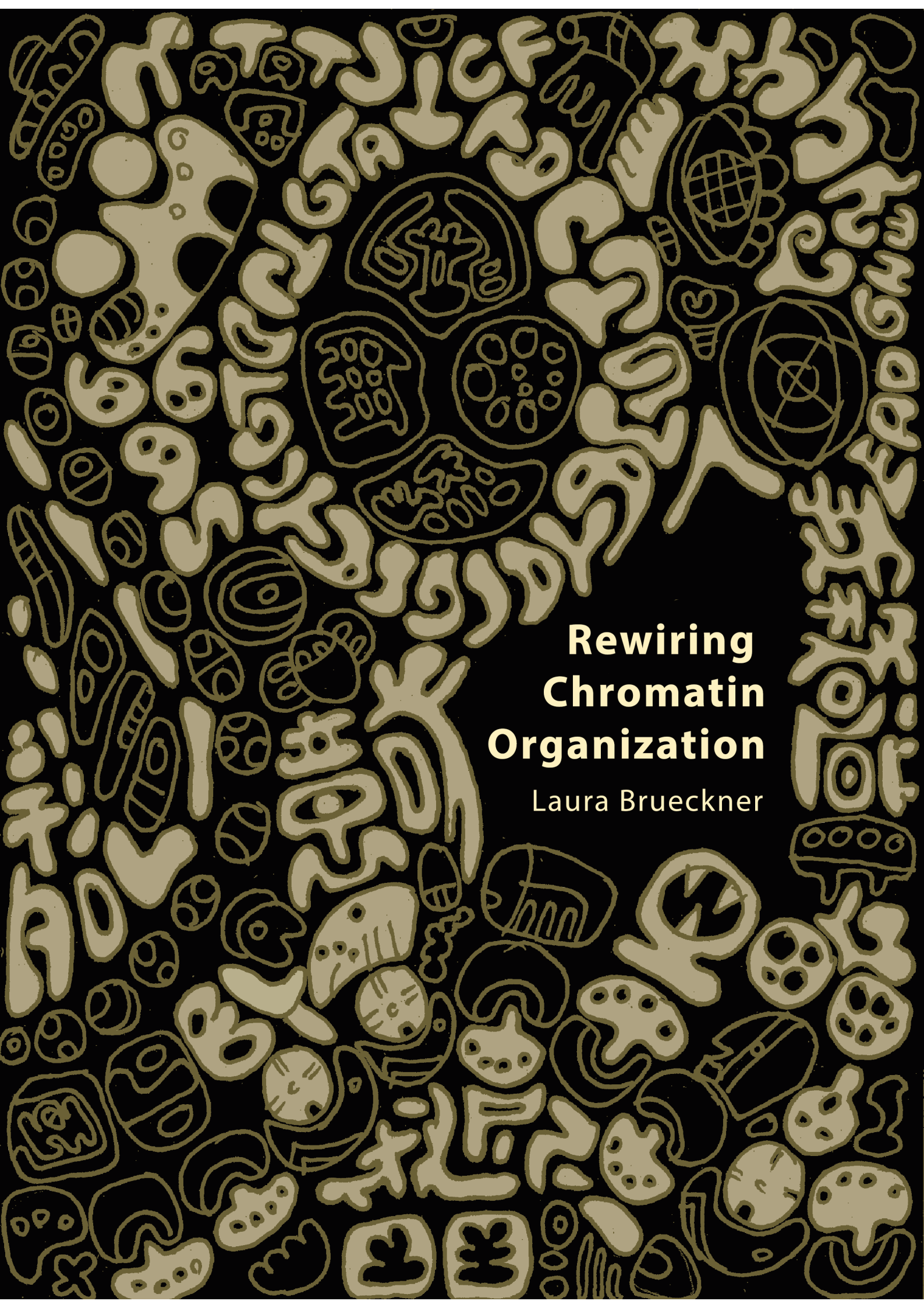


Rewiring Chromatin Organization

Laura Brueckner

Rewiring Chromatin Organization

Laura Brueckner

This research was conducted under the auspices of the Oncology Graduate School at the Netherlands Cancer Institute (NKI-AVL), Amsterdam

Rewiring Chromatin Organization

Chromatine-organisatie opnieuw bedraden

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Friday the 10th of September 2021 at 10:30 hours
by

Laura Brueckner
born in Frankfurt am Main, Germany

Erasmus University Rotterdam

Doctoral Committee:

Promotors: Prof.dr. B. van Steensel
Prof.dr. F van Leeuwen

Other members:
Prof.dr. J.N.J. Philipsen
Dr. P.J. Verschure
Prof.dr. P. Verrijzer

Table of Contents

Chapter 1

Introduction

Introduction

The organization of the genome in the nucleus is non-random

The human genome is encoded in DNA sequences with a total length of around 2 meters. These strings of information have to fit into the nucleus of each human cell which is only a few micrometers in size. It is obvious that the genomic DNA has to be compacted in order to fit into such a tiny compartment. At the same time, certain regions have to be kept accessible in order to allow gene expression. Therefore, it makes sense that the DNA is not just stuffed randomly into the nucleus. Instead it is organized in a way that allows efficient access to active genes while packing away regions where genes are either not present or inactive.

DNA is not existing in the nucleus as a naked string. In eukaryotes, DNA is wrapped around histone proteins and packed into nucleosomes. This structure of DNA associated with proteins is known as “chromatin”[1]. Genomic regions can then further be condensed into a state called “heterochromatin” which can be seen in light microscopy as dark matter in interphase cells [2,3]. One would assume that this state limits accessibility of genetic information and is therefore an obstacle to gene expression. Indeed, heterochromatic regions in the genome are mostly devoid of or low in transcription [4,5]. On the other hand, regions that are uncondensed and are called euchromatin. These are the regions where highly transcribed genes can be found.

The aforementioned histone proteins have an interesting property: They have flexible “tails” which can be chemically marked. Certain modifications correlate with active or inactive chromatin. For example, methylation on histone H3 lysine residue 9 (H3K9) is associated with heterochromatin, whereas H3K4 methylation is found on active genes [36].

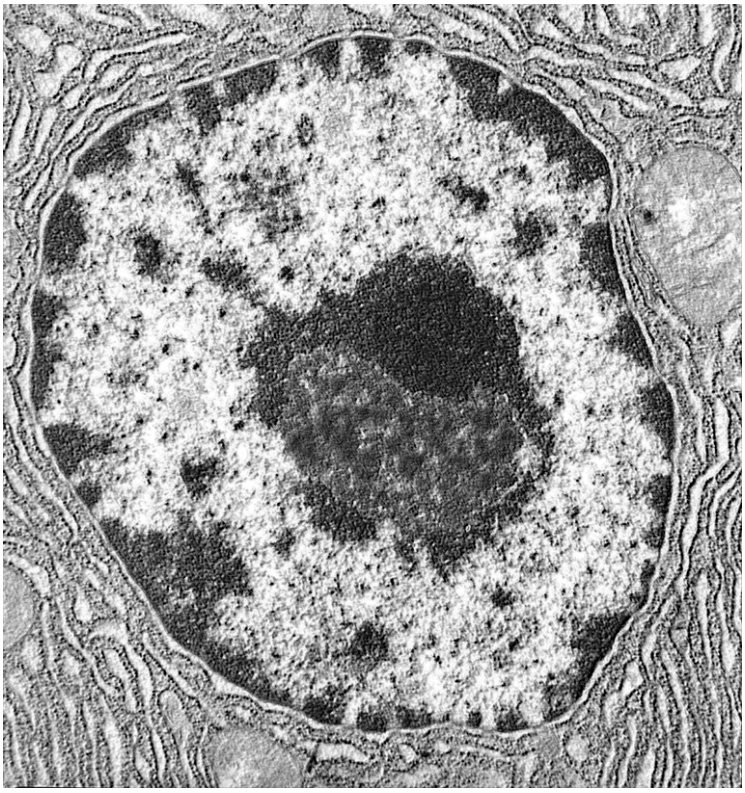


Figure 1 Electron micrograph of a mouse cell nucleus. Heterochromatin is seen as dark staining. (image provided by Kenneth M. Barth)

Interestingly, the dense heterochromatin is mostly seen close to the nuclear envelope which is covered in a fibrous structure known as the nuclear lamina (Fig. 1) [7,8]. The nuclear lamina consists of a meshwork of lamins and lamin-associated proteins and seems to act as a buffer for mechanical stress [9]. Mutations in lamin proteins lead to diseases known as laminopathies such as Hutchinson-Gilford progeria or Emery-Dreifuss muscular dystrophy. On the cellular level, lamin mutations can lead to loss of heterochromatin domains and nucleus deformation [10–14]. Notably, in rod photoreceptor cells of nocturnal and diurnal animals this architecture is inverted, with the heterochromatin localized in the nuclear center. Therefore “heterochromatin is located at the nuclear envelope” is not a strict rule [15].

In order to determine how DNA is arranged in the nucleus, the nuclear envelope can be used as a reference point. By fusing a protein of interest with a DNA-modifying enzyme such as Dam methyltransferase, genomic regions that are in contact with this protein can be chemically labelled. This technique is called DamID [16]. By performing DamID with a fusion of the Dam methyltransferase to lamin, genomic regions in proximity to the nuclear envelope can be labelled. If DNA would be randomly arranged in each cell, the DamID signal would also just be seen as random noise all over the genome. However, genomic maps generated by DamID on pools of cells revealed distinct regions preferentially located on the outer edge of the nucleus. These regions are called lamina-associated domains (LADs), are 0.1-1

megabases in size and enriched in H3K9 dimethylation as well as H3K27 trimethylation at the LAD borders [17,18]. LADs are largely consistent between single cells of the same type but change during differentiation [19,20]. Their counterpart are inter-LADs, regions that are mostly residing in the nuclear interior. Genes that become active during differentiation tend to move away from the nuclear periphery, showing a connection between gene expression and position in the nucleus [19].

Another way to look at the 3D organization of the genome is to study how DNA is looped. “3C” (chromosome conformation capture) is a technique which allows to test whether two loci are in contact with each other [21]. The technique was further advanced to “4C” which uses a “bait” region to fish for other regions in contact with it [22]. Finally, “Hi-C” was developed. The Hi-C assay first crosslinks the whole genome and then uses enzymatic digestion followed by high-throughput sequencing to map all genomic interactions [23]. This way, genome-wide maps of genome topology could be generated. Genome interactions are not random. There are regions where higher interaction frequencies are observed, showing up as distinct “pyramids” in the Hi-C profile. These are regions with a size of around 200 kb -1 Mb that preferentially interact with themselves and which are called “topologically associated domains” or “TADs” [24,25]. When zooming out and looking at whole chromosomes, larger regions can be observed that interact with each other over large distances, making up the “checkerboard” pattern of the Hi-C plot. These regions indicate two major compartments named "A" and "B" which are associated with active and inactive transcription, respectively. B compartment regions are more densely packed and mostly overlap with LADs [20,22,23,26].

The genome can be subdivided into chromatin states

The genome cannot only be studied in terms of its spatial arrangement. DNA is covered in a multitude of proteins, making up chromatin. Chromatin proteins are not distributed randomly. Instead certain proteins are frequently found in combination with others, making up distinct chromatin “states”. In a comprehensive effort to map the genomic association of individual chromatin proteins, the van Steensel lab generated maps of 52 proteins in *Drosophila* [27]. The study described the preferential association of these proteins in a model of five chromatin states based on principal component analysis. These states were named after colors YELLOW, RED, BLUE, GREEN and BLACK. YELLOW and RED are active states associated with housekeeping or developmentally regulated genes, respectively. BLUE is a repressed state associated with proteins of the well-described Polycomb machinery and GREEN demarcates heterochromatin protein 1 (HP1)-bound heterochromatin. The model

also revealed an intriguing BLACK state which is mostly devoid of chromatin marks and transcriptional activity. BLACK chromatin largely overlaps with lamina-associated regions, bringing together spatial organization and chromatin composition.

In a parallel effort, the ModEncode consortium generated maps of 18 histone modifications in *Drosophila* and defined 9 chromatin states based on Hidden Markov Model analysis which partially overlap with the five colors model but are more fine-grained. This model contains a "void" state 9 which shows low enrichment of the mapped chromatin modifications and overlaps with the BLACK state of the five colors model [28].

In mammalian cells, a state model was defined based on chromatin marks processed by a Hidden Markov Model. This ChromHMM model shows higher granularity for active regions compared to the *Drosophila* states. It describes for example different enhancer and promoter states and only contains a general "repressed" state instead of distinguishing between Polycomb- and HP1 associated chromatin [29].

Experimental tools for targeting chromatin modifiers

The association of genomic regions with certain chromatin proteins has been shown in genome-wide studies. But what are the implications of these chromatin states? What effect do these proteins have on gene expression? In order to address this question, researchers have developed methods to target a protein of choice to a certain genomic region such as a reporter gene. The following paragraphs present a brief overview of tools for targeting chromatin modifiers.

Repurposing transcription factor DNA-binding domains

Transcription factors are DNA-associated proteins that mediate gene expression. These proteins have domains which give them the ability to bind to certain DNA motifs. Several of these DNA-binding domains have been identified, prominent examples being the Lac inhibitor (LacI) protein binding the LacO sequence in *E. coli* lac operon [30] or the Gal4DBD (DNA binding domain) binding the UAS (upstream activation sequence) of the yeast Gal4 locus [31,32]. These domains have been repurposed by synthetic biologists by fusing a DNA-binding domain to a protein of interest. Hereby the protein can be targeted to a specific genetic site in the genome. This technique offers a plethora of possibilities to study protein function and gene expression.

Inducible gene activation

One example for a synthetic DNA binding protein is the Tet-On/Tet-Off system which uses the Tet-repressor (TetR) from *E. coli*. The TetR protein has been fused to an activating domain from *H. simplex* virus (VP16) generating the tetracycline transactivator (Tta) which is used for inducible gene expression. The TetR protein has a special property: it is only binding to its motif when bound (Tet-On) or not bound (Tet-Off) by tetracycline [33].

Tethering of chromatin modifiers for reverse-engineering chromatin states

Recruiting proteins to DNA is not limited to transcriptional activators. Tethering of chromatin-modifying proteins can be used to "paint" a locus of interest in a certain chromatin state. For instance, a study by Hathaway et al used recruitment of Gal4DBD-HP1 to trigger H3K9 methylation and to repress an integrated GFP reporter gene. The specialty of this study is that the HP1 recruitment could be controlled by adding a small molecule. This system was used to study the persistence of the induced chromatin modifications and gene repression in mouse embryonic stem cells and fibroblasts. The study showed that the H3K9 methylation mark and reporter gene repression could be inherited throughout mitotic divisions. The study also showed that this "memory" of repression was dependent on DNA methylation [34]. For chromatin proteins with unknown function, tethering this protein to a reporter gene can elucidate its role in gene expression. For example by recruiting the Set2 protein to a reporter gene it was shown that it has a repressive function [(35(p2)].

TALE-fusions for sequence-flexible targeting

One drawback of systems based on transcription factor fusion is that targeting is limited to a specific sequence. The target motif has to be integrated into the locus of interest by genetic manipulation which is not always desirable. More flexibility is offered by using Transcription Activator-Like Effector (TALE)-based synthetic DNA binding domains which in principle can be designed to target any DNA sequence of choice [36,37]. When fused for example to a transcriptional activator such as VP64, they can be used to upregulate an endogenous gene of interest [38,39]. Fusions with repressive domains such as KRAB or SID can be used for downregulation [40,41]. However, cloning TALE proteins specific for a sequence of choice is laborious.

CRISPR as an easy-to-use and highly flexible targeting system

An even more flexible and easy-to-use system is CRISPR-Cas9. The Cas9 protein can be targeted to a 20-nucleotide sequence using a matching sgRNA [42–44]. Designing and cloning sgRNAs is simple and cheap which makes Cas9 a fantastic tool to target the whole

genome. By fusing a transcriptional repressor or activator to a catalytically dead Cas9 protein, CRISPR has become an incredibly powerful technique to manipulate gene expression but not only that - CRISPR has also successfully been used to induce chromatin looping and is therefore an invaluable tool to study the effects of genome organization [45–49].

Chromatin at the nuclear periphery

Genomic regions associated with the nuclear lamina tend to be gene-poor and the genes that are transcribed do so at a very low level [17]. The association of inactive genome regions with the nuclear periphery poses a "chicken or egg" question: Do these regions wander towards the nuclear envelope because they are inactive or on the other hand, does the lamina actively repress any gene associated with it?

To address this question, a technique called TRIP was developed [50]. TRIP stands for "thousands of reporters integrated in parallel". In short, thousands of copies of a GFP reporter gene are integrated into the genomes of a pool of cells using a transposase. Each copy has a barcode which is part of its mRNA, making it possible to determine the expression depending on the locus of integration. Quantifying the activity of reporters in LADs showed that their activity is on average 5-fold lower compared to reporters in the nuclear interior. This experiment suggests that LADs are indeed refractory to gene expression. However, LADs have several other properties in addition to being located at the nuclear periphery such as being gene-poor or enriched for repressive chromatin marks [17].

So, can it be stated that the interaction with the nuclear lamina leads to repression? Several experiments have been performed to address this question by artificially tethering loci to the nuclear periphery but the results differ. In *Drosophila*, targeting of Lamin C to a reporter locus by using the Gal4 system led to repression [51]. The study also showed defects in muscle development upon expression of mutant Lamin C. In human cells, two different chromosomes were relocated to the nuclear periphery by using LacI fused to nuclear membrane protein Lap2beta. Some genes up to 4 Mb away from the tethering sites were downregulated while others remained active [52]. One study used a construct of LacI fused to nuclear envelope protein which was recruited to randomly integrated transgenes in mouse fibroblasts. The effect was downregulation [53]. Another study used a large 4 Mb repeat-based transgene locus tethered via an inducible LacI-mCherry-laminB1 fusion protein. The authors described that the locus retained its transcriptional competence but did not report if there was any downregulation [54]. Interestingly, the two latter studies mentioned that repositioning required transition through one cell cycle. In summary, repositioning of a

transcriptionally active locus was refractory to gene expression in most cases, showing that the spatial position of a gene can indeed influence transcription. However, some loci seemed to escape this repressive effect but these exceptions were not further explored in the aforementioned studies.

All studies mentioned so far were working with transgenes using a small set of well-known promoters. Therefore, the applicability of these results to the large variety of endogenous promoters is limited. A recent genome-wide survey of promoter autonomy showed that some lamina-associated promoters are inactive in their endogenous position but suddenly become active when moved to an episomal setting. This indicates that these promoters have the factors necessary for expression but are hindered by their position at the nuclear lamina. However other promoters seemingly have no problem expressing at the nuclear periphery. Follow-up experiments confirmed that different promoters vary in their susceptibility to being repressed at the nuclear periphery. There are gene promoters that, despite being positioned at the nuclear periphery, "escape" repression [55,56].

But what happens when a gene positioned at the nuclear lamina becomes upregulated? Almost two decades ago, a microscopy study showed repositioning of a peripheral locus to the nuclear interior upon targeting an integrated transgene with transcriptional activator VP16 [57]. Repositioning was later shown to occur 1-2 hours after targeting and to be perturbed by actin or nuclear myosin mutations [58]. Another study targeted lamina-associated regions genome-wide with VP16 and showed by microscopy that these regions subsequently moved away from the nuclear periphery. However, the observed regions only moved away slightly from the nuclear envelope, unlike the long-range movement seen in the previously mentioned studies. RNA sequencing revealed almost no changes in lamina-associated genes, indicating that detachment was not triggered by changes in transcription. However, detached regions seemed to have decreased H3K9 dimethylation, implicating that the VP16 targeting caused a change in chromatin state which led to repositioning. Nuclear lamina association could also be decreased when inhibiting H3K9 methyltransferase G9a, showing a causal role for H3K9 dimethylation in lamina association [59].

The aforementioned repositioning upon VP16 targeting is rather artificial. What about endogenous genes that change their position during development? The van Steensel lab looked at gene expression and lamina association changes during differentiation of mouse embryonic stem cells to neural precursor cells and astrocytes by DamID. Genes that moved to the nuclear interior during differentiation were frequently upregulated and vice versa.

But also, genes that did not immediately change their expression upon moving to the interior had an increased chance to become upregulated in the later differentiation stage. Movement away from the nuclear periphery seemed to prime these genes for upregulation. As with the VP16 system, these regions initially relocated without transcription as the driving force. However, upregulated genes that moved away from the nuclear lamina during differentiation mostly did so as singletons, which would indicate that transcription demarcates the region which detaches [19].

There is more experimental evidence underlining the role of chromatin state in genome architecture. Mere chromatin decondensation has been shown to trigger relocalization to the nuclear interior in several cases. Experiments were performed by targeting peripheral loci with a serendipitously found peptide which triggers chromatin decondensation but does not induce transcription. This observation was first made by recruiting the decondensation peptide to large transgene [58]. Later this finding was consolidated in a more natural setting by targeting three endogenous genes via TALE-fusion constructs combined with fluorescence in-situ hybridisation (FISH)-based imaging of the target genes. Both the decondensation peptide as well as VP16-induced transcriptional activation led to repositioning to the nuclear interior [60]. Inhibiting transcription with alpha-amanitin did not prevent relocalization, showing that repositioning can be mediated by chromatin effects only [58].

Gene position and transcriptional activity therefore do not show a unidirectional causality but instead influence each other. Changes in chromatin condensation and histone modifications have been linked to changes in spatial position but these mechanisms need to be further explored. These observations show that it is important to dissect the roles of transcription and chromatin state, e.g. under which circumstances or in which loci is transcription the driving force for repositioning rather than chromatin state and vice versa. However, separating transcription and chromatin state is a challenging task for experimenters.

HP1 as a repressive chromatin protein

HP1 is an archetypical heterochromatin chromatin protein which binds to H3K9 di- and trimethylated nucleosomes via its chromodomain to mediate chromatin condensation and transcriptional silencing [61–66]. H3K9 dimethylation is also a mark of lamina-associated domains. However, HP1 does not strictly localize to the nuclear periphery but is also prominently observed in heterochromatic foci in the nuclear interior. In the chromatin 5-state

model in *Drosophila*, HP1-associated chromatin was identified as a distinct state separate from lamina-associated chromatin [27]. However, this lack of overlap was later shown to be a specific property of the embryonic-origin Kc167 cells used in the study. In neurons, glia and fat body there was a strong overlap between HP1 and LADs, showing that is important to look at various cell types when studying chromatin states [67].

However, biochemical studies have shown that HP1 is able to bind to the lamin B receptor which indicates at least a partial mechanistic overlap between HP1-associated and lamina-associated heterochromatin [68–71]. In human cells, protein PRR14 has been found to tether peripheral HP1-H3K9me3-marked heterochromatin to the nuclear lamina, explaining how transcriptional silencing by HP1 can be linked with spatial positioning. PRR14 knockdown led to deformities of the nucleus [72]. This finding is reminiscent of observations that untethering of heterochromatin from the nuclear envelope led to increased deformability of the nucleus [73].

HP1 can oligomerize and is thought to compact genomic loci by nucleosome bridging. Findings in *Drosophila* also imply that HP1 can bridge loci over large distances of multiple megabases [74–76].

HP1 has recently been shown to drive phase-separation of chromatin which has sparked interest in liquid-liquid phase separation as a mechanism for genome organization [77–80]. But the relevance for phase separation in living cells is still under debate [81].

Chromatin remodeling by nucleosome sliding

Apart from condensation there are other ways in which chromatin can be remodeled. Nucleosomes can be ejected or slide along the DNA, e.g. to expose binding sites for transcription factors or to occlude them [82]. This kind of remodeling can be performed by ATP-dependent chromatin remodeling complexes (CRCs) which can be grouped into four major classes SWI/SNF, INO80, CHD or ISWI based on their ATPases [83]. The *Drosophila* CHRAC/ACF complex contains ISWI [84] and is the subject of chapter 3 of this thesis.

Studies on genome 3D organization mostly examine histone modifications in combination with higher-order structure. Nucleosome positioning is generally not subject of these studies. Therefore, not much is known about the contribution of ATP-dependent CRCs to higher-level genome architecture. One of the main factors regulating chromatin higher-order structure is CTCF which binds specific DNA motifs and mediates DNA loop formation [85].

Interestingly, up to 20 nucleosomes around CTCF binding sites are specifically arranged [86]. The positioning of nucleosomes around CTCF sites has been shown to be influenced by ATPases SNF2H which is part of the nucleosome remodeling factor (NURF) complex or SMARCA4 which is associated with the SWI/SNF complex [87,88]. Perturbation of SMARCA4 lead to a weakening of TAD boundary strength [88]. Snf2h knockout in mouse embryonic stem cells was recently shown to affect formation of chromatin loops and TAD insulation. Interestingly, CTCF occupancy was not affected [89].
For another subunit of the SWI/SNF complex, ARID1A, knockout led to an increase in TAD border strength and several B to A compartment switches seen in Hi-C [90].

In summary, these recent results indicate that even enzymes influencing the small-scale structure of chromatin can have effects on large-scale organization. Future studies might further elucidate the role of ATP-dependent chromatin remodelers in 3D genome organization.

Thesis outline

Within the last two decades, techniques such as DamID ad Hi-C have made it possible to visualize the 3D organization of the genome as well as the distribution of chromatin proteins. In combination with next-generation sequencing, genome-wide maps have been generated which could be further combined with transcriptome data. These studies revealed a connection between three major aspects of nuclear biology: chromatin-associated proteins, 3D architecture and transcriptional activity. As these studies are correlative in nature, it is necessary to understand causality. It has therefore become imperative to manipulate these three components individually.

Maps of chromatin composition have revealed the presence of distinct chromatin states. This raises the question:
What is the function of these chromatin states?
Or more specifically:
Does a specific protein function better in its “home” state, and, what happens if this protein is taken out of its native environment?

In **chapter 2** of this thesis, to address these questions, a classical tethering experiment using recruitment of repressive protein HP1 was elevated to a high-throughput assay. This was possible by massive, genome-wide integration of barcoded reporter gene using TRIP technology in *Drosophila* cells. These reporters were then used as a platform for targeting

HP1, an archetypical heterochromatic protein known to silence transcription. The experiments described in this chapter showed that indeed, HP1 works as a powerful repressor, downregulating the target reporter gene irrespective of its site of integration. However, we also discovered that HP1 shows a preference for a certain chromatin context. In pericentric heterochromatin, its native environment, HP1-induced repression was enhanced by twofold. This study shows that genomic environment can influence the function of chromatin modifiers and presents a flexible system to assay other proteins of interest.

Chapter 3 presents a collaborative study employing the assay described in chapter 2. High-throughput tethering of chromatin protein CHRAC/ACF revealed an inactivating function on poorly transcribed genes in repressive chromatin. This study showed a novel role for CHRAC/ACF in maintaining a repressive ground state in *Drosophila* cells.

The chapters introduced so far cover epigenetic factors in chromatin organization. But what about the central mechanism that genome organization is all about – gene transcription?

In **chapter 4** we sought to study the effect of gene expression on genome architecture by manipulating transcription at specific loci. Previous experiments have shown that genes can move to the nuclear interior upon activation. However so far these had been either genome-wide studies with low coverage for individual genes or studies on a small subset of genes using FISH technology. What was still missing is detailed information on what relocalization looks like, i.e. do genes move as singlets when activated or can they affect other genes? To what extent are flanking regions dragged along? We therefore introduced transcriptional activity either by random integration of a reporter gene or using TALE-VP64 or CRISPRa. Inversely, we abolished transcriptional activity by deletion of promoters or gene truncation. Subsequently, we generated high-resolution DamID maps to look for changes in genome architecture. We observed that all targeted genes relocated to the interior upon activation. However, this repositioning is limited to one gene at a time and does not affect neighboring genes. We also observed that genes moved to the nuclear periphery upon inactivation by promoter deletion. Truncation of a transcription unit led to attachment of the untranscribed region to the nuclear lamina. All in all, this study provides depictions of nuclear relocalization events upon manipulation of gene expression at unprecedented resolution, showing how transcription shapes 3D architecture.

In **chapter 5** the results from these studies are summarized and discussed in relation to recent developments in the field of 3D genome architecture. Open questions are addressed and potential future research strategies are presented.

References

1. Hansen JC. Conformational Dynamics of the Chromatin Fiber in Solution: Determinants, Mechanisms, and Functions. *Annu Rev Biophys Biomol Struct*. 2002;31(1):361-392. doi:10.1146/annurev.biophys.31.101101.140858
2. Heitz E. *Das Heterochromatin der Moose*. Bornträger; 1928.
3. Yunis JJ, Yasmineh WG. Satellite DNA in constitutive heterochromatin of the guinea pig. *Science*. 1970;168(3928):263-265. doi:10.1126/science.168.3928.263
4. Saksouk N, Simboeck E, Déjardin J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin*. 2015;8(1):3. doi:10.1186/1756-8935-8-3
5. Flamm WG, Walker PMB, McCallum M. Some properties of the single strands isolated from the DNA of the nuclear satellite of the mouse (Mus musculus). *J Mol Biol*. 1969;40(3):423-443. doi:10.1016/0022-2836(69)90163-6
6. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res*. 2011;21(3):381-395. doi:10.1038/cr.2011.22
7. Pueschel R, Coraggio F, Meister P. From single genes to entire genomes: the search for a function of nuclear organization. *Development*. 2016;143(6):910-923. doi:10.1242/dev.129007
8. Gerace L, Comeau C, Benson M. Organization and Modulation of Nuclear Lamina Structure. *J Cell Sci*. 1984;1984(Supplement 1):137-160. doi:10.1242/jcs.1984.Supplement_1.10
9. Sapra KT, Qin Z, Dubrovsky-Gaupp A, et al. Nonlinear mechanics of lamin filaments and the meshwork topology build an emergent nuclear lamina. *Nat Commun*. 2020;11(1):6205. doi:10.1038/s41467-020-20049-8
10. Gordon LB, Rothman FG, López-Otín C, Misteli T. Progeria: A Paradigm for Translational Medicine. *Cell*. 2014;156(3):400-407. doi:10.1016/j.cell.2013.12.028
11. McCord RP, Nazario-Toole A, Zhang H, et al. Correlated alterations in genome organization, histone methylation, and DNA–lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res*. 2013;23(2):260-269. doi:10.1101/gr.138032.112
12. Sullivan T, Escalante-Alcalde D, Bhatt H, et al. Loss of a-Type Lamin Expression Compromises Nuclear Envelope Integrity Leading to Muscular Dystrophy. *J Cell Biol*. 1999;147(5):913-920. doi:10.1083/jcb.147.5.913
13. Dahl KN, Scaffidi P, Islam MF, Yodh AG, Wilson KL, Misteli T. Distinct structural and mechanical properties of the nuclear lamina in Hutchinson–Gilford progeria syndrome. *Proc Natl Acad Sci*. 2006;103(27):10271-10276. doi:10.1073/pnas.0601058103
14. Broers JLV, Peeters EAG, Kuijpers HJH, et al. Decreased mechanical stiffness in LMNA−/− cells is caused by defective nucleo-cytoskeletal integrity: implications for the development of laminopathies. *Hum Mol Genet*. 2004;13(21):2567-2580. doi:10.1093/hmg/ddh295

15. Solovei I, Kreysing M, Lanctôt C, et al. Nuclear Architecture of Rod Photoreceptor Cells Adapts to Vision in Mammalian Evolution. *Cell*. 2009;137(2):356-368. doi:10.1016/j.cell.2009.01.052
16. Steensel B, Henikoff S. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol*. 2000;18:424-428.
17. Guelen L, Pagie L, Brasset E, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*. 2008;453(7197):948-951. doi:10.1038/nature06947
18. Poleshko A, Smith CL, Nguyen SC, et al. H3K9me2 orchestrates inheritance of spatial positioning of peripheral heterochromatin through mitosis. Aguilera A, Tyler JK, Belmont AS, eds. *eLife*. 2019;8:e49278. doi:10.7554/eLife.49278
19. Peric-Hupkes D, Meuleman W, Pagie L, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell*. 2010;38(4):603-613. doi:10.1016/j.molcel.2010.03.016
20. Kind J, Pagie L, de Vries SS, et al. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell*. 2015;163(1):134-147. doi:10.1016/j.cell.2015.08.040
21. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002;295(5558):1306-1311. doi:10.1126/science.1067799
22. Simonis M, Klous P, Splinter E, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet*. 2006;38(11):1348-1354. doi:10.1038/ng1896
23. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289-293. doi:10.1126/science.1181369
24. Dixon JR, Selvaraj S, Yue F, et al. Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature*. 2012;485(7398):376-380. doi:10.1038/nature11082
25. Nora EP, Lajoie BR, Schulz EG, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485:381-385.
26. Steensel B, Belmont A. Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell*. 2017;169:780-791.
27. Filion GJ, van Bemmel JG, Braunschweig U, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*. 2010;143(2):212-224. doi:10.1016/j.cell.2010.09.009
28. Kharchenko PV, Alekseyenko AA, Schwartz YB, et al. Comprehensive analysis of the chromatin landscape in Drosophila. *Nature*. 2011;471(7339):480-485. doi:10.1038/nature09725

29. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods*. 2012;9(3):215-216. doi:10.1038/nmeth.1906
30. Gilbert W, Müller-Hill B. ISOLATION OF THE LAC REPRESSOR. *Proc Natl Acad Sci U S A*. 1966;56(6):1891-1898.
31. Kakidani H, Ptashne M. GAL4 activates gene expression in mammalian cells. *Cell*. 1988;52(2):161-167. doi:10.1016/0092-8674(88)90504-1
32. Webster N, Jin JR, Green S, Hollis M, Chambon P. The yeast UASG is a transcriptional enhancer in human hela cells in the presence of the GAL4 trans-activator. *Cell*. 1988;52(2):169-178. doi:10.1016/0092-8674(88)90505-3
33. Das AT, Tenenbaum L, Berkhout B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther*. 2016;16(3):156-167. doi:10.2174/1566523216666160524144041
34. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. Dynamics and memory of heterochromatin in living cells. *Cell*. 2012;149(7):1447-1460. doi:10.1016/j.cell.2012.03.052
35. Strahl BD, Grant PA, Briggs SD, et al. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol*. 2002;22(5):1298-1306. doi:10.1128/mcb.22.5.1298-1306.2002
36. Cermak T, Doyle EL, Christian M, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011;39(12):e82. doi:10.1093/nar/gkr218
37. Garg A, Lohmueller JJ, Silver PA, Armel TZ. Engineering synthetic TAL effectors with orthogonal target sites. *Nucleic Acids Res*. 2012;40(15):7584-7595. doi:10.1093/nar/gks404
38. Perez-Pinera P, Ousterout DG, Brunger JM, et al. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat Methods*. 2013;10(3):239-242. doi:10.1038/nmeth.2361
39. Maeder ML, Linder SJ, Reyon D, et al. Robust, synergistic regulation of human gene expression using TALE activators. *Nat Methods*. 2013;10(3):243-245. doi:10.1038/nmeth.2366
40. Cong L, Zhou R, Kuo Y-C, Cunniff M, Zhang F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun*. 2012;3:968. doi:10.1038/ncomms1962
41. Li Y, Moore R, Guinn M, Bleris L. Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Sci Rep*. 2012;2:897. doi:10.1038/srep00897
42. Cong L, Ran FA, Cox D, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*. 2013;339(6121):819-823. doi:10.1126/science.1231143

43. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823-826. doi:10.1126/science.1232033
44. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821. doi:10.1126/science.1225829
45. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014;159(3):635-646. doi:10.1016/j.cell.2014.09.039
46. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*. 2013;154(2):442-451. doi:10.1016/j.cell.2013.06.044
47. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*. 2014;32(4):347-355. doi:10.1038/nbt.2842
48. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*. 2013;152(5):1173-1183. doi:10.1016/j.cell.2013.02.022
49. Morgan SL, Mariano NC, Bermudez A, et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat Commun*. 2017;8(1):15993. doi:10.1038/ncomms15993
50. Akhtar W, de Jong J, Pindyurin AV, et al. Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell*. 2013;154(4):914-927. doi:10.1016/j.cell.2013.07.018
51. Dialynas G, Speese S, Budnik V, Geyer PK, Wallrath LL. The role of Drosophila Lamin C in muscle function and gene expression. *Dev Camb Engl*. 2010;137(18):3067-3077. doi:10.1242/dev.048231
52. Finlan LE, Sproul D, Thomson I, et al. Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. *PLOS Genet*. 2008;4(3):e1000039. doi:10.1371/journal.pgen.1000039
53. Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. 2008;452(7184):243-247. doi:10.1038/nature06727
54. Kumaran RI, Spector DL. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol*. 2008;180(1):51-65. doi:10.1083/jcb.200706060
55. Leemans C, van der Zwalm MCH, Brueckner L, et al. Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs. *Cell*. 2019;177(4):852-864.e14. doi:10.1016/j.cell.2019.03.009

56. van Arensbergen J, FitzPatrick VD, de Haas M, et al. Genome-wide mapping of autonomous promoter activity in human cells. *Nat Biotechnol*. 2017;35(2):145-153. doi:10.1038/nbt.3754
57. Tumbar T, Belmont AS. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat Cell Biol*. 2001;3(2):134-139. doi:10.1038/35055033
58. Chuang C-H, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS. Long-range directional movement of an interphase chromosome site. *Curr Biol CB*. 2006;16(8):825-831. doi:10.1016/j.cub.2006.03.059
59. Kind J, Pagie L, Ortabozkoyun H, et al. Single-cell dynamics of genome-nuclear lamina interactions. *Cell*. 2013;153(1):178-192. doi:10.1016/j.cell.2013.02.028
60. Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science*. 2014;346(6214):1238-1242. doi:10.1126/science.1259587
61. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*. 2001;410(6824):116-120. doi:10.1038/35065132
62. Rea S, Eisenhaber F, O'Carroll D, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*. 2000;406(6796):593-599. doi:10.1038/35020506
63. Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SIS. Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. *Science*. 2001;292(5514):110-113. doi:10.1126/science.1060118
64. Bannister AJ, Zegerman P, Partridge JF, et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 2001;410(6824):120-124. doi:10.1038/35065138
65. Jacobs SA, Khorasanizadeh S. Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail. *Science*. 2002;295(5562):2080-2083. doi:10.1126/science.1069473
66. Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev*. 2003;17(15):1870-1881. doi:10.1101/gad.1110503
67. Pindyurin AV, Ilyin AA, Ivankin AV, et al. The large fraction of heterochromatin in Drosophila neurons is bound by both B-type lamin and HP1a. *Epigenetics Chromatin*. 2018;11(1):65. doi:10.1186/s13072-018-0235-8
68. Ye Q, Worman HJ. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. *J Biol Chem*. 1996;271(25):14653-14656. doi:10.1074/jbc.271.25.14653

69. Ye Q, Callebaut I, Pezhman A, Courvalin JC, Worman HJ. Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. *J Biol Chem*. 1997;272(23):14983-14989. doi:10.1074/jbc.272.23.14983
70. Makatsori D, Kourmouli N, Polioudaki H, et al. The Inner Nuclear Membrane Protein Lamin B Receptor Forms Distinct Microdomains and Links Epigenetically Marked Chromatin to the Nuclear Envelope. *J Biol Chem*. 2004;279(24):25567-25573. doi:10.1074/jbc.M313606200
71. Solovei I, Wang AS, Thanisch K, et al. LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation. *Cell*. 2013;152(3):584-598. doi:10.1016/j.cell.2013.01.009
72. Poleshko A, Mansfield KM, Burlingame CC, Andrake MD, Shah NR, Katz RA. The Human Protein PRR14 Tethers Heterochromatin to the Nuclear Lamina during Interphase and Mitotic Exit. *Cell Rep*. 2013;5(2):292-301. doi:10.1016/j.celrep.2013.09.024
73. Schreiner SM, Koo PK, Zhao Y, Mochrie SGJ, King MC. The tethering of chromatin to the nuclear envelope supports nuclear mechanics. *Nat Commun*. 2015;6(1):7159. doi:10.1038/ncomms8159
74. Csink AK, Henikoff S. Genetic modification of heterochromatic association and nuclear organization in Drosophila. *Nature*. 1996;381(6582):529-531. doi:10.1038/381529a0
75. Seum C, Delattre M, Spierer A, Spierer P. Ectopic HP1 promotes chromosome loops and variegated silencing in Drosophila. *EMBO J*. 2001;20(4):812-818. doi:10.1093/emboj/20.4.812
76. Canzio D, Chang EY, Shankar S, et al. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome bridging mechanism for heterochromatin assembly. *Mol Cell*. 2011;41(1):67-81. doi:10.1016/j.molcel.2010.12.016
77. Sanulli S, Trnka MJ, Dharmarajan V, et al. HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature*. 2019;575(7782):390-394. doi:10.1038/s41586-019-1669-2
78. Larson AG, Elnatan D, Keenen MM, et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*. 2017;547(7662):236-240. doi:10.1038/nature22822
79. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. *Nature*. 2017;547(7662):241-245. doi:10.1038/nature22989
80. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A Phase Separation Model for Transcriptional Control. *Cell*. 2017;169(1):13-23. doi:10.1016/j.cell.2017.02.007

81. Erdel F, Rademacher A, Vlijm R, et al. Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Mol Cell*. 2020;78(2):236-249.e7. doi:10.1016/j.molcel.2020.02.005
82. Cairns BR. The logic of chromatin architecture and remodelling at promoters. *Nature*. 2009;461(7261):193-198. doi:10.1038/nature08450
83. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol*. 2017;18:407-422. doi:10.1038/nrm.2017.26
84. Varga-Weisz PD, Wilm M, Bonte E, Dumas K, Mann M, Becker P. Chromatin-remodeling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature*. 1997;388:598-602. doi:10.1038/41587
85. Rao SSP, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665-1680. doi:10.1016/j.cell.2014.11.021
86. Fu Y, Sinha M, Peterson CL, Weng Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet*. 2008;4(7):e1000138. doi:10.1371/journal.pgen.1000138
87. Wiechens N, Singh V, Gkikopoulos T, Schofield P, Rocha S, Owen-Hughes T. The Chromatin Remodelling Enzymes SNF2H and SNF2L Position Nucleosomes adjacent to CTCF and Other Transcription Factors. *PLoS Genet*. 2016;12(3):e1005940. doi:10.1371/journal.pgen.1005940
88. Barutcu AR, Lajoie BR, Fritz AJ, et al. SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells. *Genome Res*. 2016;26(9):1188-1201. doi:10.1101/gr.201624.115
89. Barisic D, Stadler MB, Iurlaro M, Schübeler D. Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature*. 2019;569(7754):136-140. doi:10.1038/s41586-019-1115-5
90. Wu S, Fatkhutdinov N, Rosin L, et al. ARID1A spatially partitions interphase chromosomes. *Sci Adv*. 2019;5(5):eaaw5294. doi:10.1126/sciadv.aaw5294

Chapter 2

High-throughput assessment of context-dependent effects of chromatin proteins

Laura Brueckner, Joris van Arensbergen, Waseem Akhtar, Ludo Pagie & Bas van Steensel

Epigenetics & Chromatin **9,** 43 (2016)

Abstract

Background

Chromatin proteins control gene activity in a concerted manner. We developed a high-throughput assay to study the effects of the local chromatin environment on the regulatory activity of a protein of interest. The assay combines a previously reported multiplexing strategy based on barcoded randomly integrated reporters with Gal4-mediated tethering. We applied the assay to *Drosophila* heterochromatin protein 1a (HP1a), which is mostly known as a repressive protein but has also been linked to transcriptional activation.

