the bicoid morphogen gradient. *Elife* **6**, (2017).
 11. Schulz, K. N. *et al.* Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early *Drosophila* embryo. *Genome Res.* **25**, 1715–1726 (2015).
 12. Chen, K. *et al.* A global change in RNA polymerase II pausing during the *Drosophila* midblastula transition. *Elife* **2**, e00861 (2013).
 13. Mateo, L. J. *et al.* Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
 14. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
 15. Hou, C., Li, L., Qin, Z. & Corces, V. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* **48**, 471–484 (2012).

4. Publication: Remembering the past, mitotic bookmarking in a developing embryo

This first publication was done in the first year of my PhD and written by Dr. Mounia Lagha and myself. This review aimed at summarizing what is known about mitotic bookmarking in embryos and to ask what could be the next questions to investigate in this field. By writing this review, we realized that not much is known about the exact role of mitotic bookmarking during development, even though few studies pointed the existence of such phenomenon in embryos from *Drosophila* to mouse.

Remembering the past: Mitotic bookmarking in a developing embryo

Maelle Bellec¹, Ovidiu Radulescu² and Mounia Lagha¹

Abstract

During development, transcriptional properties of progenitor cells are stably propagated across multiple cellular divisions. Yet, at each division, chromatin faces structural constraints imposed by the important nuclear re-organization operating during mitosis. It is now clear that not all transcriptional regulators are ejected during mitosis, but rather that a subset of transcription factors, chromatin regulators and epigenetic histone marks are able to 'bookmark' specific loci, thereby providing a mitotic memory. Here we review mechanisms of mitotic bookmarking and discuss their impact on transcriptional dynamics in the context of multicellular developing embryos. We document recent discoveries and technological advances, and present current mathematical models of short-term transcriptional memory.

Addresses

¹ Institut de Genetique Moleculaire de Montpellier, University of Montpellier, CNRS, Montpellier, France

² DIMNP, UMR CNRS 5235, University of Montpellier, Montpellier, France

Corresponding author: Lagha, Mounia (Mounia.lagha@igmm.cnrs.fr)

Current Opinion in Systems Biology 2018, 11:41–49

This review comes from a themed issue on **Development and differentiation**

Edited by **Stanislav Shvartsman** and **Robert Zinzen**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 8 August 2018

<https://doi.org/10.1016/j.coisb.2018.08.003>

2452-3100/© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords

Mitotic bookmarking, Epigenetics, Transcription factors, Development, Memory, Histone modifications, Mathematical modeling, Transcription.

Introduction

During development of multicellular organisms, transcriptional programs must be faithfully transmitted during each cellular division to ensure cell fate maintenance. Two steps of the cell cycle, mitosis and replication exhibit particular topological constraints, hindering stable inheritance of transcriptional repertoires. However it is now well established that mitosis is not always accompanied by a total erasure of past chromatin states, from mother to daughter cells [1]. On the contrary,

memory of active and repressed chromatin states exists, both at the short-term between successive mitoses and in longer time scale through multiple generations [2].

In this review, we focus on recent advances in understanding how active states are transmitted through cellular divisions *in vivo*. We will discuss potential mechanisms and possible mathematical models of mitotic bookmarking and present their consequences in developing multicellular embryos.

Mechanisms of short-term memory have also been analyzed in the context of cultured cell lines and recent reviews summarize these findings [1,3]. Similarly, long-term memory of repressed states mediated by Polycomb family has been intensively reviewed and will not be discussed [4].

Mitosis, a challenging chromatin environment?

During mitosis, genome undergoes a dramatic metamorphose, resulting in clear morphological features of mitotic cells. This chromatin landscape was long thought to represent a hostile environment, incompatible with transcription and transcription factor binding [5].

However, recent studies using sensitive techniques nuance this view. While chromatin folding and compartmentalization are wiped out during mitosis [3,6], local accessibility can be maintained. ATAC-Seq experiments on mitotic *Drosophila* embryos revealed that patterns of accessibility are largely maintained through mitosis for several regulatory elements [7]. One may think that this is peculiar to the early fly embryo, where divisions are particularly fast. However genome accessibility is also widely preserved during mitosis in murine erythroblast cells, yet with local loci specific modulations [8]. Mitotic chromatin landscape has not yet been examined in vertebrate embryos, probably because of the relatively rapid loss in synchrony of divisions, operating as early as the 4-cell stage in zebrafish embryos [9]. This technical issue should be bypassed by recent developments of single cell technologies. Indeed, single cell Hi-C experiments were performed in Embryonic Stem Cells (ESCs) throughout the cell cycle, thus revealing a much more dynamic picture of chromosome organization than previous bulk experiments with unsynchronized cells [10].

Mitotic bookmarking and development

Transmission of chromatin states from mother to daughter cells can be achieved passively or *via* active mechanisms. Among known active supports of mitotic memory, stand three classes of regulators: epigenetic histone marks/chromatin regulators, general transcription factors (GTF) and sequence specific transcription factors (TF) (Figure 1).

The vast majority of experiments aiming to assess mitotic retention of these particular factors has been performed in cultured cells with more or less conflicting conclusions, partly due to the level of purity of the mitotic population, cross linking conditions and to the methods of detection (global decoration by imaging versus specific binding assessed by ChIP). In this section, rather than providing an exhaustive survey of potential mitotic bookmarkers [1], we focus on potential supports for transmission of active chromatin states in the context of multicellular developing embryos (Figure 2).

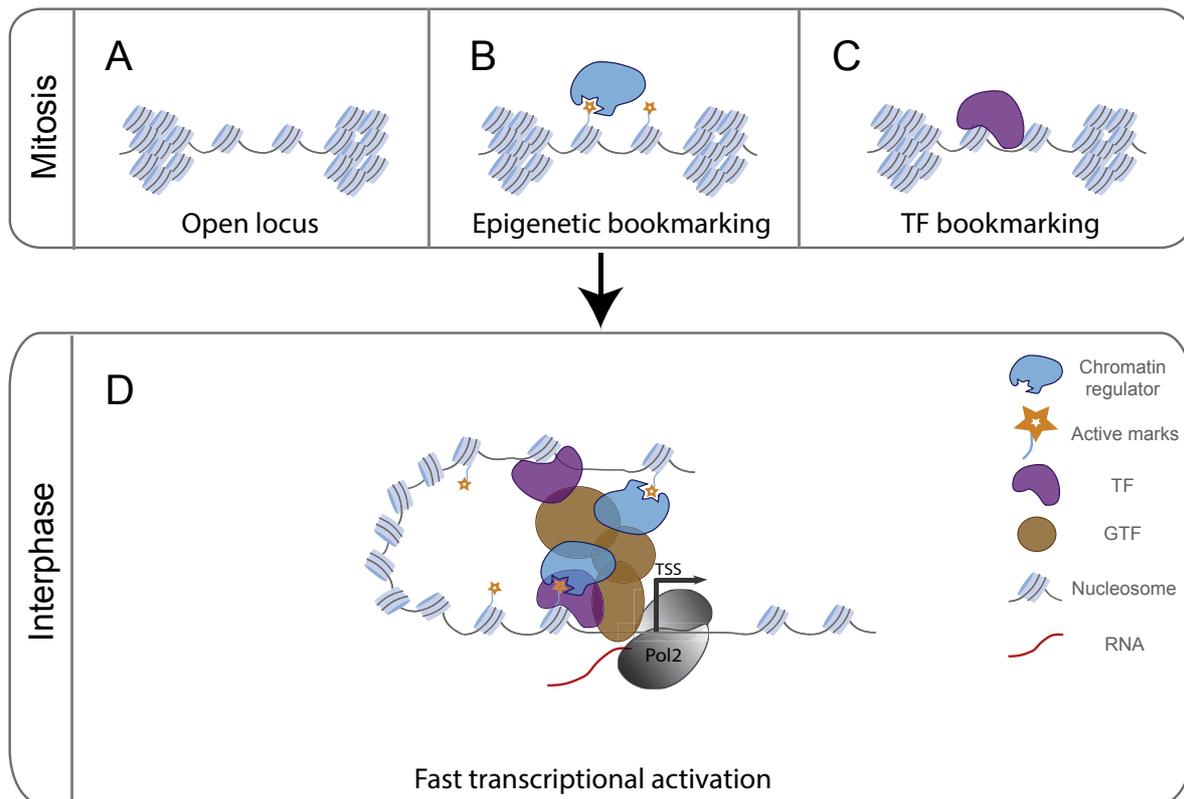
Passive mechanisms of mitotic bookmarking

So far, clear evidences of passive supports of memory are lacking. In a recent work, by monitoring transcriptional activation in living early *Drosophila* embryos, it was shown that experiencing transcription prior to mitosis does not always lead to a rapid post-mitotic activation [11]. With a mesodermal enhancer, memory of active transcriptional states is unequivocally occurring [12], while with a dorsal ectoderm enhancer, memory is not detected [11].

However, given the widespread maintenance of chromatin accessibility during mitosis in the fly embryo, a passive mechanism is plausible (Figure 1).

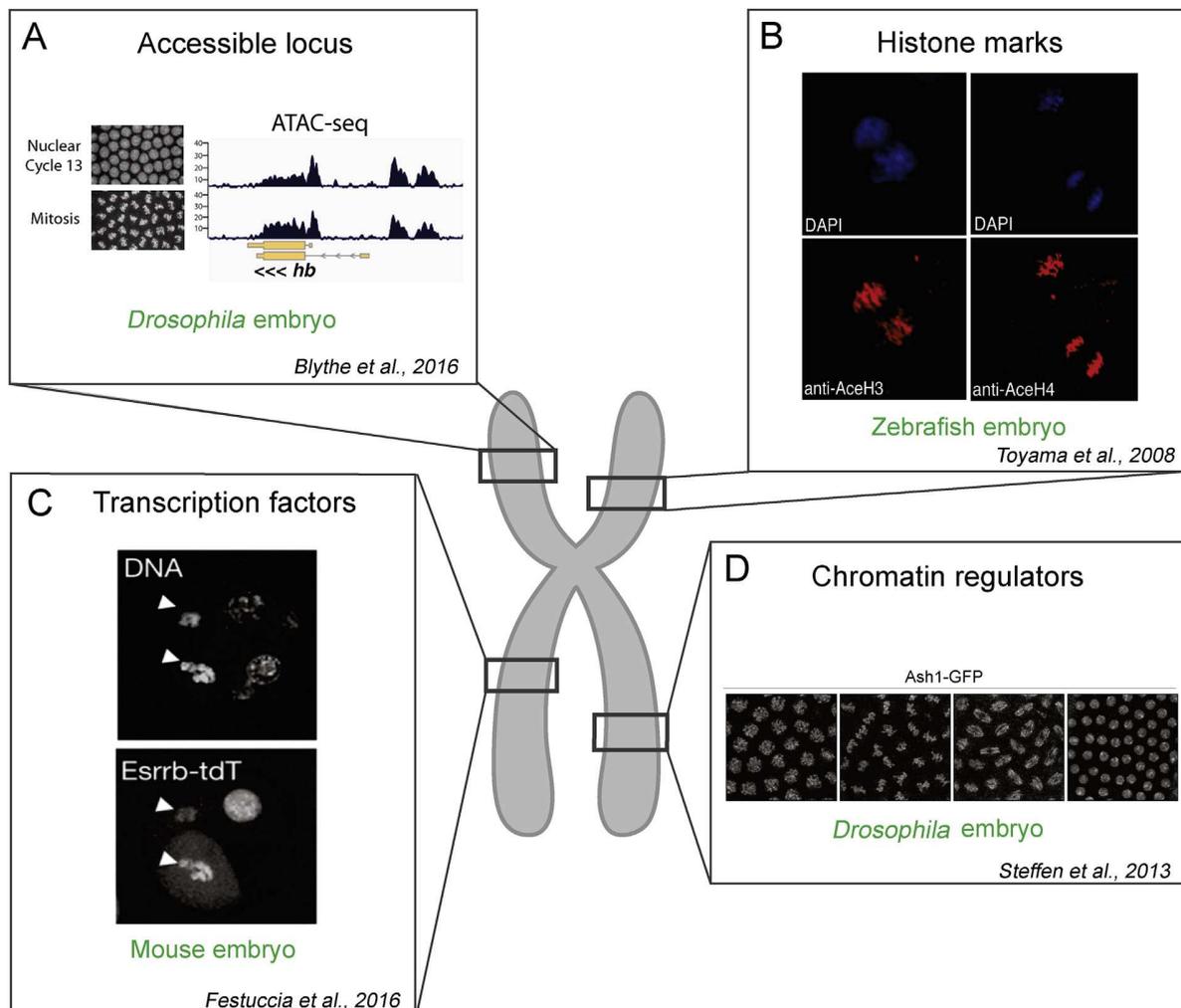
In interphase, gene expression seems to be partially regulated by local permissive or repressive environments, triggered by nuclear compartmentalization resulting from protein liquid–liquid demixing through phase separation [13] (e.g. HP1 in embryos [14]). The impact of phase separation during mitosis remains

Figure 1



Putative supports of memory of active genes/of transcriptional memory. Mutually non-exclusive mechanisms of mitotic bookmarking occurring at promoters and/or enhancers. A. Previously transcribing locus remains partially accessible during mitosis, thus facilitating post-mitotic re-activation. B. Histone marks and chromatin regulators (readers, writers) of an active chromatin state bind mitotic chromatin, thus 'bookmarking' particular loci for subsequent transcriptional activation. C. Transcription factors such as pioneer factors can associate to a subset of their targets during mitosis, consequently favoring their activation at mitotic exit. D. Through enhancer/promoter priming, all mechanisms illustrated in (A–C) lead to a rapid post-mitotic transcriptional activation.

Figure 2



Known mitotic chromatin landscapes. Potential supports of memory are illustrated by examples reported in embryos from different model organisms. A. Example of an accessible chromatin in mitotic *Drosophila* embryos, as shown by ATAC-seq experiments by Ref. [7]. B. Example of a histone modification mark, present in zebrafish embryonic dividing cells, revealed by fluorescent immunostaining (Dapi in blue, acetylated histone 3 and 4 in red) [23]. C. Representative example of mitotic chromosomal decoration by a transcription factor (Esrrb), in mouse embryo at morula stage [30]. D. Illustration of mitotic retention of a chromatin regulator, the histone methyltransferase Ash1 in living *Drosophila* embryos. Snapshot images from a time-lapse movie [20].

elusive, however mitotic structures can be assembled through phase separation in early *Caenorhabditis elegans* embryos [15]. Therefore, it is tempting to speculate that part of the mitotic maintenance of regulatory region accessibility could be due to local biochemical nuclear compartmentalization *via* phase separation. Whether this phenomenon stands as a passive or active support of mitotic memory is an open question.

Epigenetic marks, their readers and writers

Owing to decades of genetic and biochemical studies, it is well established that antagonistic actions of Polycomb group proteins (PcG) and Trithorax group proteins (TrxG) allow for dynamic regulation of developmental

genes, yet with a cellular memory [4,16]. Epigenetic transmission of active transcriptional states is supported by conserved multifaceted TrxG complexes [4,16]. Among the best-characterized TrxG function is its histone methyltransferase activity leading to the trimethylation of lysine 4 of histone H3 tails (H3K4me3).

Whether H3K4me3 qualifies as an epigenetic mark during development is debated. Indeed, using a sensitive imaging technique (proximal ligation assay) in *Drosophila* embryos, H3K4me3 does not appear to be stable through replication [17], whereas it is clearly retained at discrete loci during mitosis [18]. In another model organism, *Xenopus* embryos, H3K4me3 has been

functionally associated to memory of active transcriptional states using nuclear transfer experiments [19]. Contrary to the myriad of available data concerning PcG complexes, little is known concerning the mechanisms of TrxG recruitment and particularly during mitosis *in vivo*. Generally speaking, we can distinguish two non-mutually exclusive mechanisms of mitotic epigenetic bookmarking: either through histone-modified tails (epigenetic mark), protected or not by their ‘reader’ enzymes or through their re-establishment after mitosis by chromatin ‘writer’ enzymes (Figure 1).

In *Drosophila* embryos, Ash1, a ‘writer’ member of TrxG, decorates mitotic chromosomes [20] (Figure 2). It is not known whether the mark deposited by this enzyme, H3K36me2 is also retained during mitosis. Interestingly, Ash1 enhances the recruitment of other histone methyltransferases like Trx and its mammalian homolog MLL [21], which in turn triggers H3K4 trimethylation at promoters [16,22].