Results

Recruitment to over 1000 genomic locations revealed that HP1a is a potent repressor able to silence even highly expressing reporter genes. However, the local chromatin context can modulate HP1a function. In pericentromeric regions, HP1a-induced repression was enhanced by twofold. In regions marked by a H3K36me3-rich chromatin signature, HP1a-dependent silencing was significantly decreased. We found no evidence for an activating function of HP1a in our experimental system. Furthermore, we did not observe stable transmission of repression over mitotic divisions after loss of targeted HP1a.

Conclusions

The multiplexed tethered reporter assay should be applicable to a large number of chromatin proteins and will be a useful tool to dissect combinatorial regulatory interactions in chromatin.

Background

Eukaryotic genomes are packaged in various types of chromatin that each has specific roles in the regulation of gene expression and other nuclear functions. These chromatin types (or "states") are defined by their distinct but sometimes partially overlapping protein compositions [1,2]. One of the main challenges in chromatin biology is to understand the combinatorial logic of chromatin proteins: How does the regulatory function of each protein depend on the presence or absence of other proteins?

This question is exemplified by HP1, a key component of classical heterochromatin. HP1 binds to di- or tri-methylated lysine 9 of histone H3 (H3K9me2/3) and can package nucleosomal DNA into a conformation that is able to repress transcription [3–5]. Several observations indicate that the regulatory activity of HP1 is context-dependent.

In *Drosophila*, the archetype ortholog HP1a localizes not only to pericentromeric regions where it is thought to contribute to the silencing of transposable elements, but also to a subset of transcriptionally active genes scattered along the chromosome arms [6,7]. Certain genes even appear to be activated by HP1a [8,9]. How this context-dependency arises is largely unknown, but many proteins have been identified that either promote or counteract heterochromatin formation [10]. An example is JIL-1 kinase, which is able to phosphorylate serine 10 of histone H3, a modification that blocks the interaction of HP1 with H3K9me2/3 [11,12].

Here, we present a method to systematically study how the regulatory activity of a protein may depend on the local chromatin context. The method combines barcoded, randomly integrated reporter genes [13] with artificial tethering of a protein of interest [14–16] to these reporter genes. More than 1000 random integration sites throughout the genome offer the required statistical power to infer how various chromatin environments may influence the regulatory activity of the protein of interest. We illustrate this approach using *Drosophila* HP1a as a model.

Results

Experimental design

Our approach builds on the previously reported TRIP protocol [13], which begins with transposase-mediated random genomic integration of reporter constructs in a pool of cells. All reporters are identical except for a short random barcode sequence within the transcription unit. In the resulting pool of cells, each reporter integration is mapped by a next-generation sequencing (NGS) approach to its genomic location and linked to its unique barcode sequence. Finally, NGS-based counting of barcodes in mRNA isolated from the cell pool enables us to determine the relative expression level of each reporter. By combining this information with the location of each reporter, we can study chromatin position effects in high throughput. As an addition to this original TRIP protocol, we inserted five copies of the Gal4UAS motif upstream of the promoter of our integrated reporters. This makes it possible to tether a fusion protein consisting of GalDBD and a chromatin protein of interest, here HP1a (Fig. 1a). For brevity, Gal4DBD and HP1a will be referred to as Gal4 or HP1. Various studies in *Drosophila* have previously shown that tethered HP1 can cause silencing of a reporter gene [17–20].

We used a reporter construct consisting of a green fluorescent protein (GFP) under the control of the copper-inducible metallothionein promoter (pMT). We randomly integrated this construct via Sleeping Beauty transposition in *Drosophila* Kc167 cells. We chose this particular cell line because extensive maps of histone marks, chromatin protein binding and computationally defined chromatin states are available. In the resulting TRIP cell pool, we were able to map 1093 integrations and link them to unique barcodes.

We then induced the pMT by adding 0.5 mM $CuSO_4$ to the TRIP cell pool. Two days after induction, we transiently transfected the pool with a plasmid expressing Gal4-HP1, or unfused Gal4 or HP1 as controls. These vectors also express mCherry, which enabled us to assess transfection levels and to isolate transfected cells by fluorescence-activated cell sorting (FACS). Two days after transfection, we collected transfected cells and extracted mRNA and genomic DNA (gDNA) for barcode counting by NGS. In order to compare expression between samples and to determine absolute up- or downregulation, we used a spike-in consisting of a low-complexity independent TRIP cell library. After normalization (see “Methods”), we observed good correlations between two independent transfection experiments (supplementary data: Figure S1). For downstream analyses, we averaged the normalized expression values of these two replicates for each reporter integration.

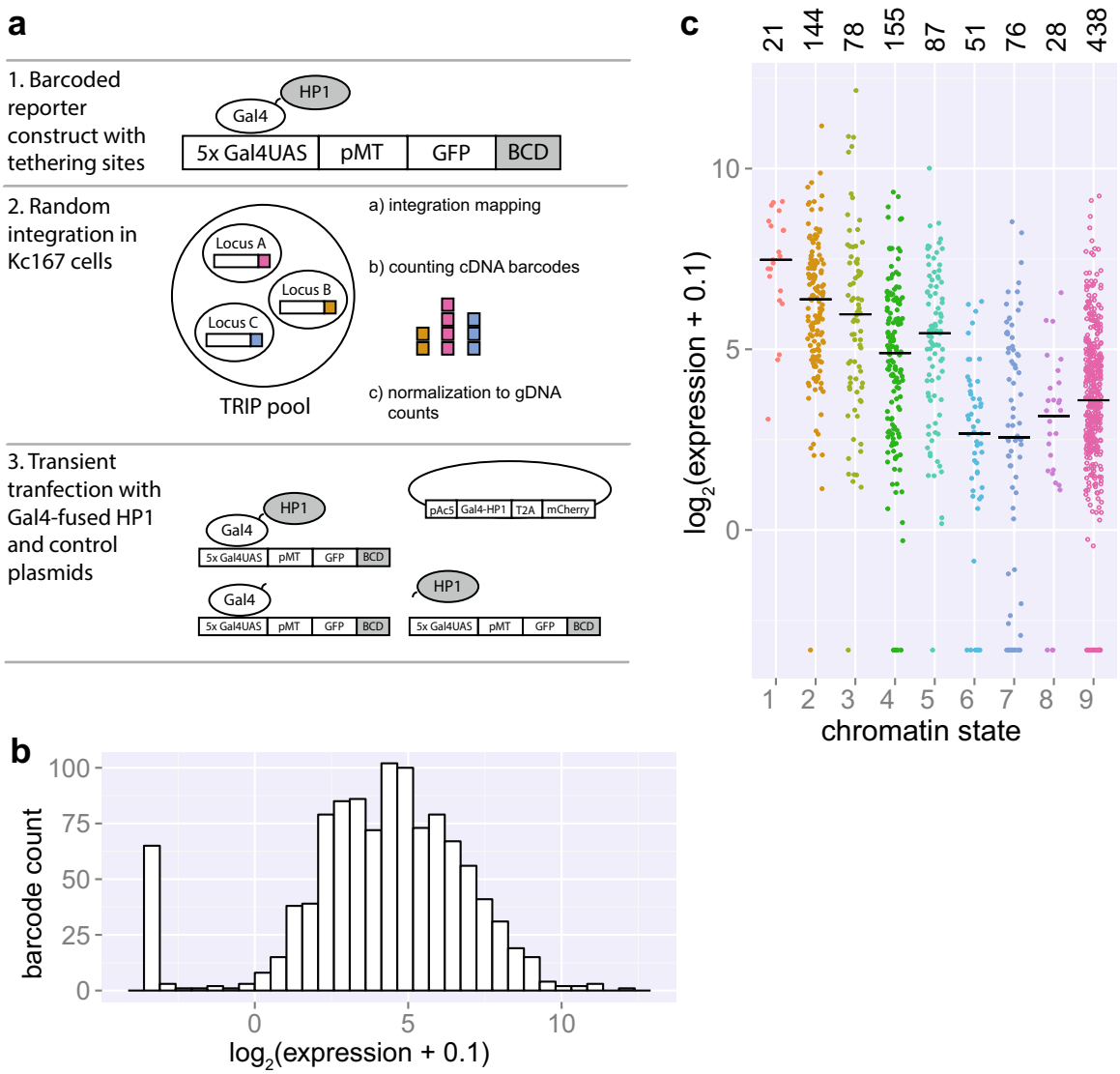


Figure 1
Chromatin effects on gene expression assessed by thousands of reporters integrated in parallel. a Principle of thousands of reporters integrated in parallel (TRIP) coupled with targeted recruitment of Gal4-fused proteins. **b** Distribution of reporter expression in the Gal4-transfected control condition as quantified by NGS. **c** Expression of integrated reporters as quantified by NGS in the Gal4-transfected control condition divided over nine chromatin states. Median values are represented by black horizontal bars. Number of reporters integrated in each state is specified above graph. Data information: In (b–c), a pseudocount of 0.1 was added to all expression values in order to be able to visualize non-expressed reporters on a log scale

Integrated reporters reflect the local chromatin state

First, we examined position effects on the expression of the integrated TRIP reporters in the absence of tethered HP1. For this, we used the Gal4-only transfected control cell pools. This revealed an approximately 1000-fold variation in reporter expression (Fig. 1b), similar to what was previously observed by TRIP in mouse cells [13]. This demonstrates strong position-dependence of reporter expression.

To investigate whether this variation in expression could be explained by differences in the local chromatin environment, we overlaid the TRIP data with a chromatin state map in which chromatin was subdivided into nine states according to combinatorial patterns of histone modifications [2]. Briefly, states 1–5 represent various chromatin states associated with active transcription; state 6 is enriched in the polycomb-associated mark H3K27me3; state 7 corresponds primarily to pericentromeric heterochromatin and is highly enriched in HP1a, H3K9me2/me3 and the corresponding histone methyltransferase Su(var)3-9; state 8 describes heterochromatin-like regions present on autosomal arms and with lower HP1 and Su(var)3–9 occupation than state 7; and finally, state 9 covers 40 % of the genome and is mostly devoid of histone marks [2].

We observed more than 30-fold differences in the average expression levels of reporters across the nine chromatin types (Fig. 1c). Mean reporter expression was highest in states 1 and 2 which describe transcriptionally active non-intronic regions. Expression was lowest in polycomb-marked state 6 and heterochromatin states 7 and 8. These observations are in accordance with the previously described association of these chromatin types with active or inactive transcription [2] and indicate that integrated reporters are strongly influenced by the state of the surrounding chromatin. We obtained similar results with an alternative five-state chromatin model [1], which showed up to tenfold differences in mean reporter expression depending on the chromatin state (supplementary data: Figure S2).

Tethering HP1 results in global downregulation of integrated reporters

Next, we determined the global effect of tethering of HP1 to the integrated reporters. We first measured GFP expression by flow cytometry on days 2, 4 and 6 after transfection with Gal4-HP1, Gal4 or HP1. GFP levels were visibly reduced by day 4 after Gal4-HP1 transfection compared to both controls and stayed low until day 6 (supplementary data: Figure S3). We decided to focus on day two after transfection, because at later time points the HP1-induced repression might have become too strong to observe position-dependent differences. We measured bulk GFP expression at the mRNA level across all reporters in the cell pool by conventional RT-qPCR analysis. This yielded ratios of 0.33 for Gal4-HP1/Gal4,

0.30 for Gal4-HP1/HP1 and 0.92 for HP1/Gal4. Thus, tethering of HP1 results in an overall downregulation of reporter RNA levels, whereas expression of untethered HP1 does not have a major effect.

We then investigated the effect of tethered HP1 on each individual reporter by comparing the normalized barcode expression levels in Gal4-HP1 expressing cells to those of Gal4 control cells. This revealed that most reporters are downregulated upon HP1 tethering (Fig. 2a, d). Seventy percent of reporters exhibited a greater than twofold reduction in expression, and 18 % of reporters with detectable expression in the control condition were completely silenced. For reporters with detectable expression in both conditions, the average fold change Gal4-HP1/Gal4 was 0.23 ± 0.25 (mean ± standard deviation). Transfection with HP1 alone may have a mild effect on a subset of reporters (mean fold change HP1/Gal4 0.90 ± 0.98) (Fig. 2b, e), but the repressive effect of Gal4-HP1 was consistently much stronger (mean fold change Gal4-HP1/HP1 0.35 ± 0.54) (Fig. 2c, f), which indicates that transfection with untethered HP1 does not have a major effect on bulk reporter expression. Western blot analysis indicated that the strong decrease in expression with Gal4-HP1 was not due to higher expression of the fusion protein compared to HP1 only (supplementary data: Figure S4). This analysis also indicates that expression of endogenous HP1 was not affected by transfection with Gal4-HP1 or HP1.
Only 5.1 % of reporters were unaffected by HP1 tethering, as defined by a less than twofold change in expression. Finally, 0.27 % of all reporters were upregulated more than twofold upon HP1 tethering. Such rare events could be due to technical noise rather than a biological effect. We conclude that HP1 represses transcription when tethered upstream of a promoter, in the vast majority of genomic locations.

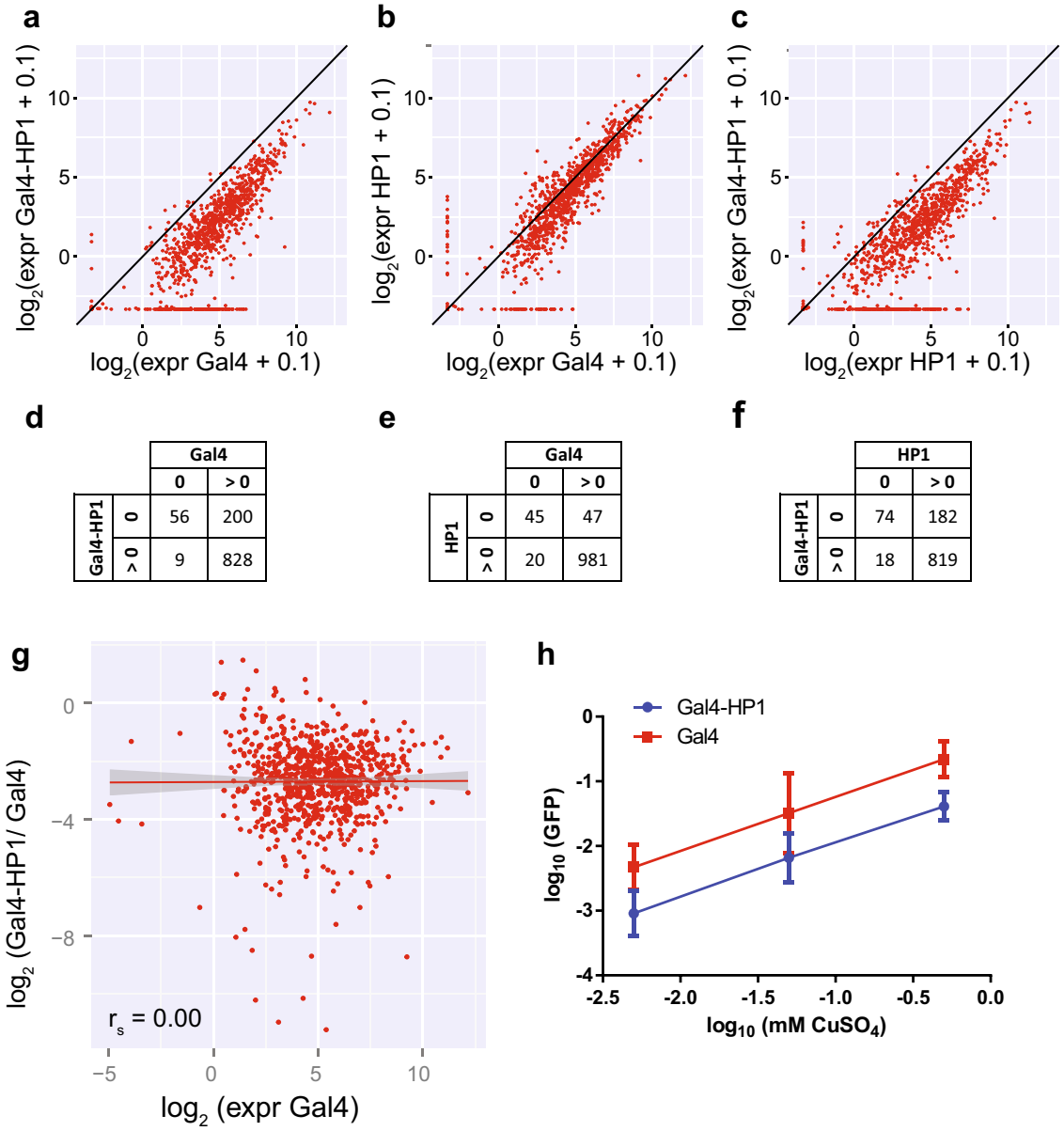


Figure 2

Tethering of Gal4-HP1 results in global downregulation of integrated reporters. a–c Reporter expression compared between TRIP pool transfected with Gal4-HP1 or controls Gal4 or HP1 only as quantified by NGS 2 days after transfection. **d–f** Quantification of reporters showing expression detectable by at least one sequencing read. **g** Correlation (Spearman) between expression in the Gal4-transfected control condition and downregulation upon HP1 tethering with linear regression (red) and standard error. Downregulation is quantified by ratio of expression in the Gal4-HP1-transfected condition over expression in the Gal4-transfected control condition; therefore, only reporters with detectable expression in both conditions are shown. **h** GFP reporter expression in a cell line with single integration induced with increasing levels of CuSO4 and transfected with Gal4-HP1 or Gal4 as quantified by qPCR. Amount of repression upon Gal4-HP1 tethering is constant over expression range. Data information: In (a–c), a pseudocount of 0.1 was added to all expression values in order to be able to visualize non-expressed reporters on a log scale

Gal4-HP1-induced silencing is not correlated with transcription levels in the absence of tethering

Because we transfected the TRIP cell pool 2 days after induction of reporter expression, HP1 has to compete with the transcription machinery in order to establish a heterochromatin state. We therefore wondered whether the degree of repression was related to the initial expression level of the reporter gene. We used the Gal4-transfected control cells to estimate this initial expression (leaving out reporters without any detectable initial expression) and compared it to the extent of silencing observed with Gal4-HP1 (Fig. 2g). Strikingly, there was no correlation between expression levels of reporters and the extent of downregulation observed with Gal4-HP1 (Spearman's $\rho = 0.00$).

We took a complementary approach to confirm that repression by HP1 is independent of the initial transcriptional activity of a gene. We generated cells with a pMT-driven GFP reporter that was integrated in a single euchromatic locus. We then induced transcription from this promoter by adding various concentrations of $CuSO_4$ and determined reporter expression levels by RT-qPCR. Compared to bulk expression of TRIP reporters, we obtained 6.2-fold lower expression with the lowest level of induction and 7.5-fold higher expression with the highest $CuSO_4$ concentration. Thus, we could test the effect of HP1 on a single gene over a broad range of expression levels. We then tethered Gal4-HP1 and observed a consistent fold reduction in expression of 0.22 ± 0.12, 0.28 ± 0.24 and 0.21 ± 12 from lowest to highest induction level (Fig. 2h). In conclusion, Gal4-HP1-induced silencing is equally efficient over a wide range of reporter expression levels.

Gal4-HP1-induced silencing is increased in pericentromeric heterochromatin

Next, we were curious whether the extent of silencing upon HP1 tethering is linked to different chromatin environments. We therefore analyzed the Gal4-HP1/Gal4 fold change in reporter expression as a function of the chromatin types according to the nine-state model (Fig. 3a, b; supplementary data: Figure S5). This revealed up to 3.5-fold variation between chromatin states, with the least repression occurring in states 1 and 2 and the strongest repression in state 7. This suggests that the local chromatin environment can modulate the ability of tethered HP1 to repress transcription.

State 7 is of particular interest because it coincides with pericentromeric regions that are densely occupied by endogenous HP1 and Su(var)3–9. Reporters in state 7 exhibited a significantly stronger repression by tethered HP1 than reporters in the other chromatin states (2.0-fold difference in median repression levels; Wilcoxon's test $p = 2.3 \times 10^{-7}$).

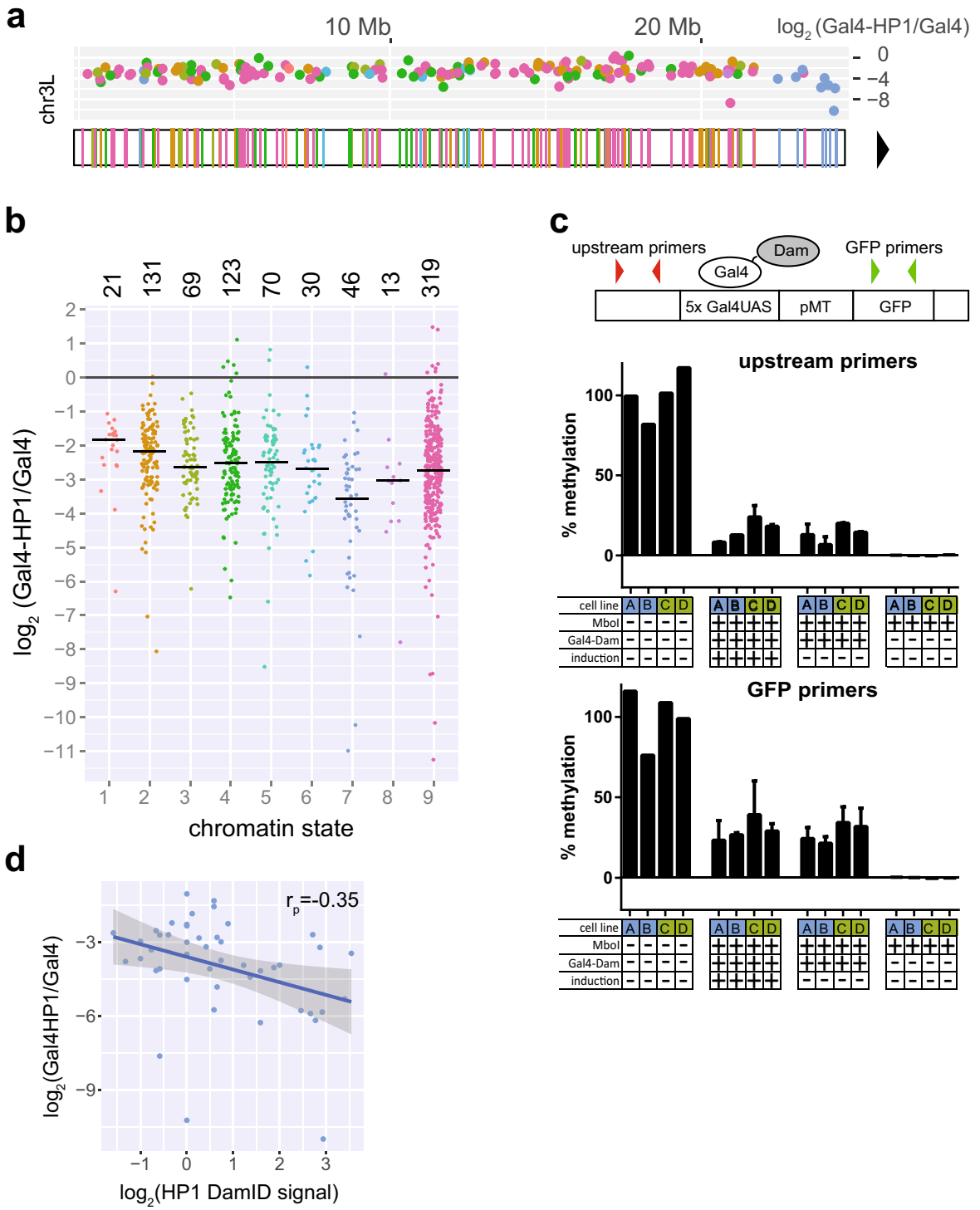


Figure 3

Repression upon HP1 tethering is modulated by the local chromatin environment. a Positions of the integration sites of TRIP reporters on chromosome 3L. Colors represent chromatin state at the integration site according to the nine-state model (as labeled in Fig. 3b). Centromere position is indicated by black triangle. Plot above the chromosome shows fold downregulation upon Gal4-HP1 tethering. All chromosomes are shown in supplementary data: Figure S4. **b** Fold downregulation upon Gal4-HP1 tethering segregated by chromatin state. Median values are indicated by black horizontal bars. Number of reporters integrated in each state is specified above graph. **c** Quantification of reporter accessibility via tethering of Gal4-Dam in four cell lines each with a single integration. Dam-mediated adenine methylation was assessed by MboI digestion of unmethylated GATC sequences followed by qPCR. Position of qPCR primer pairs is indicated on schematic representation of reporter. Bar graphs show methylation levels measured in two cell lines (labeled A, B) with reporter integration in state 7 heterochromatin (blue) and two cell lines

(labeled C, D) with integration in state 3 euchromatin (green). Error bars represent standard deviation of two independent transfections. Measurements were taken with or without reporter induction. **d** HP1 occupancy levels at reporter integrations sites in state 7 domains as quantified by DamID correlate with downregulation upon Gal4-HP1 recruitment (Pearson). Linear regression (blue) with standard error

As shown above, transcriptional activity in general did not correlate with the extent of downregulation observed upon HP1 recruitment, but since there are large differences in reporter expression between chromatin states we wanted to exclude it as possible explanation for the observed difference in silencing. We therefore picked reporters from other chromatin types with expression levels most closely matching those of the reporters in state 7 and compared the Gal4-HP1/Gal4 fold change in expression (we excluded four reporters in state 7 that were outside of the expression range of the remaining reporters and could therefore not be matched). Again, we observed a significantly stronger repression for reporters in state 7 (2.2-fold difference in median repression levels, Wilcoxon's test $p = 2.8 \times 10^{-5}$). Repression by Gal4-HP1 was also significantly stronger in state 7 than in other states when normalized to untethered HP1 (Wilcoxon's test $p = 0.0077$, supplementary data: Figure S6). We conclude that state 7 chromatin provides a favorable environment for repression by tethered HP1, irrespective of the initial activity of the reporter. Although heterochromatin is generally thought to be less accessible, we checked whether the targeting of Gal4 fusion proteins was somehow more efficient in state 7 chromatin. We tested this by DamID of Gal4 using a previously reported qPCR-based readout [21]. We performed this assay in two cell lines with a single reporter integration in state 7 heterochromatin and two cell lines with integrations in state 3 euchromatin (Fig. 3c). As expected, state 7 integrations did not show higher accessibility for Gal4 binding as measured by methylation levels.

To explain the enhanced downregulation in state 7, we examined the occupancy of endogenous HP1 at reporter integrations sites by DamID profiles. HP1 is markedly enriched in state 7 pericentromeric heterochromatin, with fourfold higher levels compared to all other states combined. Within state 7, the fold downregulation upon tethering correlates with local HP1 binding levels (Pearson's $r = -0.35$, $p = 0.018$) (Fig. 3d) but not in other states. We also confirmed by immunofluorescence microscopy that Gal4-HP1 as well as control HP1 accumulated in the chromocenter that contains high concentrations of endogenous HP1 (supplementary data: Figure S7). This homing behavior in combination with high availability of endogenous HP1 in pericentric heterochromatin might facilitate the repressive action.

Gal4-HP1-induced silencing is less effective in chromatin associated with elongating transcription

Silencing by tethered HP1 was least effective in chromatin states 1 and 2, with median repression reduced by 1.8- or 1.4-fold, respectively, compared to reporters in other chromatin states. State 2 gave a significant test result versus other chromatin states (Wilcoxon's test $p = 0.00022$) and also when comparing reporters of matching expression ($p = 0.0094$). State 1 resulted in a p value of 0.016 when comparing downregulation with all other chromatin types but did not give a significant result in the expression-matched test ($p = 0.48$). We therefore focused on state 2. This state is typically present on exonic regions of transcribed genes and associated with high levels of H3K36me3 and H3K79me2 [2]. Indeed, we observed significantly reduced repression of reporters integrated in H3K36me3-bound regions (1.4-fold difference in median repression, $p = 0.00028$, $N = 118$). H3K36me3 also significantly correlated with fold change Gal4-HP1/Gal4 when using continuous ChIP-seq data ($p = 2.3 \times 10^{-5}$, Spearman's $\rho = 0.15$). We obtained similar results for H3K79me2 (1.3-fold difference in median repression, $p = 0.00060$, $N = 135$). It will be interesting to investigate whether these histone marks are directly responsible for the inhibitory effect on HP1-induced silencing, or mere indirect correlates. Finally, we also found a significant correlation with ChIP scores for JIL-1, a H3S10 kinase known to limit heterochromatin spreading ($p = 0.0014$, Spearman's $r\rho = 0.11$) [22].

Variegation upon HP1 tethering

HP1-containing heterochromatin is known for its ability to cause variegating patterns of gene repression, with the target gene being either “on” or “off” [23]. The degree of repression of a particular TRIP reporter may therefore reflect the proportion of cells in the “off” state, rather than downregulation along a continuous scale. Because TRIP cannot discriminate between these two different modes of repression, we tested the effect of Gal-HP1 tethering by FACS analysis of two cell lines with stably integrated single-copy pMT-driven GFP reporters (Fig. 4). As in the TRIP experiments, we first activated the pMT promoter and then transfected the cells with Gal-HP1 or control plasmids. After 6–8 days, both lines showed bimodal GFP expression distributions after Gal-HP1 tethering, with the highest expression mode roughly coinciding with a single peak of expression observed in control cells that were transfected with Gal4 alone. However, the two cell lines differed in the proportion of cells in the “off” state, and this proportion was also dependent on the time point at which the cells were analyzed. This result indicates that tethered HP1 causes variegating expression, with the balance between “on” and “off” depending on the local chromatin environment. Considering the strong similarity of the reporter construct used for this experiment and the one used in the TRIP studies, it is likely that variegating expression also occurs in the case of TRIP reporters after Gal-HP1 tethering.

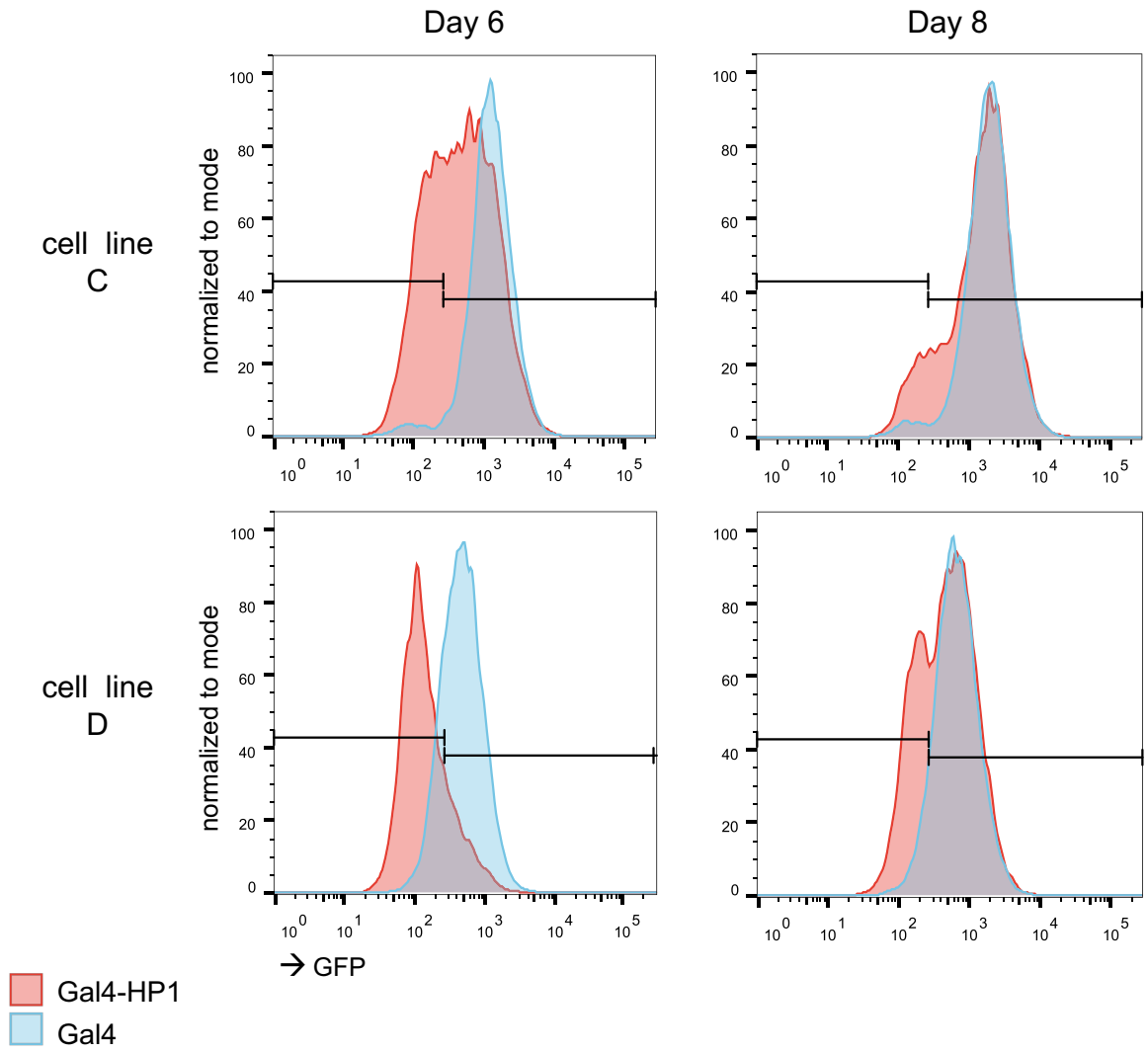


Figure 4
Variegation upon HP1 targeting. GFP expression distributions in two single-integration cell lines upon Gal4-HP1 targeting (red) compared to Gal4 only (blue) as quantified via FACS by 6 and 8 days after transient transfection

Gal4-HP1-induced silencing does not result in permanent memory

Targeting of exogenous HP1 can induce a mitotically heritable heterochromatin state in mammalian cells [24,25]. We were therefore curious whether we could observe heritability of Gal4-HP1-induced silencing after loss of tethered HP1 in the around 1000 reporters in our *Drosophila* cell system. To test this, we kept the TRIP cell pool transiently transfected with Gal4-HP1 and control plasmids in culture until the cells had lost plasmid expression. Initial transfection efficiencies, as quantified by FACS, were approximately 50 % for all plasmids (Fig. 5a). By day 6, we still measured 40–55 % transfected cells. At this time point, in the Gal4-HP1-transfected replicates 7.8 and 8.0 % of the cells were GFP positive, compared to 12–13 % in the Gal4- and HP1-transfected control samples. This confirmed that tethered HP1 induced silencing of GFP expression. By day 16, we measured 1.3–2.1 % remaining mCherry-positive cells, indicating that the TRIP cell pool had mostly lost the transfected plasmids. At this time point, 15 % of cells in the duplicate cultures that were initially transfected with Gal4-HP1 were GFP positive, which is comparable to 13–17 % GFP-positive cells in the control samples. This indicated that most integrated reporters had recovered from Gal4-HP1-induced silencing. We collected these cell pools and quantified individual barcode expression in mRNA by NGS (Fig. 5b). The average fold expression change for all barcodes was 0.92 ± 0.53 for Gal4-HP1/Gal4 (mean ± standard deviation), 1.2 ± 0.72 for Gal4-HP1/HP1 and 0.89 ± 0.55 for HP1/Gal4. Moreover, statistical analysis using *limma* [26] did not detect any individual reporters with significant repression after 16 days, nor did we observe significant differences in Gal4-HP1/Gal4 ratios (Wilcoxon's test) when the data were aggregated by the nine chromatin states. We therefore conclude that there is no stable mitotic transmission of Gal4-HP1-induced silencing by 16 days after transient transfection.