Another mitotically retained member of TrxG family is the bromodomain-containing protein 4 (BRD4), shown to coat mitotic chromosomes of zebrafish embryos [23], even before zygotic transcriptional activation. Although with little validation in embryos, extensive work in mammalian cultured cells shows that BRD4 is a pleiotropic transcriptional and epigenetic regulator [24]. Interestingly H4K5ac, a mark recognized by BRD4 is detectable at particular promoters during mitosis in mammalian cells [25] as well as acetylated H4 in zebrafish embryos [23] and could thus be considered as a true ‘epigenetic’ bookmark (Figure 2).

Mitotic bookmarking by RNA Pol II machinery

Very few studies directly examined the localization of Pol II machinery during mitosis in embryos. The early fly embryo exhibits synchronous mitotic waves, whereby various steps of mitosis can be visualized in a single embryo. Using this ideal context, it was recently shown that active Pol II (Pol II-Ser5P) is present during prophase but is largely evicted at metaphase [7]. During mitosis, the clear majority of transcription ceases. However recent studies that employ sensitive techniques (for example pulse-labeling nascent RNA) nuance this statement and reveal low-level mitotic transcription in cell culture [26]. These paradigm-shifting findings would need to be confirmed in physiological conditions *in vivo*. However, they are in agreement with mitotic retention of GTF as the TATA binding protein, TBP, shown to bind mitotic chromosomes in ES [27] cells and to decorate chromatin in dividing mouse blastocysts [28].

Mitotic bookmarking by transcription factors

In developing embryos, very few sequence specific transcription factors have been shown to remain

associated to their targets during mitosis. However recent whole-genome profiling and live imaging revealed that several pluripotency TFs have the ability to bind mitotic chromosomes in cultured dividing ES cells [1,29]. This mitotic retention might also occur in mouse embryos but remains to be demonstrated. So far, among the few reported mitotically bound TF during development, stand *Essrb* (Figure 2) and *Klf4* in dividing mouse blastomeres [30,31] and *HNF1beta*, retained during renal development in mice (Pontoglio lab, personal communication). Future investigations regarding TF retention during mitosis in embryos would require testing multiple fixation procedures and cross-validations with live imaging approaches. Indeed, formaldehyde fixation can lead to the artifactual displacement of some TFs from mitotic chromosomes [32,33].

Interestingly, the majority of the known mitotically retained TFs, for example some pluripotency factors (e.g. *Oct4/Sox2/Klf4*) function as pioneer factors [34]. Pioneer factors are a particular class of TFs, able to engage their target DNA at compacted nucleosomal regions, thereby fostering subsequent access to classical non-pioneer TF to their targets [35]. It is thus legitimate to ask whether pioneer factors might have intrinsic properties to bind chromatin during mitosis. However, little is known regarding the functional importance of pioneer factors for mitotic memory in developing multicellular organisms. Intriguingly, the pioneer factor *Zelda*, an essential activator of the early zygotic genome in *Drosophila* embryos, is not retained during mitosis and does not contribute to mitotic memory [11]. This could reflect that mitotic bookmarking is not a general feature of all pioneer factors, or alternatively, that *Zelda* is not a canonical pioneer factor.

Mitotic retention: stable versus dynamic bookmarking

Whole-genome kinetics of nucleosome turnover is much more dynamic than what conceptual ideas of stable epigenetic inheritance of chromatin states may provide. In cultured cells, histone modifications can be often erased and re-established several times during the cell cycle [36]. This finding may not be contradictory with mitotic bookmarking, when the binding properties of chromatin regulators are examined with advanced fluorescent microscopy methods (FCS, FRAP or SPT [37]).

In early *Drosophila* embryos, quantitative *in vivo* analysis of Ash1 dynamic properties demonstrate that this mitotically retained histone methyltransferase engages chromatin dynamically with estimated residence times in the order of seconds [20].

Similarly, in ES cells, pluripotency transcription factors bind mitotic chromosomes dynamically (e.g. residence time of *Sox2* ~ 10 s) [1,29,38].

Of note, dynamic binding could be attributed to either sequence specific binding events, or to non sequence specific DNA binding [38]. The long-lived interactions, in the order of seconds generally reflect sequence specific bindings [29]. However, evidence for non-specific DNA binding during mitosis exists (e.g. FoxA) [39].

In order to clearly distinguish specific from non-specific binding events, one should ideally use mitotic Chip-seq experiments with quantitative imaging of DNA binding mutant versions of the TF assessed [37].

In sharp contrast to the dynamic mode of binding described earlier, some factors can bind to mitotic chromosomes in a stable manner. This is exemplified with TBP, which exhibits an average residence time in the order of minutes, in living dividing ES cells [27]. In conclusion, mitotic retention can occur through dynamic or stable binding, *via* sequence specific binding and non-specific binding.

Consequences of mitotic binding on transcriptional dynamics

Mitotic retention does not directly bestow a bookmarking function. To firmly prove that a mitotically retained factor is involved in 'memory', one need to examine the consequences of its transient mitotic-specific depletion on dynamics of post-mitotic transcriptional activation.

These experiments are challenging and have thus far never been performed in embryos. However knockdown experiments of candidate bookmarking factors have been performed in synchronized mitotic cells [26,27,38] and in *Dictyostelium* [40]. Generally two types of outputs are assessed: whole-genome approaches to examine bulk transcriptional levels or live imaging following transcriptional reporters in single cells.

Mitotic-specific depletion of the hematopoietic transcription factor GATA1 in erythroid precursors delays the reactivation of key lineage specifying mitotically bookmarked genes [41]. Similarly, TBP mitotic transient depletion in ES cells delays the reactivation of the global ES cell transcriptional programme [27]. While displaying an overall genome-wide picture, these ensemble approaches lack temporal resolution. The MS2/MCP system allows to directly monitor actively transcribing loci (tagged with MS2 repeats) in living cells [42]. This technique not only measures the dynamics of transcriptional activation but also provides access to the lineage and thus to mitotic memory of active genes.

Using such approaches, memory of active genes has been first visualized in *Dictyostelium* [40], in mammalian cells [25] and more recently in *Drosophila* [12]. Using a stochastically expressed transgene, Ferraro *et al.* [12]

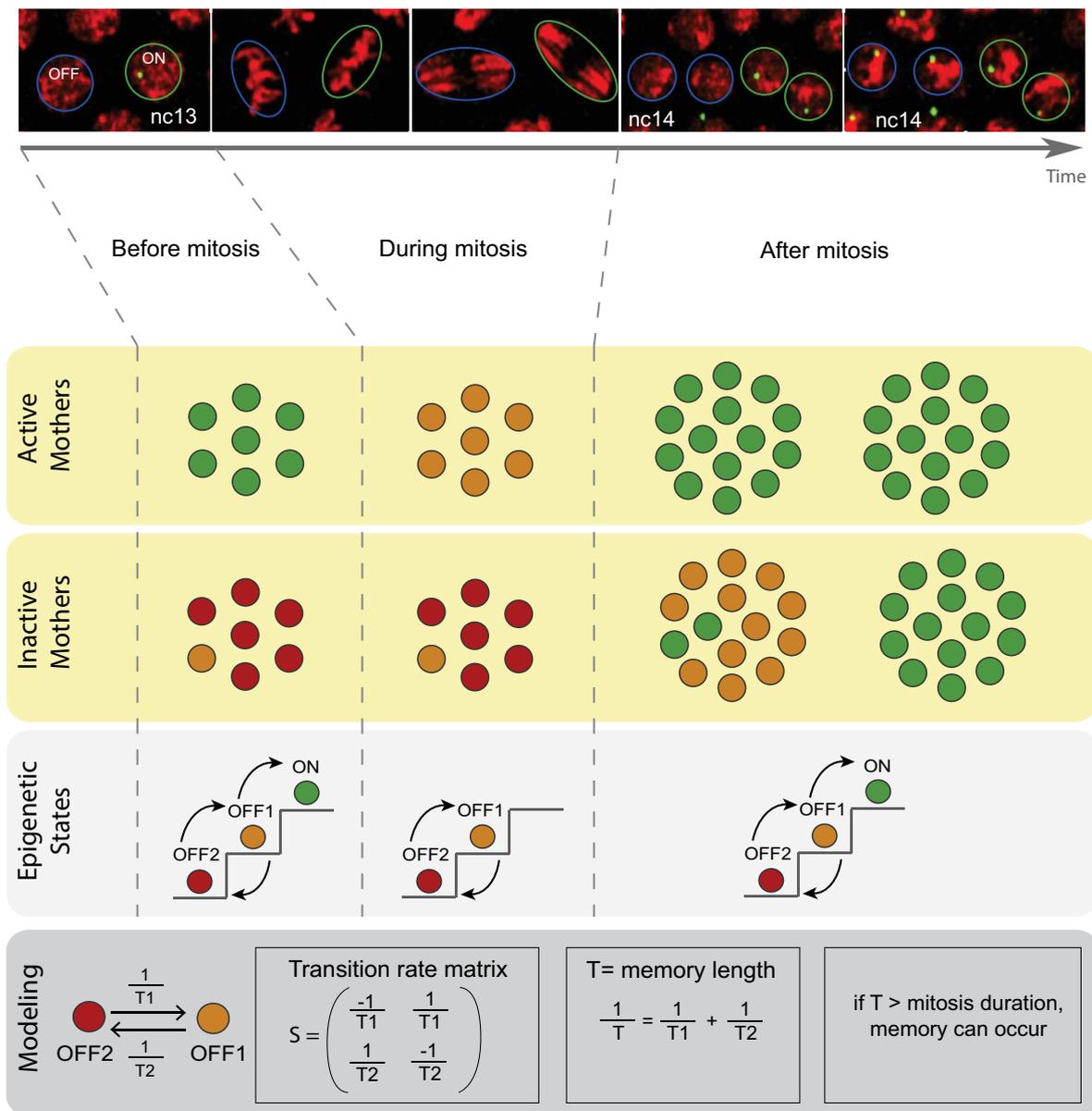
could distinguish transcriptionally active mother nuclei from their inactive neighbors. Quantitative analyses reveal that there is a higher probability for rapid reactivation after mitosis, when the mother was active. This bias corresponds to transcriptional mitotic memory. In the future, generalization of live imaging of transcription methods should allow for memory detection in other model organisms.

Mathematical modeling of memory

A quantitative measure of mitotic memory is the difference in the timing of post-mitotic re-activation between descendants of active and inactive mothers. This difference can be explained by the fact that these two populations 'travel' through different states prior to activating transcription after mitosis. We can distinguish one active (ON) state and various inactive states (OFF), interpreted as more or less favorable chromatin landscapes [43]. The number of rate limiting kinetic steps leading from an inactive to an active state and *vice versa* can be higher than one, with lifetimes of inactive states ranging from one to several hundreds of minutes [43]. By stabilizing a competent yet inactive state (OFF1), mitotic bookmarking could prevent the decline of a mother cell to a less permissive state (OFF2), allowing a faster transcriptional recovery of the daughters (Figure 3).

A rather general class of mathematical models of memory can be based on Markov chains, used to explain transcriptional memory in *Drosophila* embryos [11] (Figure 3). Originally, Markov chains were used to describe memoryless processes. However, these models can acquire memory by an operation called "projection". By projection, instead of following the dynamics of all the states of the system, one observes only the result of this "hidden" dynamics, namely the mRNA production. The projected dynamics exhibit memory in the sense that the future of the observed variable (mRNA production) depends not only on its present, but also on all past events that occurred prior to a time T (T is called memory length). This method, known as Mori-Zwanzig projection operator technique [44], has been recently applied to analyze propagation of signals in gene networks, as those occurring during vertebrate neural tube patterning [45]. In the model presented here (Figure 3), memory results from hidden states preservation, whereas memory length corresponds to the required time to forget differences between initial values of these states. For a finite-state, continuous-time Markov chain model, this time can be easily computed as the inverse of the smallest, non-zero eigenvalue of the transition rate matrix, in absolute value [46]. In the case of two inactive states, characterized by lifetimes T_1 and T_2 , respectively (Figure 3), memory length (T) is the harmonic mean of the states lifetimes, which is long when both lifetimes are long. Thus, this model suggests that

Figure 3



Modeling memory of active states in an embryo. A. Snapshots of live imaging of an MS2-reporter transgene in a *Drosophila* embryo at nuclear cycle (nc) 13, during mitosis and at nc14. Nuclei are labeled with a Histone-RFP transgene and nascent MS2-mRNAs with MCP-GFP in green spots (Images from M. Lagha laboratory). B. Model of memory with two inactive states OFF1 (inactive but competent) and OFF2 (inactive and repressed). A third state ON is transcribing. The lifetimes of states OFF1, OFF2 are T_1 , T_2 , and correspond to the transition time from OFF1 to OFF2, and from OFF2 to OFF1, respectively. T_3 is the time to go from the state OFF1 to the active state. The transition rates $1/T_1$, $1/T_2$, $1/T_3$ are reciprocals of the transition times. The memory length T , computed from the eigenvalues of the transition rate matrix, is $T_1 T_2 / (T_1 + T_2)$. Prior to mitosis, active mother nuclei (bookmarked) will be in state ON, whereas the population of inactive mothers will be distributed among states OFF1 and OFF2. During mitosis, the ON state is no longer accessible (transcription ceases). Therefore, previously active nuclei will be simply downgraded to state OFF1 while, provided that mitosis length is short compared to T , inactive nuclei will keep their states. After mitosis, daughters will be able to switch from OFF1 to ON or travel progressively from OFF2 to OFF1 to ON. The path to activation is shorter and consequently, the probability per unit of time to activate transcription after mitosis is increased for descendants of active mothers. Within a developmental pattern, this memory bias could favor the temporal coordination in gene activation (i.e. synchrony), with consequences in terms of cell fate.

efficient memory requires slow dynamics. Slow dynamics could in principle be favored by the presence of mitotic bookmarking, but this remains to be demonstrated.

Consequences of memory during development

In the context of a developing multicellular embryo, mitotic memory can lead to a multitude of consequences, sometimes with opposite outcomes.

Inheritance of active transcriptional states ensures fidelity of transcriptional programs during their propagations through mitosis. Combined with other priming mechanisms during interphase, such as local chromatin opening at enhancers and promoter poising with paused polymerase, mitotic memory results in augmented transcriptional precision with less inter-cellular variability in levels of expression [47]. In sharp contrast, by allowing some cells to re-activate transcription faster than others, mitotic memory can enhance transcriptional noise [48]. From a theoretical point of view, memory can be seen as a ‘low pass filter’, buffering fast fluctuations, while allowing longer-lived fluctuations. Depending on memory length relative to mitosis duration and the architecture of the gene regulatory network (e.g. existence of feedback loops), mitotic memory can thus create precision and increase noise, but both are valuable during development [47]. For example, temporal precision in gene expression has been shown to be essential to gastrulation in the fast developing *Drosophila* embryo [49]. However, in slower developing embryos such as the mouse blastocyst, heterogeneity in gene expression has been proposed as advantageous since it allows for mis-patterning corrections [50].

Mitotic memory thus allows for a spectrum of consequences ranging from stability to flexibility and plasticity. However, for the particular developmental contexts of trans-differentiation or reprogramming, transcriptional states must be erased rather than memorized. Accordingly, demethylation of H3K4 facilitates reprogramming of *Xenopus* embryos [19] and mouse epiblast cells [22].

Prospects

The vast majority of our knowledge regarding memory stems from studies performed in drug-synchronized cells in culture and has not yet been mirrored in multicellular developing embryos. This is mainly attributed to the dual technical challenge of profiling small number of cells/embryos and imaging living organisms.

With the recent technological advances in gene editing, signal amplification, microscopy combined to the possibility to probe the entire genome of limited material [51], developmental biology is embracing a new exciting era.

Synergizing quantitative biological data with synthetic biology frameworks [52] and mathematical modeling opens promising avenues for a better understanding of mechanisms and functional relevance of memory.

Recent combination of optogenetics with concomitant detection of transcriptional readouts in living embryos [53] should greatly facilitate testing the direct impact of candidate bookmarking factors on kinetics of transcriptional post-mitotic reactivation.

Decoding the role of mitotic bookmarking in developing model organisms endows two obvious consequences for human therapy. Indeed defective mitotic bookmarking has been associated to human pathological conditions, as exemplified by HNF1beta [54]. Moreover understanding how cells remember their past should greatly facilitate the design of reprogramming strategies for gene therapy.

Conflict of interest statement

Nothing declared.

Acknowledgements

We thank Bernd Schuttengruber, Jeremy Dufourt, Youssa Yahia and Cyril Esnault for comments on the manuscript.

Research in the M.L. laboratory is supported by the Centre National de la Recherche Scientifique, the European Research Council (Sync_DEV). O.R. thanks CNRS and LABEX Epigenmed for support.