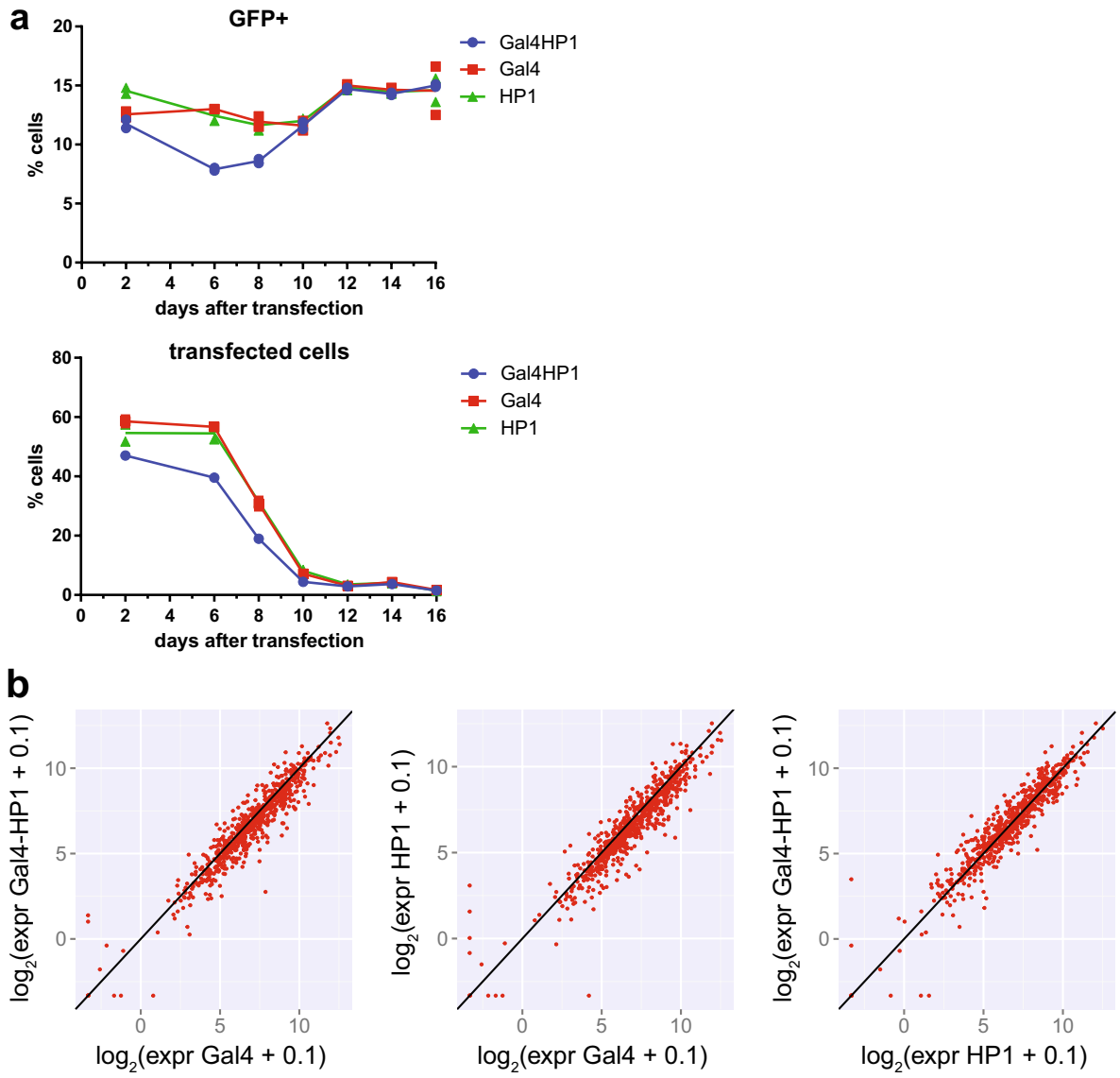


Figure 5
Tethering of HP1 does not result in permanent repression. a GFP expression of TRIP pool and amount of transfected cells over time after transient transfection with Gal4-HP1 or control plasmids as quantified by FACS. Graphs show data from two replicate experiments with line connecting means. **b** Reporter expression compared between TRIP pool transfected with Gal4-HP1 or controls Gal4 or HP1 only as quantified by NGS 16 days after transfection

Discussion

HP1 as a potent and universal transcriptional repressor

Our results indicate that HP1a is a potent transcriptional repressor that can reduce the activity of integrated reporters in most genomic contexts. In previous studies employing targeted recruitment of HP1 to integrated reporters in *Drosophila*, tethering did not result in silencing in a number of cases. In one study, three Gal4 binding sites were inserted between the *white* and *lacZ* reporter genes and targeted by Gal4-HP1 expressed under heat-shock promoter HSHP1-83C. In this setup, the white reporter was silenced at only one of six genomic locations tested and the one site that supported silencing was adjacent to repetitive sequences [20]. In another study, LacI-HP1 was expressed under the Hsp70 promoter and tethered to 256 lac repeats which resulted in silencing of the white reporter gene in 25 out of 26 cases as observed by white eye color [17]. In our system, Gal4-HP1 is expressed under the control of the Actin5c promoter and tethered to five repeats of the Gal4UAS motif. We did not identify a genomic environment that completely inhibited HP1. Use of a strong constitutive promoter for Gal4-HP1 expression could explain why we observe this highly efficient silencing.

Several reports have linked HP1 in *Drosophila* with an activating effect on transcription [8,27–29]. In our study, we observed upregulation greater than twofold in only three of 1093 reporters which did not allow us to draw general conclusions. It is possible that HP1 can activate transcription in an indirect fashion or that this effect depends on HP1 localization within gene bodies, which we cannot test in our current system.

Silencing by tethered HP1 does not depend on transcription level of the target gene

Surprisingly, the magnitude of the repression by HP1 (fold change in reporter activity) is unaffected by even high levels of transcription of the reporter. This argues against a competition model in which the transcriptional machinery, if active enough, can overrule heterochromatin. Rather, HP1a turns down transcription by a nearly constant factor, irrespective of the initial transcription levels. Although the TRIP readout is not suitable to detect cell-to-cell variation in expression, our analysis of single integrated GFP reporters indicates that HP1a causes variegation. Reduced expression as detected by TRIP should therefore be considered to reflect that the reporters are turned “off” in a proportion of the cells. This proportion is then largely independent of the initial transcription activity of the reporter.

Chromatin environments modulate HP1 action

While the activity of the reporter itself did not affect HP1 action, we did observe quantitative differences in HP1 effects between chromatin states. Repression by HP1 was slightly but significantly less efficient in chromatin state 2, which is an exon-biased chromatin environment associated with transcriptional elongation and high levels of H3K36me3. Because H3K36 methylation is deposited co-transcriptionally [30,31], it is difficult to determine by correlative studies whether the histone mark, RNA Pol II or other factors inhibit HP1 action. One candidate is cyclin-dependent kinase 12 (CDK12), a transcription elongation-associated RNA Pol II kinase that was recently reported to antagonize heterochromatin in *Drosophila* [32]. We also found reduced silencing in regions bound by the H3S10 kinase JIL-1, in agreement with earlier observations [22].

Our data indicate that pericentric heterochromatin facilitates repression by tethered HP1. Cooperative interactions with endogenous HP1 or its partner proteins are likely to account for this effect, because the repression by Gal-HP1 correlates with heterochromatin domain size and local density of endogenous HP1. Notably, in this study we only covered uniquely mappable reporters. It might be interesting to examine integrations in repetitive sequences.

Absence of epigenetic memory after transient HP1 recruitment

Induction of a mitotically stable heterochromatin state by tethering of HP1 has been observed in mammalian cells [24,25]. In fission yeast, transient overexpression of Swi6, the yeast homolog of HP1 resulted in mitotically and meiotically stable silencing at the mating-type locus [33]. H3K9 methylation nucleated at this locus could be heritably maintained after loss of the nucleation center but only in the presence of Swi6 [34]. In other studies in fission yeast, a silenced chromatin state triggered by Gal4-mediated targeting of H3K9 methylation could be maintained over mitotic and meiotic divisions. Heritability required deletion of the corresponding demethylase and the presence of Swi6 [35,36]. In our experiments, HP1 targeting did not result in heritable silencing in any chromatin environment, as reporters ultimately regained expression levels equivalent to the unsilenced control. In murine embryonic stem cells, induction of high levels of DNA methylation was necessary for maintaining transcriptional silencing after loss of HP1 targeting and required 4 weeks of continuous HP1 recruitment. In the absence of stable tethering, HP1 was ultimately overcome by transcription [24]. We introduced HP1 via transient transfection which resulted in tethering for approximately 1 week. The short duration of HP1 targeting and the absence of canonical DNA methyltransferases Dnmt1 or Dnmt3 might explain the lack of stable

silencing in our *Drosophila* system. It will be interesting to conduct tethered TRIP experiments in mammalian cells to study the chromatin context effects on such epigenetic phenomena.

Conclusions

Tethering TRIP as a tool to study context-dependent effects of regulatory proteins

This study provides proof of principle for the use of TRIP in combination with protein tethering to investigate how specific regulatory proteins interact with their local environment to control gene activity. As Gal4 fusions of a protein of interest can easily be generated and transfected, a wide variety of chromatin proteins may be studied by this approach. Potential differences in transfection levels between individual cells are compensated for by FACS-based isolation of 10×10^6 cells which represents a theoretical coverage of around 300× of the initial barcode complexity.

The near-random integration of Sleeping Beauty transposons [37] combined with the multiplexed barcode readout makes it possible to survey most of the commonly occurring chromatin states with sufficient statistical power. Furthermore, TRIP is compatible with many reporter designs that may be used to probe a variety of functions [13]. For example, by varying the spacing between the Gal4-binding sites and the promoter of the reporter it may be possible to test the distance over which a tethered protein can exert its regulatory effects (e.g., through spreading of a chromatin state), and how this distance may depend on the local chromatin environment.

We cannot rule out that the integrated reporter in some instances alters the local chromatin state of the integration site. However, the ~1000-fold range in expression levels of the untethered reporter indicates that the local environment strongly controls the reporter. It is thus likely that the tethered protein is exposed to the same local influences. We expect that TRIP in combination with protein tethering will be a useful tool to further explore the context-dependent functions of regulatory proteins.

Methods

Cell culture

Kc167 cells were cultured at 23.5 °C in Shields and Sang M3 Insect Medium (Sigma-Aldrich) with 0.25 % Bacto peptone (BD), 0.1 % yeast extract (BD), 5 % heat-inactivated FBS (Thermo Scientific) and 1 % penicillin/streptomycin (Thermo Scientific). For pMT induction, sterile-filtered $CuSO_4$ (Sigma-Aldrich) dissolved in H_2O was added to 0.5 mM final concentration unless mentioned otherwise. The original source of the Kc167 cells used in this study cannot be traced, but the cells have been used in our laboratory for about 15 years, e.g., in [1,38–40].

Constructs

The TRIP vector was derived from Addgene plasmid #65488 by exchanging pHsp with 5x Gal4UAS—pMT via restriction enzyme cloning with EcoRV/EcoRI. pMT was derived from a commercially available pMT/V5-His/lacZ plasmid (Thermo Scientific), and five Gal4UAS sites (CGGAGTACTGTCCTCCGAG) were added as an oligonucleotide. Additionally, a second I-CeuI cutting site was integrated upstream of the IR-DR(R) sequence by PCR using primers 160JvA and 161JvA, followed by self-ligation.

Plasmids for expressing the Gal4-HP1 fusion and controls as well as Gal4-Dam were derived from the STABLE 2 vector for bicistronic expression via the T2A peptide, a gift from Jim Sutherland [41]. The NeoR gene in the template was replaced with mCherry and the EGFP gene with Gal4DBD-V5-HP1a, Gal4DBD-V5, V5-HP1a or Gal4DBD-Myc-Dam by Gibson assembly generating pAc5-Gal4-V5-HP1a-T2A-mCherry, etc. Gal4DBD (1-147) was copied from Addgene plasmid #43969, mCherry from mCherry-G9a [42], HP1 and Myc-Dam from pDamHP1 [21]. V5 tag (GGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACG) was added as an oligonucleotide.

The plasmid encoding Sleeping Beauty transposase is deposited as Addgene #65487. Plasmids expressing Dam-HP1 or Dam were previously reported [21].

Establishment of TRIP plasmid library

A 21-nt random barcode was added to the TRIP vector by PCR amplification with Phusion polymerase (2 U, Thermo Fisher) in GC buffer with 10 ng template plasmid, 500 nM forward and reverse primer (JvA168 and JvA169) and dNTPs (250 µM each) in a total volume of 100 µl. PCR conditions were 1 min at 98 °C (1×), 15 s at 98 °C, 30 s at 55 °C, 4 min at 72 °C (30×). PCR products were treated with T4 polymerase (9 U, NEB) with added dNTPs (250 µM each) for 20 min at 12 °C to complement barcode sequences. The reaction was stopped by adding EDTA to a final concentration of 10 mM and heat inactivation for 20 min

at 75 °C. Products were purified using ISOLATE II PCR and Gel kit (Bioline) according to the manufacturer's instructions and digested with DpnI (20 U, NEB) in buffer 4 for 1 h at 37 °C in a volume of 50 µl to remove the template. The reaction was terminated by heat inactivation for 20 min at 75 °C, and barcoded fragments were circularized over night at 12 °C with T4 ligase (50 U, NEB) in T4 ligase buffer in a total volume of 900 µl. After ligation, remaining unligated fragments were removed by treatment with Plasmid-Safe DNase (40 U, Epicentre) with added ATP (final concentration 1 mM) in the manufacturer-supplied reaction buffer in a total volume of 1500 µl for 5 h at 37 °C. The reaction was terminated by heat inactivation for 30 min at 70 °C. Products were purified by two times phenol/chloroform extraction followed by ethanol precipitation.
70 ng barcoded plasmids were transformed into 50 µl megacompetent cells (MegaX DH10B™ T1R Electrocomp™ Cells, Invitrogen) by electroporation (settings: 2 kV, capacitance = 25, capacitance extension = 250, pulse control = 200) and purified using a Genopure Maxi Kit (Roche) according to the manufacturer's instructions.

TRIP cell pool establishment and mapping of integrations

The barcoded reporters were integrated by transfecting 1×10^6 Kc167 cells with 1 µg barcoded plasmid library and 1 µg plasmid encoding Sleeping Beauty transposase. Transposase expression was induced by four heat-shock treatments of 2.5 h at 37 °C distributed over 36 h. Transfected cells were expanded until Sleeping Beauty expression was lost. For the resulting TRIP pool, we determined the number of integrations to be 0.3 per cell based on qPCR on gDNA for GFP integrations compared to a cell line with a single integration. A subpool of 30,000 cells was taken from the TRIP pool and expanded to limit the library complexity to a maximum of 10,000 integrations.
To map the reporter integrations, gDNA was extracted from 20×10^6 cells using ISOLATE II Genomic DNA Kit (Bioline) and digested with NlaIII (40 U, NEB) in buffer 4 supplemented with BSA in a total volume of 100 µl for 2 h at 37 °C. The reaction was terminated by heat inactivation for 20 min at 65 °C. Fragments were circularized over night at 12 °C using T4 ligase (100 U, NEB) in T4 ligase buffer in a total volume of 1600 µl. Products were precipitated with ethanol and purified with ISOLATE II PCR and Gel Kit (Bioline). Unligated fragments were removed by digestion with Plasmid-Safe DNase (20 U, Epicentre) with added ATP (final concentration 1 mM) in the manufacturer-supplied reaction buffer in a total volume of 100 µl for 5 h at 37 °C. The reaction was terminated by heat inactivation for 30 min at 70 °C and purified using ISOLATE II PCR and Gel Kit (Bioline). To eliminate any remaining unintegrated or fully integrated plasmid and linearize the template, products were digested with I-CeuI (5 U, NEB) in buffer 4 supplemented with BSA in a total volume of 60 µl

for 2 h at 37 °C. The reaction was terminated by heat inactivation for 20 min at 65 °C and purified using ISOLATE II PCR and Gel Kit (Bioline). Products were amplified in triplicate reactions by inverse PCR using Phusion polymerase (2 U, Thermo Fisher) in GC buffer with 20 µl of I-CeuI-digested DNA, 500 nM forward and reverse primer (151AR and 219AR) and dNTPs (200 µM each) in a total volume of 100 µl. PCR conditions were 1 min at 98 °C (1×), 30 s at 98 °C, 30 s at 60 °C, 45 s at 72 °C (22×). Products were purified using ISOLATE II PCR and Gel Kit (Bioline). To add indices and adapters for next-generation sequencing, one-third of PCR products was amplified in a second PCR using Phusion polymerase (1 U, Thermo Fisher) in GC buffer with 500 nM forward and reverse primer (151AR and iPCR indexing primer) and dNTPs (200 µM each) in a total volume of 50 µl. PCR conditions were 1 min at 98 °C (1×), 30 s at 98 °C, 30 s at 60 °C, 45 s at 72 °C (10×). Products were purified using ISOLATE II PCR and Gel Kit (Bioline) and prepared for next-generation sequencing as described below.

Establishment of integrase-mediated cassette exchange in Kc167 cell clones

The vector p81_JvA_DTR_IMCE to generate clonal cassette exchange Kc cell lines was constructed using standard molecular biology techniques and is available upon request. It is based on the TRIP vector backbone in which the sequence in between the Sleeping Beauty inverted repeats has been replaced by two full AttP sites which were PCR amplified from Addgene plasmid #13843, a gift from Ting Wu [43], in a “head-to-head” orientation to facilitate cassette exchange with a AttB-containing vector (see below). In between the AttP sites, a bicistronic neomycin/diphtheria toxin selection cassette was placed which is based on the Ac5-STABLE1-neo plasmid [41] and in which we replaced the GFP ORF by the ORF of the human diphtheria toxin receptor. To generate stable clones, 20×10^6 Kc167 cells were electroporated with 20 µg of Sleeping Beauty expression vector (Addgene plasmid #65487) and 4 µg of p81_JvA_DTR_IMCE and cells were expanded for 3 weeks. Selection with 40 µl/ml G418 (Sigma-Aldrich) was started after the first week. Clones were then established by limited dilution, the number of integrations was analyzed by qPCR, and clones with more than one integration were excluded from further analysis. To identify the genomic locus of the cassette exchange site in each clone, an iPCR strategy was used. Briefly, 100 ng of gDNA was digested in a volume of 25 µl with 5 units of Nla III (NEB) for 30 min and heat inactivated for 20 min at 65 °C. Of this reaction, 5 µl was self-ligated in a volume of 40 µl with 1 unit of ligase (Roche) at 16 °C overnight. On 5 µl of the ligation material, 34 PCR cycles (15 s at 98 °C; 15 s at 60 °C; 30 s at 72 °C) were performed using primer JvA45 and JvA94. The PCR product was purified using ISOLATE II PCR and Gel Kit (Bioline), Sanger sequenced using primer 45JvA and mapped to the genome (Table 1).

Table 1 Integration sites in Kc167 cell clones

Cell line	Integration locus
A	chr2RHet: 3094936–3094937
B	chr2L: 16340822–16340823
C	chrX: 5784726–5784727
D	chr2R: 5436488–5436489

The donor cassette vector containing the AttB sites is based on the backbone of the pMT/V5-His *Drosophila* expression vector (ThermoFisher Scientific) to which we added the full AttB sites from piB-GFP (Addgene plasmid #13844), the metallothionein promoter of the pMT/V5-His *Drosophila* expression vector, 5× Gal4UAS sites and the GFP ORF of the TRIP vector.

Cassette exchange was performed by electroporation of 1×10^6 cells with 1 µg of donor cassette and 1 µg of pBS130 (a kind gift from Tom Clandinin, Addgene plasmid # 26290). After 1 week, recombined cells were selected for 1 week by adding 0.5 µg/ml diphtheria toxin (Sigma-Aldrich) every 48 h. For the inducible expression experiment and variegation experiment, we wanted to exclude cells that might have escaped selection due to epigenetic silencing of the diphtheria toxin receptor gene. We therefore isolated GFP expressing cells of clones “C” and “D” by FACS 2 days after inducing GFP expression with 0.5 mM $CuSO_4$.

Transfection

Kc167 cells were transfected by electroporation of 20×10^6 cells with 20 µg plasmid at 1000 µF/250 V or 1 Mio cells with 1 µg plasmid at 450 µF/86 V using a Gene Pulser II (BioRad).

Flow cytometry analysis and sorting

Flow cytometry quantitative analysis was performed on an LSR FORTESSA (BD Biosciences) and processed with FlowJo software. Cells were gated based on forward and side scatter for single viable cells. Remaining cells were analyzed for GFP and mCherry levels with gates set according to wild-type cells. mCherry+ cells were sorted by using a Moflo Astrios (Beckman Coulter) and immediately resuspended in TRIsure (Bioline).

RNA isolation and cDNA generation

RNA was extracted from TRIsure-resuspended samples by chloroform extraction and precipitated with isopropanol. For TRIP samples, the polyadenylated fraction was isolated using the Oligotex Kit (Qiagen) according to the manufacturer's instructions. RNA samples were treated with RNase-free DNase I (2 U, Roche) in DNase buffer in a total volume of 20 µl for 30 min at 37 °C. The reaction was terminated by adding 1 µl of 25 mM EDTA and incubation at 70 °C for 15 min. cDNA was generated using Tetro cDNA Synthesis Kit (Bioline) according to the manufacturer's instructions. For TRIP samples, a GFP-specific primer (AR152) was used instead of OligoDT (Table 2).

Table 2 Oligonucleotides

Name	Sequence (5′–3′)
GFP_f	AGGACAGCGTGATCTTCACC
GFP_r	CTTGAAGTGCATGTGGCTGT
upstream_f	CGTACTCCACCTCACCCATC
upstream_r	TTCATCGATACCGTCGACCT
TSR_f	CAAAGAAGCAAAAGCTGTTCCTTA
TSR_r	GCTGGAGTACAACATCTTCTTCTTGAC
160JvA	CTAAGGTAGCGAAGGCAATGCTACCAAATAC
161JvA	GACCGTTATAGTTATTTAAATTGTTTAACTTGG
AR152	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
168_JvA	/5′phos/NNNNNNNNNNNNNNNNNNNNGATCATGCTAGTTGTGGTTTGTC
169_JvA	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTCAGTGAAAAAAATGCTT
151AR	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT CCGATCT
219AR	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAAGGTGTATGTAAACT TCCGACTTCAACTG
iPCR indexing	CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCAGACGTG TGCTCTTCCGATC
Library indexing	CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCACGCCTTCAAGACCCCCATCGCC
DamID adapter top	CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA
DamID adapter bottom	TCCTCGGCCGCG
Adr-PCR-Rand1	NNNNGTGGTCGCGGCCGAGGATC
Y-adaptor top	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Y-adaptor bottom	/5′phos/GATCGGAAGAGCACACGTCT
Illumina index	CAAGCAGAAGACGGCATACGAGNNNNNNNNGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT
P5-Illumina-2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT CCGATCT
45JvA	ATTCTGATATTTCACATTCTTAAAATAAAGTGG
94JvA	ACCGTTATAGTTATTAACTTGGGTCAAACATTT

qPCR

qPCR was performed using SensiFAST SYBR No-ROX Kit (Bioline) with 300 nM forward and reverse primer in 10 µl volume. Detection was performed in a LightCycler 480 (Roche) under the following qPCR conditions: 95 °C for 2 min (1×), 95 °C for 5 s, 60 °C for 10 s, 72 °C for 10 s (45×). Ct values were normalized to a housekeeping gene (*tsr*).

Spike-in for expression normalization

An independent TRIP pool with a complexity of approximately 200 barcodes was suspended in TRIsure (Bioline) and added to TRIP cells collected for cDNA extraction to an amount of 2.5 % based on cell counts.

Library preparation for next-generation sequencing

For cDNA reads, RNA from 10×10^6 transfected cells isolated by FACS sorting was extracted and transcribed into cDNA. One-fourth of total cDNA was used in a PCR to add sequencing adapters. Before sequencing library preparation, the appropriate PCR cycle number was determined by qPCR to avoid overamplification. For gDNA reads, gDNA was isolated from 10×10^6 unsorted TRIP cells using ISOLATE II Genomic DNA Kit (Bioline). 500 ng was used in a PCR to add sequencing adapters and amplified over 25 cycles. PCR was performed using Phusion polymerase (2 U, Thermo Scientific) in GC buffer with 500 nM forward and reverse primer (library indexing primers and 151AR) and dNTPs (200 µM each) in a total volume of 100 µl. PCR conditions were 1 min at 98 °C (1×), followed by 15 s at 98 °C, 30 s at 68 °C and 30 s at 72 °C (cycle number as determined by qPCR) followed by 3 min at 72 °C. PCR products were purified using the ISOLATE II PCR and Gel Kit (Bioline).

Next-generation sequencing

Before next-generation sequencing, product size selection was performed using 2 % E-Gel SizeSelect Gels (Thermo Scientific). For expression and normalization reads, 6 cDNA samples and 6 gDNA samples were multiplexed in one lane and sequenced with 65-bp single reads using a HiSeq2500 (Illumina) yielding 210×10^6 reads. For integration mapping, triplicate indexed PCR samples were pooled and sequenced with 75-bp paired-end reads using a MiSeq (Illumina) yielding 24×10^6 reads. Reads were mapped to *Drosophila melanogaster* genome release dm3 using Bowtie.

Data analysis

Extraction of 21-nt barcode reads from fastq files and alignment of TRIP mapping reads was performed using the TRIP script available at http://trip.nki.nl. We set a Hamming

distance of 2 for removing mutated barcodes, a maximum distance of 500 nt on the forward read and 20 nt on the reverse read to cluster the positions during mapping and a minimum read number of 1. We obtained a total of 27/28 × 10^6 barcoded cDNA reads and 41/46 × 10^6 barcoded gDNA reads for two replicate experiments of day 2 after transfection. Spike-in library gDNA was sequenced separately to determine the spike-in barcode sequences. A list of the most abundant spike-in barcodes with over 1000 counts in gDNA was used for extracting spike-in reads from the TRIP samples. cDNA read counts per barcode were divided by the sum of cDNA read counts for all spike-in barcodes in that sample to correct for sequencing depth and standardize expression between samples. gDNA read counts were transformed to counts per million (cpm) to correct for sequencing depth. Spike-in corrected cDNA counts were then divided by gDNA cpm counts which yields the normalized expression per barcode. We set a cutoff for barcodes that had at least 100 gDNA reads in all samples to ensure that each barcode was sufficiently represented in the pool. After cutoff for gDNA counts, we worked with an average of 1200/1500 cDNA reads per barcode in replicate 1/2 of the Gal4-HP1-transfected samples and 920/1400 and 1700/1900 for the Gal4- and HP1-transfected control samples. For gDNA counts, we obtained an average of 1900/2800 reads per barcode of the Gal4-HP1-transfected samples and 2900/2800 and 2200/2500 for the Gal4- and HP1-transfected controls. To ensure reliable allocation of barcode integrations, we worked with a total of 1093 barcodes that fit the following criteria: (1) more than 2 reads in both forward and reverse mapping, (2) 80 % of forward and reverse mapping reads matched with the first mapping location and less than 10 % of forward and reverse mapping reads matched with the second mapping location and (3) mapq score of 10 or higher for forward and reverse mapping. Sleeping Beauty is known not to generate tandem integrations; moreover, by exclusion of integrations with non-concordant forward and reverse mapping reads we exclude potential tandem integration events. As we observed good correlation between replicate experiments, we performed all further analysis on the mean normalized expression per barcode. For barcodes with 0 reads in one replicate, we set the mean normalized expression to 0. Data analysis was done in RStudio, R version 3.2.1.

Immunofluorescence microscopy

Kc167 cells were collected 2 days after transfection, resuspended in serum-free medium and allowed to settle on coverslips coated with poly-D-lysine (Sigma-Aldrich) for 1 h. Cells were washed with PBS and fixed with 2 % paraformaldehyde (Sigma-Aldrich) for 10 min. After washing with PBS, cells were permeabilized with 0.5 % NP-40 in PBS for 10 min and washed with PBS, and coverslips were transferred to a wet chamber. Coverslips were

blocked with PBG (0.5 % of 0.2 % cold water fish gelatin with 0.2 % sodium azide in PBS) for 15 min and then incubated with anti-HP1a primary antibody at 386 ng/ml concentration (C1A9, Developmental Studies Hybridoma Bank of the University of Iowa) for 1 h. After four 5-min washes with PBG, coverslips were incubated with 1:200-diluted DyLight 594-coupled secondary antibody (715-515-150, Jackson ImmunoResearch) for 1 h. After washing with PBG and PBS, coverslips were mounted to objective slides in Vectashield mounting medium with DAPI (Vector Laboratories). Images were taken using an Axio Observer (Zeiss) with a 40× objective and oil immersion and processed with Zen software.

Western blot

Cells transfected with Gal4-HP1, HP1 or Gal4 plasmids were isolated by FACS sorting by 2, 4 and 6 days after transfection. Cells were kept on ice and lysed for protein extraction in lysis buffer (10 mM KCl, 1.5 mM MgCl2, 10 mM Tris, 10 % SDS) supplemented with proteinase inhibitor (cOmplete, Roche). 40 µg per sample was diluted in Laemmli buffer (50 mM Tris–HCl pH 6.8, 2 % SDS, 10 % glycerol, 1 % beta-mercaptoethanol, 12.5 mM EDTA, 0.02 % bromophenol blue), denatured at 90 °C for 5 min, separated on a 16 % SDS-PAGE gel and blotted on nitrocellulose membrane (Amersham Protran 0.45 µM). HP1a was detected using monoclonal antibody C1A9 obtained from the Developmental Studies Hybridoma Bank of the University of Iowa at 386 ng/ml concentration. H3 was detected using antibody ab1791 obtained from Abcam at 33 ng/ml concentration. Quantitative western blot analysis was performed using the LI-COR Odyssey IRDye® IR imager (Biosciences), IRDye secondary antibodies and the Odyssey LI-COR software.

MboI assay for reporter accessibility

1×10^6 cells of each single-integration cell line with GFP reporter integrations were transfected with Gal4-Dam and incubated for 2 days. gDNA was extracted using ISOLATE II Genomic DNA Kit (Bioline) according to manufacturer's instruction, and 500 ng was digested with 10 units of each MboI and NaeI for 1 h at 37 °C in CutSmart buffer (New England Biolabs) in a total volume of 25 µl. 20 ng digested DNA was used in a qPCR with primers "upstream," "GFP" and "TSR." Amplicons of "upstream" and "GFP" primers each cover one GATC site and can therefore be used to detect DAM methylation levels, whereas the amplicon of primers "TSR" does not contain GATC and was therefore used for normalization. Methylation levels were calculated as percentage of averages of negative and positive controls.

DamID-seq

20×10^6 wild-type Kc167 cells were transfected in duplicates with 20 µg plasmids expressing Dam-HP1 or Dam only and collected 2 days after transfection. gDNA from 5×10^6 cells was isolated using ISOLATE II Genomic DNA Kit (Bioline) according to manufacturer's instruction. 500 ng gDNA was digested with DpnI (10 U, New England Biolabs) in CutSmart buffer in a total volume of 20 µl at 37 °C for 8 h. Reaction was terminated by heat inactivation at 80 °C for 20 min. Fragments were ligated to 12.5 pmol DamID adapters using T4 ligase (2.5 U, New England Biolabs) in T4 ligase buffer in a total volume of 25 µl incubated at 16 °C for 16 h. The reaction was heat inactivated for 10 min at 65 °C. Products were then digested with DpnII to remove partially methylated fragments. DpnII buffer and DpnII (10 U, New England Biolabs) were added in a total volume of 80 µl and incubated at 37 °C for 1 h. 20 µl of DpnII-digested products was amplified by PCR with MyTaq Red Mix (Bioline) and 2.5 µM primers Adr-PCR-Rand1 in a total volume of 80 µl. PCR settings were 8 min at 72 °C (1×) followed by 20 s at 94 °C, 30 s at 58 °C, 20 s at 72 °C (15×) and 2 min at 72 °C (1×). PCR products were cleaned up using the ISOLATE II PCR and Gel Kit (Bioline) according to the manufacturer's instructions and eluted in 26 µl H_2O. Fragment ends were blunted using the End-It™ DNA End-Repair Kit (Epicentre) in a 50-µl reaction according to the manufacturer's instructions. Products were cleaned up using the ISOLATE II PCR and Gel Kit (Bioline) and eluted in 26 µl H_2O. For adding a 3′ overhang, fragments were treated with Klenow fragment (3′-5′ exo-, 25 U, New England Biolabs) in buffer 2 with 200 µM dATP (Thermo Fisher Scientific) in a 50-µl reaction incubated at 37 °C for 30 min. Reaction was terminated by heat inactivation at 75 °C for 20 min. Products were purified using CleanPCR magnetic beads (CleanNA) according to the manufacturer's instructions and eluted in 20 µl H_2O. 260 ng fragments were ligated to 2.5 µM Y-adaptors using T4 ligase (2.5 U, New England Biolabs) in T4 ligase buffer in a total volume of 10 µl incubated at 16 °C for 16 h. Reaction was terminated by heat inactivation at 65 °C for 10 min. Products were purified using CleanPCR magnetic beads (CleanNA) according to the manufacturer's instructions and eluted in 20 µl H_2O. Adapters and indices for next-generation sequencing and multiplexing were added by PCR using 8 µl of purified fragments in MyTaq Red mix together with 250 nM Illumina index primers and 250 nM primer P5-Illumina-2 in a total volume of 20 µl. PCR settings were 1 min at 94 °C (1×) followed by 30 s at 94 °C, 30 s at 58 °C, 30 s at 72 °C (11×) and 2 min at 72 °C (1×). PCR products were pooled at equal amounts, purified using CleanPCR magnetic beads (CleanNA) and eluted in 20 µl H_2O. Pooled samples of two replicates transfected with Dam-HP1 or Dam only were sequenced with 69×10^6 reads in total.

Reads were filtered for sequences containing DamID adapter sequence using cutadapt, aligned to *Drosophila* genome release dm3 using Bowtie 2 and matched with GATC-flanked fragments which yielded 6–11 × 10^6 fragments per sample. As replicate fragment counts were highly correlated (r_p = 0.99), we worked with mean counts of two Dam-HP1 replicates normalized to mean of Dam only replicates for further analysis.

Inducible expression experiment with single-integration cell line

GFP reporter expression in single-integration cell line "C" was induced by adding $CuSO_4$ to a final concentration of 0.005, 0.05 and 0.5 mM. The experiment was performed in three replicates. Two days after induction, 20 × 10^6 cells were transfected with plasmids encoding Gal4-HP1 or Gal4 only. Two days after transfection, transfected cells were isolated by FACS sorting for mCherry. GFP expression was measured by qPCR using primers "GFP" and normalized to housekeeping gene *tsr* using primers "TSR."

Abbreviations

FACS:	fluorescence-activated cell sorting
Gal4DBD:	Gal4 DNA-binding domain
gDNA:	genomic DNA
GFP:	green fluorescent protein
HP1:	heterochromatin protein 1
NGS:	next-generation sequencing
pMT:	metallothionein promoter
TRIP:	thousands of reporters integrated in parallel

Authors' contributions

LB and BvS designed experiments with advice from WA. LB performed experiments and analyzed the data. JvA generated cell lines and helped in optimizing the protocol. LP and WA contributed to data analysis. LB and BvS wrote the manuscript. All authors (except WA) read and approved the final manuscript.

Acknowledgements

We thank the NKI Genomics Core Facility and FACS Facility for technical support; members of our laboratory for helpful suggestions; Gillaume Filion and Arantxa Rosado for sharing unpublished protocols; Maarten van Lohuizen for supporting WA.
Waseem Akhtar: Deceased on October 14, 2015.

Competing interests

The authors declare that they have no competing interests.

Availability of data and material

The TRIP expression and DamID-seq data described here are available from Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo), accession GSE83715. For matching TRIP integrations with protein occupancy or epigenetic marks, we used publicly available maps generated by DamID and data available from the modENCODE consortium. DamID maps were downloaded from GEO accession GSE36175. Genomic annotation of chromatin states according to the five-state model was downloaded from GEO (GSE22069). Annotation according to the nine-state model (accession # modENCODE_21955), ChIP-seq peak data for H3K36me3 (modENCODE_302) and JIL-1 (modENCODE_3037) were downloaded from modMine (http://intermine.modencode.org v33 28JUL2014).

Funding

This work was supported by ERC Advanced Grant 293662, NIH U54 Grant 1U54DK107965-01 and by an NWO-ALW grant to WA and M. van Lohuizen.

References

1. Filion GJ, van Bemmel JG, Braunschweig U, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*. 2010;143(2):212-224. doi:10.1016/j.cell.2010.09.009
2. Kharchenko PV, Alekseyenko AA, Schwartz YB, et al. Comprehensive analysis of the chromatin landscape in Drosophila. *Nature*. 2011;471(7339):480-485. doi:10.1038/nature09725
3. Eissenberg JC, James TC, Foster-Hartnett DM, Hartnett T, Ngan V, Elgin SC. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. In: *Proc Natl Acad Sci USA*. Vol 87. ; 1990:9923-9927.
4. Verschure PJ, Kraan I, Leeuw W, et al. In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. *Mol Cell Biol*. 2005;25:4552-4564.
5. Canzio D, Chang EY, Shankar S, et al. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome bridging mechanism for heterochromatin assembly. *Mol Cell*. 2011;41(1):67-81. doi:10.1016/j.molcel.2010.12.016
6. Yin H, Sweeney S, Raha D, Snyder M, Lin H. A high-resolution whole-genome map of key chromatin modifications in the adult Drosophila melanogaster. *PLoS Genet*. 2011;7:e1002380.
7. Wit E, Greil F, Steensel B. Genome-wide HP1 binding in Drosophila: developmental plasticity and genomic targeting signals. *Genome Res*. 2005;15:1265-1273.
8. Yasuhara JC, Wakimoto BT. Oxymoron no more: the expanding world of heterochromatic genes. *Trends Genet*. 2006;22:330-338.
9. Kwon SH, Workman JL. The changing faces of HP1: from heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. *BioEssays*. 2011;33:280-289.
10. Schotta G, Ebert A, Dorn R, Reuter G. Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. *Semin Cell Dev Biol*. 2003;14:67-75.
11. Fischle W, Tseng BS, Dormann HL, et al. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature*. 2005;438:1116-1122.
12. Wang C, Cai W, Li Y, et al. The epigenetic H3S10 phosphorylation mark is required for counteracting heterochromatic spreading and gene silencing in Drosophila melanogaster. *J Cell Sci*. 2011;124(Pt 24):4309-4317.