We apologize to colleagues whose work could not be cited due to space constraints.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Festuccia N, Gonzalez I, Owens N, Navarro P: **Mitotic bookmarking in development and stem cells.** *Development* 2017, **144**:3633–3645.
 2. Zenk F, *et al.*: **Germ line–inherited H3K27me3 restricts enhancer function during maternal-to-zygotic transition.** *Science (80-)* 2017, **357**:212–216.
 3. Oomen ME, Dekker J: **Epigenetic characteristics of the mitotic chromosome in 1D and 3D.** *Crit Rev Biochem Mol Biol* 2017, **52**:185–204.
 4. Piunti A, Shilatifard A: **Epigenetic balance of gene expression by polycomb and compass families.** *Science (80-)* 2016, **352**.
 5. de Castro IJ, Gokhan E, Vagnarelli P: **Resetting a functional G1 nucleus after mitosis.** *Chromosoma* 2016, **125**:607–619.
 6. Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM: **Chromatin architecture emerges during zygotic genome activation independent of transcription.** *Cell* 2017, **169**:216–228.e19.
 7. Blythe SA, Wieschaus EF: **Establishment and maintenance of heritable chromatin structure during early drosophila embryogenesis.** *Elife* 2016, **5**:1–21.
- Using ATAC-Seq in single *Drosophila* embryos, the authors revealed that patterns of chromatin accessibility are maintained through mitosis.
8. Hsiung CCS, *et al.*: **Genome accessibility is widely preserved and locally modulated during mitosis.** *Genome Res* 2015, **25**:213–225.
 9. Olivier N, *et al.*: **Cell lineage reconstruction of early.** *Science (80-)* 2010, **70**:967–971.

10. Nagano T, *et al.*: **Cell-cycle dynamics of chromosomal organization at single-cell resolution.** *Nature* 2017, **547**:61–67.
11. Dufourt J, *et al.*: **Temporal control of transcription by Zelda in living *Drosophila* embryos.** *bioRxiv* 2018, <https://doi.org/10.1101/282426>.
12. Ferraro T, *et al.*: **Transcriptional memory in the *Drosophila* embryo.** *Curr Biol* 2016, **26**:212–218.
Using MS2-MCP system in the early *Drosophila* embryo, the authors visualized transcriptional memory for the first time in a living multicellular embryo.
13. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA: **A phase separation model for transcriptional control.** *Cell* 2017, **169**:13–23.
14. Strom AR, *et al.*: **Phase separation drives heterochromatin domain formation.** *Nature* 2017, **547**:241–245.
15. Woodruff JB, *et al.*: **The centrosome is a selective condensate that nucleates microtubules by concentrating tubulin.** *Cell* 2017, **169**:1066–1077.e10.
16. Schuettengruber B, Bourbon HM, Di Croce L, Cavalli G: **Genome regulation by polycomb and trithorax: 70 years and counting.** *Cell* 2017, **171**:34–57.
17. Petruk S, *et al.*: **TrxG and PcG proteins but not methylated histones remain associated with DNA through replication.** *Cell* 2012, **150**:922–933.
18. Black KL, *et al.*: **Chromatin proteins and RNA are associated with DNA during all phases of mitosis.** *Cell Discov* 2016, **2**.
19. Hörmanseder E, *et al.*: **H3K4 methylation-dependent memory of somatic cell identity inhibits reprogramming and development of nuclear transfer embryos.** *Cell Stem Cell* 2017, **21**:135–143.e6.
20. Steffen PA, *et al.*: **Quantitative in vivo analysis of chromatin binding of polycomb and trithorax group proteins reveals retention of ASH1 on mitotic chromatin.** *Nucleic Acids Res* 2013, **41**:5235–5250.
Using time-lapse microscopy on living EGFP tagged Ash1 transgenic *Drosophila* embryos, the authors showed that Ash1 decorates mitotic chromosomes. Moreover, they employed quantitative imaging methods to reveal that Ash1 binds to chromatin dynamically.
21. Blobel GA, *et al.*: **A reconfigured pattern of MLL occupancy within mitotic chromatin promotes rapid transcriptional reactivation following mitotic exit.** *Mol Cell* 2009, **36**:970–983.
This paper was among the firsts to show that mitotic binding by MLL at specific loci accelerates their post-mitotic transcriptional reactivation.
22. Zhang H, *et al.*: **MLL1 inhibition reprograms epiblast stem cells to naive pluripotency.** *Cell Stem Cell* 2016, **18**:481–494.
23. Toyama R, Rebbert ML, Dey A, Ozato K, Dawid IB: **Brd4 associates with mitotic chromosomes throughout early zebrafish embryogenesis.** *Dev Dynam* 2008, **237**:1636–1644.
24. Devaiah BN, Geronne A, Singer DS: **Bromodomain 4: a cellular Swiss army knife.** *J Leukoc Biol* 2016, **100**:679–686.
25. Zhao R, Nakamura T, Fu Y, Lazar Z, Spector DL: **Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation.** *Nat Cell Biol* 2011, **13**:1295–1304.
This paper demonstrated the consequences of mitotic bookmarking (by BRD4 and H4K5 acetylation) on post-mitotic transcriptional dynamics in mammalian cells.
26. Palozola KC, *et al.*: **Mitotic transcription and waves of gene reactivation during mitotic exit.** *Science (80-)* 2017, **358**:119–122.
27. Teves SS, An L, Bhargava-Shah A, Xie L, Darzacq X, Tjian R: **A stable mode of bookmarking by TBP recruits RNA polymerase II to mitotic chromosomes.** *Elife* 2018, **7**:1–22, <https://doi.org/10.7554/eLife.35621>.
28. Sun SC, *et al.*: **TBP dynamics during mouse oocyte meiotic maturation and early embryo development.** *PLoS One* 2013, **8**:1–6.
29. Raccaud M, Suter DM: **Transcription factor retention on mitotic chromosomes: regulatory mechanisms and impact on cell fate decisions.** *FEBS Lett* 2017, **592**:878–887.
30. Festuccia N, *et al.*: **Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network.** *Nat Cell Biol* 2016, **18**:1139–1148.
By using live imaging and ChIP experiments in ES cells, this paper identified Esrrb as a potential mitotic bookmarker. A subset of interphase-bound targets, remain bound by Esrrb during mitosis and correspond to key regulatory regions (super-enhancers).
31. Liu Y, *et al.*: **Widespread mitotic bookmarking by histone marks and transcription factors in pluripotent stem cells.** *Cell Rep* 2017, **19**:1283–1293.
32. Teves SS, *et al.*: **A dynamic mode of mitotic bookmarking by transcription factors.** *Elife* 2016, **5**.
This paper demonstrated that the majority of transcription factors examined remain associated to mitotic chromosomes in ES cells. A quantitative study on Sox2 revealed that this pluripotency factor exhibits a dynamic mode of mitotic binding. This paper points to the artificial mitotic eviction of some transcription factors triggered by formaldehyde fixation.
33. Pallier C, *et al.*: **Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes.** *Mol Biol Cell* 2003, **14**:2372–2384.
34. Soufi A, *et al.*: **Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming.** *Cell* 2015, **161**:555–568.
35. Iwafuchi-Doi M, Zaret KS: **Cell fate control by pioneer transcription factors.** *Development* 2016, **143**:1833–1837.
36. Deal RB, Henikoff JG, Henikoff S: **Genome-wide kinetics of nucleosome.** *Science (80-)* 2010, **328**:1161–1165.
37. Mueller F, Stasevich TJ, Mazza D, McNally JG: **Quantifying transcription factor kinetics: at work or at play?** *Crit Rev Biochem Mol Biol* 2013, **48**:492–514.
38. Deluz C, *et al.*: **A role for mitotic bookmarking of SOX2 in pluripotency and differentiation.** *Genes Dev* 2016, **30**:2538–2550.
In this paper, the authors profiled the mitotic binding for Sox2 and Oct4 and quantified their binding dynamics with quantitative imaging methods. By analyzing the consequences of Sox2 mitotic depletion, they demonstrated Sox2 bookmarking activity.
39. Caravaca JM, *et al.*: **Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes.** *Genes Dev* 2013, **27**:251–260.
This study identified two modes of mitotic bookmarking for the pioneer factor FoxA: via sequence specific binding and via nonspecific binding.
40. Muramoto T, Müller I, Thomas G, Melvin A, Chubb JR: **Methylation of H3K4 is required for inheritance of active transcriptional states.** *Curr Biol* 2010, **20**:397–406.
This paper was the first to directly visualize the inheritance of transcriptionally active states down cell lineages in *Dictyostelium*. By using various genetic mutants, the authors demonstrated that this memory can be partially attributed to H3K4 methylation.
41. Kadauke S, *et al.*: **Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1.** *Cell* 2012, **150**:725–737.
42. Bertrand E, *et al.*: **Localization of ASH1 mRNA particles in living yeast.** *Mol Cell* 1998, **2**:437–445.
43. Lionnet T, Singer RH: **Transcription goes digital.** *EMBO Rep* 2012, **13**:313–321.
44. Forster D. Hydrodynamic fluctuations, broken symmetry, and correlation functions. (*Frontiers in Physics* 47, XIX, 326 S, 1975).
45. Herrera-Delgado E, Perez-Carrasco R, Briscoe J, Sollich P: **Memory functions reveal structural properties of gene regulatory networks.** *PLoS Comput Biol* 2018, **14**.
46. Karlin S, Taylor HE: *A first course in stochastic processes*. 2nd ed. 2012, <https://doi.org/10.1016/C2009-1-28569-8>.

47. Bentovim L, Harden TT, DePace AH: **Transcriptional precision and accuracy in development: from measurements to models and mechanisms.** *Development* 2017, **144**:3855–3866.
48. Chubb JR: **Gene regulation: stable noise.** *Curr Biol* 2016, **26**:R61–R64.
49. Lagha M, *et al.*: **XPaused Pol II coordinates tissue morphogenesis in the drosophila embryo.** *Cell* 2013, **153**:976–987.
50. Simon CS, Hadjantonakis A-K, Schröter C: **Making lineage decisions with biological noise: lessons from the early mouse embryo.** *Wiley Interdiscip Rev Dev Biol* 2018, **e319**, <https://doi.org/10.1002/wdev.319>.
51. Skene PJ, Henikoff JG, Henikoff S: **Targeted in situ genome-wide profiling with high efficiency for low cell numbers.** *Nat Protoc* 2018, **13**:1006–1019, <https://doi.org/10.1038/nprot.2018.015>.
52. Bintu L, *et al.*: **Dynamics of epigenetic regulation at the single cell level.** *Science (80-.)* 2016, **351**:720–724.
53. Huang A, Amourda C, Zhang S, Tolwinski NS, Saunders TE: **Decoding temporal interpretation of the morphogen bicoid in the early drosophila embryo.** *Elife* 2017, **6**.
54. Lerner J, *et al.*: **Human mutations affect the epigenetic/bookmarking function of HNF1B.** *Nucleic Acids Res* 2016, **44**:8097–8111.

5. Publication: Zelda le maestro du réveil du génome zygotique

This publication was the first in which I participated mainly by doing figures. This was a review in a French journal following the publication of the laboratory (Dufourt et al. 2018) about the role of the pioneer factor Zelda on zygotic transcription activation.

NOUVELLE

Zelda, le maestro du réveil du génome zygotique

Jeremy Dufourt¹, Maëlle Bellec¹, Olivier Messina¹, Antonio Trullo¹, Cyril Favard², Ovidiu Radulescu³, Mounia Lagha¹

> Chez de nombreux animaux, bien que la fertilisation aboutisse rapidement à la fusion des génomes maternels et paternels, l'activation du génome zygotique ne débute pas dès les premières heures du développement. En effet, il existe un stock d'ARN et de protéines dans le gamète maternel qui va régir l'embryogenèse précoce. Durant cette première période critique du développement, les cellules totipotentes issues de l'œuf fécondé vont rapidement se différencier pour pouvoir ensuite former les trois feuillettes embryonnaires fondamentales : ectoderme, endoderme, et mésoderme.

La transition materno-zygotique

Chez tous les métazoaires, le passage graduel du contrôle maternel au contrôle zygotique du développement embryonnaire se déroule lors d'un événement appelé transition materno-zygotique (MZT), durant laquelle, de manière concomitante à la dégradation des ARN et protéines provenant du gamète maternel, le génome de l'embryon est activé. Cette activation est caractérisée par une vague progressive de transcription, qui coïncide avec le ralentissement des cycles cellulaires dans de nombreuses espèces. Le démarrage de cette vague de transcription ainsi que le nombre de cycles de division qu'elle accompagne varient considérablement d'une espèce à l'autre. Cependant, pour chaque espèce, le déroulement temporel de la MZT est précis et très reproductible d'un embryon à l'autre. Chez les espèces au développement rapide, telles que le nématode (*Caenorhabditis elegans*), le xénope

(*Xenopus laevis*), le poisson zèbre (*Danio rerio*) et la drosophile (*Drosophila melanogaster*), la MZT s'achève quelques heures seulement après la fécondation, et est suivie du premier événement de la morphogénèse : la gastrulation. En revanche, chez les espèces à développement plus lent, telles que la souris (*Mus musculus*) et l'homme (*Homo sapiens*), la MZT peut durer plusieurs jours [1].

Chez la drosophile, l'embryogénèse commence par 14 cycles nucléaires (nc) rapides au sein d'un syncytium, contrôlés principalement par l'apport en ARN et protéines du gamète maternel. Environ deux heures après la fécondation, le génome zygotique s'active transcriptionnellement et le syncytium se sépare en cellules individuelles (cellularisation) comportant chacune un seul noyau (Figure 1A). Le réveil transcriptionnel du génome zygotique nécessite des facteurs de transcription spécifiques, dont la protéine Zelda, qui joue un rôle essentiel dans ce processus [2]. Zelda possède plusieurs caractéristiques des facteurs de transcription dits « pionniers » : 1) sa liaison aux éléments *cis*-régulateurs avant l'activation des gènes, établissant une pré-compétence pour l'adoption de certains destins cellulaires, et 2) sa capacité à ouvrir la chromatine pour faciliter la liaison de facteurs de transcription classiques. De plus, la distribution homogène de Zelda dans l'embryon permet de potentialiser l'action locale de morphogènes (par exemple, *bicoid* et *dorsal*) dans différentes régions de l'embryon [3-5] : la réponse de gènes-cibles

à la concentration de morphogène est renforcée par la liaison de Zelda à ces gènes. Cependant, le rôle de ce facteur « pionnier » dans le contrôle temporel de l'activation transcriptionnelle n'a été élucidé que très récemment.

Zelda favorise la coordination temporelle de la transcription

Dès le début de la MZT, le gène *snail* (*sna*) est transcrit dans les cellules du futur mésoderme de l'embryon de drosophile. Il code un facteur de transcription qui va permettre la spécification du mésoderme et contrôler la transition épithélio-mésenchymateuse lors de la gastrulation. Dès les premiers cycles nucléaires, la protéine Zelda se lie à deux *enhancers* (distal et proximal) de *sna*, qui contrôlent spatialement et temporellement son expression.

Pour tenter de comprendre le rôle de Zelda dans l'activation transcriptionnelle au cours de la MZT, nous avons mesuré les temps d'activation de la transcription, dans les cellules du mésoderme en formation, de plusieurs transgènes ayant une version tronquée de l'*enhancer* distal de *sna* (*snaE*) et possédant un nombre croissant de sites de fixation de la protéine Zelda. Le suivi de l'activation de la transcription dans des embryons vivants est possible grâce à l'ajout, dans une partie non-codante du transgène rapporteur, de séquences de fixation de la protéine du bactériophage MS2 à l'ARN [6]. Lors de la transcrip-

¹Institut de génétique moléculaire de Montpellier, université de Montpellier, CNRS-UMR 5535, 1919 route de Mende, 34293 Montpellier Cedex 5, France.

²Institut de recherche en infectiologie de Montpellier, CNRS, université de Montpellier UMR 9004, 1919 route de Mende, 34293 Montpellier Cedex 5, France.

³LPHI, UMR CNRS 5235, université de Montpellier, place E. Bataillon - Bât. 24 cc 107, 34095 Montpellier, Cedex 5, France.