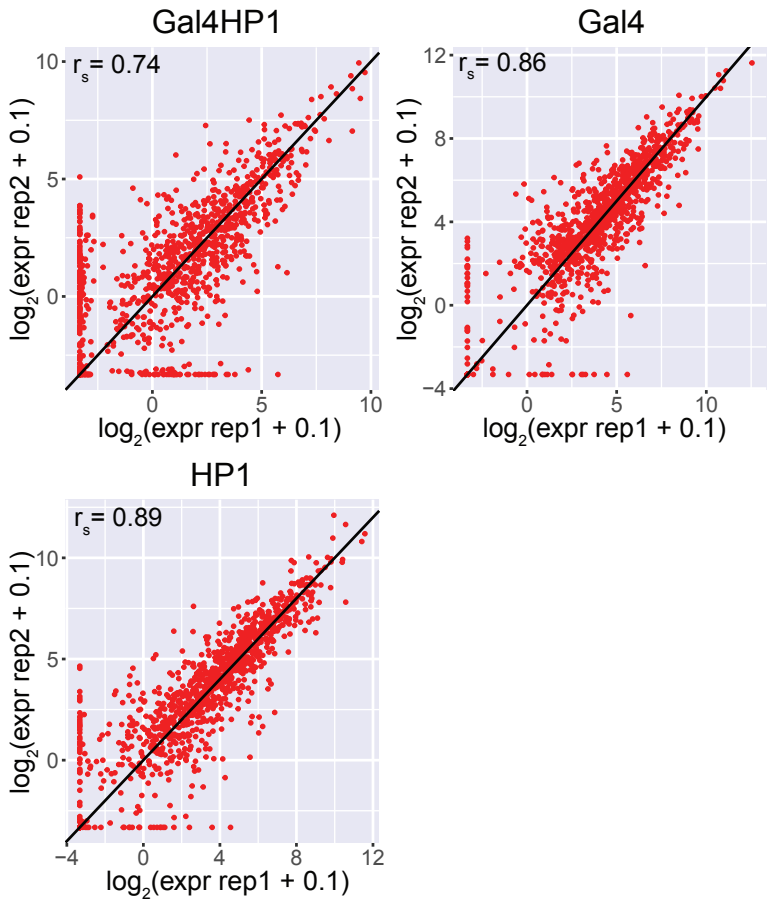
13. Akhtar W, de Jong J, Pindyurin AV, et al. Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell*. 2013;154(4):914-927. doi:10.1016/j.cell.2013.07.018
14. Pan T, Coleman JE. Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor. In: *Proc Natl Acad Sci USA*. Vol 86. ; 1989:3145-3149.
15. Sadowski I, Ma J, Triezenberg S, Ptashne M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature*. 1988;335:563-564.
16. Margolin JF, Friedman M JR, WK V, H T, HJ R, F.J. Krüppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci USA*. 1994;91:4509-4513.
17. Li Y, Danzer A JR, P B, AS W, L.L. Effects of tethering HP1 to euchromatic regions of the Drosophila genome. *Development*. 2003;130:1817-1824.
18. Dialynas G, Speese S, Budnik V, Geyer PK, Wallrath LL. The role of Drosophila Lamin C in muscle function and gene expression. *Dev Camb Engl*. 2010;137(18):3067-3077. doi:10.1242/dev.048231
19. Danzer W JR, L.L. Mechanisms of HP1-mediated gene silencing in Drosophila. *Development*. 2004;131:3571-3580.
20. Seum C, Delattre M, Spierer A, Spierer P. Ectopic HP1 promotes chromosome loops and variegated silencing in Drosophila. *EMBO J*. 2001;20(4):812-818. doi:10.1093/emboj/20.4.812
21. Steensel B, Henikoff S. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol*. 2000;18:424-428.
22. Zhang W, Deng H, Bao X, et al. The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in Drosophila. *Development*. 2006;133:229-235.
23. Elgin SCR, Reuter G. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harb Perspect Biol*. 2013;5:a017780.
24. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. Dynamics and memory of heterochromatin in living cells. *Cell*. 2012;149(7):1447-1460. doi:10.1016/j.cell.2012.03.052
25. Ayyanathan K, Lechner MS, Bell P, et al. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev*. 2003;17:1855-1869.
26. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. Published online 2015.

27. Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S. Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. *J Cell Biol*. 2003;161:707-714.
28. Cryderman DE, Grade SK, Li Y, Fanti L, Pimpinelli S, Wallrath LL. Role of Drosophila HP1 in euchromatic gene expression. *Dev Dyn*. 2005;232:767-774.
29. Lundberg LE, Stenberg P, HP1a LJ. Su(var)3–9, SETDB1 and POF stimulate or repress gene expression depending on genomic position, gene length and expression pattern in Drosophila melanogaster. *Nucleic Acids Res*. 2013;41:4481-4494.
30. Yoh SM, Lucas JS, Jones KA. The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev*. 2008;22:3422-3434.
31. Sun X-J, Wei J, Wu X-Y, et al. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J Biol Chem*. 2005;280:35261-35271.
32. Pan L, Xie W, Li K-L, et al. Heterochromatin remodeling by CDK12 contributes to learning in Drosophila. In: *Proc Natl Acad Sci USA*. Vol 112. ; 2015:13988-13993.
33. Nakayama J, Klar AJ, Grewal SI. A chromodomain protein, Swi6, performs imprinting functions in fission yeast during mitosis and meiosis. *Cell*. 2000;101:307-317.
34. Hall IM, Shankaranarayana GD, Noma K-I, et al. Establishment and maintenance of a heterochromatin domain. *Science*. 2002;297:2232-2237.
35. Audergon PNCB, Catania S, Kagansky A, et al. Restricted epigenetic inheritance of H3K9 methylation. *Sci-Ence*. 2015;348:132-135.
36. Ragunathan K, Jih G, Moazed D. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science*. 2014;348(1258699).
37. Jong J, Akhtar W, Badhai J, et al. Chromatin landscapes of retroviral and transposon integration profiles. *PLoS Genet*. 2014;10:e1004250.
38. Greil F, Kraan I, Delrow J, et al. Distinct HP1 and Su(var)3–9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev*. 2003;17:2825-2838.
39. Tolhuis B, Wit E, Muijrers I, et al. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. *Nat Genet*. 2006;38:694-699.
40. Bemmel J, Filion G, Rosado A, et al. A network model of the molecular organization of chromatin in Drosophila. *Mol Cell*. 2013;49:759-771.
41. González M, Martín-Ruíz I, Jiménez S, Pirone L, Barrio R, Sutherland JD. Generation of stable Drosophila cell lines using multicistronic vectors. *Sci Rep*. 2011;1(75).

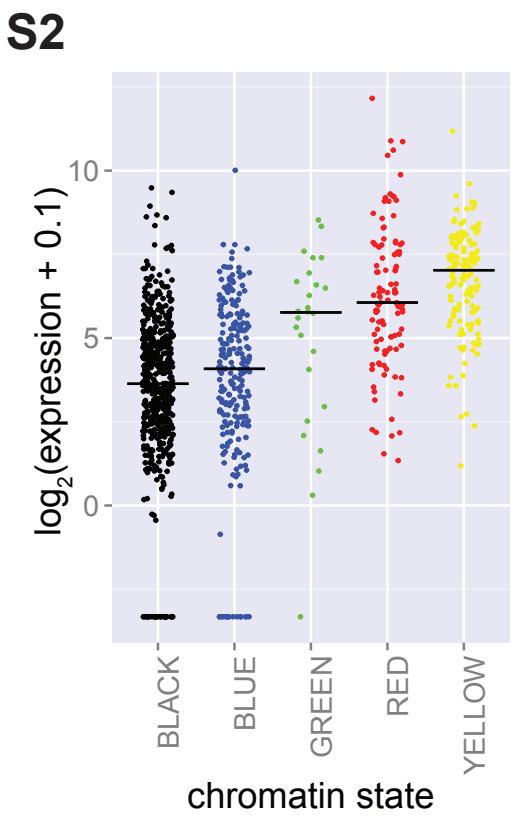
42. Kind J, Pagie L, Ortabozkoyun H, et al. Single-cell dynamics of genome–nuclear lamina interactions. *Cell*. 2013;153:178-192.
43. Bateman L JR, AM W, C. Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. *Genetics*. 2006;173:769-777.

Supplementary Data

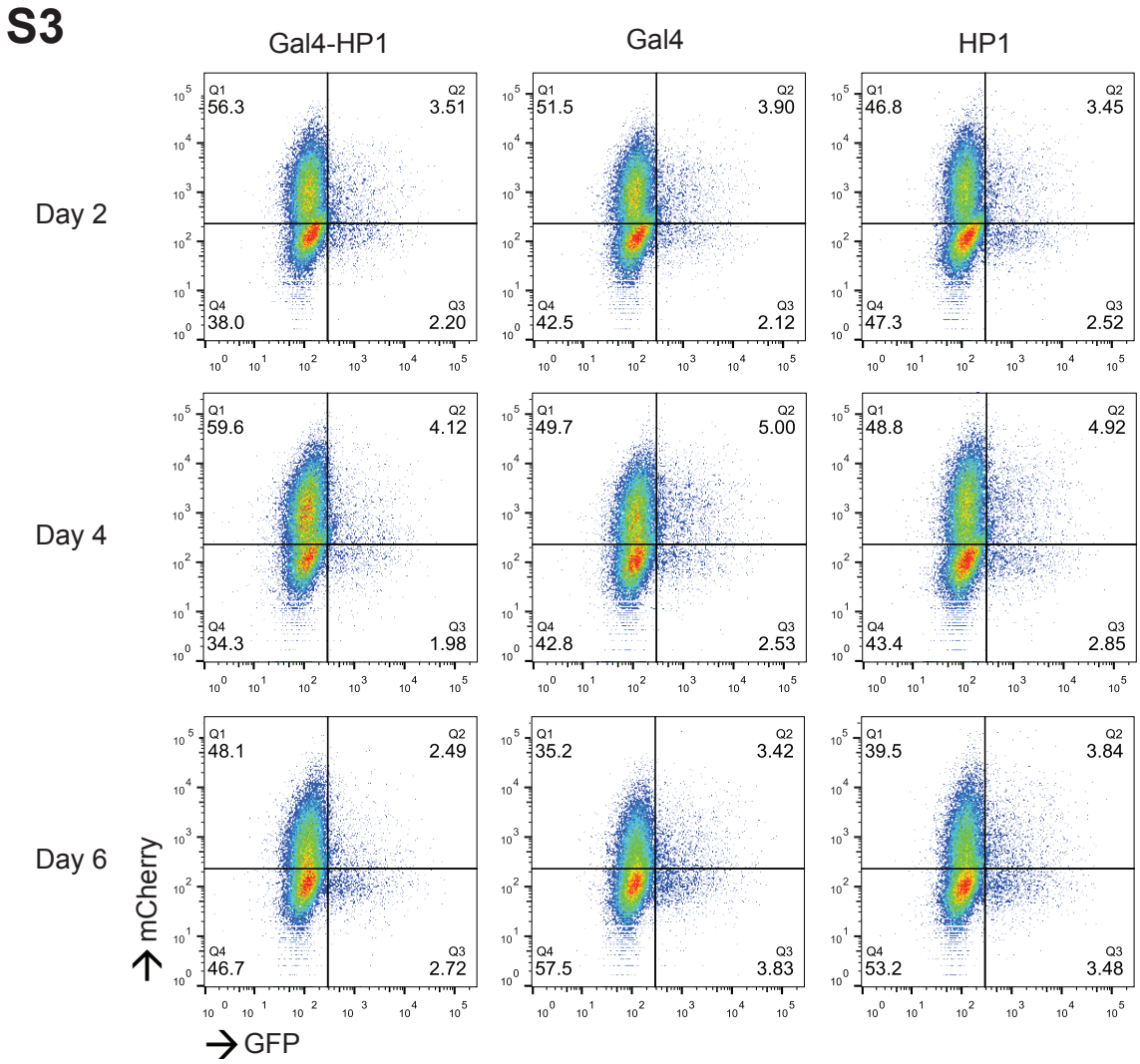
S1



Supplementary Figure S 1 Correlation of normalized expression between replicate experiments.

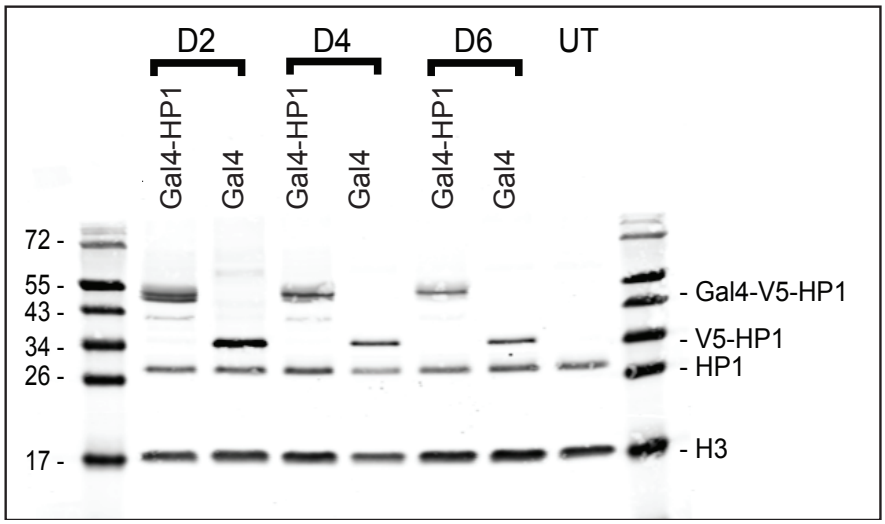


Supplementary Figure S 2 Expression of integrated reporters as quantified by NGS in the Gal4-transfected control condition divided over five chromatin states. Median values are represented by black horizontal bars. Number of reporters integrated in each state is specified above graph.



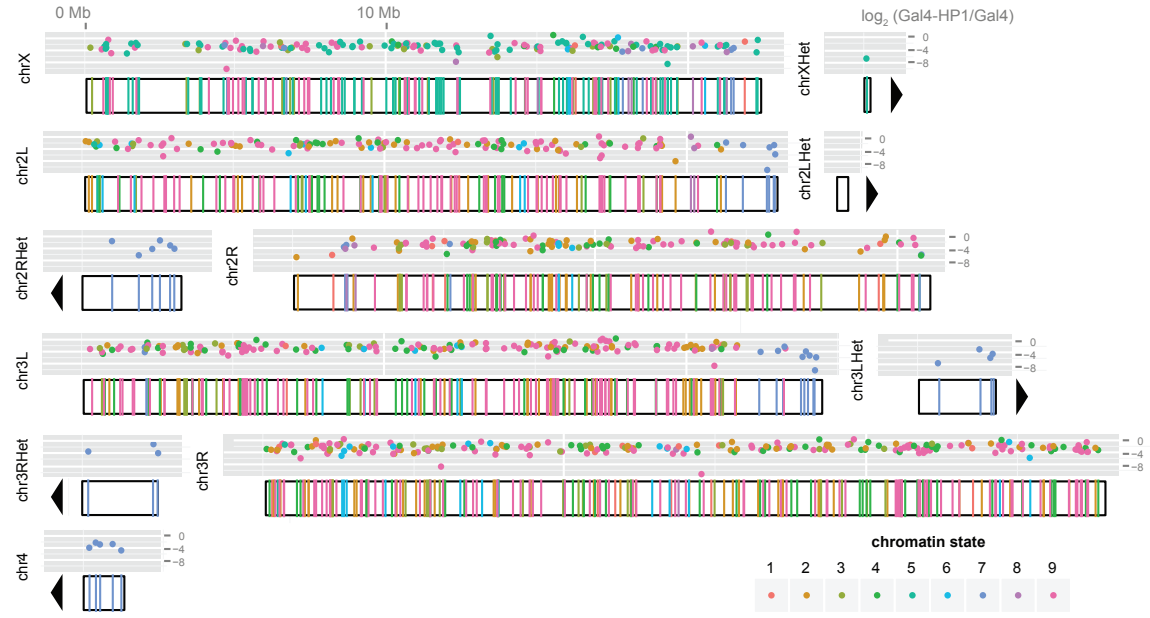
Supplementary Figure S 3 GFP expression and transfection rate (measured via mCherry) of TRIP pool as quantified by FACS by days 2, 4 and 6 after transfection.

S4

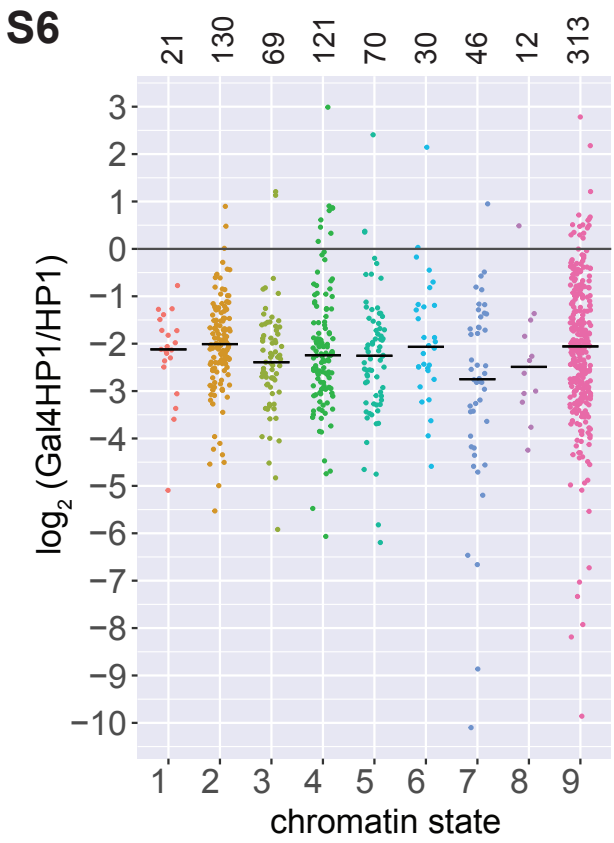


Supplementary Figure S 4 Western blot of Kc167 cells transiently transfected with Gal4-V5-HP1 and V5-HP1 or untransfected control (UT) stained with anti-HP1a and anti-H3.

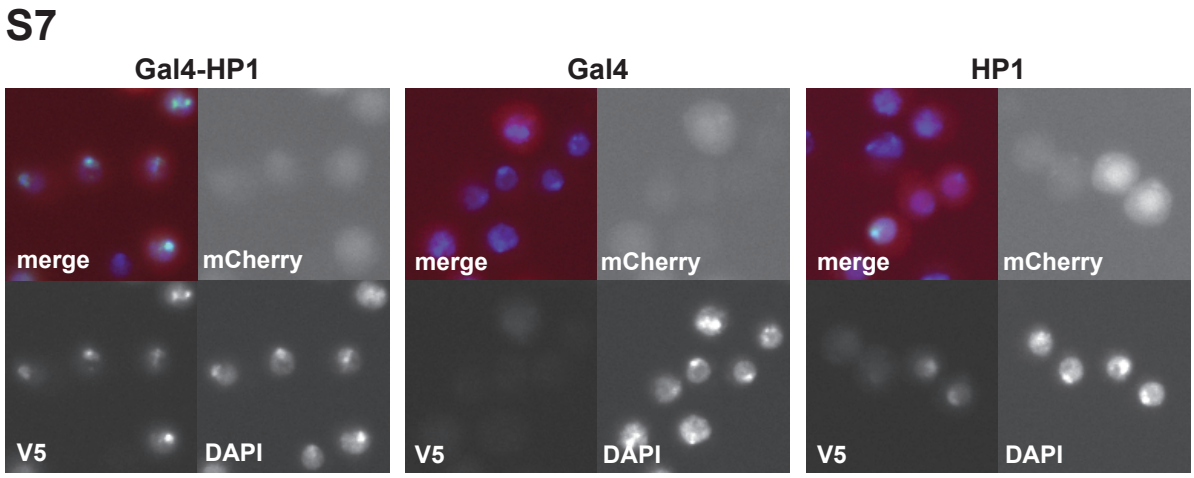
S5



Supplementary Figure S 5 Chromosome ideograms of TRIP reporter integration sites. Colors represent chromatin state at the integration site according to the nine-state model. Centromere position is indicated by black triangle. Scatter plot above ideogram shows fold downregulation upon Gal4-HP1 tethering.



Supplementary Figure S 6 Fold change in reporter expression Gal4-HP1/HP1 as quantified by NGS divided over nine chromatin states. Median values are represented by black horizontal bars. Number of reporters integrated in each state is specified above graph.



Supplementary Figure S 7 Localization of transiently transfected constructs. Gal4-V5-HP1, Gal4-V5 and V5-HP1 in Kc167 cells as detected by immunofluorescence with anti-V5 (green) and DAPI (blue). Transfected cells express mCherry (red).

Chapter 3

CHRAC/ACF contribute to the repressive ground state of chromatin

Alessandro Scacchetti, Laura Brueckner, Dhawal Jain, Tamas Schauer, Xu Zhang, Frank Schnorrer, Bas van Steensel, Tobias Straub, Peter B Becker

Life Science Alliance **1,** 1 (2018)

Abstract

The chromatin remodeling complexes chromatin accessibility complex and ATP-utilizing chromatin assembly and remodeling factor (ACF) combine the ATPase ISWI with the signature subunit ACF1. These enzymes catalyze well-studied nucleosome sliding reactions in vitro, but how their actions affect physiological gene expression remains unclear. Here, we explored the influence of *Drosophila melanogaster* chromatin accessibility complex/ACF on transcription by using complementary gain- and loss-of-function approaches. Targeting ACF1 to multiple reporter genes inserted at many different genomic locations revealed a context-dependent inactivation of poorly transcribed reporters in repressive chromatin. Accordingly, single-embryo transcriptome analysis of an *Acf* knock-out allele showed that only lowly expressed genes are derepressed in the absence of ACF1. Finally, the nucleosome arrays in *Acf*-deficient chromatin show loss of physiological regularity, particularly in transcriptionally inactive domains. Taken together, our results highlight that ACF1-containing remodeling factors contribute to the establishment of an inactive ground state of the genome through chromatin organization.

Introduction

The chromatin accessibility complex (CHRAC) and the related ATP-utilizing chromatin assembly and remodeling factor (ACF) are prototypic nucleosome sliding factors purified originally from extracts of *Drosophila melanogaster* embryos [1,2]. ACF consists of ISWI, an ATPase of the helicase superfamily 2, and a large subunit, ACF1. ACF associates with two histone-fold subunits, CHRAC-14 and CHRAC-16, to form CHRAC [3]. Both complexes have very similar nucleosome sliding activity in vitro [4]. Because ISWI is present in several other nucleosome remodelers [5], ACF1 serves as the signature regulatory subunit for the two complexes.

The mechanism of nucleosome sliding has been well described by biochemical and biophysical studies. ISWI and ACF1 bind target nucleosomes and flanking linker DNA. Substrate binding and ATP hydrolysis cycles trigger conformation changes in the remodeler that disrupt histone–DNA interactions and eventually displace the intact histone octamer along the DNA, effectively sliding a nucleosome [6–12] (for review, see reference [13]).

Nucleosome sliding may theoretically affect transcription through local and global mechanisms [14]. Locally, nucleosomes could be slid off promoters, exposing binding sites for transcription factors. Conversely, a remodeler might push nucleosomes to occlude regulatory sequences. The yeast Isw2 complex, which is related to the metazoan CHRAC complexes, has been shown to slide nucleosomes toward promoters [15,16]. Alternatively, nucleosome sliding factors may influence transcription by globally affecting the tightness of DNA packaging in chromatin. Nucleosome sliding factors may improve the regularity of nucleosome arrays by closing gaps ("nucleosome spacing"), thus minimizing the level of accessible DNA [1,6,17]. In vitro, regularly spaced nucleosome arrays readily fold into "30 nm"–type fibers, a process that has been suggested to promote the formation of "higher order," repressive chromatin structures [14].

A role for CHRAC/ACF in establishing such repressive chromatin had been derived from early studies of *Acf* mutant embryos that documented defects in nucleosome spacing, in the formation of repressive pericentric heterochromatin and polycomb-mediated silencing [18,19]. A more direct role for ACF1 in the repression of wingless target genes has also been described [20]. The phenotypic oogenesis defects observed in *Acf* mutants [21] may be explained by either mechanism. These early studies based their conclusions on the analysis of Acf^{1} and Acf^{2} alleles, which were later shown in the context of oogenesis not to deliver a complete loss-of-function genotype because they still express the C-terminal PHD/bromo domains of ACF1 [21]. Indeed, some oogenesis phenotypes observed in Acf^{1} and Acf^{2} could not be reproduced with a larger gene deletion (Acf^{7}, considered a true null allele) or under

RNAi conditions, as shown in reference [21]. Therefore, the consequences of a complete loss of ACF1 (and thus the remodeling complexes it defines) are still unknown.
Clues about ACF1 functions at specific loci may be derived from mapping the chromosomal binding sites of the remodeler by chromatin immunoprecipitation (ChIP). Unfortunately, despite many efforts, we were not able to map ACF1 binding sites by ChIP, presumably because the interaction of the remodeler is too transient and dynamic to be efficiently cross-linked [22].
To unequivocally clarify the effect of ACF1-containing remodelers on transcription, we performed two key experiments. In a gain-of-function approach, we artificially targeted ACF1 to reporter genes integrated at many different chromatin loci and monitored the consequences for reporter gene transcription. Furthermore, using a null allele, we compared the transcriptome of individually staged null mutants to that of matched wild-type embryos. Both approaches suggest that the main effect of ACF1 on transcription is that it participates in the silencing of genes in inactive chromosomal domains. Importantly, derepression in mutant embryos correlates with defects in nucleosome spacing. Hence, we conclude that ACF1-containing remodelers contribute to a repressed ground state of the genome through chromatin organization.

Results

Artificial ACF1 tethering leads to context-dependent repression

To investigate potential effects of ACF1 on transcription, we first applied an established approach involving the ectopic targeting of ACF1 to a reporter gene locus. We used a fly line with a defined genomic insertion of a reporter gene cassette consisting of *lacZ*/*mini-white* genes and 5′ UAS^{Gal} sequences [23] (Fig 1A). We generated flies expressing N- or C-terminal fusions of ACF1 to the DNA-binding domain of the yeast activator GAL4 (GAL4DBD) under the control of the endogenous Acf promoter, which assured expression at levels comparable to endogenous ACF1 (Fig 1A and B). A control line harbored a construct expressing only the GAL4DBD. Mating the two types of fly lines yields offspring in which ACF1 is recruited to the UAS^{Gal} element.

Successful tethering of ACF1-GAL4DBD in early embryos was confirmed by ChIP–quantitative PCR (qPCR) (Fig 1C). The entire remodeler seems to be recruited because its catalytic partner ISWI was also detected at the UAS sites (Fig S1A). ACF1 targeting resulted in an about twofold reduced LacZ transcription relative to the GAL4DBD only (mean fold-change = 0.58 for GAL4DBD-ACF1 and mean fold-change = 0.39 for ACF1-GAL4DBD) (Fig S1B). However, no obvious changes in nucleosome positions over and around the reporter locus could be scored by MNase-seq (Fig S1C and D).

These experiments provided the proof of principle that a functional ACF1-GAL4DBD fusion protein could be recruited to UAS^{Gal} elements integrated in the fly genome, but these lacked the necessary generalization to document the presumed repressive effect. Therefore, we used a previously characterized library of several hundred barcoded reporter genes that had been randomly integrated into the genome of *Drosophila* Kc167 cells. We previously tethered heterochromatin protein 1 (HP1) as a GAL4DBD fusion to these sites and determined how the chromatin environment modulated HP1 repression [24]. HP1, a known repressor, provides a convenient reference for ACF1 in this system. In parallel transient transfections, we introduced the various constructs into Kc167 cells and confirmed their expression by Western blotting and immunofluorescence microscopy: GAL4-ACF1, tagged ACF1 lacking a GAL4DBD, ACF1 lacking any tag, and a tagged GAL4DBD (Fig S2A–C).
As in the case of the single-reporter system, recruitment of ACF1 resulted in a general down-regulation (median log2 fold-change = −1.24), almost comparable to HP1 (median log2 fold-change = −1.46) (Fig 2A), which served as a positive control. Expression of ACF1 lacking the GAL4DBD had a much weaker effect than its tethered counterpart (median log2 fold-change = −0.24). Interestingly, the extent of ACF1-induced repression inversely

correlated with the mean expression levels of the reporters: the repressive effect was less pronounced for reporters with a high expression level (Fig 2B). A similar correlation could be also observed for the untethered ACF1 (Fig 2C). The tethered HP1 showed instead just a small correlation between down-regulation and reporter expression (Fig 2D), significantly different from GAL4-ACF1 (Fig S2D).

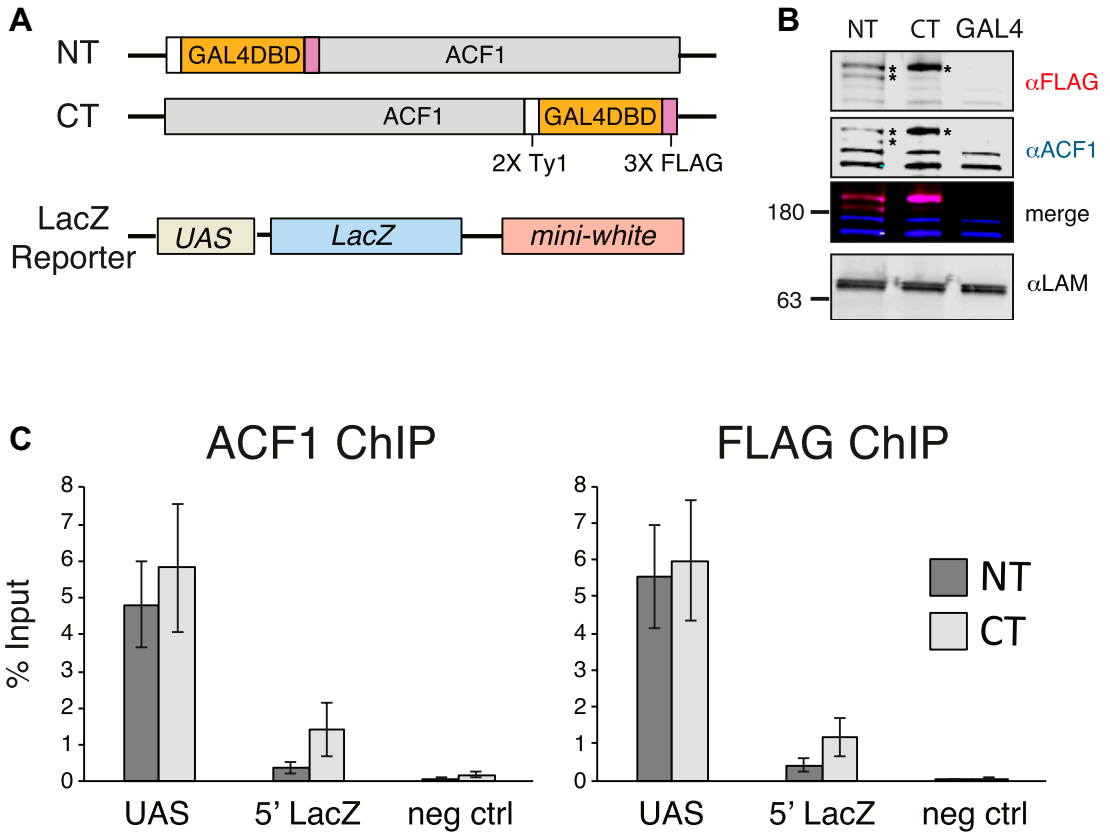


Figure 1
ACF can be tethered to a reporter locus through GAL4-DBD.
(A) Schematic illustration of the transgenes used for testing the effects of ACF1 recruitment. ACF1 is fused to the GAL4 DNA-binding domain (GAL4DBD) at either the N-terminus or the C-terminus. A transgene containing the GAL4DBD alone (GAL4) is used as a negative control. The reporter transgene contains five UASGal4 (UAS) 5′ of LacZ and mini-white genes. (B) Western blot detection of ACF1 in an embryo nuclear extract (0–16 h AEL). Endogenous and fusion proteins were detected with a specific ACF1 antibody (blue channel); the ACF1-GAL4 fusions are FLAG-tagged and detected with an anti-FLAG antibody (magenta channel). Asterisks indicate the expressed transgenic ACF1-GAL4 fusions. Embryos containing a transgene coding for the GAL4DBD alone (GAL4) are included as a negative control. Lamin serves as a loading control. (C) ChIP-qPCR monitors the recruitment of ACF1 to UAS in 0- to 12-h embryos. The immunoprecipitation was conducted using ACF1 and FLAG antibodies. "UAS" and "5′ LacZ" denote the regions amplified by qPCR. Bars denote average % Input enrichment (n = 3 biological replicates) ± SEM. "neg ctrl" represents a negative control locus (encompassing the Spt4 gene).

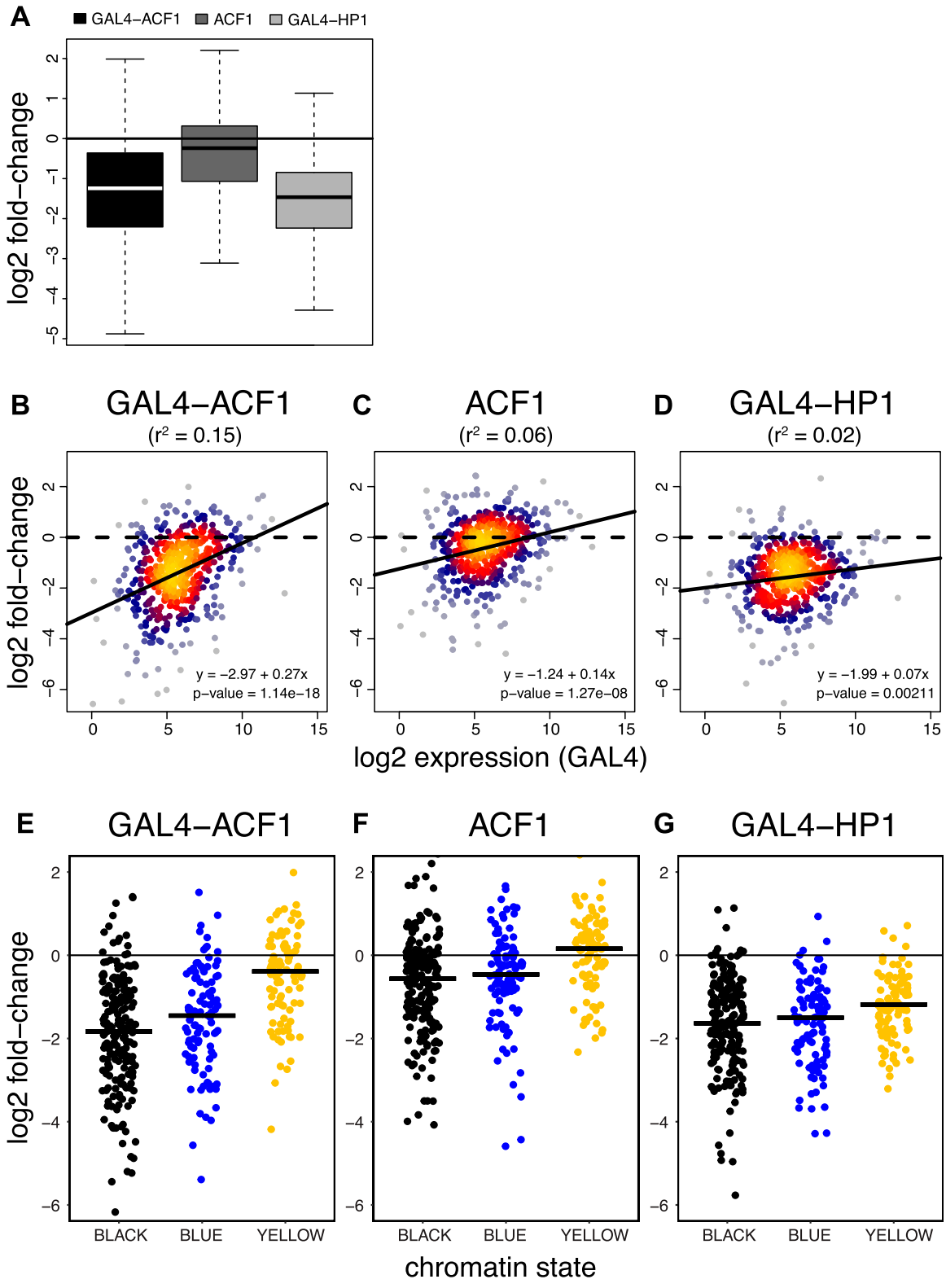
To explore whether the chromatin environment in which the individual reporter genes are integrated modulates ACF1-mediated repression, we referred to the five-state model of chromatin [25]. In this model, YELLOW and RED represent constitutively and developmentally regulated active chromatin domains, respectively; GREEN corresponds to HP1-marked

heterochromatic domains; and BLACK and BLUE correspond to inactive and polycomb-repressed domains, respectively. We added a sixth state, GRAY, to refer to reporters integrated in genomic regions not defined by any of the original five states. We found that reporters integrated in BLACK and BLUE chromatin domains are strongly repressed upon ACF1 targeting (BLACK: median log2 fold-change = −1.83; BLUE: median log2 fold-change = −1.45), whereas the ones integrated in YELLOW are only slightly affected (median log2 fold-change = −0.39) (Fig 2E). Similarly, down-regulation of reporters in RED (median log2 fold-change = −0.70) and GREEN (median log2 fold-change = −0.97) states results smaller than the one in BLACK and BLUE domains (Fig S2E). Recruitment of HP1, instead, shows a more general repressive effect compared with that of ACF1, which does not correlate with any type of chromatin (Figs 2G and S2G). Interestingly, expression of untethered ACF1 also shows a very mild context-dependent repression, reminiscent of its tethered counterpart (Figs 2F and S2F).

In summary, the tethering approach suggested that ACF1-containing remodelers have a repressive function. In contrast to repression by HP1, which served as a positive control for repression, the effect of ACF1 strongly depends on the chromatin context and is particularly robust in lowly expressed genes in overall inactive chromatin domains.

Figure 2

ACF1 represses multiple reporters in a context-dependent manner.
(A) Boxplots represent log2 fold-change distribution upon ACF1 tethering (GAL4-ACF1), ACF1 overexpression (ACF1), or HP1 tethering (GAL4-HP1), compared with the control (N = 492). (B) Log2 fold-change for each reporter in relation to its mean log2 expression upon ACF1 tethering versus the control (GAL4-ACF1) (N = 492). Black lines represent linear regression fit. r2 values derived from the linear model are shown in parentheses. Equations of the regression lines are displayed in the plot. P-value refers to the significance of this relationship (slope). (C) Same as (B) but for ACF1 overexpression (ACF1). (D) Same as (B) but for HP1 tethering (GAL4-HP1). (E) Jitter plots represent the distribution of log2 fold-changes of reporters integrated in BLACK (N = 197), BLUE (N = 102), and YELLOW (N = 94) chromatin domains for the case of tethered ACF1 (GAL4-ACF1). Black horizontal bars represent median. (F, G) Same as (E) but for ACF1 overexpression (ACF1) and HP1 tethering (GAL4-HP1), respectively.


A
GAL4–ACF1
ACF1
GAL4–HP1
log2 fold–change
B
GAL4–ACF1
(r² = 0.15)
y = –2.97 + 0.27x
p–value = 1.14e–18
C
ACF1
(r² = 0.06)
y = –1.24 + 0.14x
p–value = 1.27e–08
D
GAL4–HP1
(r² = 0.02)
y = –1.99 + 0.07x
p–value = 0.00211
log2 expression (GAL4)
E
GAL4–ACF1
F
ACF1
G
GAL4–HP1
BLACK
BLUE
YELLOW
chromatin state

Transcriptome analysis of Acf-deficient Drosophila embryos

The tethering experiment suggested that ACF1 may not affect gene transcription like a classical corepressor, but at a more fundamental level. However, given the artificial nature of the approach with its uncertainties about the functionality of the DBD fusion protein and the remodeling activity of the reconstituted complex, we sought to test the hypothesis of a context-dependent repressive effect of ACF1 by using a loss-of-function approach in a physiological system. A transcriptome analysis for an ACF1 deficiency has not been reported so far. Conceivably, a function of ACF/CHRAC may be best observed during early embryogenesis in *Drosophila* because ACF1 expression peaks during these stages and both CHRAC and ACF have been originally identified in embryos. Early on, embryogenesis defects had been noted for the **Acf**^{1} allele [18]. However, this allele only deletes an N-terminal fragment of the **Acf** gene and still allows the expression of a C-terminal "stub" containing a PHD/bromo domain module that may interfere with relevant interactions. We later concluded that the more extensive deletion of the **Acf**^{7} allele most likely represents a clean loss of function [21]. In parallel with these earlier studies, we generated a clean **Acf** gene deletion using a CRISPR/Cas9-based engineering approach (**Acf**^{C}). Expression of ACF1 is not detectable by Western blotting in homozygous embryos for the **Acf**^{C} or **Acf**^{7} alleles (Fig S3A) [21]. **Acf**^{C} and **Acf**^{7} embryos show a slightly lower hatching rate compared with their wild-type counterparts (Fig S3B; unpublished observation), but the survivors develop normally into viable and fertile flies (unpublished observation). We were concerned that **Acf**^{C} mutants might develop slower than wild-type embryos and hence did not rely on simple developmental stage timing for proper transcriptome comparison. Rather, we selected single embryos either before zygotic genome activation (ZGA) or after ZGA based on morphological hallmarks (Fig 3A; see the Materials and Methods section) and determined their transcriptome by RNA-seq analysis. Principal component analysis (PCA) showed no strong differences between **Acf**^{C} and wild type in both developmental stages (Fig 3B) with clear transition from the maternal to the zygotic RNA pool (Fig S3C). Differential gene expression analysis revealed a relatively small number of genes significantly affected by ACF1 loss at both stages (Fig 3C and D), but without a clear direction (activation or repression) and without a uniquely defined gene ontology enrichment (Fig S3D).