Jeremy.dufourt@igmm.cnrs.fr

Mounia.lagha@igmm.cnrs.fr



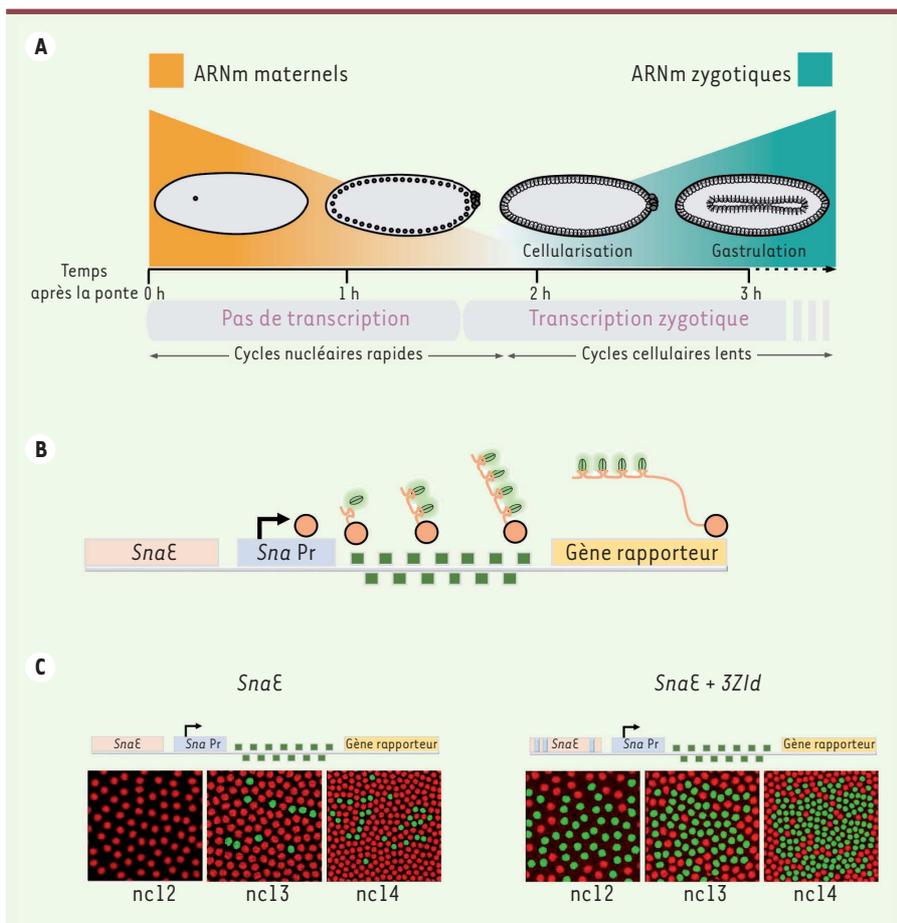


Figure 1. A. Représentation schématique de la transition materno-zygotique (MZT) chez la drosophile. B. Schéma du système MS2/*MS2 coat protein* (MCP) couplée à la *green fluorescent protein* (GFP) permettant de visualiser la transcription génique en temps réel. C. Représentation schématique des deux transgènes comportant soit la version tronquée de l'*enhancer* distal de *sna* (*SnaE*) ne comportant aucun site de liaison à Zelda, soit cet *enhancer* avec la présence additionnelle de trois sites de liaison à Zelda (*SnaE* + *3Zld*), et imagerie en temps réel de leur activité de transcription au cours de trois cycles nucléaires (nc12, nc13, nc14). Chaque *enhancer* (rectangle orange) contrôle le promoteur minimal de *sna* (*SnaPr*, rectangle bleu), qui permet la transcription d'un gène rapporteur (rectangle jaune) comportant 24 répétitions du motif de fixation de la protéine MCP (carrés verts) dans sa partie 5' non codante. Les noyaux qui présentent une activité transcriptionnelle sont détectés grâce à la fluorescence de la GFP et colorés en vert. Ils sont beaucoup plus nombreux lorsque l'*enhancer* distal contient des sites de liaison à Zelda qu'en leur absence.

tion, ces étiquettes forment des « tiges-boucles » avec l'ARNm, auxquelles va rapidement se lier une protéine de fusion fluorescente (*MS2 coat protein* couplée à la *green fluorescent protein*), ce qui permet de visualiser ces nouveaux transcrits (Figure 1B). Nous avons montré que la protéine Zelda accélérât la coordination de l'activation de la transcription (synchronie) de *sna* au sein du mésoderme en formation (Figure 1C).

Zelda masque la mémoire transcriptionnelle

La possibilité de suivre l'activation de la transcription dans des embryons vivants permet aussi de déterminer l'effet de Zelda sur l'activité transcriptionnelle au cours de multiples cycles nucléaires. Nous avons ainsi récemment documenté l'existence d'une mémoire mitotique transcriptionnelle, en montrant que le statut transcriptionnel des noyaux

mères lors d'un cycle nucléaire donné influençait l'activation de la transcription, au cycle suivant, dans les noyaux qui en dérivent [7] (Figure 2A). Les cellules vont se diviser de nombreuses fois au cours du développement, et la régulation génique des cellules mères doit être transmise aux cellules filles afin de maintenir l'identité tissulaire. Cependant, à chaque division cellulaire (mitose), la transcription s'arrête, les chromosomes se condensent, et la plupart des facteurs de transcription se dissocient de la chromatine. Afin de transmettre l'information d'identité cellulaire, certains mécanismes garantissent la réactivation post-mitotique de la transcription de certains gènes d'identité dans les cellules filles. Ce processus est appelé « bookmarking » mitotique en référence aux facteurs qui restent associés à la chromatine durant la mitose et servent

de « marque-pages » [8]. Les facteurs de transcription « pionniers », outre leur propriété de se lier aux nucléosomes, restent généralement associés aux chromosomes mitotiques. La rétention mitotique de ces facteurs de transcription en fait des candidats idéaux pour la transmission de l'identité cellulaire au cours des divisions mitotiques. Sous l'effet de l'activation de ses *enhancers*, l'expression de *sna* est très rapide, ce qui empêche d'analyser le rôle de Zelda dans le tempo d'activation de *sna*. Nous avons donc utilisé la version tronquée de l'*enhancer* distal de *sna* (*snaE*) qui ne contient plus de sites de fixation de la protéine Zelda, et lui avons ajouté, ou pas, des sites de fixation de Zelda. L'activation de ce transgène est stochastique, mais toujours restreinte aux noyaux du mésoderme en formation. Ce dispositif expérimental permet de suivre les temps d'activa-