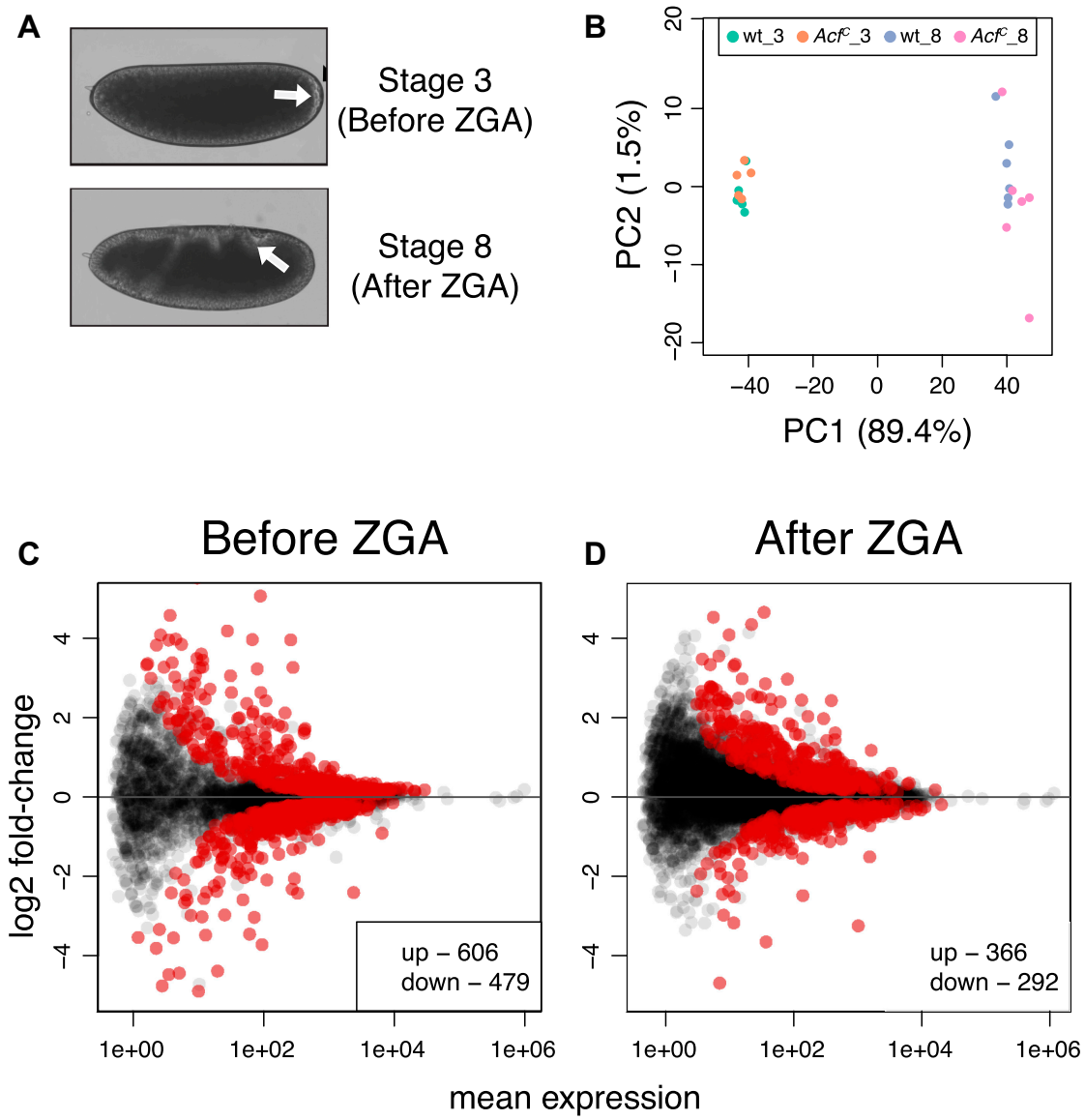


Figure 3

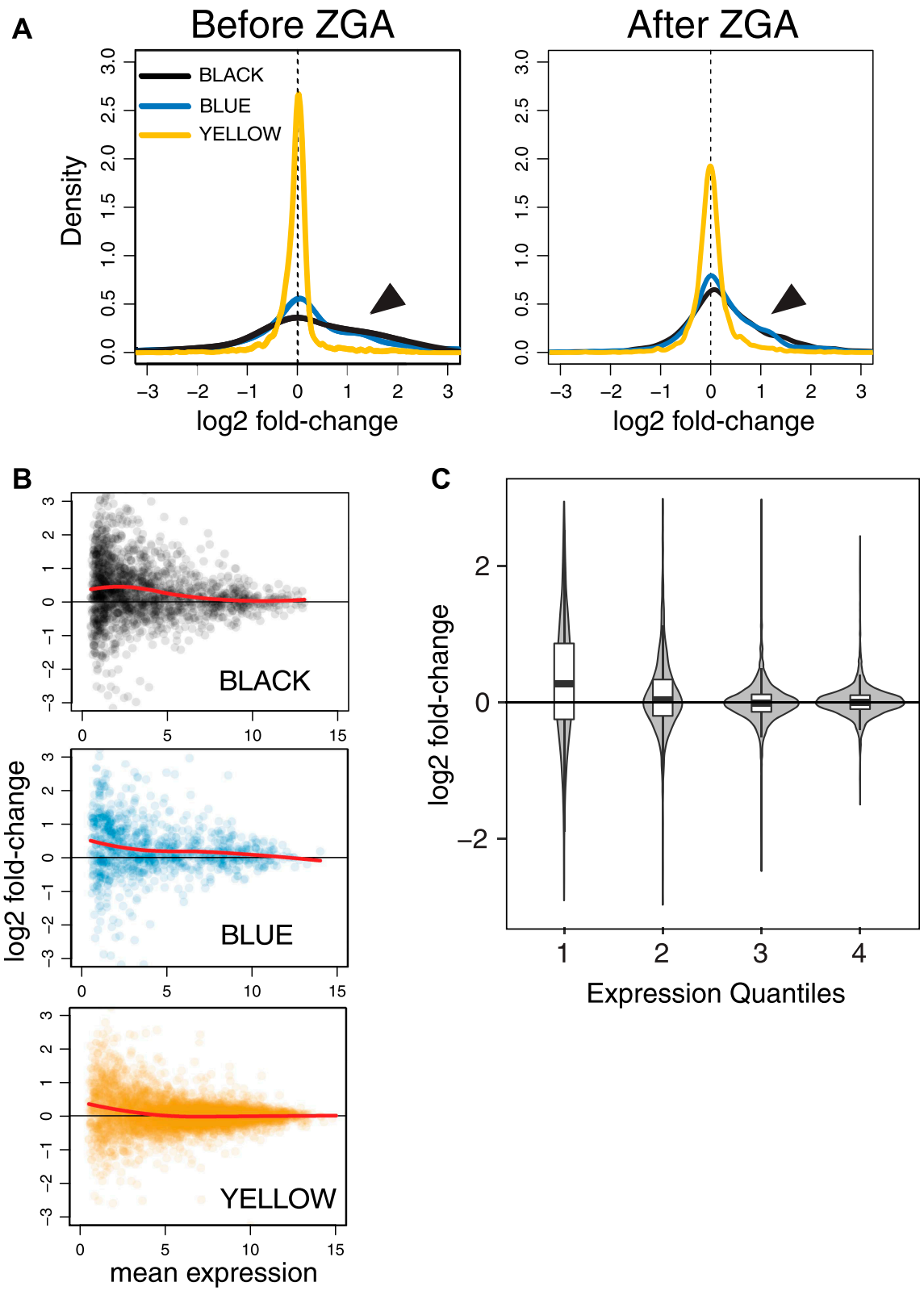
ACF1 loss perturbs gene expression in early embryos.
(A) Embryo stages selected for transcriptome analysis. In Bownes Stage 3, zygotic transcription is not established yet. Bownes Stage 8 shows robust zygotic transcription. Arrows highlight morphological features of the corresponding stages (appearance of pole cells in Stage 3 and germ band elongation in Stage 8). (B) PCA of single-embryo transcriptomes. Each dot represents a single replicate for the corresponding genotype/condition. _3 and _8 indicate embryos before and after ZGA, respectively. (C) Differential gene expression analysis of coding genes from RNA-seq data. Wild-type and AcfC transcriptomes were compared before ZGA (N = 7,585). Scatter plots represent log2 fold-change of AcfC over wild type for each gene in relation to its mean expression (mean of normalized counts). Red dots represent significant (q-value < 0.1) up- or down-regulated genes. (D). Same as (C) but after ZGA (N = 10,088).

Deletion of the Acf gene leads to relaxation of the repressive ground state of chromatin in early embryos

The relatively small number of differentially expressed genes upon **Acf** deletion may be explained by functional redundancy with other remodelers. However, the observation of context-dependent effects of ACF1 tethering prompted us to relate the transcription effects in embryos to the chromatin state of genes and to their transcriptional activity.
Evaluating the differences between **Acf**^{C} and wild-type embryos in the context of the five-state model of chromatin organization, we observed a small increase in expression of genes in inactive BLACK and BLUE chromatin domains in **Acf**-deficient embryos. In contrast, genes in the YELLOW domains (Fig 4A) and in GREEN or RED chromatin (Fig S4A) were largely unaffected by the loss of ACF1. We correlated our data with the modENCODE histone modification data [26] that had been obtained from 2- to 4-h embryos, very close to the zygotic stage analyzed in our study. The derepression of transcription upon ACF1 loss correlates with the absence of defined chromatin marks (BLACK), with the presence of H3K27me3 (BLUE), and with the absence of H3K36me3 (YELLOW) (Fig S4B, top panels). No clear correlation was observed for H3K9me3 (GREEN) or H3K4me3 (RED) (Fig S4B, bottom panels). The ACF1-dependent effect was most pronounced for lowly expressed genes, not only in the BLACK and BLUE domains but also in active YELLOW chromatin (Fig 4B). Indeed, the extent of derepression in **Acf** mutant embryos correlates generally with low expression levels, regardless of the chromatin domain a gene resides in (Fig 4C).
We conclude that loss of ACF1 leads to a widespread derepression of genes that are characterized by low levels of transcription in wild type. This supports our earlier conclusion derived from the ACF1 tethering experiments.

Figure 4

Loss of ACF1 affects transcription prominently in inactive chromatin.
(A) Comparison of wild-type and AcfC transcriptomes in relation to the five-state chromatin model. Plots represent the distribution of log2 fold-changes for genes belonging to the YELLOW, BLUE, and BLACK chromatin domains before and after ZGA. Arrows indicate the differences between BLACK/BLUE and YELLOW. (B) Each scatter plot represents log2 fold-change for each gene of the indicated chromatin state in relation to its mean expression (after ZGA only). Colors match the chromatin domains as described in the five-state model. Red lines represent local regression fit. (C) Violin plots represent log2 fold-change distributions for each given expression quartile (after ZGA only), regardless of the chromatin state. Boxplots are overlapped to show median values.


A
Before ZGA
After ZGA
BLACK
BLUE
YELLOW
Density
log2 fold-change
B
BLACK
BLUE
YELLOW
log2 fold-change
mean expression
C
log2 fold-change
Expression Quantiles

CHRAC/ACF repress inactive chromatin by maintaining nucleosome regularity

Nucleosome sliding by CHRAC/ACF improves the regularity of nucleosome arrays in vitro and hence optimizes the packaging of DNA. Lack of nucleosome spacing activity in vivo leads to irregular chromatin, which may explain the observed derepression phenotype in **Acf** mutants. To test this hypothesis, we analyzed high-quality nucleosome occupancy maps obtained from wild-type and **Acf**^{7} embryos [27] and assessed global chromatin regularity by applying an autocorrelation function to the nucleosome dyad density patterns. Briefly, the function calculates the correlation between nucleosome dyad signals in an array of nucleosomes with a stepwise-shifted copy of itself. The calculated correlation coefficients for each shift (lag) are then plotted as a function of the shift (lag) length. Autocorrelation has previously been applied to score nucleosome repeat lengths [28] and promoter architecture [29]. Applied to nucleosome maps, the analysis reveals periodic oscillations, in which the amplitude and the decay rate provide information about regularity, whereas the maxima reveal the distance between adjacent nucleosomes. For **Acf**^{7} embryos, the analysis documents a genome-wide decay of autocorrelation amplitude together with a trend toward increased nucleosome repeat length (wild type = 188.4 ± 0.7 bp, **Acf**^{7} = 195.3 ± 1.5 bp) (Fig 5A). Evidently, loss of ACF1 globally affects the regularity and spacing of nucleosome arrays. To determine whether this global trend applied to the five chromatin states, the autocorrelation analysis was repeated for each of the chromatin domains. First, we found a dampening of the autocorrelation function upon ACF1 loss in the BLACK and BLUE inactive domains, which was not evident for the YELLOW active domains (Fig 5B–D). Second, we found a context-dependent increase in nucleosome repeat length in BLACK (wild type = 192.6 ± 0.5 bp, **Acf**^{7} = 206.0 ± 0.6 bp) and BLUE (wild type = 192.2 ± 0.6 bp, **Acf**^{7} = 201.0 ± 2.5 bp) but not in YELLOW (wild type = 182.0 ± 0.5 bp, **Acf**^{7} = 184.0 ± 1.2 bp) domains.

The correlation between the decay of physiological chromatin regularity and derepression of transcription in *Acf* mutant embryos suggests that the reduced stringency of DNA packaging in the absence of prominent spacing factors perturbs the repressed ground state of the genome installed by nucleosome arrays.

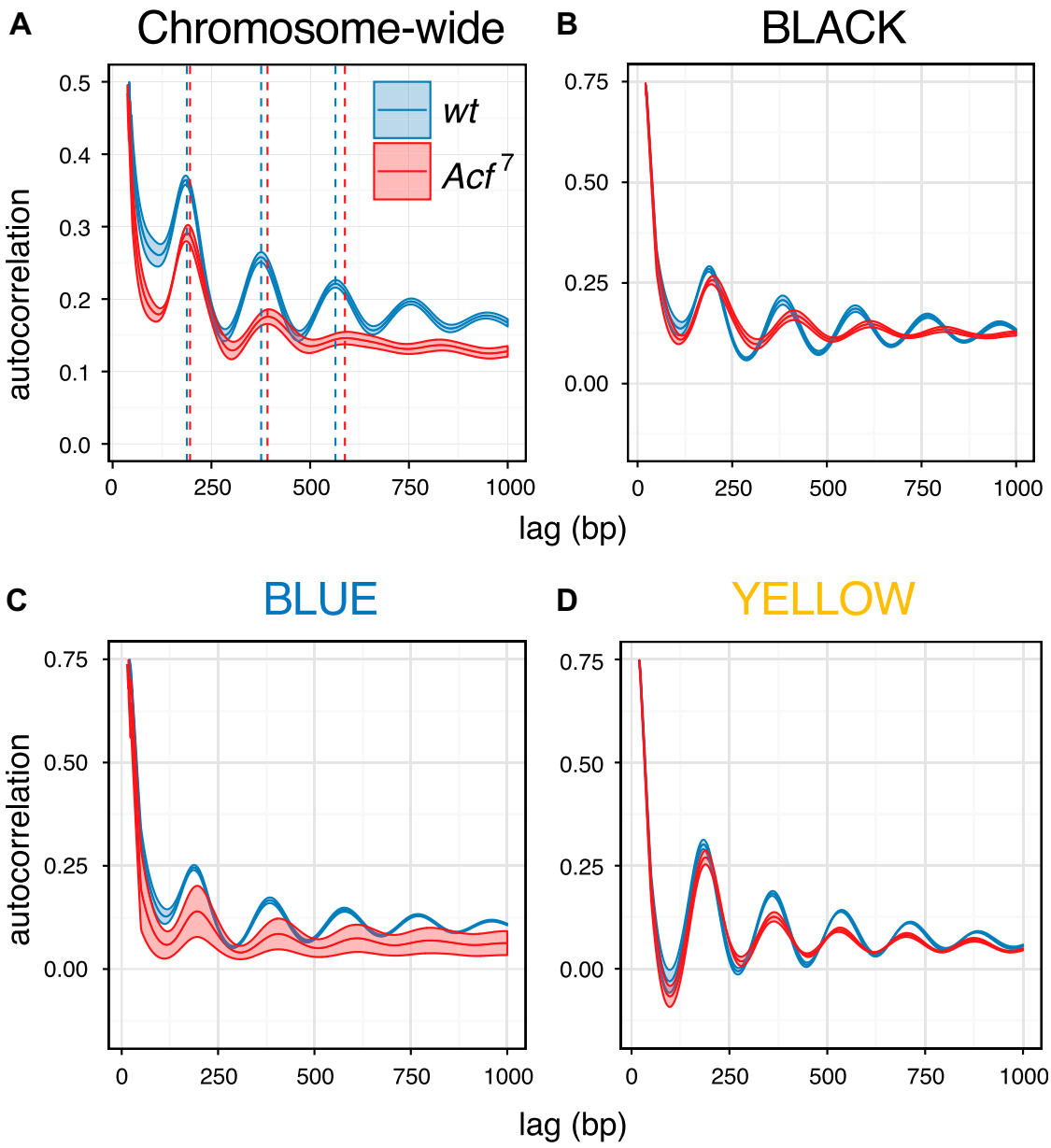


Figure 5

Global and context-dependent decrease in nucleosome regularity as a consequence of ACF1 loss.
(A) Changes in nucleosome periodicity on chromosomes 2 and 3 are estimated using the autocorrelation function. The correlation coefficients for the nucleosome occupancy values are plotted against the relative shifts (lag). The mean and SEM of replicate samples (n = 5 for wt and n = 3 for Acf7) are displayed. Dashed lines indicate the centers of nucleosome positions derived from the autocorrelation peaks. (B) Changes in nucleosome periodicity in BLACK chromatin domains estimated by the autocorrelation function. The mean and SEM of replicate samples are displayed. (C, D) Same as (B) but for BLUE and YELLOW chromatin domains, respectively.

Discussion

CHRAC was identified two decades ago following a biochemical activity that increased the accessibility of DNA in in vitro assembled chromatin [1]. Further characterization revealed that CHRAC did not destroy chromatin, but, to the contrary, improved the regularity of nucleosome fibers, identifying a first nucleosome spacing factor. This conundrum was resolved by the discovery that ISWI-containing remodeling factors catalyze nucleosome sliding [6]. ACF, which slides nucleosomes such as CHRAC, was originally purified searching for chromatin assembly factors [17,30]. The first genetic analyses of **Acf** deficiencies highlighted defects in pericentric heterochromatin and suppression of variegation, supporting the idea that both "higher order" chromatin structures and gene silencing rely on proper chromatin organization [18,19]. To date, however, a systematic assessment of the contribution of ACF1-containing remodelers to transcription has not been performed.

Our current study now clarifies this open issue. In our experimental design, we avoided several potential pitfalls. (i) We used clear **Acf** gene deletion with a clean null phenotype. Previous studies used **Acf**^{1} and **Acf**^{2} alleles that were later shown to yield oogenesis phenotypes that were clearly distinct from true loss-of-function phenotypes [21]. (ii) To assure that the transcriptome analysis was not flawed by a delay in the development of mutant embryos, we hand-selected mutant and wild-type embryos of matched age and determined their transcriptomes individually. (iii) We used two orthogonal approaches, each avoiding the technical or conceptual shortcomings of the other.

Although we could confirm a function of CHRAC/ACF in gene silencing, the extent of transcriptional repression scored in our tethering system was much stronger compared to the one in developing embryos. However, the consequences of the genetic deficiency may be masked by functional redundancy. For example, the ISWI-containing RSF remodeling complex [31,32] possesses similar nucleosome assembly and spacing activities as CHRAC/ACF. The targeting of ACF1 via an ectopic DNA-binding domain is expected to locally increase the ACF1 concentration around the tethering site, allowing effects to be scored above the background activities of endogenous factors. Regardless of magnitude, both types of experiments yielded highly complementary results.

The high-throughput targeting system we employed has previously been validated for HP1 [24], a well-known repressor, which provided an important benchmark. The repression induced by ACF1 recruitment was of the same order of magnitude as the effect of HP1 tethering determined in parallel. However, the repression mediated by targeted ACF1 was strongly modulated by the chromatin environment, with an obvious effect in overall inactive chromatin domains and lowly expressed genes. Nucleosome remodelers can work if

tethered [33], but given the dynamic nature of DNA interactions observed with most transcription factors, we think that the tethering rather increased the local concentration of the factor around the UAS^{Gal} site. In support of this notion, the overexpression of ACF1 lacking a DNA-binding domain had a similar effect, but milder. Importantly, the selective effect of ACF1 on poorly transcribed genes was similar, whether the ACF1 concentrations were increased globally or locally. This context dependence of CHRAC/ACF repression was confirmed by studying **Acf** deficiency in developing embryos.

Notwithstanding possible functional redundancies, we detected a major and global impact of ACF1 on physiological nucleosome regularity by applying an autocorrelation function to genome-wide nucleosome dyad maps. The impact of ACF1 depletion was more evident for inactive chromatin domains, establishing a clear correlation between the extent of physiological chromatin regularity and general repression, which we suggest is of a causal nature.

Various ISWI-type nucleosome sliding factors have very different functions. NURF [30,34], for example, is recruited by sequence-specific transcription factors to promoters of certain gene classes, where it serves as a coactivator [35]. CHRAC/ACF, in contrast, are most likely not targeted to promoters and enhancers [22]. Conceivably, these remodelers might establish the regularity of the nucleosome fiber in the context of replication [19,36] and/or DNA repair [37,38] and may exert a general “surveillance” function in search for gaps in the nucleosome array to be closed. We propose that their action establishes a repressive ground state of chromatin, rendering the genome inaccessible through optimal nucleosome packaging. Any further regulation, such as the specific activation of genes by recruitment of histone modifiers and more dedicated remodelers as well as the targeting of silencing machineries, happens on top of the general naive infrastructure provided by regular nucleosome arrays. In support of this idea, **Acf** depletion did not significantly affect the expression of transposable elements (unpublished observation), which are silenced through heterochromatinization. CHRAC/ACF and related factors are to be considered the caretakers of this genomic infrastructure. Their important and global role in generating a basal level of genome-wide repression can only be appreciated in regions that are devoid of all other, more potent, targeted and specific regulatory mechanisms.

Materials and Methods

Drosophila strains and genetics

The ACF1-GAL4 fusion constructs were generated by recombineering [39]. Briefly, a fosmid containing the genomic region of **Acf** (pflyfos021945) was recombined in **Escherichia coli** with a combinatorial tag cassette consisting of 2x-TY1-GAL4DBD(1-147)-3XFLAG to tag ACF1 either at its N-terminus or its C-terminus, or to entirely replace its coding sequence to serve as a control. Fosmids were inserted into attp40 (**yw**; attP40, locus 25C7, chr2L) (Genetic Services Inc.) to generate fly lines with ACF1 transgenes in chromosome 2L. The mosaic F0 generation was crossed with **w1118** and progeny flies from generation F1 onward were screened for dsRed phenotype (red eye fluorescence). Homozygous stocks were established by tracking eye fluorescence and the expression of ACF1 constructs was confirmed by Western blotting. ACF1-GAL4 transgenic flies were crossed to N1 flies (containing the **UAS-LacZ-mini-white** reporter [23]), to generate the final tethering system. The **Acf^{7}** allele had been described earlier [21]. It contains a deletion of most of the **Acf1** coding sequences and, to the best of our knowledge, corresponds to a loss-of-function phenotype.

Generation of the AcfC mutant allele

Predicted single guide RNA (sgRNA) targeting sequences for 5′ and 3′ ends of ACF1 were obtained from the Zhang lab CRISPR resource (a total eight for 5′ end and four for 3′ end) (http://crispr.mit.edu/). The 20-bp targeting sequences were inserted into the framework of primer-1 (5′-TAATACGACTCACTATAG-(targeting sequence)-GTTTTAGAGCTAGAAATAGC-3′) in the 5′ to 3′ direction. Using a scaffold primer (5′-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC-3′) and a universal reverse primer (5′-AAAAGCACCGACTCGGTGCC-3′), a final DNA was assembled for in vitro transcription by PCR. The PCR product was purified using the GeneElute PCR cleanup kit (Sigma, Cat. No. NA1020). In vitro transcription was performed using the T7 MEGAshortscript kit (Ambion, Cat. No. AM1354) and purified RNA was assessed by agarose gel electrophoresis. Efficiency of the RNA-mediated cleavage was assessed by transfecting 1 μg sgRNA to 7 × 10^{5}/ml SL2 cells (clone Hgr14 stably expressing Cas9) [40] in 2 ml final volume (24-well plate). Genomic DNA was prepared after 48 h. An ~600-bp region surrounding the selected guide RNA (gRNA) sequences was amplified, and the PCR product was melted at 95°C for 5 min and then cooled slowly at the ramp rate of 0.1°C/s [41]. gRNA cleavage frequently gives rise to mismatched base pairs around the cutting site, which were detected by T7 endonuclease (M0302S, NEB) cleavage and agarose gel electrophoresis. gRNA combinations that lead to T7 endonuclease cleavage were selected.

Genomic DNA 1.3 kb upstream and 1.5 kb downstream of gRNA sequences for ACF1 were amplified using a high-fidelity PCR system. These homology arms excluded the sgRNA sites. The homology arms and 3XP3-dsRed fly selection cassette (obtained from pJet1.2 [41]) were assembled together in a pBS donor backbone by using the Golden Gate cloning strategy. The final clone was validated by sequencing.

The purified plasmid and sgRNA for 5′ and 3′ ends of the **Acf** gene were co-injected into blastoderm embryos of **yw; Cas9; lig^{4169}** genotype [42]. The F0 mosaic males were crossed with **w1118** females and F1 transformants were screened for red fluorescence eye phenotype. The flies were backcrossed to the **yw** strain for four subsequent generations and rendered homozygous. Deletion of the locus was screened by PCR and loss of protein was assessed on Western blot. Final deletion of ACF1 encompasses around 4 Kb from the 562nd base onward, removing most of the gene except its 5′ and 3′ UTRs. No ACF1 protein could be detected in the newly generated **Acf^{C}** mutant (see the Results section), similarly to that observed in the previously analyzed **Acf^{7}** allele (generated by imprecise p-element excision) [21].

For hatching assays, 0- to 16-h embryos were collected on apple juice agar plates and allowed to develop for an additional 25 h at 25°C. Hatched larvae were counted.

Nuclei isolation and Western blot

For isolation of nuclei, embryos were collected overnight (0–16 h after egg laying [AEL]) onto apple juice agar plates and dechorionated in 25% bleach for 5 min. After extensive washes with PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4), the embryos were transferred to 1.5-ml tubes, resuspended in NB-0.3 (15 mM Tris–Cl, pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, pH 8, 0.3 M sucrose, 0.2 mM PMSF, 1 mM DTT, and Roche cOmplete protease inhibitor without EDTA), and homogenized using a metal pestle (LLG Labware, Cat. No. 9.314.501). The homogenate was collected and carefully layered on top of a biphasic solution consisting of NB-1.4 (15 mM Tris–Cl, pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, pH 8, and 1.4 M sucrose) and NB-0.8 M sucrose. After spinning at 13 krpm for 10 min (4°C), the nuclei pellet was collected and washed twice with NB-0.3 (spinning at 5,000 rpm for 5 min at 4°C in between washes). For Western blot analysis, the nuclei were suspended in 5× Laemmli sample buffer (250 mM Tris–HCl, pH 6.8, 10% w/v SDS, 50% v/v glycerol, 0.1% w/v bromophenol blue, and 10% β-mercaptoethanol) and boiled at 96°C for 8 min. The following antibodies were used for Western blots: αACF1 8E3 [22] (1:5), αFLAGm2 (1:1,000, Sigma, Cat. No. F1804), and α Lamin T40 (1:1,000, a kind gift from H. Saumweber).

ChIP-qPCR

For ChIP analysis, embryos were collected 0–12 h AEL and dechorionated in 25% bleach for 3 min. After extensive washes with water, the embryos were transferred to 15-ml tubes and weighted. Between 0.5 and 1 g of embryos were washed with 50 ml of PBS/0.01% Triton X-100 and then resuspended in 9 ml of fixing solution (50 mM Hepes, pH 7.6, 100 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA)/3.7% formaldehyde (Merck, Cat. No. 1040031000). Then 30 ml of n-heptane was added and the tubes were shaken for 1 min, followed by 13.5 min of incubation on a rotating wheel (18°C). The embryos were pelleted at 3,000 rpm for 1 min, resuspended in 50 ml of PBS/0.01% Triton X-100/125 mM glycine, and incubated at RT for 5 min. After two washes with PBS/0.01% Triton X-100, the embryos were frozen in liquid nitrogen and stored at −80°C until further processing. The frozen embryos were resuspended in 5 ml of RIPA buffer (10 mM Tris–Cl, pH 8, 1 mM EDTA, 140 mM NaCl, 1% Triton X-100, 0.1% SDS, and 0.1% sodium deoxycholate/1 mM DTT/0.2 mM PMSF/Roche cOmplete Protease inhibitor without EDTA) and dounced 10 times using a loose pestle and 10 times with a tight pestle. The homogenate was transferred to a 15-ml tube and spun at 170 g for 10 min at 4°C. The nuclei were resupended in 5 ml of RIPA/gram of embryos and split into 1-ml aliquots. Chromatin was sonicated using a Covaris S220 (100 W Peak Power, 20% Duty Factor, 200 Cycles/Burst, 15 min total time) and insoluble material was removed by centrifugation at 13.2 krpm for 20 min (4°C). Soluble chromatin was pre-cleared by adding RIPA-equilibrated 50% slurry of protein A+G (1:1) sepharose beads and rotating at 4°C for 1 h. Then 200 µl of chromatin was incubated overnight with 4 µl of the respective antibody: αACF1 Rb2 [22], αFLAGm2 (Sigma, Cat. No. F1804), and αISWI Rb1 (Becker Lab, unpublished). Then 30 µl protein A+G (1:1) 50% slurry was added and the tubes were rotated for 3 h at 4°C. After five washes with RIPA buffer, RNase-A was added (10 µg/100 µl; Sigma, Cat. No. R4875) and the tubes were incubated at 37°C for 20 min. Subsequent protease digestion (using 250 ng/µl Proteinase K; Genaxxon, Cat. No. M3036.0100) and cross-link reversal were performed simultaneously at 68°C for 2 h. DNA was purified using 1.8× Agencourt AMPure XP beads (Beckman Coulter, Cat. No. A63880) following a standard protocol and eluted in 50 µl of 5 mM Tris–Cl, pH 8. Purified DNA was used for standard qPCR analysis at 1:2 dilution. The primers are listed in Table S1.

RT–qPCR

For **LacZ** expression analysis, embryos were collected 2–8 h AEL and dechorionated in 25% bleach for 3 min. After extensive washes with PBS, the embryos were transferred into a 1.5-ml tube, resuspended in 300 µl of QIAzol (QIAgen, Cat. No. 79306), and homogenized using a metal pestle. After addition of 700 µl of QIAzol, the samples were snap-frozen in liquid

nitrogen and stored at −80°C until further processing. RNA was extracted using the standard protocol provided by QIAgen. The Superscript III First Strand Synthesis System (Invitrogen, Cat. No. 18080051, random hexamer priming) was used to generate cDNA starting from 1.5 µg of total RNA. cDNA was used for standard qPCR analysis at 1:10 dilution. The primers are listed in Table S1.

Immunofluorescence microscopy

For immunofluorescence of *Drosophila* Kc167 cells, 200 µl of cells (>10^6 cell/ml) were transferred onto polylysine–coated three-well depression slides (Thermo Scientific, Cat. No. 631-0453) and incubated for 1.5 h at 26°C. The cells were washed with PBS and fixed in PBS/3.7% formaldehyde for 10 min. After two washes with PBS, the cells were permeabilized in ice-cold PBS/0.25% Triton X-100 for 6 min. The cells were washed twice with PBS and blocked with PBS/0.1% Triton X-100/5% normal donkey serum (Jackson Immuno Research)/5% nonfat milk for 2 h. After a brief wash with PBS, the cells were incubated overnight at RT with primary antibodies αV5 (1:1,000, GenScript, Cat. No. A00623) and αmCherry [43] (1:20). The cells were washed twice with PBS/0.1% Triton X-100 and incubated with secondary antibodies donkey-αrat-Cy3 (1:500, Jackson Immuno Research) and donkey-αrabbit-Alexa488 (1:300, Jackson Immuno Research) for 2 h at RT. The cells were washed twice with PBS/0.1% Triton X-100, incubated with 1:500 DAPI for 10 min at RT, and washed again with PBS. Coverslips were mounted using Vectashield mounting medium (Vector Laboratories, Cat. No. H-1000) and sealed with nail polish. Pictures were acquired on a Leica Sp5 confocal microscope using the same settings for all the constructs.

Artificial tethering of ACF1 to multiple reporters in** Drosophila **cells

Kc167 cells containing the barcoded reporter library were generated as previously described [24]. Plasmids for the expression of GAL4-ACF1 fusion and controls were derived from pAc5-Gal4-V5-HP1a-T2A-mCherry [24] by Gibson assembly. All constructs were validated using DNA sequencing and restriction digestion analysis. The artificial tethering, including sample preparation, sequencing, and data processing/analysis, was performed as described in reference [24], including GAL4-HP1 construct as a positive control. Two biological replicates for each condition were analyzed. Reporters with normalized counts equal to zero in at least one condition were discarded. Linear models were calculated using the **lm** function in R. For transient expression of Gal4-ACF1 fusion and controls, 3 × 10^6 Kc167 cells were transfected with 1 µg of the corresponding plasmid using the X-tremeGENE HP transfection reagent (Sigma, Cat. No. 6366236001) following the standard

protocol (4.5:1 transfection reagent:DNA ratio). 3 d after transfection, 1 ml cells was collected, spun at 800 g for 5 min, resuspended in 20 µl of 5× Laemmli sample buffer per 10^6 cells, and boiled at 95°C for 10 min.

Single-embryo RNA-seq

Before RNA-seq, the **Acf^c** flies were backcrossed with the wild-type **OrR** strain for eight generations. Embryos were collected 0–45 min AEL and allowed to develop at 25°C until approximately 30 min before the desired stage (around 1 h for Bownes Stage 3 and 4 h for Bownes Stage 8). Without prior dechorionation, the embryos were hand-picked and submerged into a drop of Voltalef 10 S halocarbon oil (Lehman and Voss Co.) placed on a microscope slide. After about 5 min, the embryonic structures became visible under the stereomicroscope. Embryos were allowed to develop further under the halocarbon oil until the desired stage. Single embryos were picked and crushed with a 26-G needle into 200 µl of lysis buffer (supplemented with Proteinase K) from the Agencourt RNAdvance Tissue kit (Beckman Coulter, Cat. No. A32645). After the addition of 10 µl of 1:100 ERCC Spike-in RNA mix (Ambion, Cat. No. 4456740), the samples were incubated at 37°C for 20 min, snap-frozen in liquid nitrogen, and stored at −80°C. Total RNA was extracted from the single-embryo homogenate using the same Agencourt RNAdvance Tissue kit, following standard protocol but using half of the volumes recommended. RNA integrity was checked on a Bioanalyzer 2100 instrument (Agilent). Ribosomal RNA depletion was achieved using an rRNA depletion kit (human/mouse/rat) (New Englands Biolab, Cat. No. E6310) and the rRNA-depleted RNA was stored at −80°C until further processing. Nondirectional libraries were prepared using an NEBnext Ultra RNA Library Prep kit for Illumina (New Englands Biolab, Cat. No. E7530S) following standard protocol. Six replicates per genotype and stage were sequenced on an Illumina HiSeq1500 instrument. Paired-end RNA-seq reads were mapped against the reference genome (FB2016_01 dmel_r6.09 with selected chromosomes) using STAR (version 2.5.0a) with **quantMode GeneCounts** for counting reads per gene [44]. One replicate from the **Acf^c** genotype (Stage 3) was excluded because of improper staging (data not depicted). Size factors for normalization were calculated by DESeq2 [45]. PCA was carried out on selected genes with variance across samples between the 85th and 99th percentiles. Genes with a read count equal to zero in at least half of the samples were filtered out for further analysis. Differential expression (DESeq2) analysis (mutant versus wild type) was carried out by fitting negative binomial GLM independently for the two developmental stages [45]. Cutoffs for adjusted **P**-values were defined at the 0.1 level. Gene ontology analysis on significantly different genes was performed on the FlyMine online database [46]. Genes were assigned to five-state chromatin domains [25] by the **nearest** method

from the GenomicRanges Bioconductor packages. Trends on MA plot were visualized by local polynomial regression fitting (**loess**). modENCODE histone modification signals (smoothed M-values) [26] were averaged over genes and low/high levels were distinguished by a cutoff based on the local minimum in the density of the H3K36me3 levels. Genes were classified as marked/unmarked based on whether they carry high/low histone modification levels in all four marks investigated in the analysis.

Nucleosome mapping and autocorrelation

For mapping nucleosomes, embryos were collected 2–8 h AEL. The embryos (between 0.2 and 0.5 g per replicate and genotype) were dechorionated and fixed as described in the ChIP-qPCR section.

For nuclei isolation, the embryos were slowly thawed and dounced using a glass homogenizer (Schubert, Cat. No. 9164693) with 20 strokes each of the A and B pestles in ice-cold NX-I buffer (15 mM Hepes, pH 7.6, 10 mM KCl, 2 mM $MgCl_2$, 0.5 mM EGTA, 0.1 mM EDTA, 350 mM sucrose, 1 mM DTT, 0.2 mM PMSF, and Roche cOmplete Protease inhibitor without EDTA). Nuclei were subsequently pelleted at 3,500 rpm for 10 min at 4°C. For MNase digestion, the nuclei were suspended in the RIPA buffer supplemented with 2 mM $CaCl_2$. The nuclei were digested with 13 units of MNase per gram of starting embryos (Sigma, Cat. No. N5386) for 15 min at 37°C while shaking at 500 rpm. The reaction was stopped by adding 0.5 M EDTA (pH 8.0) to a final concentration of 10 mM and the tubes were quickly transferred to ice for 5 min. The nuclei were spun at 12.5 krpm for 10 min at 4°C. The supernatants containing most of the DNA were collected and the residual RNA was digested by RNase-A (50 µg/ml; Sigma, Cat. No. R4875) at 37°C for 30 min. Protein digestion and cross-linking reversal were performed as previously described in the ChIP-qPCR section. DNA was purified using 1.8× Agencourt AMPure XP beads (Beckman Coulter, Cat. No. A63880) following standard protocol and eluted in 50 µl of 5 mM Tris–Cl, pH 8. Recovered DNA was quantified using the Qubit dsDNA HS assay kit (Life Technologies, Cat. No. Q32851) and sequencing libraries were prepared using a custom-made protocol available upon request. Libraries were sequenced on a HiSeq 1500 (Illumina) instrument.

Paired-end reads were mapped to *Drosophila* genome version dm6. We used Bowtie v1.1.1 with "-X 750" parameter setting. Dyad coverage vectors were obtained by size-selecting fragments of length >120 and <200 bp and resizing their length to 50 bp fixed at the fragment center.

The nucleosome dyad maps used for autocorrelation were generated, validated, and interpreted by Jain et al [27]. The autocorrelation function was calculated for the dyad

coverage vectors obtained for the entire genome and for the five-state domains described by Filion et al [25]. The vectors for the last cases represent head-to-tail concatemerized regions of given annotation. The function was run for the lag length of 1,000 bp. Nucleosomal repeat lengths were obtained by linear regression of the first and second autocorrelation peak positions with zero intercept. The slope of the regression was defined as repeat length. Values reported in the text correspond to average nucleosomal repeat length (between biological replicates) ± SEM.

Accession codes

Sequencing data have been deposited in the Gene Expression Omnibus under accession numbers GSE106759 (artificial tethering of ACF1 in Kc167 cells) and GSE106733 (nucleosome maps and single-embryo RNA-seq).

Author Contributions

A Scacchetti: Conceptualization, formal analysis, investigation, methodology, writing—original draft, review, and editing.

L Brueckner: Formal analysis, investigation, methodology, writing—review and editing. D Jain: Formal analysis, investigation, methodology, writing—review and editing.

T Schauer: Data curation, formal analysis, methodology, writing—review and editing.

X Zhang: Methodology, writing—review and editing.

F Schnorrer: Supervision, funding acquisition, writing—review and editing.

B van Steensel: Supervision, funding acquisition, writing—review and editing.

T Straub: Data curation, formal analysis, methodology, writing—review and editing.

PB Becker: Conceptualization, supervision, funding acquisition, writing—original draft, project administration, writing—review and editing.

Acknowledgements

We thank K Förstemann for sharing reagents and cell lines for gRNA testing. We thank S Krebs and H Blum and the Netherlands Cancer Institute Genomics Core Facility for outstanding sequencing service. PB Becker and A Scacchetti are funded by the European Research Council (ERC), MSCA-ITN-2014-ETN No. 642934. F Schnorrer is funded by the ERC [(FP/2007–2013); grant 310939], the Centre Nationnal de la Recherche Scientifique, the excellence initiative Aix-Marseille University AMIDEX, the Agence Nationale de la Recherche, and the LabEX-INFORM. B van Steensel is supported by National Institutes of Health grant U54 DK107965.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

1. Varga-Weisz PD, Wilm M, Bonte E, Dumas K, Mann M, Becker P. Chromatin-remodeling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature*. 1997;388:598-602. doi:10.1038/41587
2. Ito T, Levenstein ME, Fyodorov DV, Kutach AK, Kobayashi R, Kadonaga JT. ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev*. 13:1529-1539. doi:10.1101/gad.13.12.1529
3. Corona DFV, Eberharter A, Budde A, et al. Two histone fold proteins, CHRAC-14 and CHRAC-16, are developmentally regulated subunits of chromatin accessibility complex (CHRAC. *EMBO J*. 2000;19:3049-3059. doi:10.1093/emboj/19.12.3049
4. Hartlepp KF, Fernandez-Tornero C, Eberharter A, Grune T, Muller CW, Becker PB. The histone fold subunits of Drosophila CHRAC facilitate nucleosome sliding through dynamic DNA interactions. *MolCell Biol*. 2005;25:9886-9896. doi:10.1128/mcb.25.22.9886-9896.2005
5. Becker PB, Workman JL. Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol*. 2013;5. doi:10.1101/cshperspect.a017905
6. Längst G, Bonte EJ, Corona DFV, Becker PB. Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell*. 1999;97:843-852. doi:10.1016/s0092-8674(00)80797-7
7. Eberharter A, Ferrari S, Längst G, et al. Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *EMBO J*. 2001;20:3781-3788. doi:10.1093/emboj/20.14.3781
8. He X, Fan HY, Narlikar GJ, Kingston RE. Human ACF1 alters the remodeling strategy of SNF2h. *J Biol Chem*. 2006;281:28636-28647. doi:10.1074/jbc.m603008200
9. Blosser TR, Yang JG, Stone MD, Narlikar GJ, Zhuang X. Dynamics of nucleosome remodelling by individual ACF complexes. *Nature*. 2009;462:1022-1027. doi:10.1038/nature08627
10. Leonard JD, Narlikar GJ. A nucleotide-driven switch regulates flanking DNA length sensing by a dimeric chromatin remodeler. *Mol Cell*. 2015;57:850-859. doi:10.1016/j.molcel.2015.01.008
11. Mueller-Planitz F, Klinker H, Becker PB. Nucleosome sliding mechanisms: new twists in a looped history. *Nat Struct Mol Biol*. 2013;20:1026-1032. doi:10.1038/nsmb.2648
12. Racki LR, Yang JG, Naber N, et al. The chromatin remodeller ACF acts asa dimeric motor to space nucleosomes. *Nature*. 2009;462. doi:10.1038/nature08621

13. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol*. 2017;18:407-422. doi:10.1038/nrm.2017.26
14. Korber P, Becker PB. Nucleosome dynamics and epigenetic stability. *Essays Biochem*. 2010;48. doi:10.1042/bse0480063
15. Fazzio TG, Tsukiyama T. Chromatin remodeling in vivo: Evidence for a nucleosome sliding mechanism. *Mol Cell*. 2003;12:1333-1340. doi:10.1016/s1097-2765(03)00436-2
16. Whitehouse I, Rando OJ, Delrow J, Tsukiyama T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature*. 2007;450:1031-1035. doi:10.1038/nature06391
17. Ito T, Michael B, Pazin MJ, Kobayashi R, Kadonaga JT. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell*. 1997;90:145-155. doi:10.1016/s0092-8674(00)80321-9
18. Fyodorov DV, Blower MD, Karpen GH, Kadonaga JT. Acf1 confers unique activities to ACF/CHRAC and promotes the formation rather than disruption of chromatin in vivo. *Genes Dev*. 2004;10(1101):170-183. doi:10.1101/gad.1139604
19. Chioda M, Vengadasalam S, Kremmer E, Eberharter A, Becker PB. Developmental role for ACF1-containing nucleosome remodellers in chromatin organisation. *Development*. 2010;137:3513-3522. doi:10.1242/dev.048405
20. Liu YI, Chang MV, Li HE, et al. The chromatin remodelers ISWI and ACF1 directly repress Wingless transcriptional targets. *Dev Biol*. 323:41-52. doi:10.1016/j.ydbio.2008.08.011
21. Boerner K, Jain D, Vazquez-Pianzola P, et al. A role for tuned levels of nucleosome remodeler subunit ACF1 during Drosophila oogenesis. *Dev Biol*. 2016;411:217-230. doi:10.1016/j.ydbio.2016.01.039
22. Jain D, Baldi S, Zabel A, Straub T, Becker PB. Active promoters give rise to false positive Phantom Peaks in ChIP-seq experiments. *NucleicAcids Res*. 2015;43:6959-6968. doi:10.1093/nar/gkv637
23. Prestel M, Feller C, Straub T, Mitlohner H, Becker PB. The activation potential of MOF is constrained for dosage compensation. *Mol Cell*. 2010;38:815-826. doi:10.1016/j.molcel.2010.05.022
24. Brueckner L, van Arensbergen J, Akhtar W, Pagie L, van Steensel B. High-throughput assessment of context-dependent effects of chromatin proteins. *Epigenetics Chromatin*. 2016;9(1):43. doi:10.1186/s13072-016-0096-y

25. Filion GJ, van Bemmel JG, Braunschweig U, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*. 2010;143(2):212-224. doi:10.1016/j.cell.2010.09.009

26. Celniker SE, Dillon LA, Gerstein MB, et al. Unlocking the secrets of the genome. *Nature*. 2009;459:927-930. doi:10.1038/459927a

27. Jain D, Baldi S, Zabel A, Straub T, Becker PB. *Genome-Wide Analysis of Phased Nucleosomal Arrays Revals the Functional Characteristic of the Nucleosome Remodeler ACF*.; 2016. bioRxiv.doi:10.1101/093666

28. Braunschweig U, Hogan GJ, Pagie L, Steensel B. Histone H1 binding is inhibited by histone variant H3.3. *EMBO J*. 2009;28. doi:10.1038/emboj.2009.301

29. Wan J, Lin J, Zack DJ, Qian J. Relating periodicity of nucleosome organization and gene regulation. *Bioinformatics*. 2009;25. doi:10.1093/bioinformatics/btp323

30. Hamiche A, Sandaltzopoulo R, Gdula DA, Wu C. ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell*. 1999;97:833-842. doi:10.1016/s0092-8674(00)80796-5

31. LeRoy G, Orphanides G, Lane W, Reinberg D. Requirement of RSF and FACT for transcription of chromatin templates in vitro. *Science*. 1998;282:1900-1904. doi:10.1126/science.282.5395.1900

32. Hanai K, Furuhashi H, Yamamoto T, Akasaka K, Hirose S. RSF governs silent chromatin formation via histone H2Av replacement. *PLoSGenet*. 2008;4:1000011. doi:10.1371/journal.pgen.1000011

33. McKnight JN, Jenkins KR, Nodelman IM, Escobar T, Bowman GD. Extranucleosomal DNA binding directs nucleosome sliding by Chd1. *MolCell Biol*. 2011;31:4746-4759. doi:10.1128/mcb.05735-11

34. Badenhorst P, Voas M, Rebay I, Wu C. *Biological Functions of TheISWI Chromatin Remodeling Complex NURF.Genes Dev16: 3186–3198.Doi:10.1101/Gad.1032202*.; 2002.

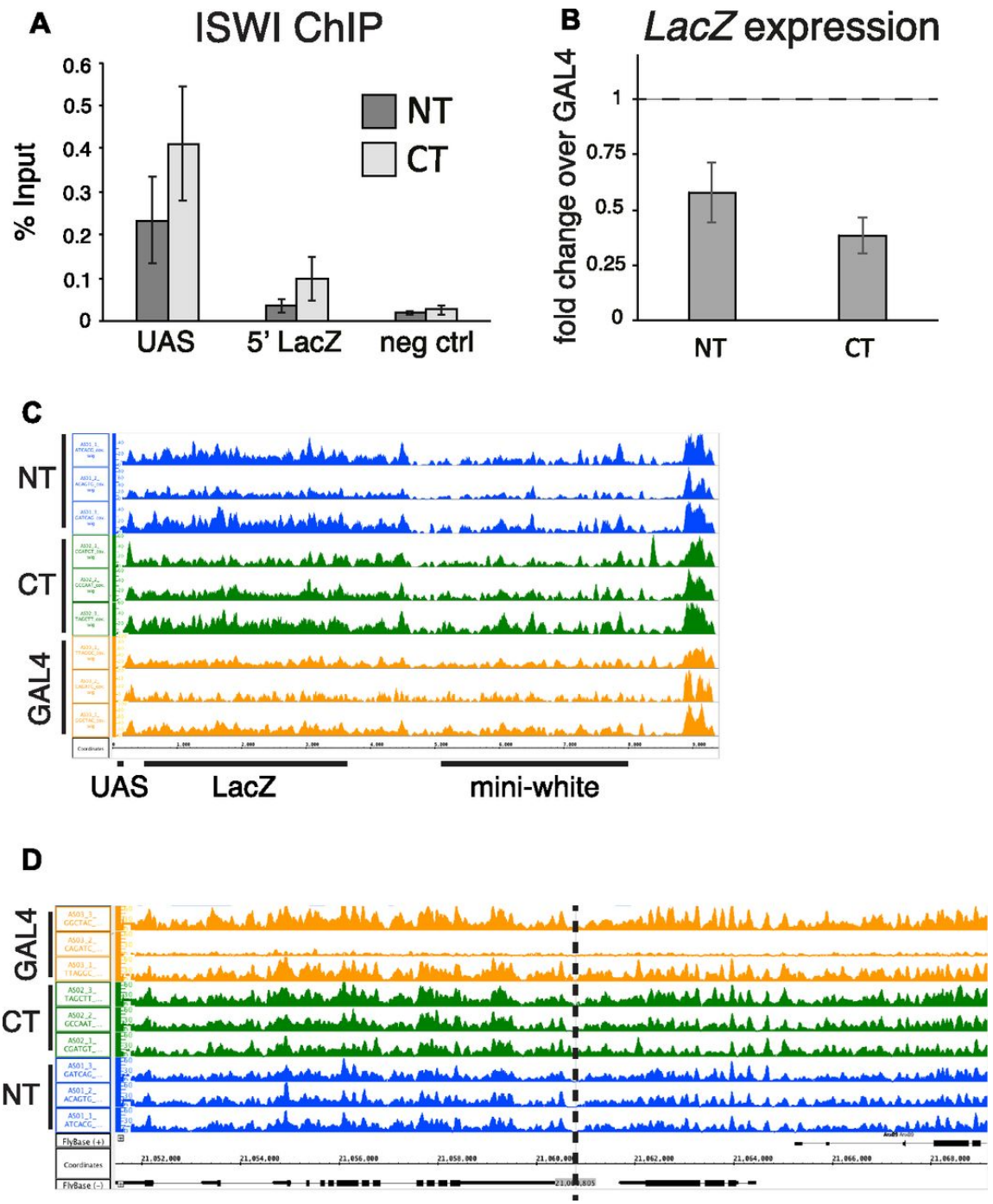
35. Kwon SY, Grisan V, Jang B, Herbert J, Badenhorst P. Genome-wide mapping targets of the metazoan chromatin remodeling factor NURF reveals nucleosome remodeling at enhancers, Core promoters and gene insulators. *PLoS Genet*. 2016;12:1005969. doi:10.1371/journal.pgen.1005969

36. Collins N, Poot RA, Kukimoto I, Garcia-Jimenez C, Dellaire G, Varga-Weisz, P.D. An ACF1-ISWI chromatin-remodeling complex is required forDNA replication through heterochromatin. *Nat Genet*. 2002;32. doi:10.1038/ng1046

37. Sanchez-Molina S, Mortusewicz O, Bieber B, et al. Role for hACF1 in the G2/M damage checkpoint. *Nucleic Acids Res*. 2011;39:8445-8456. doi:10.1093/nar/gkr435

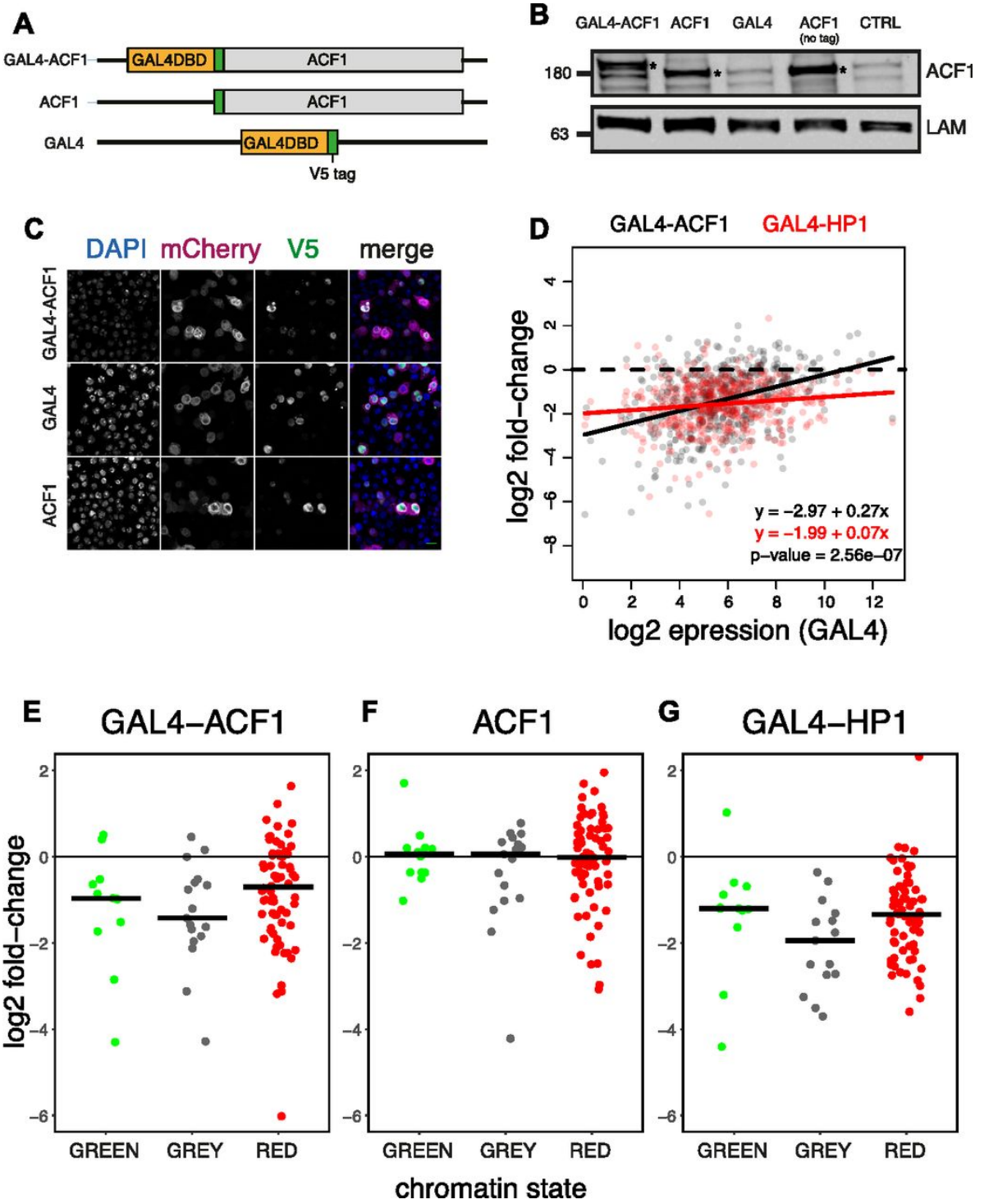
38. Lan L, Ui A, Nakajima S, et al. The ACF1 complex is required for DNA double-strand break repair in human cells. *Mol Cell*. 2010;40. doi:10.1016/j.molcel.2010.12.003
39. Ejsmont RK, Sarov M, Winkler S, Lipinski KA, Tomancak P. A toolkit for high-throughput, cross-species gene engineering in Drosophila. *NatMethods*. 2009;6:435-437. doi:10.1038/nmeth.1334
40. Bottcher R, Hollmann M, Merk K, et al. Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. *Nucleic Acids Res*. 2014;42:e89. doi:10.1093/nar/gku289
41. Zhang X, Ferreira IR, Schnorrer F. A simple TALEN-based protocol for efficient genome-editing in Drosophila. *Methods*. 2014;69. doi:10.1016/j.ymeth.2014.03.020
42. Zhang X, Koolhaas WH, Schnorrer F. *A versatile two-step CRISPR-and RMCE-based strategy for efficient genome engineering in Drosophila*. Vol G3.; 2014. doi:10.1534/g3.114.013979
43. Rottach A, Kremmer E, Nowak D, Leonhardt H, Cardoso MC. Generation and characterization of a rat monoclonal antibody specific for multiple red fluorescent proteins. *Hybridoma*. 2008;27. doi:10.1089/hyb.2008.0031
44. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15-21. doi:10.1093/bioinformatics/bts635
45. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15. doi:10.1186/s13059-014-0550-8
46. Lyne R, Smith R, Rutherford K, et al. FlyMine: an integrated database for Drosophila and Anopheles genomics. *Genome Biol*. 2007;8:129. doi:10.1186/gb-2007-8-7-r129

Supplementary Data



Supplementary Figure S 1

(A) ChIP-qPCR to verify the recruitment of ISWI to the UAS sites. "UAS" and "5′ LacZ" denote the regions amplified by qPCR. Bars denote average % input enrichment (n = 3 biological replicates) ± SEM. "neg ctrl" represents a negative control locus (encompassing the Spt4 gene). (B) RT–qPCR assesses expression changes of the LacZ gene in early embryos (2–8 h) upon ACF1 tethering. Bars denote average fold-change over the control (GAL4DBD alone) (n = 3 biological replicates) ± SEM. (C) Screenshot of the genome browser representing the LacZ/mini-white reporter locus. Each track shows nucleosome dyad densities for each replicate of the corresponding genotype (as described in Fig 1A). (D) Same as (C) but depicting the genomic region surrounding the integration site of the LacZ reporter transgene (dashed line).



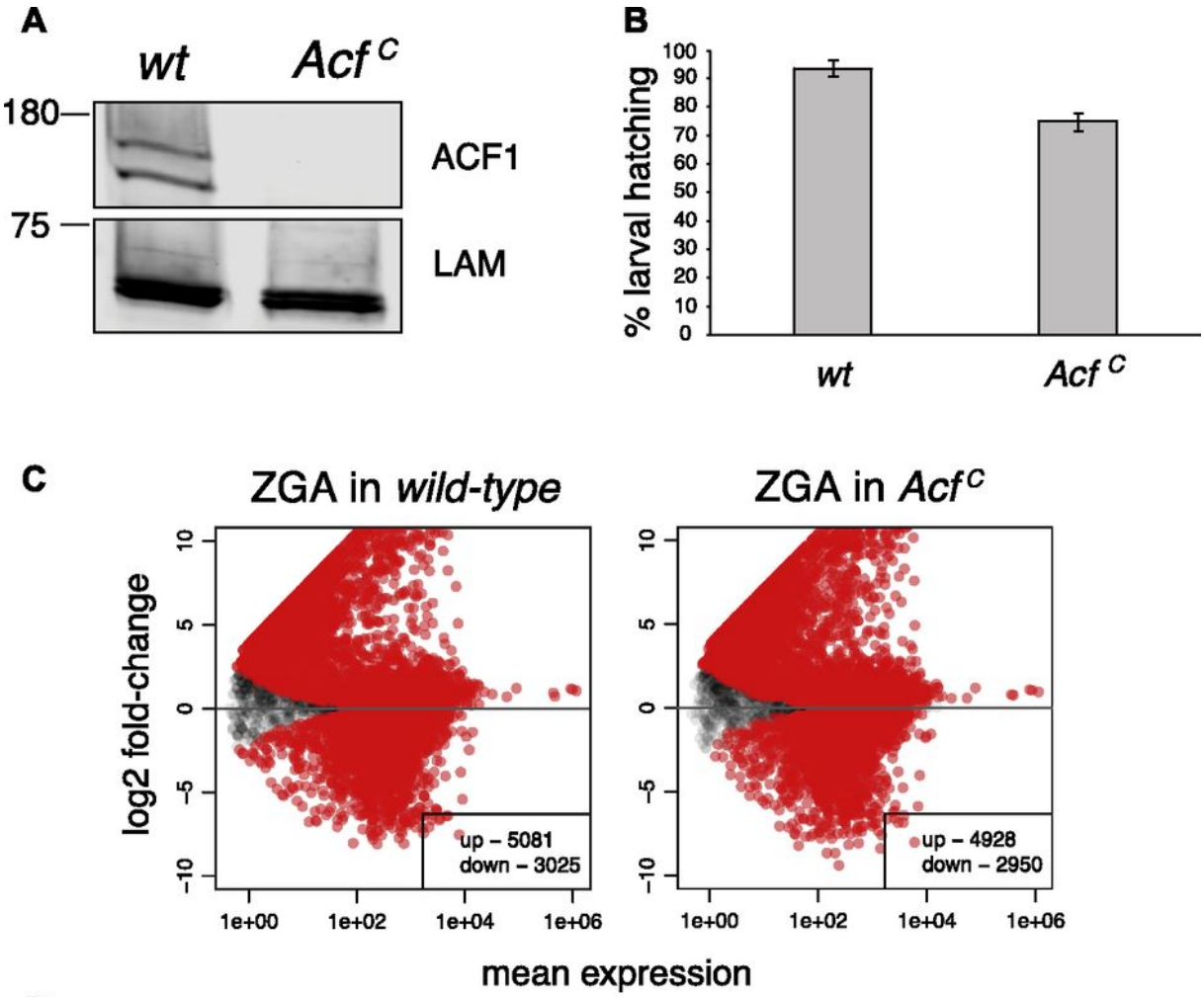
Supplementary Figure S 2

(A) Schematic illustration of the transgenes used for testing the effects of ACF1 recruitment to multiple reporters in Kc167 cells. ACF1 is fused to the GAL4 DNA-binding domain (GAL4 DBD) at the N-terminus. A transgene containing the GAL4 DBD alone is used as a control. All constructs contain a V5 tag. (B) Western blot of whole cell extracts from transiently transfected Kc167 cells. ACF1 antibody was used for the detection of ACF1 transgenes and Lamin (LAM) was used for the loading control. The first three lanes represent ACF1 targeting construct and controls; all of them possess a V5 tag. “ACF1 (no tag)” is a construct that expresses ACF1 without the V5 tag and “CTRL” lane refers to cells transfected with water. Asterisks indicate expressed transgenic products. (C) Immunofluorescence microscopy of the cells transiently transfected as in (A). The cells were stained with DAPI and V5 antibody for the detection of the constructs. mCherry antibody was used as a control. Green horizontal bar indicates the scale (10 μm). (D) Comparison of the correlation between repression and reporter expression upon ACF1 (GAL4-ACF1, black) or HP1 tethering (GAL4-HP1,

red). Lines represent linear regression. Equations describe the fitted lines. P-value refers to the significance of the difference in the slope between the two regression lines. (E) Jitter plots represent the distribution of log2 fold-changes of reporters integrated in GREEN (N = 12), GRAY (N = 17), and RED (N = 70) chromatin domains for the case of tethered ACF1 (GAL4-ACF1). Black horizontal bars represent median. (F) Same as (E) but for ACF1 overexpression (ACF1). (G) Same as (E) but for HP1 tethering (GAL4-HP1).

Supplementary Figure S 3

(A) Western blot probing nuclear extract from wild-type (wt) and the newly generated ACF1 mutant (AcfC) embryos (0–16 h). No ACF1 was detected in mutant tissue using the antibody against ACF1. Lamin (LAM) serves as a loading control. (B) Comparison of larval hatching rate between wild-type (wt) and AcfC. Bars represent average % larval hatching (n = 3 biological replicates of N > 47 embryos each) ± SEM. (C) Differential gene expression analysis of coding genes from RNA-seq data. Wild-type (left panel) and AcfC (right panel) transcriptomes were separately analyzed to assess gene expression changes during maternal to zygotic transition. Scatter plots represent log2 fold-change for each gene between the two different developmental stages in relation to its mean expression (mean of normalized counts) (N = 9,235 for wild type, N = 9,163 for AcfC). Red dots represent statistically significant (q-value < 0.1) up- or down-regulated genes. (D) Gene ontology analysis of statistically significant (q-value < 0.1) up- or down-regulated genes in AcfC compared to wild-type in the two developmental stages. Top five gene ontology terms are displayed.



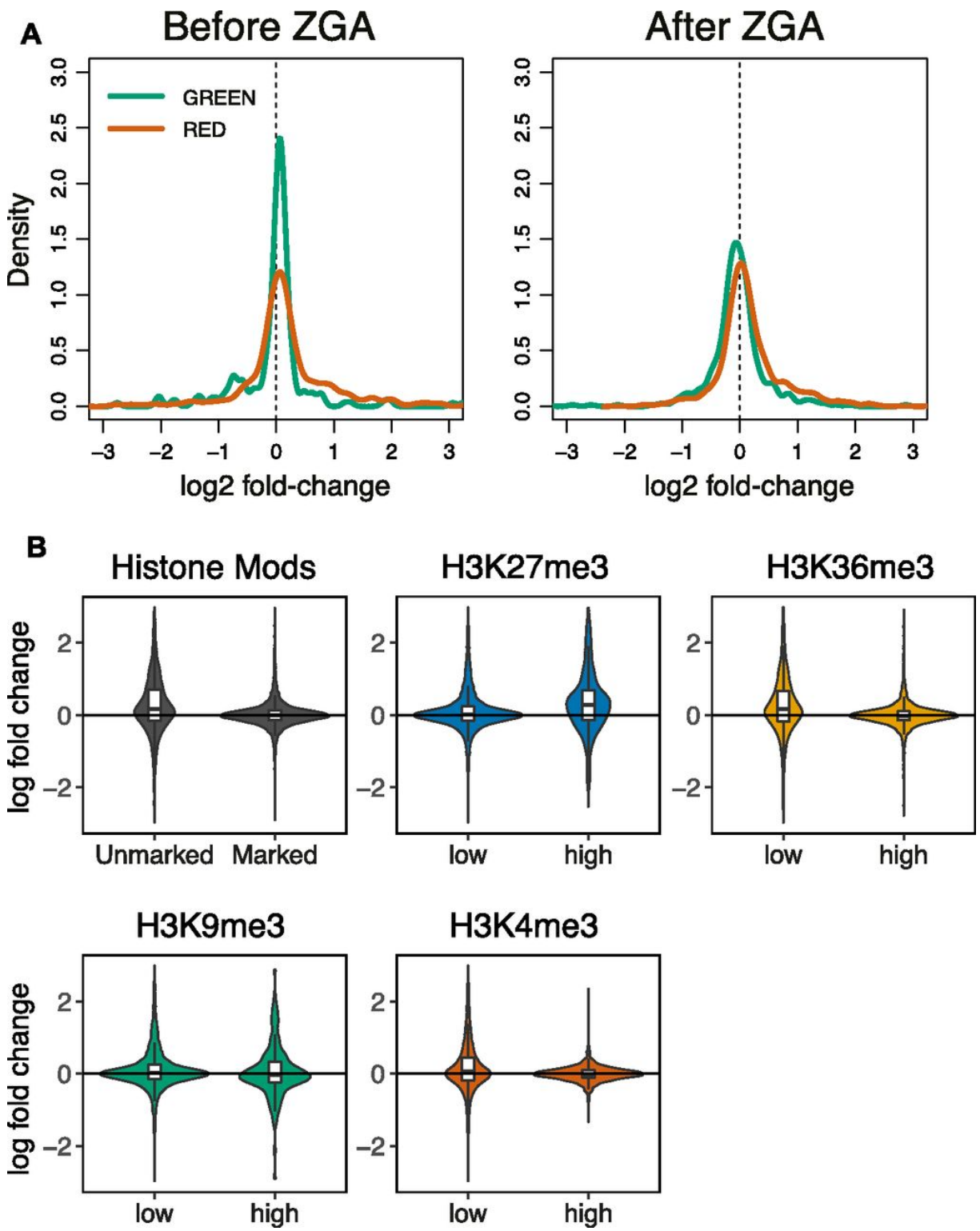
D

Before ZGA

GO term	p-value	GO
organic cyclic compound metabolic process	2.90E-10	GO:1901360
cellular aromatic compound metabolic process	4.38E-10	GO:0006725
heterocycle metabolic process	2.16E-09	GO:0046483
cellular process	5.83E-09	GO:0009987
cellular developmental process	6.68E-09	GO:0048869

After ZGA

GO term	p-value	GO
cell differentiation	2.94E-18	GO:0030154
cellular developmental process	3.72E-18	GO:0048869
system development	8.49E-17	GO:0048731
developmental process	1.60E-13	GO:0032502
single-organism developmental process	4.60E-13	GO:0044767



Supplementary Figure S 4

(A) Comparison of wild-type and AcfC transcriptomes in relation to the chromatin five-state model. Plots represent the distribution of log2 fold-changes for genes belonging to the GREEN and RED chromatin domains before or after ZGA. (B) Violin plots represent the distribution of log2 fold-changes for genes belonging to genomic regions marked or unmarked by the histone marks analyzed (top left panel, Histone Mods), or to genomic regions with high or low occupancy of H3K27me3, H3K36me3, H3K9me3, and H3K4me3 (remaining panels). Boxplots are overlapped to show median values.

Table S 1 List of primers used in this study.

Primer	Sequence	Usage
Acf_Rec_NF	GCCATTTAAACTTAAGGACATTCAAGAGCAAAAGGAAAACACCAAACATGGAAGTGCATACCAATCAGGAC	recombineering
Acf_Rec_NR	TTCTTGCCCTCCTTCTGATTCAGGTCGAATCCTTCCCGCTTGCAAATGGGCTTGTCGTCGTCATCCTTGTA	recombineering
Acf_Rec_CF	TGCAGCTACCGTTTAGGCCTAGCGATATGAACGGGGAAGTCAAAGCTTGCGAAGTGCATACCAATCAGGAC	recombineering
Acf_Rec_CR	GTAGACTAACTAATTTACACGATAGCTGGTGGAGATCAGCGTCCGGCTCACTTGTCGTCGTCATCCTTGTA	recombineering
UAS_ChIP_f1	TCCGAGCGGAGACTCTAGC	ChIP-qPCR, UAS
UAS_ChIP_r1	TTTGCTTGTTTGAATTGAATTGTC	ChIP-qPCR, UAS
UAS_ChIP_f2	TAGGGAATTGGGAATTCGAG	ChIP-qPCR, 5' of LacZ
UAS_ChIP_r2	TTAGCAGGCTCTTTCGATCC	ChIP-qPCR, 5' of LacZ
Spt4_fw	GCTCCGATTCATAAGCCCAG	ChIP-qPCR, negative control region
Spt4_rv	GCCTCTTTCGGAGCAGCTTT	ChIP-qPCR, negative control region
LacZ_RT_f1	CGCTAGAGTCGACCAATTCC	RT-qPCR
LacZ_RT_r1	GGCAACGAAAATCACGTTCT	RT-qPCR
RpII-140_fw	GCGCTATGGGTAAGCAAGCT	RT-qPCR, control
RpII-140_rv	TCACAAGTGGCTTCATCGGA	RT-qPCR, control
Acf1_crispr_LHS6	TTAGCGCAGCGTCCACCTAA	CRISPR/Cas9, gRNA
Acf1_crispr_RHS2	GGGCGCCTACCAGCTAAACG	CRISPR/Cas9, gRNA

Chapter 4

Local rewiring of genome–nuclear lamina interactions by transcription

Laura Brueckner, Peiyao A Zhao, Tom van Schaik, Christ Leemans, Jiao Sima, Daniel Peric-Hupkes, David M Gilbert, Bas van Steensel

EMBO Journal **39,** 6 (2020)

Abstract

Transcriptionally inactive genes are often positioned at the nuclear lamina (NL), as part of large lamina-associated domains (LADs). Activation of such genes is often accompanied by repositioning toward the nuclear interior. How this process works and how it impacts flanking chromosomal regions are poorly understood. We addressed these questions by systematic activation or inactivation of individual genes, followed by detailed genome-wide analysis of NL interactions, replication timing, and transcription patterns. Gene activation inside LADs typically causes NL detachment of the entire transcription unit, but rarely more than 50–100 kb of flanking DNA, even when multiple neighboring genes are activated. The degree of detachment depends on the expression level and the length of the activated gene. Loss of NL interactions coincides with a switch from late to early replication timing, but the latter can involve longer stretches of DNA. Inactivation of active genes can lead to increased NL contacts. These extensive datasets are a resource for the analysis of LAD rewiring by transcription and reveal a remarkable flexibility of interphase chromosomes.

Introduction

In metazoan cell nuclei, large chromatin domains are associated with the nuclear lamina (NL) [1–5]. Mammalian genomes have roughly one thousand of such lamina-associated domains (LADs), which are typically hundreds of kb or even a few Mb in size. The NL contacts of some LADs are highly consistent between cell types, while other LADs interact in cell-type-specific (facultative) manners with the NL. How LAD-NL contacts are regulated is poorly understood.

Most genes inside LADs have very low transcriptional activity [6–8]. When cells differentiate, detachment of genes from the NL often coincides with transcriptional activation, while increased NL interactions correlate with reduced transcription [7,9–11]. These observations raise the interesting possibility that the NL helps to establish a repressive environment. In support of this notion, depletion of lamins can lead to derepression of specific genes (primarily in *Drosophila*) [12–14]; transfer of human inactive promoters from LADs to a neutral chromatin environment can lead to activation of these promoters [8]; and artificial tethering of some genes to the NL can reduce their activity [15–18].

This, however, does not rule out that the contacts of genes with the NL are the *consequence* of a lack of transcriptional activity, and vice versa, that genes detach from the NL *in response* to their activation. This was initially suggested by experiments with fluorescently tagged lacO arrays that were integrated in a locus near the NL. Tethering of the transcriptional activator peptide VP16 to these arrays caused repositioning away from the NL [19]. Similar observations were made when VP64 (a tetramer of VP16) was tethered to promoters of three distinct genes in LADs in mouse embryonic stem (mES) cells [20]. Another study found that activation of the long non-coding RNA gene *ThymoD* in mouse T-cell progenitors contributed to the detachment of the neighboring gene *Bcl11b* from the NL [21]. The molecular signals that cause detachment of a locus from the NL are still poorly understood.

Analysis of NL detachment that follows forced activation of a gene has so far been limited to a handful of loci. It is thus unclear whether the observed detachment from the NL after transcription activation is universal, or limited to genes with particular features. For example, do the size of the gene and its level of transcription matter? Moreover, the previous studies of individual loci have only been based on microscopy-based assays such as fluorescence *in situ* hybridization (FISH) or LacO tagging and have only visualized the targeted genes themselves, but not the flanking DNA sequences. It has therefore remained unclear what the

impact of these repositioning events is on the surrounding chromosomal regions. One possible scenario is that activation of a single gene inside a LAD leads to movement of the whole surrounding LAD to the nuclear interior. Alternatively, detachment could be restricted to the target gene itself or only affect some of its flanking regions. Possibly, detachment of one locus from the NL could be compensated by increased NL contacts of another locus nearby. To investigate this, high-resolution maps of NL interactions after manipulation of the activity of individual genes are needed.

Nuclear lamina interactions have also been associated with the timing of DNA replication during S-phase. LADs typically coincide with late-replicating domains, but the overlap is not complete, particularly at the edges of LADs [6,7,22]. These local discrepancies are still poorly understood, but may provide important clues about the interplay between the mechanisms that establish LADs and late-replicating domains. Above-mentioned activation of genes in LADs with TALE-VP64 was accompanied by a switch from late to early replication; however, it was not analyzed how far this switch extends across the locus and how well it tracks with the changes in NL contacts [20].

To study these issues, we took three complementary approaches. First, we used two VP16-tethering methods to activate a total of 14 different genes inside LADs, querying a variety of gene contexts. Second, we inactivated or truncated selected genes genetically to test whether they would re-attach to the NL. Third, we integrated an active transgene driven by a strong promoter into multiple LADs and tested how this altered NL interactions of the integration sites and the flanking regions. In each instance, we used DamID to map NL interactions, enabling us to visualize the extent of NL detachment in detail along entire chromosomes. We also compared the changes in NL interactions to changes in replication timing.

Results

Detachment of genes from the NL upon activation by TALE-VP64

We first employed a previously reported system in mouse embryonic stem (ES) cells, in which individual NL-associated genes are upregulated by means of TALE-VP64 fusion proteins that target the promoters [20]. In this system, relocation of the activated genes from the NL toward the nuclear interior was observed by FISH [20]. However, it is not known how much of the flanking DNA is involved in this detachment from the NL. We therefore repeated these experiments, but now we employed DamID mapping of lamin B1 interactions. This method has repeatedly been shown to correspond well with FISH microscopy [6,7,11,23,24], but it provides much more detailed maps of NL interactions.

We focused on two previously studied genes, *Sox6* and *Nrp1* [20]. In line with the reported FISH results, we observed clear detachment of each gene from the NL, when activated by the corresponding TALE-VP64 construct (Fig 1A and B middle panels). To assess the statistical significance of these changes, we compared their magnitude to those observed throughout the remainder of the genome. Because the size of the affected region is *a priori* not known, we calculated this comparison for various window sizes between about 30 kb and 1 Mb. This resulted in domainograms [25,26] that depict the genome-wide ranking of displacement magnitudes as a function of window position as well as window size (Fig 1A and B top panels; see Appendix Fig S1 for an explanation of domainograms). We regard displacements that rank above the 95th percentile or below the 5th percentile (marked in shades of blue and red for decreased and increased NL interactions, respectively) and that occur locally near the targeted gene, to be highly likely due to direct effects. We note that some indirect displacements elsewhere in the genome may be expected, because the perturbations of *Sox6* and *Nrp1* may have secondary effects on gene expression.

The domainograms indicate that the displacements of the respective targeted genes were among the most extreme throughout the genome. For *Sox6,* the NL detachment included the entire gene, but it was more pronounced near the promoter than toward the 3′ end (Fig 1A). Upstream of the promoter the detachment extended over ~50–100 kb, up to the LAD border. Downstream of the gene, a modest detachment was observable that tapered off over ~300 kb. Interestingly, about 0.5 Mb upstream, across the LAD border, also some reduction in NL interactions is visible. For Nrp1, the detachment also involved the entire gene body but did not extend much beyond it (Fig 1B).