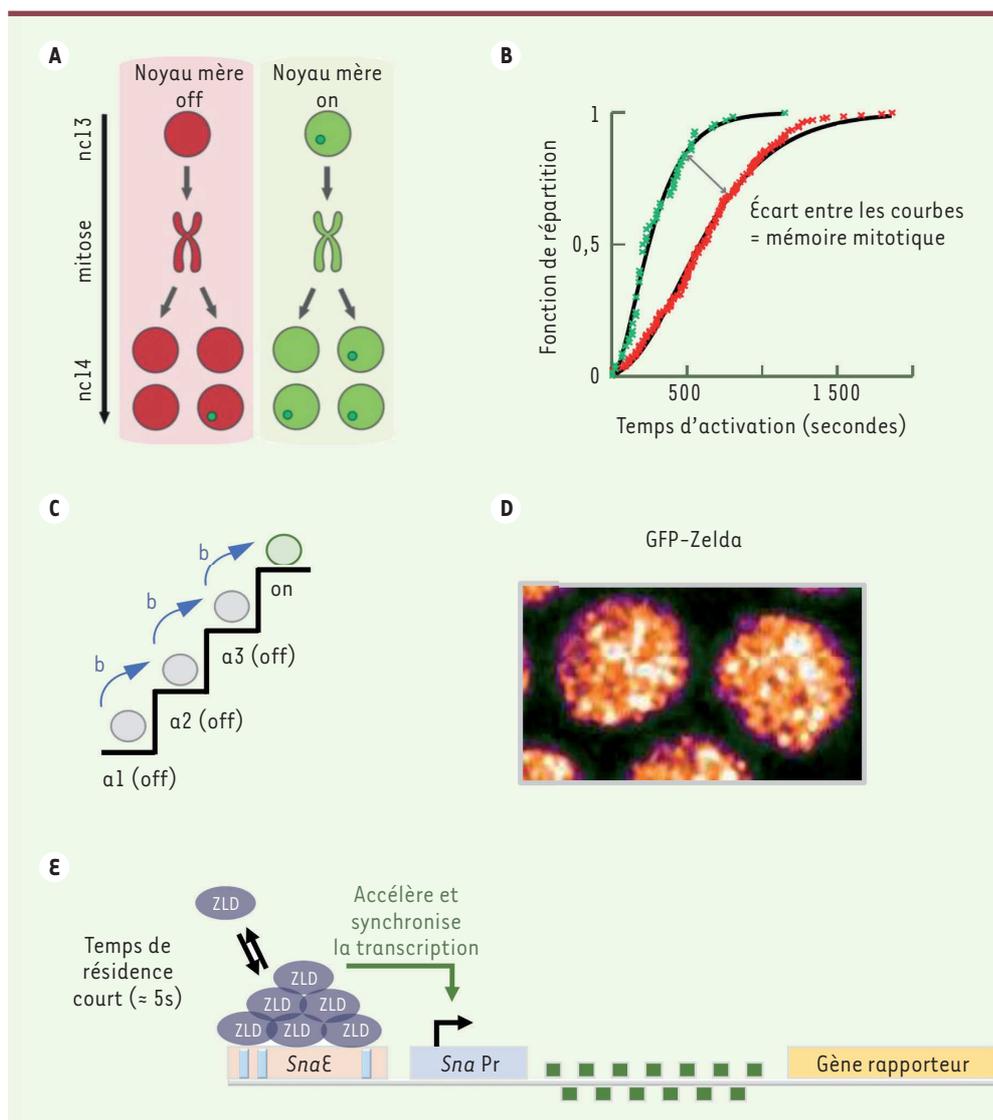


Figure 2. A. Schéma représentant la mémoire mitotique entre deux cycles nucléaires consécutifs (nc 13 et nc14) de l'embryon de drosophile. Les noyaux issus de noyaux mères dans lesquelles la transcription génique est active (on, en vert) s'activent en moyenne plus rapidement que ceux issus de noyaux mères dans lesquelles la transcription génique est inactive (off, en rouge), au cycle nucléaire suivant (nc14).

B. Fonction de répartition des temps d'activation au cycle nucléaire 14. Les valeurs des temps d'activation transcriptionnelle des noyaux issus de noyaux mères « actifs » sont représentées en vert, et celles des noyaux issus de noyaux mères « inactifs » sont représentées en rouge. Les paramètres du modèle mathématique sont estimés à partir de courbes ajustées à ces valeurs expérimentales (courbes noires). **C.** Représentation schématique des états épigénétiques discrets nécessaires pour atteindre l'activation de la transcription (état ON). La durée de chaque transition d'un état à l'autre est fournie par le paramètre noté « b ».

D. Image de microscopie confocale de la protéine Zelda étiquetée avec une GFP (par la technique CRISPR/Cas9) montrant la présence de « microenvironnements nucléaires » dans lesquels la protéine s'accumule. **E.** Schéma représentant le rôle de Zelda dans l'activation synchrone de la transcription de multiples gènes zygotiques. ZLD : Zelda ; GFP : *green fluorescent protein*.

tion transcriptionnelle des noyaux filles provenant de noyaux mères transcriptionnellement « actifs » ou « inactifs » (Figure 2B) : ces temps d'activation reflètent la mémoire transcriptionnelle à travers la mitose [7]. Nous avons pu ainsi montrer que Zelda accélérât l'activation transcriptionnelle des deux sous-populations de noyaux filles. Par une accélération générale de la transcription, Zelda tend ainsi à diminuer le biais de mémoire mitotique. Cependant ce biais de mémoire est récupéré

lorsqu'on diminue l'expression de ce facteur dans l'embryon, ce qui suggère l'implication d'autres facteurs dans le processus de mémoire mitotique. Pour affiner notre compréhension du rôle biologique de Zelda dans l'activation du génome zygotique, nous avons formalisé nos données en développant un modèle mathématique, qui permet de mieux appréhender le nombre et la durée des étapes nécessaires à l'activation de la transcription. Dans ce modèle, l'activation de la transcription est atteinte

après avoir parcouru une série de transitions aléatoires entre états épigénétiques discrets (Figure 2C). Le modèle prédit que le type de distribution du temps d'activation (loi exponentielle, loi gamma, etc.) dépend du nombre d'états discrets. L'analyse des résultats expérimentaux en utilisant ce modèle indique qu'en moyenne, les noyaux provenant de noyaux mères, dans lesquels la transcription est active, doivent parcourir un nombre plus petit de transitions jusqu'à l'activation, et suggère

que Zelda agit principalement en diminuant la durée (et non le nombre) des étapes de pré-initiation de la transcription, révélant ainsi son rôle accélérateur dans l'embryon précoce de drosophile.

Zelda se fixe transitoirement à la chromatine

Compte tenu des caractéristiques des facteurs pionniers, nous nous attendions à ce que Zelda joue un rôle dans la conservation de la mémoire transcriptionnelle durant la mitose. Cependant, nos données génétiques et notre modélisation indiquent que Zelda n'est pas le support de cette mémoire. Nous avons analysé la localisation de Zelda dans des noyaux d'embryons vivants et avons montré que Zelda quitte le noyau à chaque mitose, mais y retourne rapidement en fin d'anaphase. Afin de préciser les propriétés dynamiques de cet activateur du génome, nous avons eu recours à des expériences d'imagerie quantitative (*fluorescence recovery after photobleaching* et *fluorescence correlation spectroscopy*) sur embryon vivant. Nous avons montré que ce facteur de transcription « pionnier » ne se liait à la chromatine que très transitoirement, avec un temps de résidence estimé de l'ordre de quelques secondes. Cependant, ce faible temps de résidence de la protéine Zelda sur la chromatine pourrait être compensé par une augmentation locale de sa concentration. En effet, nous avons montré que Zelda s'accumu-

lait au sein de « microenvironnements » nucléaires (*hubs*) (Figure 2D), ce qui a aussi été rapporté par une autre équipe en utilisant une technique de microscopie à haute résolution (*lattice light sheet*) permettant de suivre des particules uniques [4]. L'accumulation de Zelda dans ces « microenvironnements » nucléaires pourrait favoriser une coopérativité entre plusieurs facteurs de transcription [4] ou le rapprochement de plusieurs segments d'ADN cibles (*enhancers*) pour favoriser l'activation transcriptionnelle des gènes zygotiques (Figure 2E).

Ces travaux de recherche [9] ouvrent de nouvelles pistes pour mieux comprendre le rôle des facteurs de transcription « pionniers » (e.g. *Oct4*, *Pou5f3*, *Sox2*), l'activation des gènes et l'organisation nucléaire durant la période d'activation du génome zygotique chez les vertébrés. Les gènes codant des facteurs de transcription « pionniers », tels que ceux de la famille *Forkhead box* (*Fox*), sont impliqués dans de nombreuses maladies génétiques complexes telles que les cancers, la maladie de Parkinson ou les troubles du spectre autistique [10]. De ce fait, la compréhension de l'activité spatio-temporelle de ces facteurs de transcription « pionniers » permettrait de progresser dans la physiopathologie de ces maladies dans le but d'identifier de nouvelles stratégies thérapeutiques. ♦

Zelda, maestro of the zygotic genome awakening

REMERCIEMENTS

Nous remercions Matthieu Dejean pour sa lecture critique du manuscrit. Ces travaux ont été soutenus par l'ERC SyncDev et une subvention HFSP-CDA.

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development* 2009 ; 136 : 3033-42.
2. Hamm DC, Harrison MM. Regulatory principles governing the maternal-to-zygotic transition: insights from *Drosophila melanogaster*. *Open Biol* 2018 ; 8 : 180183.
3. Yamada S, Whitney PH, Huang SK, et al. The drosophila pioneer factor Zelda modulates the nuclear microenvironment of a dorsal target enhancer to potentiate transcriptional output. *Curr Biol* 2019 ; 29 : 1387-93.
4. Mir M, Stadler MR, Ortiz SA, et al. Dynamic multifactor hubs interact transiently with sites of active transcription in *Drosophila* embryos. *Elife* 2018 ; 7 (10.7554/eLife.40497).
5. Mir M, Reime A, Haines JE, et al. Dense Bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev* 2017 ; 31 : 1784-94.
6. Bertrand E, Chartrand P, Schaefer M, et al. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 1998 ; 2 : 437-45.
7. Ferraro T, Esposito E, Mancini L, et al. Transcriptional memory in the drosophila embryo. *Curr Biol* 2016 ; 26 : 212-8.
8. Bellec M, Radulescu O, Lagha M. Remembering the past: mitotic bookmarking in a developing embryo. *Curr Opin Syst Biol* 2018 ; 11 : 41-9.
9. Dufourt J, Trullo A, Hunter J, et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat Commun* 2018 ; 9 : 5194.
10. Golson ML, Kaestner KH. Fox transcription factors: from development to disease. *Development* 2016 ; 143 : 4558-70.