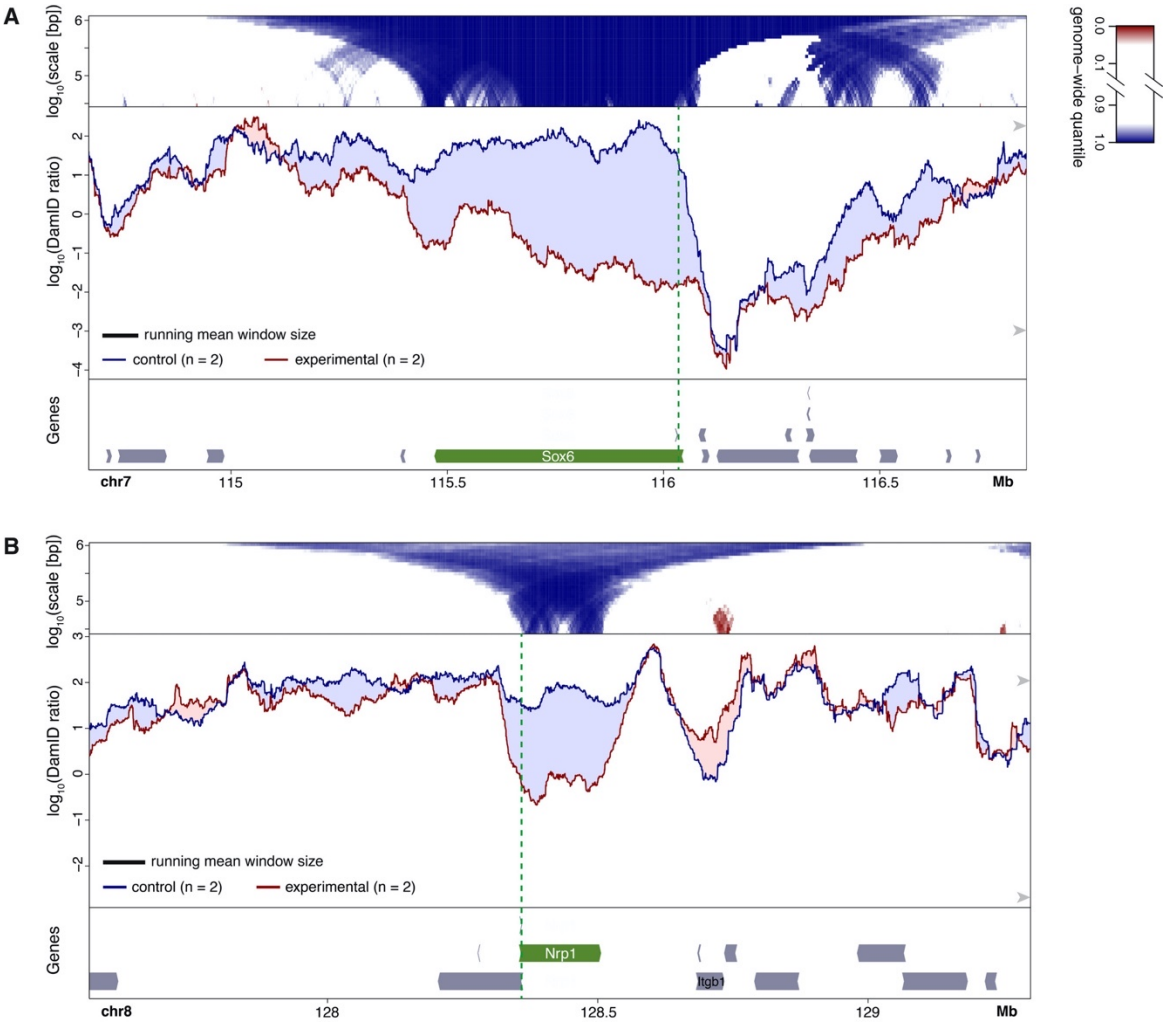


Figure 1

Changes in NL interactions of Sox6 (A) and Nrp1 (B) in mES cells after activation by TALE-VP64

A, B. Bottom panels: gene annotation track (mm10); the activated gene is marked in green and the location of the TALE-VP64 target sequence is shown by the vertical dashed green line. Middle panels: DamID tracks of NL interactions in control cells ("control", blue line) and cells expressing TALE-VP64 ("experimental", red line). n indicates the number of independent biological replicates that were combined. Noise was suppressed by a running mean filter of indicated window size. Shading between the lines corresponds to the color of the sample with the highest value. Arrowheads on the right-hand side mark the 5th and 95th percentiles of genome-wide DamID values. Top panels: domainograms; for every window of indicated size (vertical axis) and centered on a genomic position (horizontal axis), the pixel shade indicates the ranking of the change in DamID score (experimental minus control) in this window compared to the genome-wide changes in DamID scores across all possible windows of the same size. Blue: DamID score is highest in control samples; red: DamID score is highest in experimental samples (color key on the right of panel (A)). In (A) activation of Sox6 was the experimental perturbation, activation of Nrp1 (which is located on a different chromosome) served as control; in (B) activation of Nrp1 was the experimental condition and activation of Sox6 served as control.

We note that the loss of DamID signal along active transcription units cannot be attributed to the RNA polymerase machinery blocking access to Dam methylation, because all DamID data are normalized to a Dam-only control that corrects for such accessibility biases [27], and the same experimental design has successfully detected interaction of specific proteins with actively transcribed regions [28]. About 180 kb downstream of *Nrp1*, the gene *Itgb1* showed a modest increase in NL interactions. This will be discussed below. Together, these data show that activation of two genes inside LADs of mES cells results in detachment from the NL along the entire gene body, possibly with some subtler involvement of flanking regions.

Detachment span is linked to transcript length

To extend this analysis to a larger number of genes, we switched to a more flexible gene activation system that does not require a custom-made TALE for every promoter of interest. We chose a previously established human RPE-1 cell line that stably expresses the SunCas-CRISPRa system [29,30], in which multiple copies of VP64 can be targeted to a promoter of interest by a single sgRNA. We first used this system to activate *NLGN1*, a gene of 885 kb that is located in a LAD. Transfection with a sgRNA targeting the promoter caused ~80-fold upregulation (Appendix Fig S2) and resulted in clear detachment from the NL (Fig 2A). Relocalization primarily affected the NLGN1 gene itself, with a mild 5′ to 3′ gradient inside the gene body and gradually decreasing along ~100 kb of flanking DNA. We also activated the SOX6 gene in RPE-1 cells. Here, we activated one of the known alternative promoters located internally in the gene. Although the magnitude of the detachment was more modest, this gene also showed loss of NL interactions, but only downstream of the activated promoter (Fig 2B).

We applied this analysis in RPE-1 cells to 12 individual genes (Table 1) for which activation by the SunCas system could be achieved, as determined by RT–qPCR (Appendix Fig S2) or RNA-seq (Fig EV1). We chose genes of a wide variety of lengths, from ~2 kb to ~1.5 Mb. In three cases (*ABCB1*, *SLC35F3, and SOX6*), we targeted a known alternative promoter located in the middle of the gene, instead of the promoter located most 5′. Strikingly, for all 12 activated genes we observed detachment of the entire region extending from the activated promoter to the 3′ end of the gene (Fig 2C). For most of the tested genes, detachment did not extend more than several tens of kb upstream of the activated promoter. A clear exception to this is *PTN*, which exhibited upstream detachment over nearly 0.5 Mb (see below). Likewise, for most activated genes the detachment did not extend more than 50–100 kb downstream of the 3′ end, although the precise range varied.

Table 1 Genes and sequences targeted by CRISPRa in RPE-1 cells

Target gene	Target sequence	Target site (hg19)
ABCB1	GGGCCGGGAGCAGTCATCTG	chr7:87230290-87230310
ABCB4[a]	TGCAACGGTAGGCGTTTCCC	chr7: 87105074-87105093
ADAM22	CGGGCGACAAGAGCTCGGCA	chr7:87563472-87563492
CCSER1	GTGCGCGGAGTGTGACTGTG	chr4:91048592-91048612
GRID2	CAAAAGCATCCTGCAGCCTG	chr4:93225024-93225044
MLK4	AGGGCGGAATGAACCTGGAG	chr1:233463313-233463333
NLGN1	TGAAGGGTCAACCCTCCGCG	chr3:173115478-173115498
PCNX2	TCCCTCCTTAGCCTTCGCTG	chr1:233431545-233431565
PTN	GAGCAGAGGAAAATCCAAAG	chr7:137028354-137028374
RUNDC3B[a]	GCTGCTTTAAAAGGTCCGCG	chr7: 87257590-87257609
SLC35F3	TAAAGGGCTTCTCAGAGAGG	chr1:234349808-234349827
SOX6	GCTCCCCTCCCAGACAACAC	chr11:16629348-16629368
TRAM1L1	AGAATTCAGGAGCATCTTGG	chr4:118006859-118006879
ZNF804B	AGGCGCGGGTACCCATCGTC	chr7:88388864-88388884

[a] ABCB4 and RUNDC3B were only targeted in combination with ADAM22 and ABCB1.

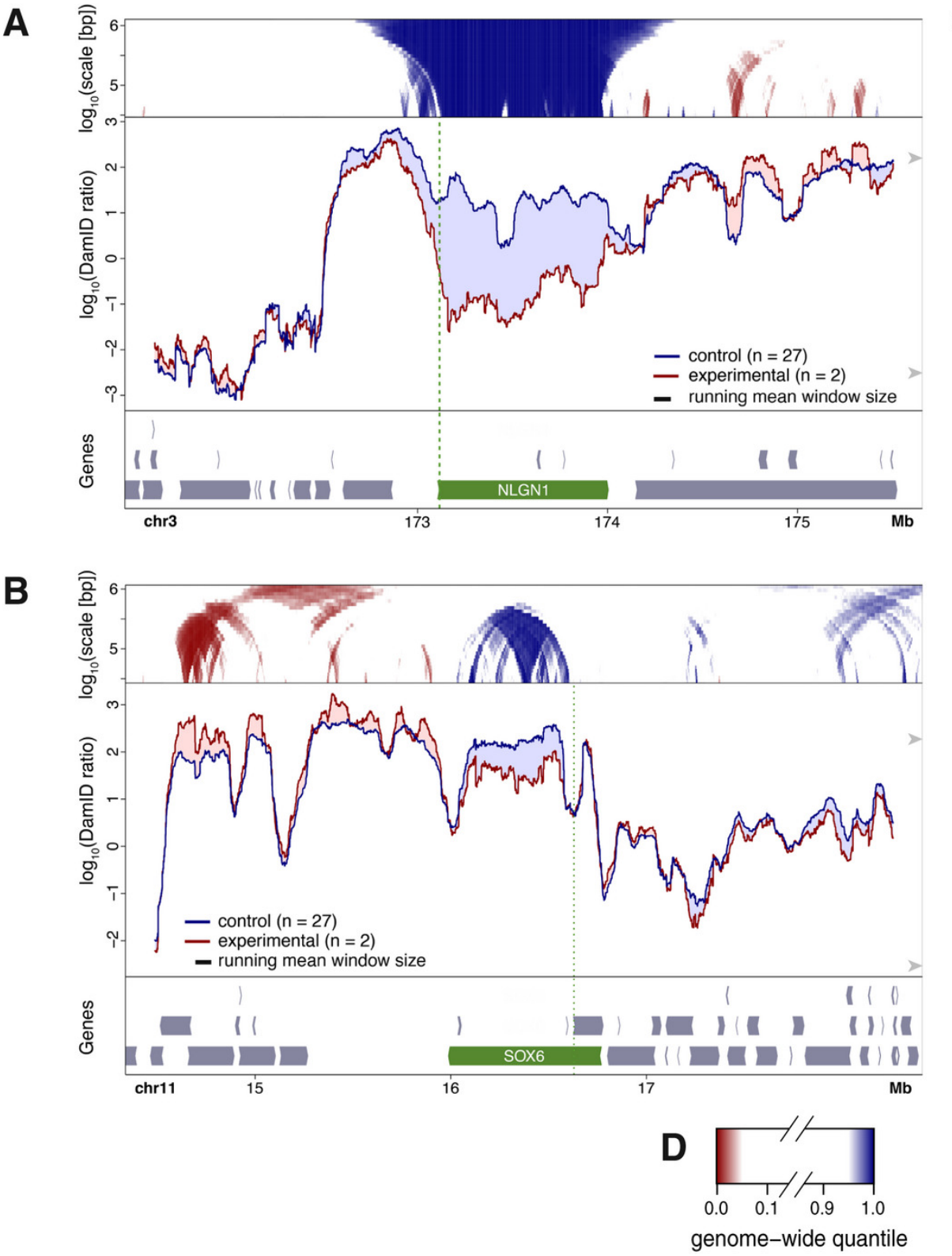
Figure 2

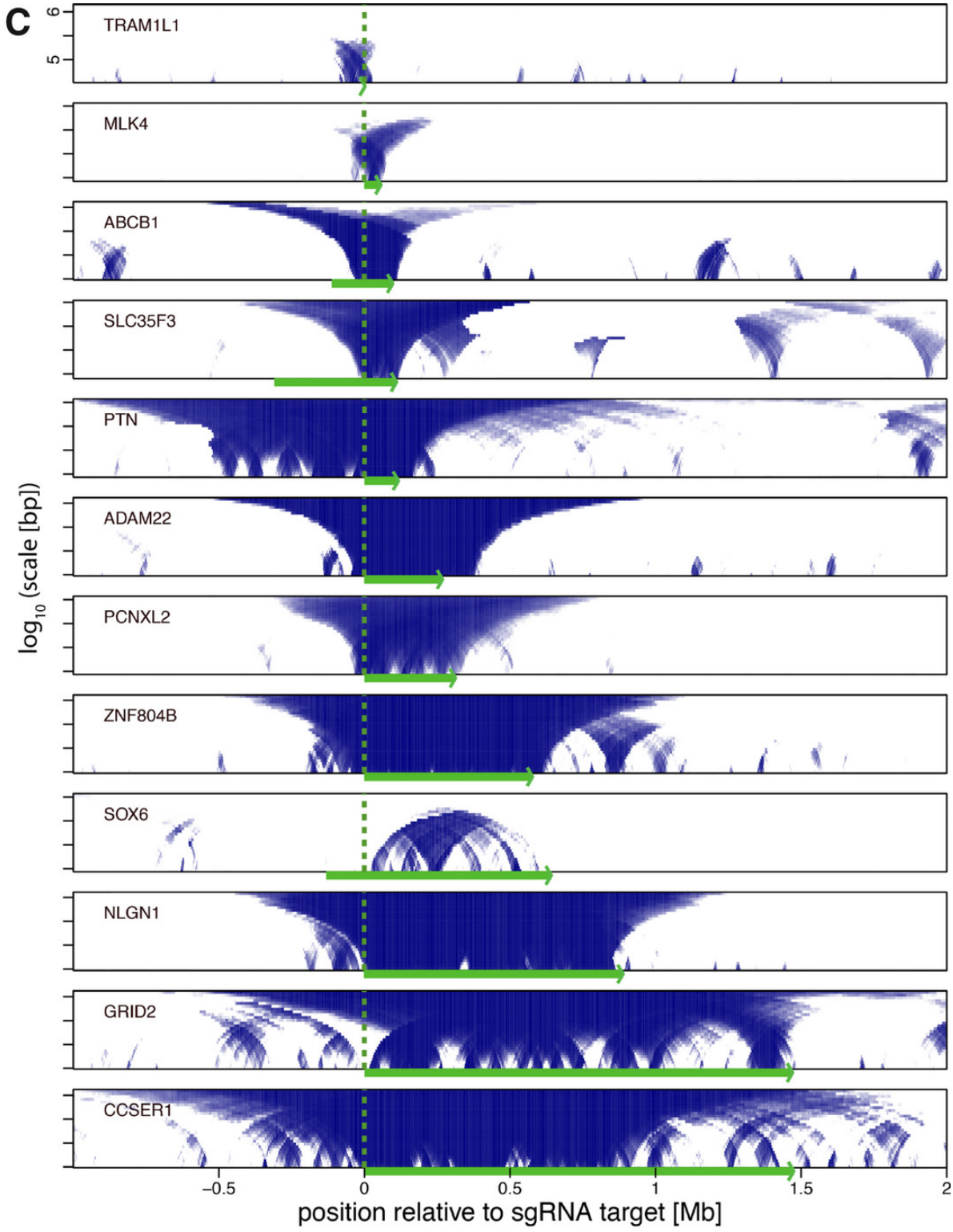
Local NL detachment caused by gene activation by CRISPRa in human RPE-1 cells

A, B. Plots as in Fig 1, showing changes in lamin B1 DamID signals upon CRISPRa activation of NLGN1 (A) and SOX6 (B). Control cells were treated either without sgRNA or with one of various sgRNAs targeting promoters on different chromosomes. Vertical green dotted lines mark the position of the sgRNA target sequence.

C. Domainograms showing regions with reduced NL interactions around 12 genes individually activated by CRISPRa. Genomic regions are aligned by the respective sgRNA target positions and oriented so that the activated genes are all transcribed from left to right. Corresponding DamID traces are shown in Figs 2A and B, and 4, 5, EV2, EV4.

D. Color key of domainograms in (A–C). Increases in NL interactions (red) are not shown in (C).


A
log10(scale [bp])
log10(DamID ratio)
control (n = 27)
experimental (n = 2)
running mean window size
Genes
NLGN1
chr3
173
174
175
Mb
B
log10(scale [bp])
log10(DamID ratio)
control (n = 27)
experimental (n = 2)
running mean window size
Genes
SOX6
chr11
15
16
17
Mb
D
0.0
0.1
0.9
1.0
genome–wide quantile


C
TRAM1L1
MLK4
ABCB1
SLC35F3
PTN
ADAM22
PCNXL2
ZNF804B
SOX6
NLGN1
GRID2
CCSER1
log10 (scale [bp])
−0.5
0
0.5
1
1.5
2
position relative to sgRNA target [Mb]

Quantitative link between gene expression level and NL detachment

We wondered whether the degree of NL detachment of a gene is quantitatively linked to the transcription level. To measure gene activity accurately, we performed RNA-seq after activation of 8 of the 12 genes and in the untreated parental cell line (Fig EV1). Comparing RNA levels and average DamID scores across the SunCas-activated genes before and after upregulation revealed a strong negative correlation (Fig 3).

Thus, there is a remarkably quantitative inverse link between expression levels and NL interaction frequencies.

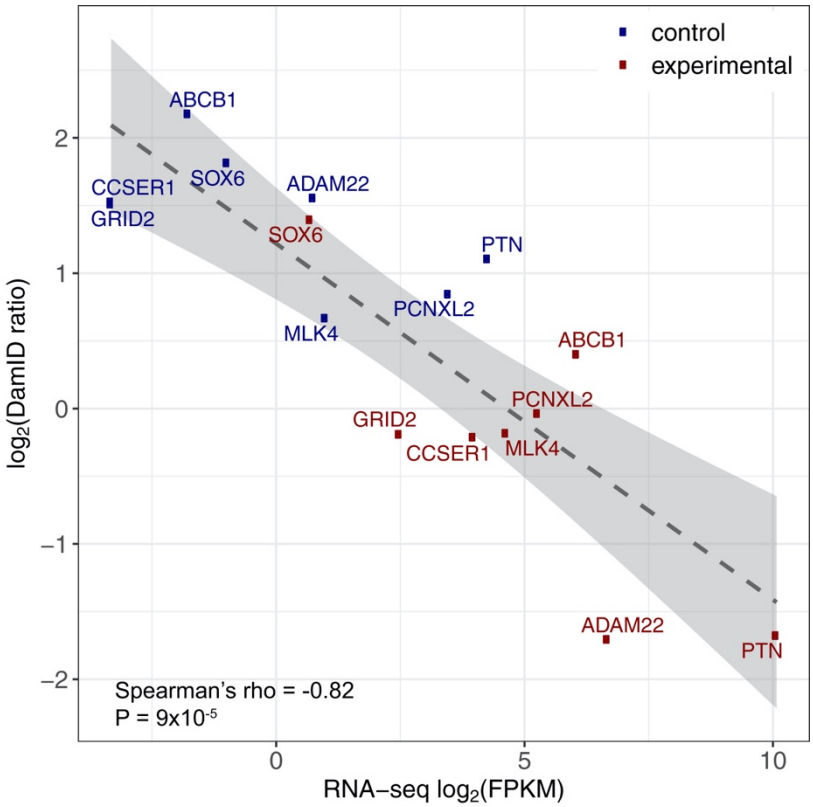


Figure 3

Inverse correlation between NL interaction and gene expression level in human RPE-1 cells

Average DamID values plotted against average expression levels of eight genes activated by CRISPRa (red; n = 2) or in control cells that were treated either without sgRNA or with one of various sgRNAs targeting promoters on different chromosomes (blue; n ranging from 19 to 27).

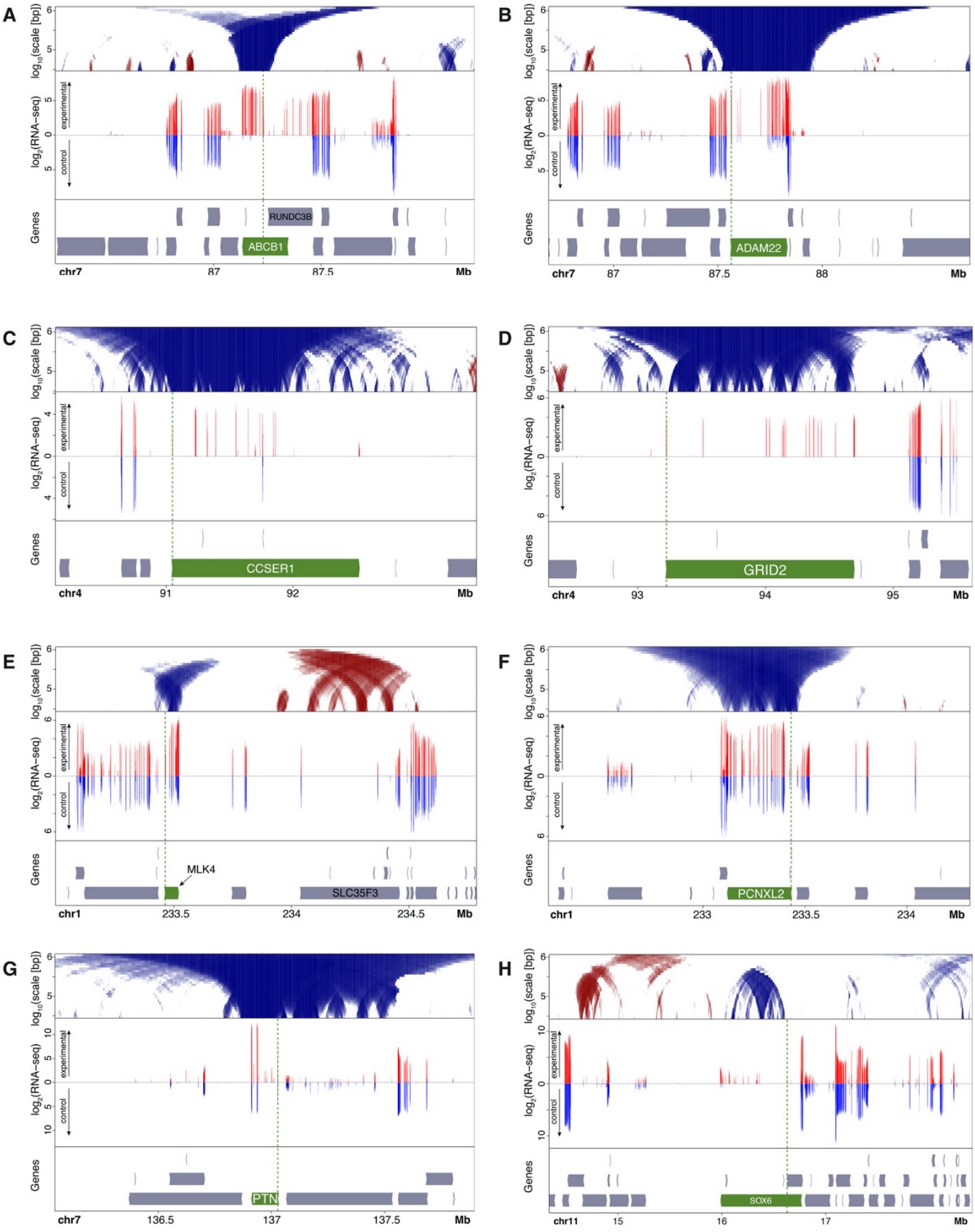

A
B
C
D
E
F
G
H
log10(scale [bp])
log2(RNA−seq)
experimental
control
Genes
RUNDC3B
ABCB1
ADAM22
CCSER1
GRID2
MLK4
SLC35F3
PCNXL2
PTN
SOX6
chr7
chr4
chr1
chr11
Mb

Figure EV 1

Expression levels of CRISPRa-targeted genes and neighboring genes in human RPE-1 cells

A–H. Top panels: domainograms showing changes in NL interactions around genes activated by CRISPRa. Domainograms are same as in Fig 2C. Middle panels: log2 mRNA expression (normalized read counts in 50 bp bins) as determined by mRNA-seq of CRISPRa-activated cells (red, upward y-axis) compared to untransfected control cells (blue, downward y-axis). Note that only introns give detectable reads. Data are average of two independent replicate experiments each. Bottom panels: gene annotation track. Each CRISPRa-targeted gene is highlighted in green; targeted promoters are marked by vertical green dotted line.

Neighboring genes of targeted genes do not generally show altered expression

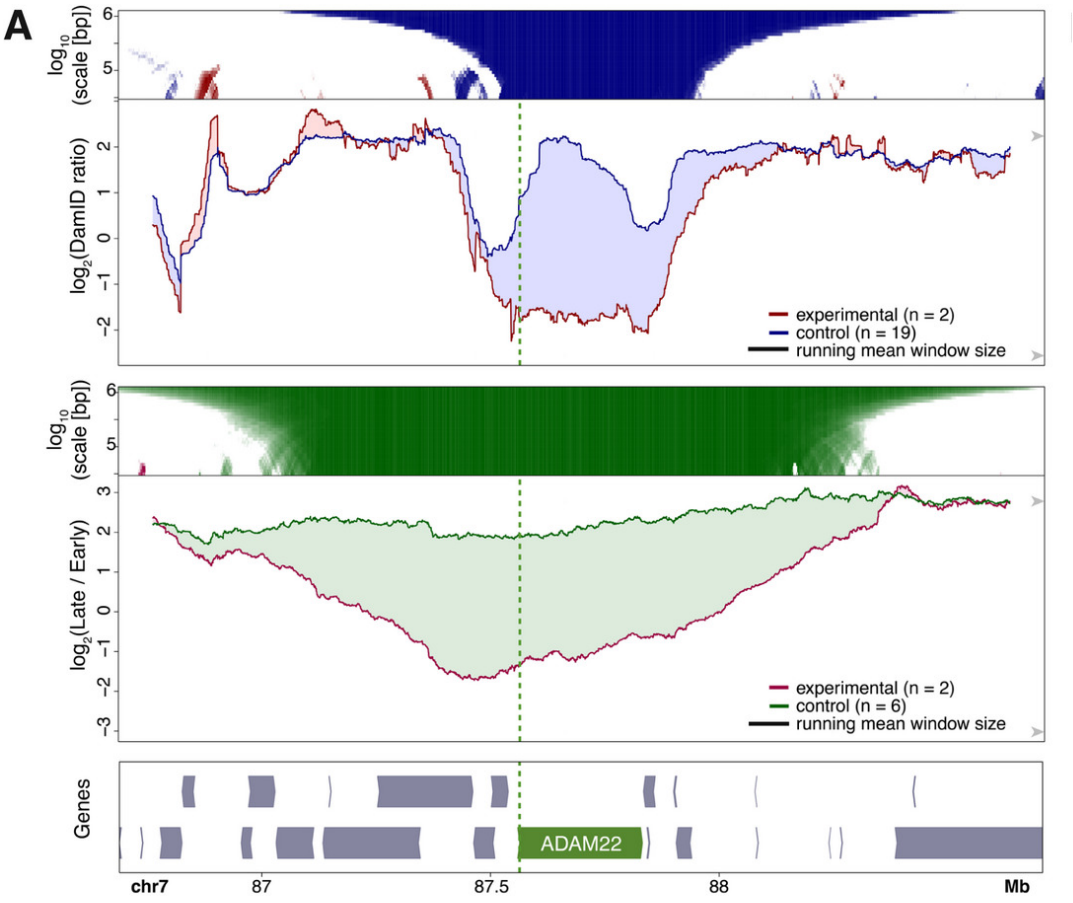
We also queried our RNA-seq data for neighboring genes of the activated genes. We examined genes within ~1 Mb of our targets and generally could not detect substantial up- or downregulation (Fig EV1A–H). A notable exception is the gene *RUNDC3B* that is partially overlapping and antisense to the activated gene *ABCB1*. *RUNDC3B* is co-activated but does not show measurable detachment from the NL (Fig EV1A). Another, minor, exception is the gene *STEAP4* nearby the activated ADAM22 gene (Fig EV1B). In this case, the absolute expression level of this co-activated gene remained much lower than its CRISPRa-targeted neighbor. Thus, strong SunCas-induced upregulation is in most cases restricted to the targeted gene.

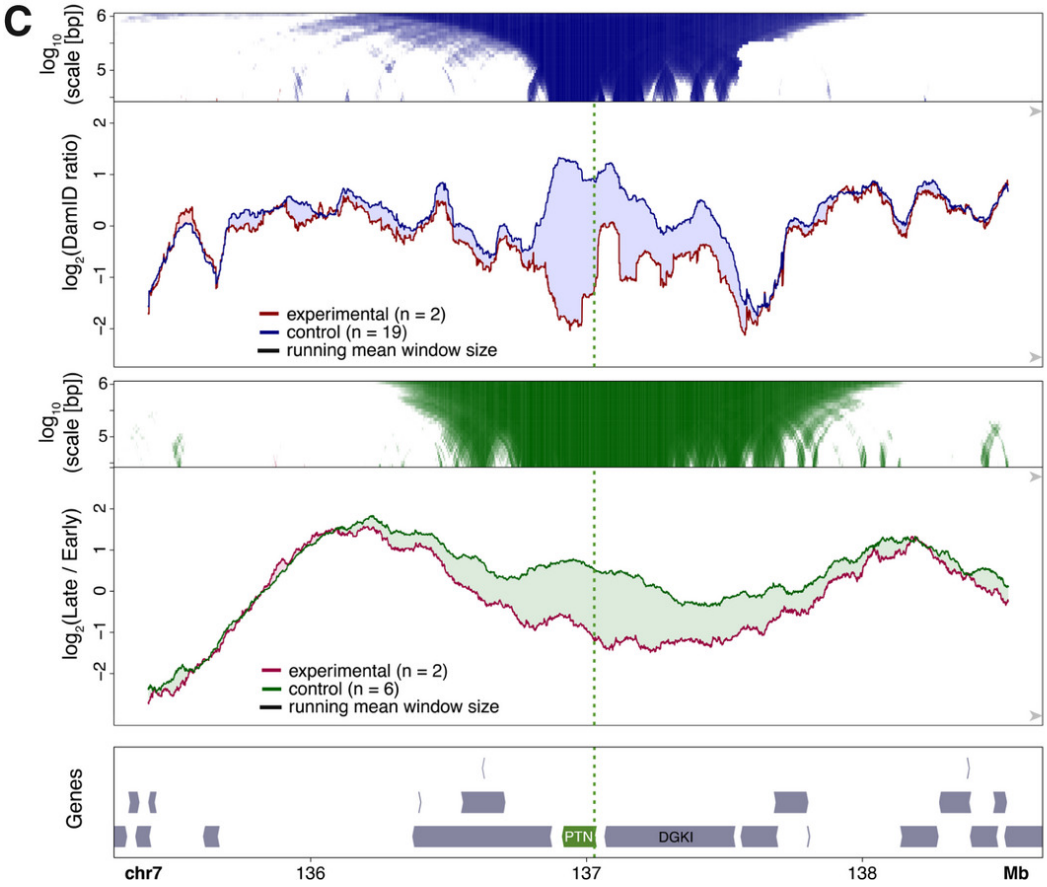
Some flanking genes co-detached from the NL together with the activated gene, but showed no detectable change in expression. The most striking example is the gene *DGKI* that flanks *PTN*. Much of this ~0.5 Mb gene shows reduced NL interactions upon activation of *PTN*, but *DGKI* does not undergo a detectable upregulation (Fig EV1G). We conclude that CRISPRa activation and the ensuing changes in NL contacts generally do not have substantial effects on the expression of nearby genes.

NL detachment partially overlaps with changes in replication timing

Next, we investigated the link between changes in NL interactions and replication timing. We applied Repli-seq [31] to visualize replication timing, revisiting five genes that exhibited NL detachment in RPE-1 cells upon activation by CRISPRa. For all five activated genes, we observed a clear shift toward earlier replication. When the activated genes were relatively small (*ADAM22*, *ABCB1*, *PTN*), this shift was more or less symmetrical around the activated promoter and extended about 0.4–0.8 Mb on each side, i.e., well beyond the activated transcription units and also beyond the changes in NL interactions (Fig 4A–C). With longer activated genes (*CCSER1*, *GRID2,* both about 1.5 Mb long), again the shift in replication timing was strongest around the targeted promoter and extended about 0.6 Mb upstream (Fig EV2). Downstream of these promoters, the shift declined gradually toward the end of the gene, similar to the detachment from the NL.

We investigated whether the changes in NL interactions and replication timing were linked to topologically associated domains (TADs) as detected by the Hi-C technology [32,33]. For this, we analyzed previously reported Hi-C data from wild-type RPE-1 cells [34,35]. We focused on the *PTN* locus, which showed the most striking difference between the changes in NL contacts and Repli-seq patterns after CRISPRa. Interestingly, the change in Repli-seq pattern appeared not visibly linked to the (pre-existing) TAD pattern, while the changes in NL interactions showed a partial correlation with the TAD organization (Fig EV3). Together, these results reveal that changes in replication timing only partially overlap with changes in NL interactions (see Discussion).


A
log10 (scale [bp])
log2(DamID ratio)
experimental (n = 2)
control (n = 19)
running mean window size
log10 (scale [bp])
log2(Late / Early)
experimental (n = 2)
control (n = 6)
running mean window size
Genes
ADAM22
chr7
87
87.5
88
Mb


C
log10 (scale [bp])
log2(DamID ratio)
experimental (n = 2)
control (n = 19)
running mean window size
log10 (scale [bp])
log2(Late / Early)
experimental (n = 2)
control (n = 6)
running mean window size
Genes
PTN
DGKI
chr7
136
137
138
Mb

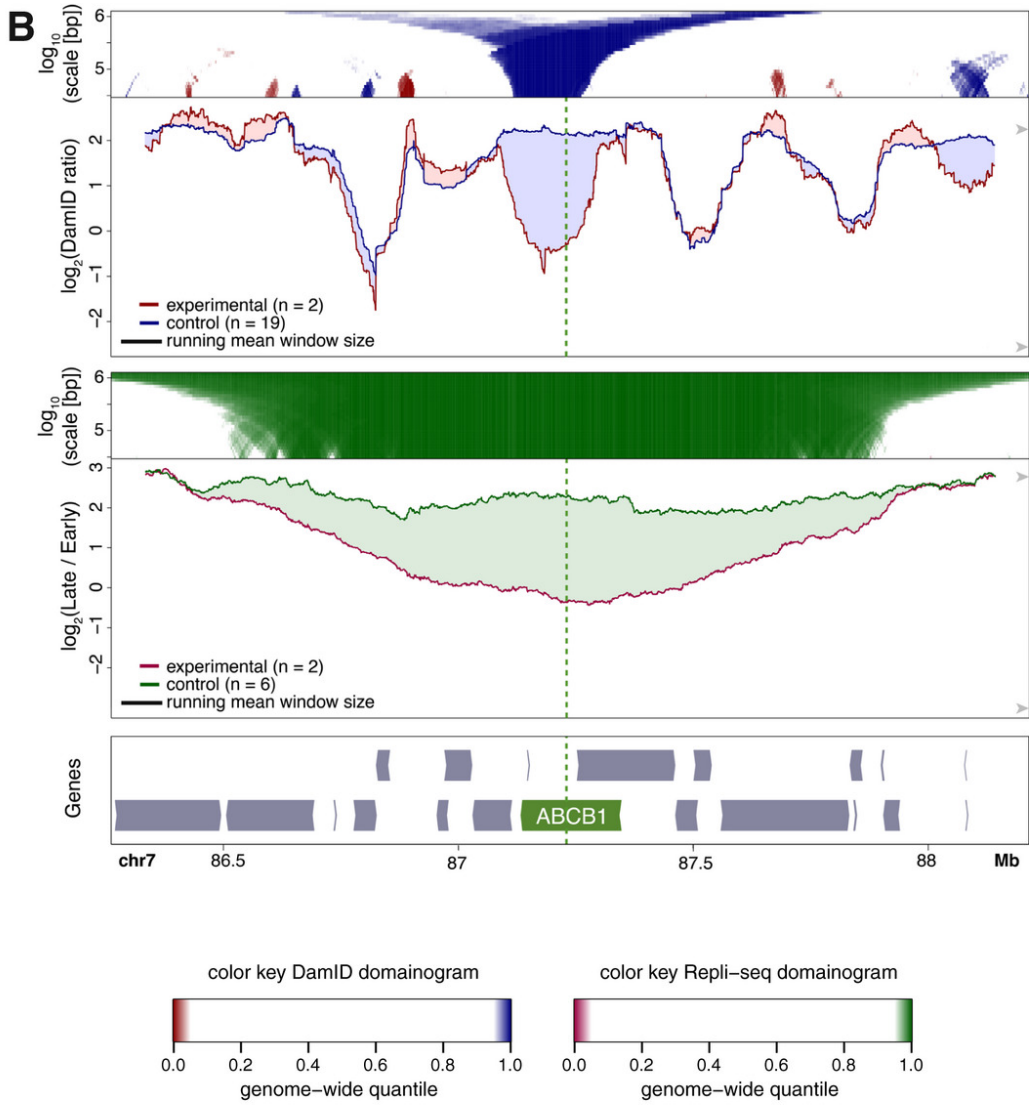


Figure 4

NL interactions and replication timing around activated genes

A–C. CRISPRa activation of ADAM22 (A), ABCB1 (B), and PTN (C) in human RPE-1 cells. Top panels visualize DamID data similar to Fig 2A and B. Middle panels show maps of replication timing at the same resolution and in the same plotting style as panels, except that different colors are used as indicated. Bottom panels show gene track, with activated gene highlighted in green.

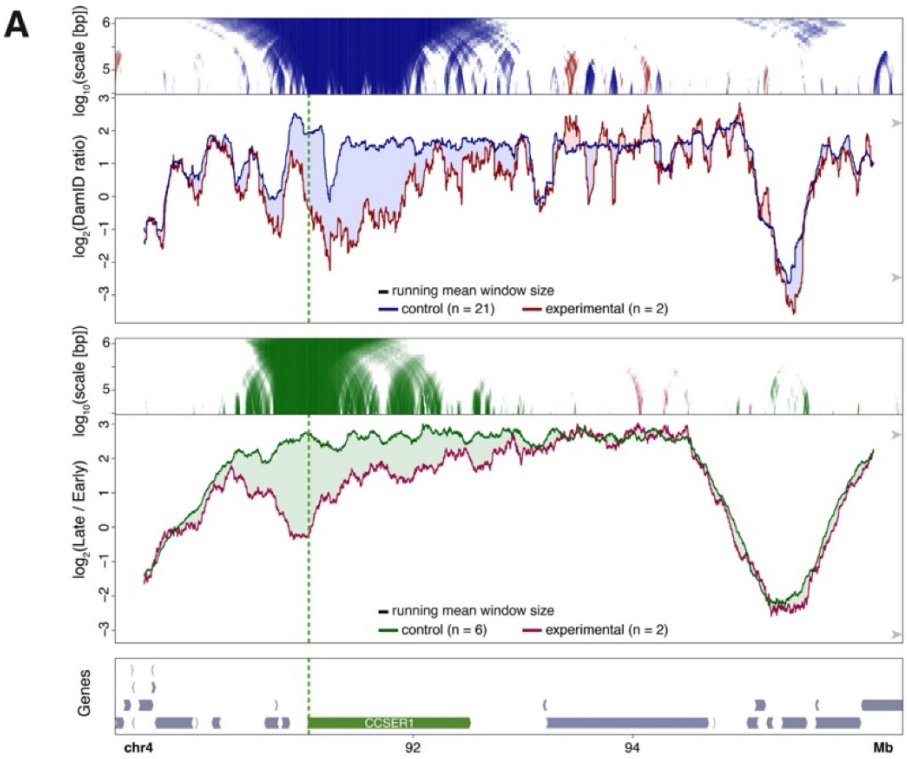
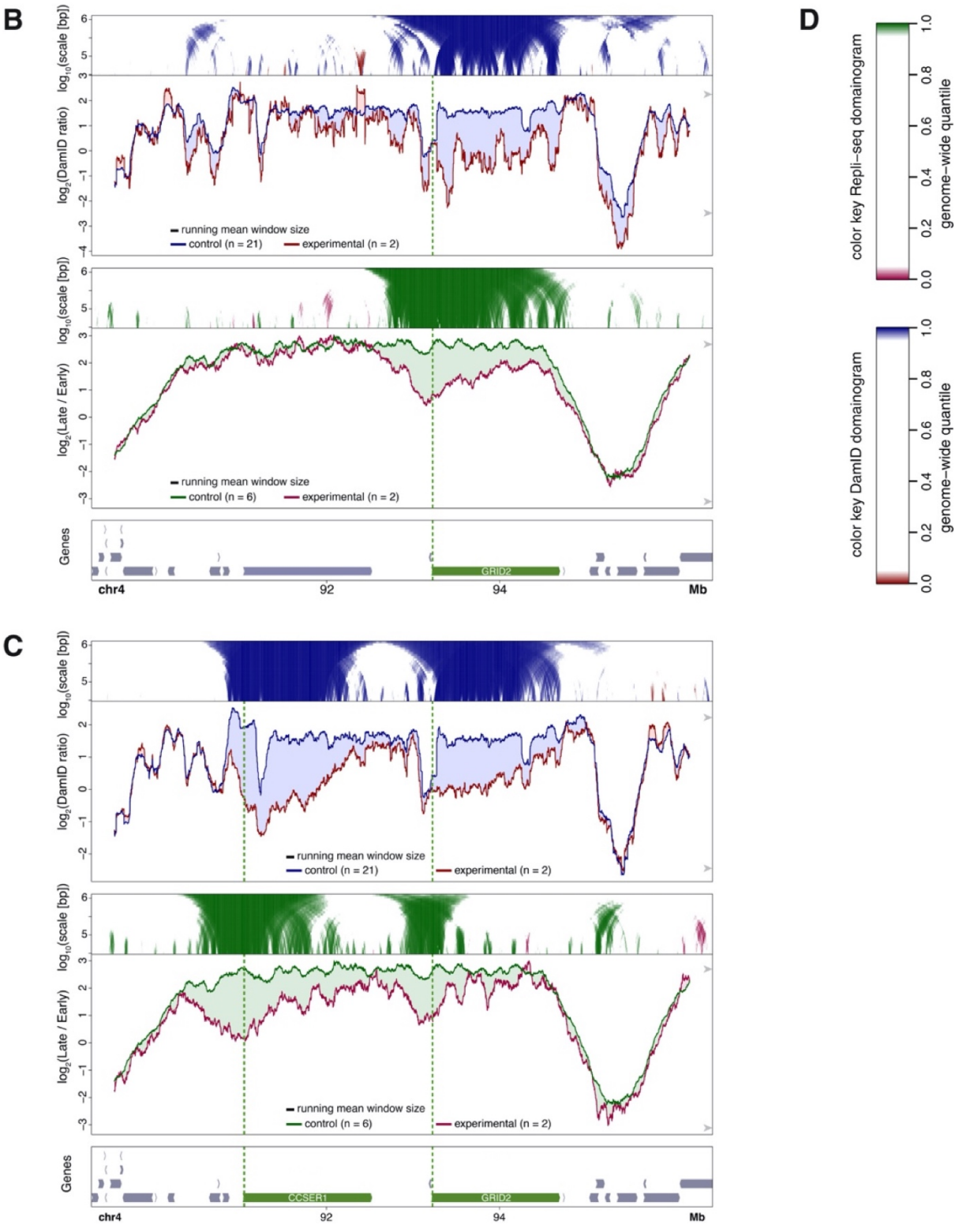


Figure EV 2

Changes in NL interactions and replication timing in the CCSER1/GRID2 locus

A–D. Effects of CRISPRa activation of CCSER1 (A), GRID2 (B) or both genes (C) in human RPE-1 cells. Top panels show DamID data similar to Fig 2A and B. Middle panels show maps of replication timing at the same resolution and in the same plotting style as for DamID, except that different colors are used as indicated in (D). Bottom panels show gene annotation; activated gene(s) highlighted in green; targeted promoters are marked by vertical green dotted line. DamID data are the same as in Figs 2C and 6A.


B
log10(scale [bp])
log2(DamID ratio)
running mean window size
control (n = 21)
experimental (n = 2)
log2(Late / Early)
control (n = 6)
experimental (n = 2)
Genes
GRID2
chr4
92
94
Mb
C
CCSER1
D
color key Repli-seq domainogram
genome-wide quantile
color key DamID domainogram

Possible compensatory movements

In a few instances, we observed that the loss of NL interactions of the activated gene was accompanied by a gain of NL interactions of a nearby region. This was particularly notable for a region ~0.6 Mb downstream of the activated MLK4 gene (Fig 5). This region coincides approximately with gene *SLC35F3*. The expression of *SLC35F3* is reduced by ~30% (P = 0.02, DESeq2 analysis) when *MLK4* was activated (Fig EV1E). Possibly, detachment of *MLK4* leads to compensatory movement of *SLC35F3* toward the NL, which in turn may contribute to slightly stronger repression of *SLC35F3*. Forced activation of *SLC35F3* caused its own NL detachment as expected (Fig 2C), but it did not alter the NL interactions of *MLK4* (Fig EV4A). This suggests that the putative compensatory relationship is not reciprocal, but we note that this latter experiment was done only once and should therefore be interpreted with caution.

We also observed moderately enhanced NL interactions of a region ~1.3 Mb downstream of the activated *SOX6* gene (Fig 2B). This region coincided with the promoters of two divergent genes that were not significantly up- or downregulated (Fig EV1H). Likewise, in case of *Nrp1* activated by TALE-VP64, the gene *Itgb1* (about 180 kb downstream of *Nrp1*) showed a modest increase in NL interactions (Fig 1B), but its expression was not found to be detectably altered by TALE-VP64 targeting of *Nrp1* [20]. We found also minor local increases in NL interactions within ~2 Mb of the activated genes *NLGN1* (Fig 2A), T*RAM1L1* and *ZNF804* (Fig EV4B and C). However, because of their modest magnitude, we did not further investigate these movements. In summary, possible compensatory changes in NL interactions around activated genes are relatively modest and may only anecdotally affect gene expression, at least in the cell systems we studied. These increases in NL interactions may reflect compensatory movement to fill up space at the NL vacated by the activated genes, but other secondary mechanisms cannot be ruled out.

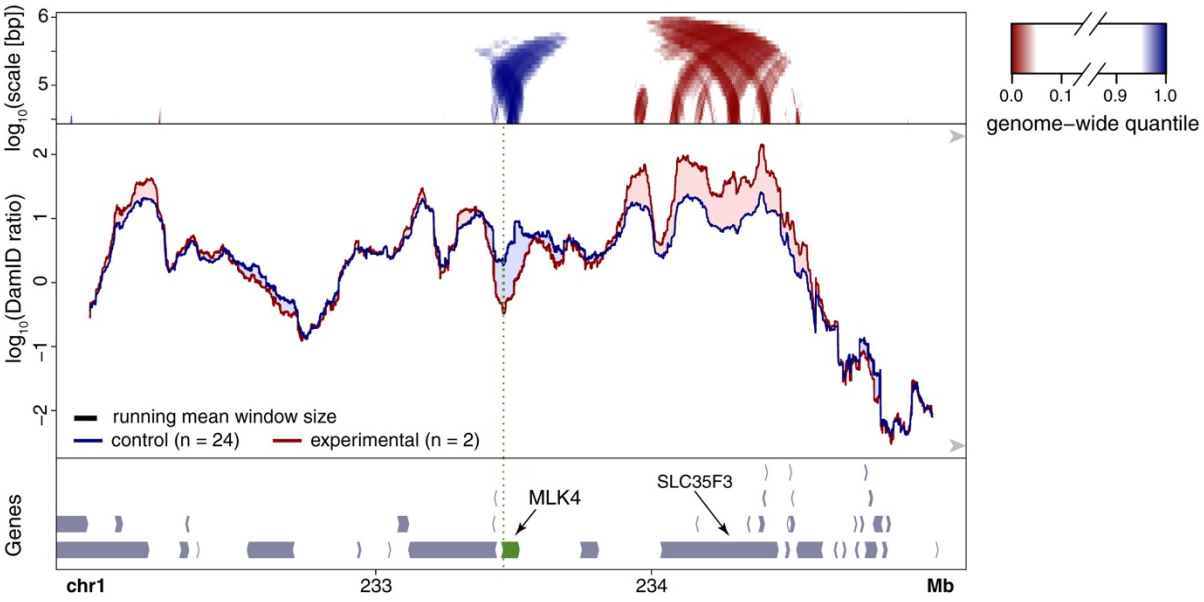


Figure 5

Putative compensatory movement near the activated MLK4 gene Changes in NL interactions after CRISPRa activation of MLK4 in human RPE-1 cells. Visualization of DamID data as in Fig 2A and B.

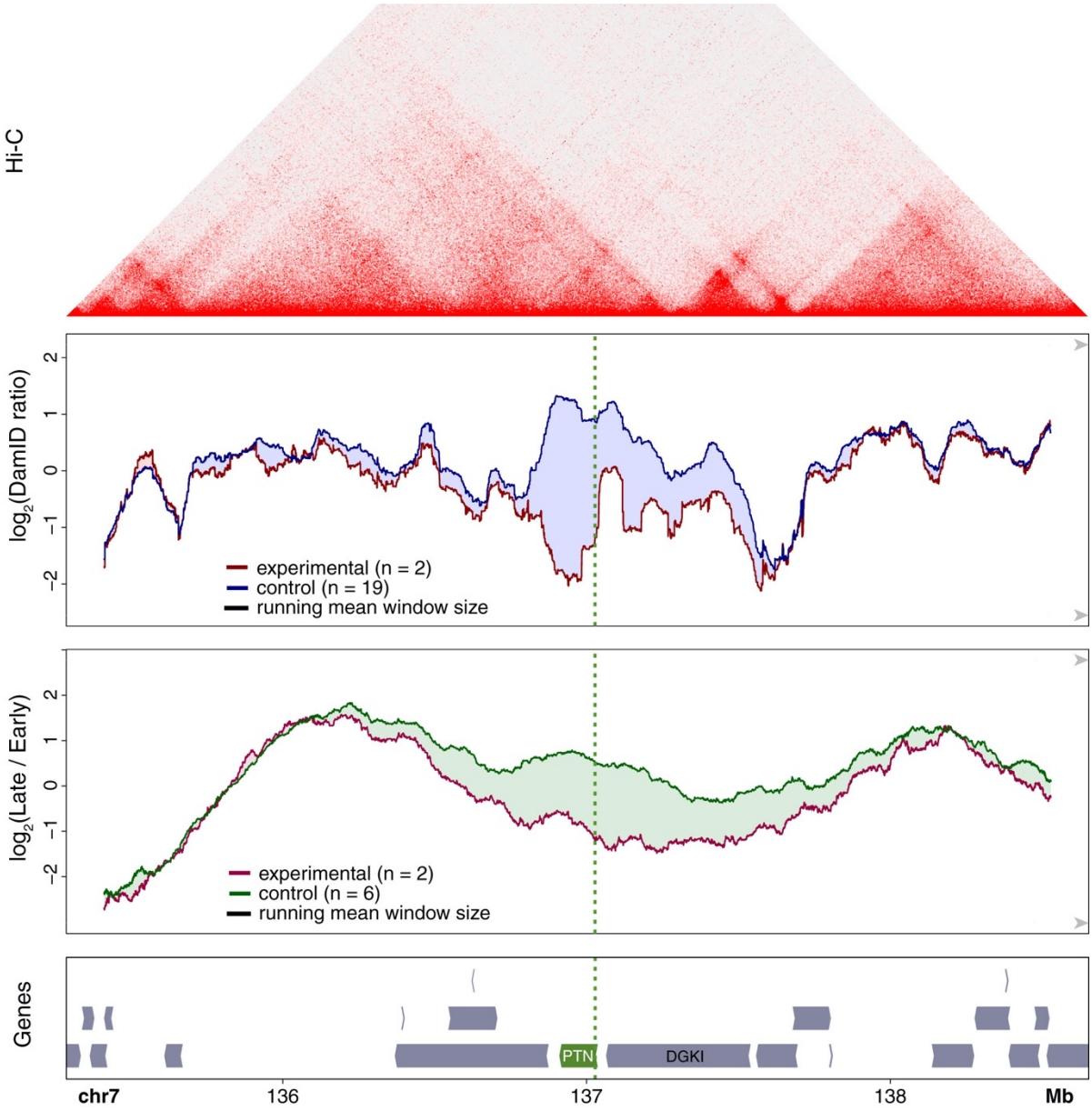


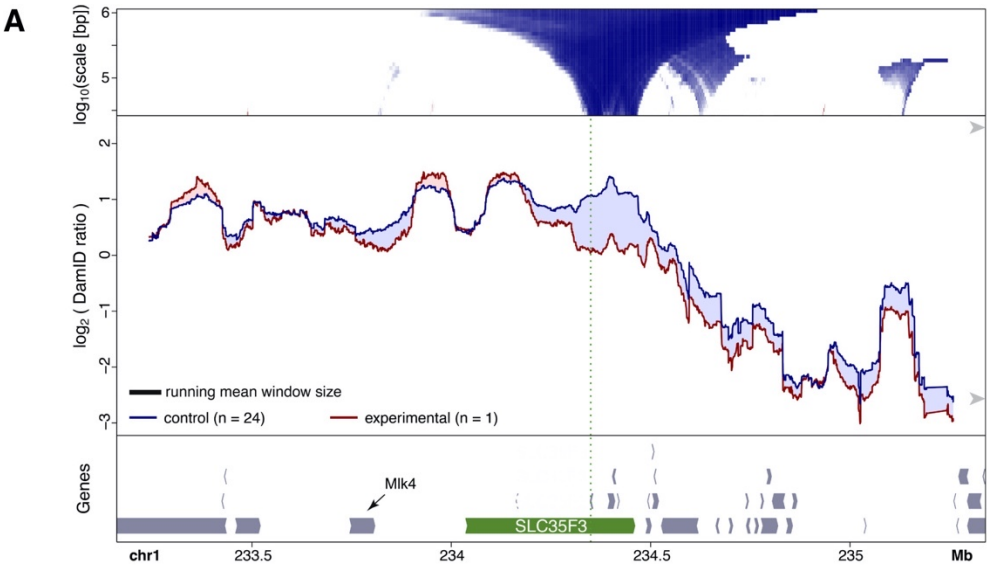
Figure EV 3

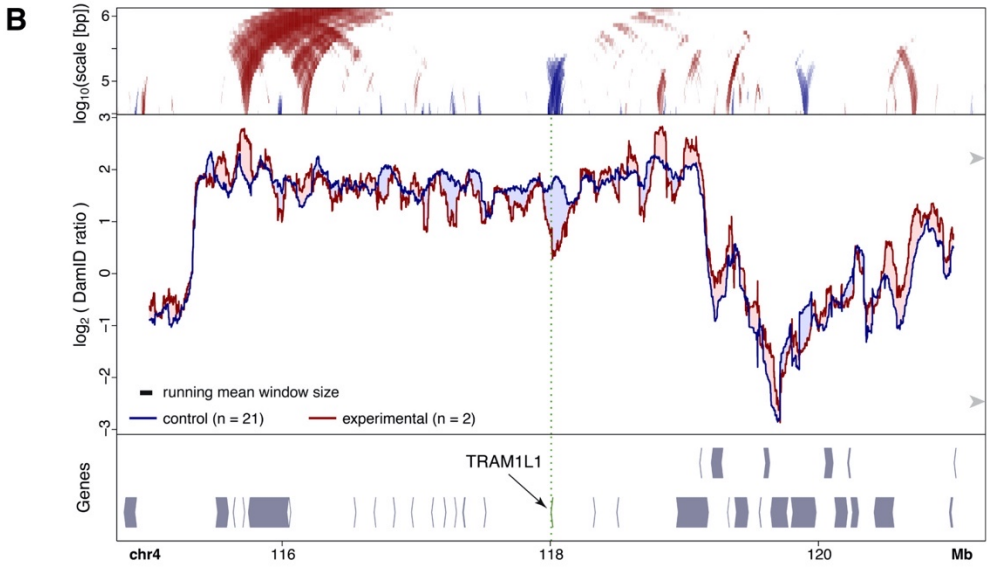
Comparison of DamID and replication timing patterns around the PTN gene to Hi-C data DamID data and gene annotation track of Fig 4C (three bottom panels) aligned to Hi-C data from untreated human RPE-1 cells (top panel). Hi-C data are from (Darrow et al, 2016) (Data ref: Darrow et al, 2016).

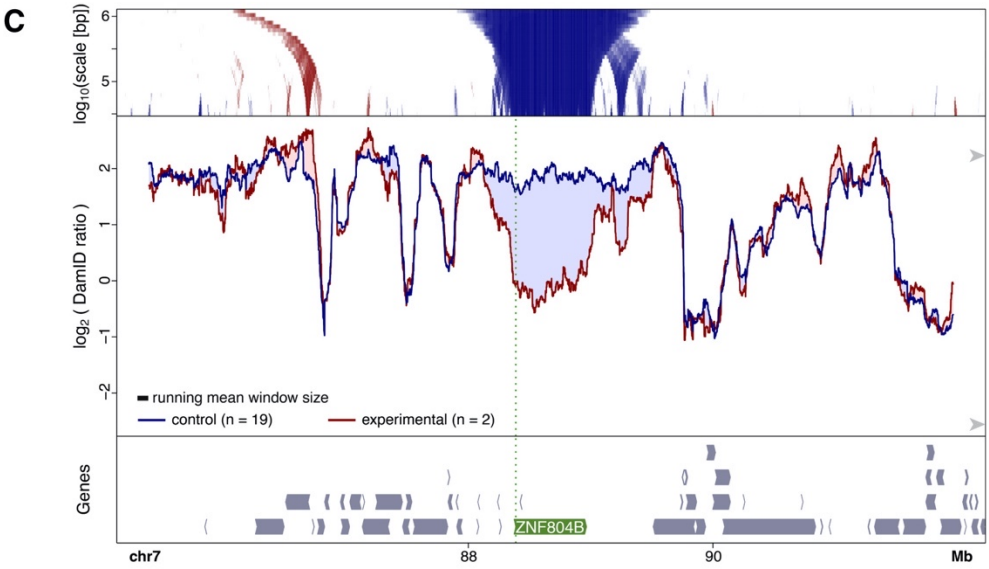
Figure EV 4

Changes in NL interactions of genes upregulated by CRISPRa

A–C. DamID data obtained after CRISPRa in human RPE-1 cells of SLC35F3 (A), TRAM1L1 (B), ZNF804B (C). Domainograms are the same as in Fig 2C, but additionally show increased NL contacts in red.


A
log10(scale [bp])
log2 (DamID ratio)
running mean window size
control (n = 24)
experimental (n = 1)
Genes
Mlk4
SLC35F3
chr1
233.5
234
234.5
235
Mb


B
log10(scale [bp])
log2 (DamID ratio)
running mean window size
control (n = 21)
experimental (n = 2)
Genes
TRAM1L1
chr4
116
118
120
Mb


C
log10(scale [bp])
log2 (DamID ratio)
running mean window size
control (n = 19)
experimental (n = 2)
Genes
ZNF804B
chr7
88
90
Mb

Detachment remains local even if multiple genes are activated

We next explored whether it is possible to detach a whole LAD by activating multiple genes in the domain. We first tested this for two neighboring large genes, *CCSER1* (1,475 kb) and *GRID2* (1,468 kb). Activation of each gene individually caused clear detachment from the NL (Figs 6A and EV2A and B). When both genes were activated simultaneously (by co-transfection of the respective sgRNAs), both genes detached, but the intervening ~700 kb region showed no significant reduction in NL association (Fig 6A). Under this double-activating condition, the intervening region also continued to be replicated a bit later in S-phase than the two activated genes (Fig EV2C).

We also applied CRISPRa simultaneously to the much more closely spaced genes *ABCB4*, *ABCB1*, *ADAM22,* and *RUNDC3B* by co-transfection of four sgRNAs. All four genes were induced to varying levels (Fig EV5A and B). However, the activity of both *ABCB4* and *RUNDC3B* remained rather low under this quadruple activation condition, compared to their levels after CRISPRa of *ABCB1* alone (*cf*. Fig EV1A). This may be due to competition of the multiple sgRNAs for the available SunCas. We compared the resulting DamID maps to those obtained after activation of *ABCB1* or *ADAM22* alone (Figs 6B and EV5C–E). While the single gene activations resulted in selective detachment of the respective genes, the quadruple activation caused detachment of each gene, with the degree of detachment roughly corresponding to the activity of each gene after activation (Fig EV5A and E). There was no indication that the observed detachment of the four genes involved a more extensive region than a simple combination of their independent detachments. However, it is noteworthy that the very modest activity of *ABCB4* and *RUNDC3B* after quadruple CRISPRa is sufficient to partially detach these two genes from the NL. This suggests that the detachment of these genes is facilitated by the detachment of the nearby *ABCB1* and *ADAM22* genes. Together, these results indicate that co-activation of multiple neighboring genes may lead to more efficient detachment of moderately active genes, but not to a broader detachment of flanking regions.

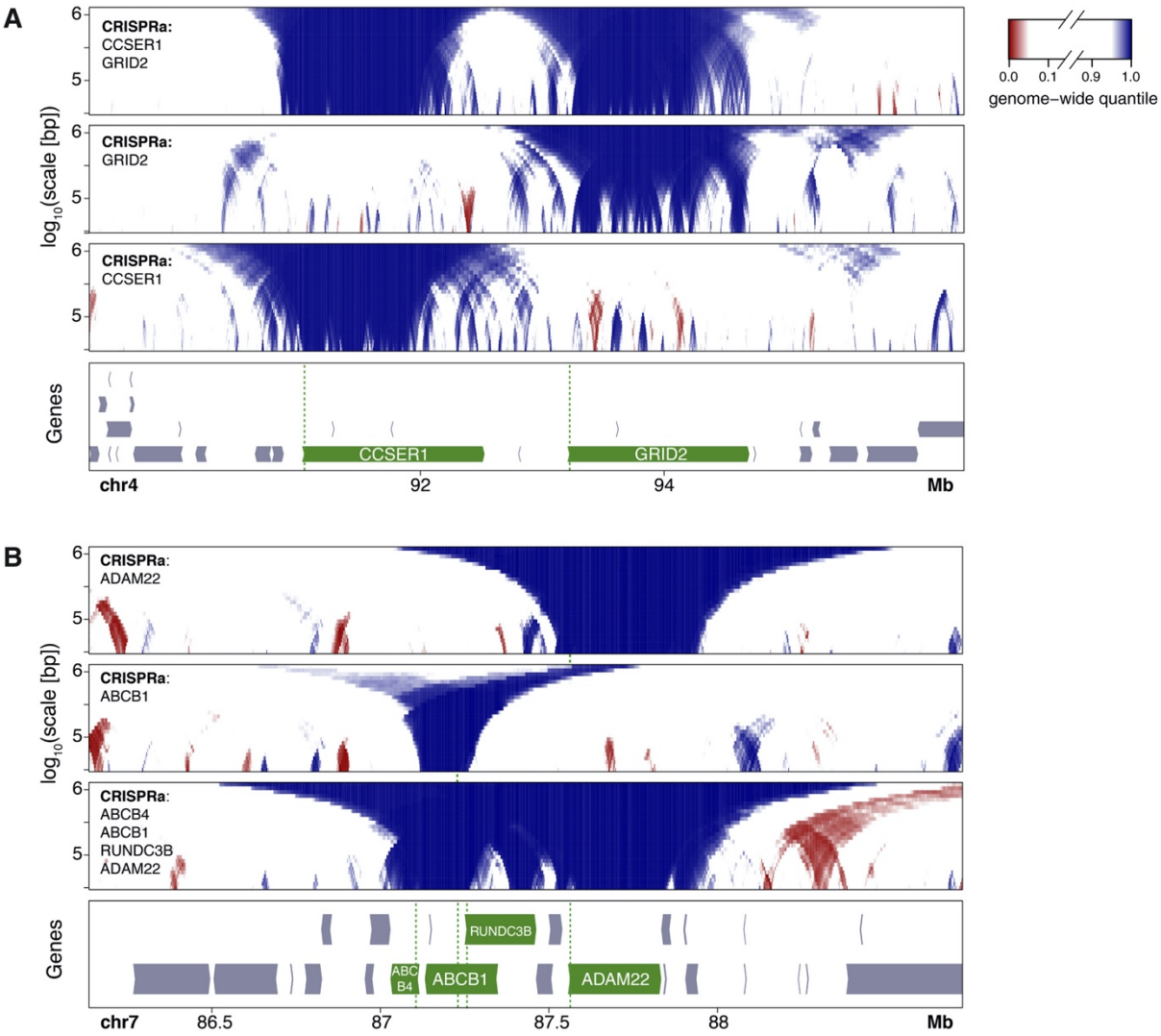


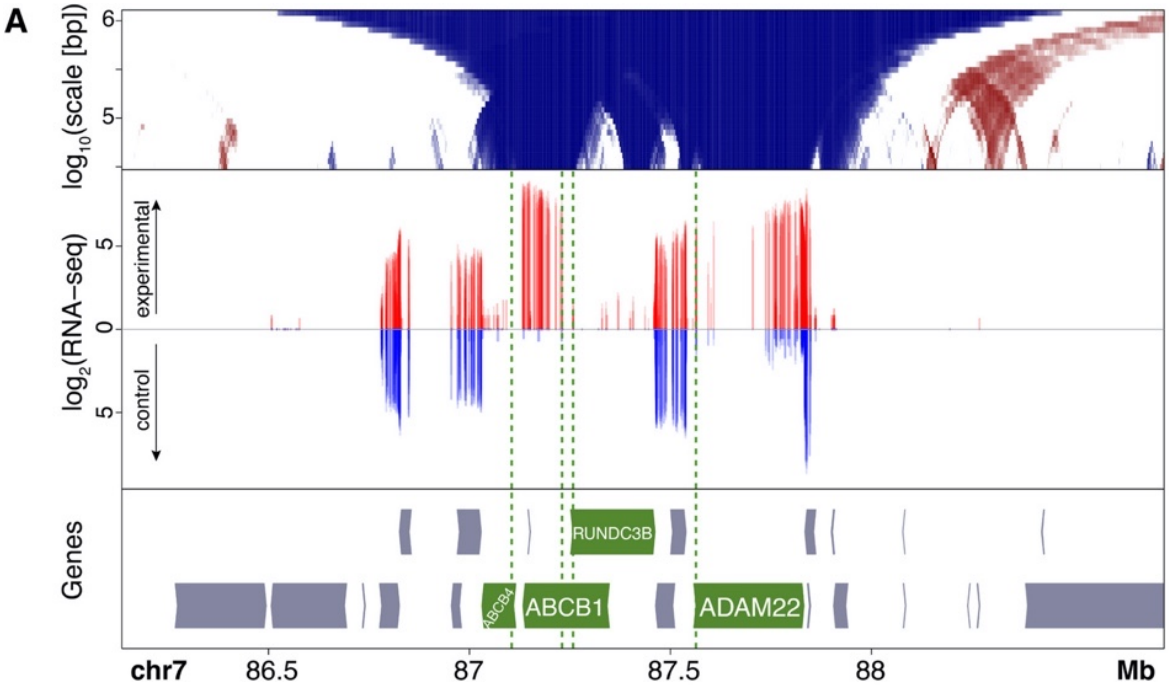
Figure 6

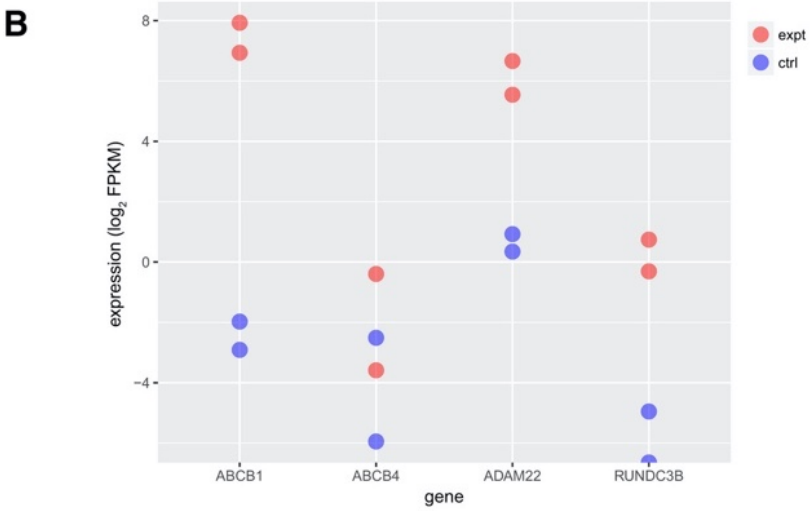
Effects of activation of multiple neighboring genes on NL interactions

A. DamID domainograms of NL interactions after activation of CCSER1, GRID2, or both.

B. DamID domainograms after activation of ADAM22, ABCB1, or the genes ABCB4, ABCB1, ADAM22, and RUNC3B simultaneously.

Data information: All data are from human RPE-1 cells. DamID data of activation of ADAM22 and ABCB1 alone are same as in Fig 4A and B. See Figs EV2 and EV5 for additional visualization of these DamID data.


A
log10(scale [bp])
log2(RNA−seq)
experimental
control
Genes
RUNDC3B
ABCB4
ABCB1
ADAM22
chr7
86.5
87
87.5
88
Mb


B
expt
ctrl
expression (log2 FPKM)
ABCB1
ABCB4
ADAM22
RUNDC3B
gene

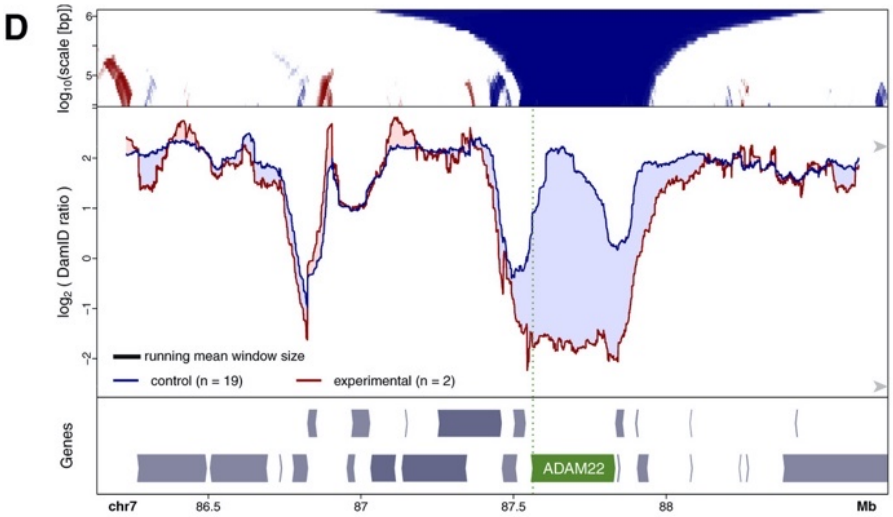

D
log10(scale [bp])
log2 (DamID ratio)
running mean window size
control (n = 19)
experimental (n = 2)
Genes
ADAM22
chr7
86.5
87
87.5
88
Mb

Figure EV 5

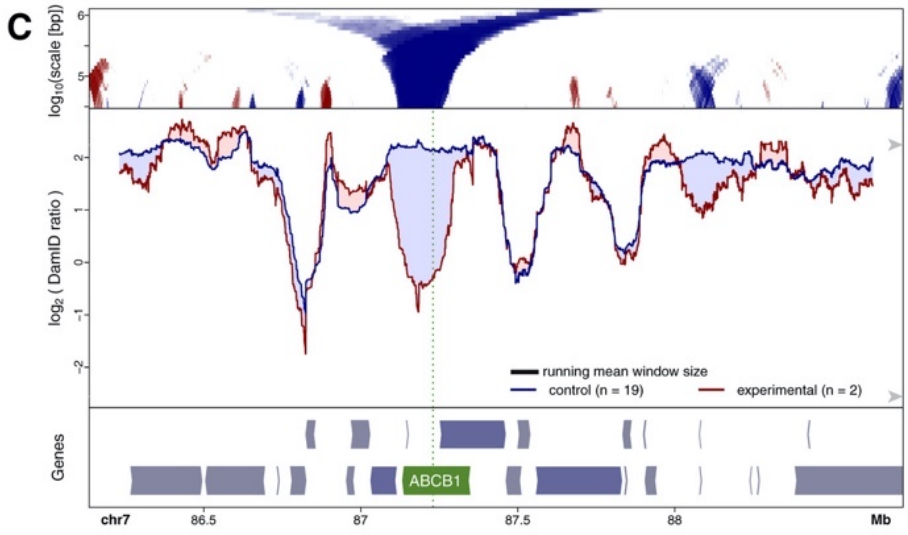
Expression and NL interactions of ABCB4, ABCB1, RUNDC3B, and ADAM22 after combined activation

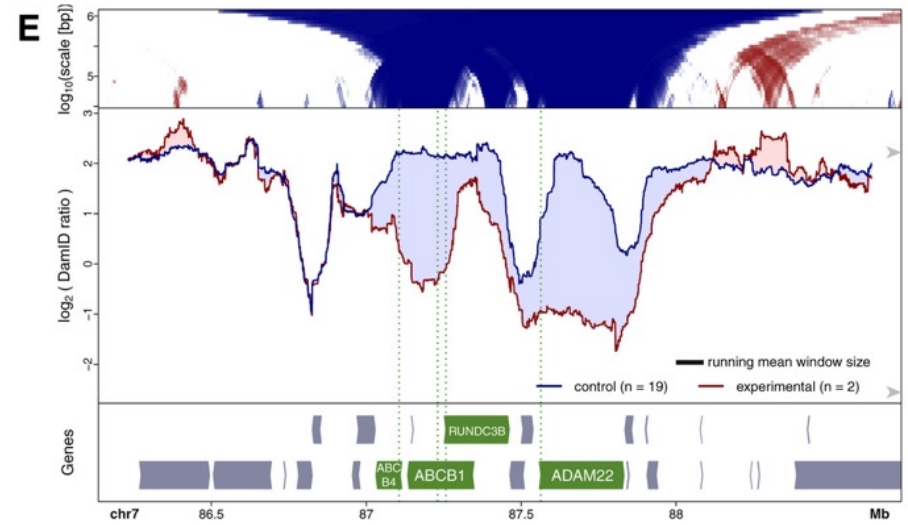
A. Visualization of mRNA expression (middle panel), together with DamID domainogram (top panel) and gene annotation track (bottom panel) in human RPE-1 cells. CRISPRa was done using four sgRNAs simultaneously, one for each promoter. Middle panel shows log2 mRNA expression (normalized read counts in 50 bp bins) as determined by mRNA-seq of CRISPRa-activated cells (red, upward y-axis) compared to untransfected control cells (blue, downward y-axis). Note that only introns give detectable reads. Data are average of two independent replicate experiments each.

B. Quantification of mRNA expression of the four genes, same mRNA-seq data as in (A).

C. DamID profiles after activation of ABCB1 (C), ADAM22 (D), or ABCB4, ABCB1, ADAM22, and RUNDC3B simultaneously (E). Data visualization as in Fig 2A and B.

Data information: In (A, C–E), the CRISPRa-targeted genes are highlighted in green, and the sgRNA target locations are marked by vertical green dotted lines.





Inactivation of genes can promote NL interactions

Our observations so far strongly suggested that the act of transcription is a driving force that localizes genes to the nuclear interior. To test this further, we set out to block transcription by two complementary genetic strategies.

In the first strategy, we aimed to disrupt all transcription in an inter-LAD region (iLAD), to test whether this would lead to increased NL interactions of the entire region. We focused on the mouse genes *Dppa2, Dppa4, Morc1,* and *Morc* (a shorter form of *Morc1* that initiates from an alternative transcription start site). In ES cells, these genes are localized in an approximately 500 kb-sized iLAD. However, this region is NL-associated in mouse neural precursor cells [7], and therefore, it has the potential to become a LAD. We used recently reported [36] F1 hybrid *Cast/129Sv* mES cell clones (named E2 and A6) with a heterozygous triple deletion of the promoters of *DppA2*, *Morc1,* and *Morc* on the *129Sv*-derived chr16. These deletions also stop transcription of the *Dppa4* gene and therefore essentially abolish transcriptional activity in the whole iLAD [36]. Owing to the high density of SNPs that differ between the *129Sv* and *Cast* genomes, we could generate allele-specific DamID maps, enabling us to compare the mutated and wild-type chromosomes.

DamID profiles of the locus revealed that *Morc1* and *Morc* on the mutated chromosome had moved toward the NL in the mutant cells, when compared to control cells carrying an unrelated mutation on a different chromosome (Fig 7A). This effect was not observed for the wild-type Cast chr16 in the same cells (Fig 7B). Interestingly, the region containing *Dppa2* and *Dppa4* was unaffected and clearly remained detached from the NL. This suggests that a transcription-independent detachment mechanism may exist in addition to a transcription-linked mechanism. To determine whether ablation of the most prominent transcript would be sufficient to induce attachment, we also tested a single deletion of the *Morc* promoter (clones A12 and B11), which reduces transcription of *Morc1* by ~2-fold and presumably ablates expression of *Morc,* but does not alter expression of *Dppa2* and *Dppa4* [36]. In the mutated *129Sv*-derived locus, this perturbation resulted in a more restricted increase in NL interactions of *Morc* while the 5′ end of *Morc1* was much less affected (Fig 7C). Again, in the wild-type Cast-derived locus only minor changes were observed between mutated and control clones (Fig 7D). These data show that inactivation of one or more genes in a facultative iLAD can lead *in cis* to locally increased NL interactions of the inactivated genes.

In the second genetic approach, we aimed to truncate a single transcript, to test directly whether transcription elongation is required for detachment from the NL. We chose the 228 kb *Cobl* gene, which is active and locally detached from the NL in mES cells but

inactive and NL-associated in neuronal precursor cells (NPCs), indicating that its detached state is facultative and linked to transcription. We created a heterozygous truncation of the *Cobl* transcription unit in F1 hybrid *Cast/129* mES cells by insertion of a polyadenylation sequence (PAS) in the *129* allele of *Cobl*, 89 kb downstream of the TSS. Analysis of *Cobl* allelic sequence variants in mRNA-seq data confirmed the premature termination of transcription at the *129Sv* allele (Appendix Fig S3). Allele-specific DamID profiles show increased NL interactions of the *129Sv* allele of *Cobl*, particularly downstream of the PAS integration (Fig 7E). This did not occur at the unmodified CAST allele, although some modest changes in NL interactions were detected in the surrounding region (Fig 7F). We conclude that blocking of *Cobl* transcription elongation causes local increases in NL interactions.

Figure 7

Increased NL interactions upon allele-specific transcription inactivation in F1 hybrid mES cells

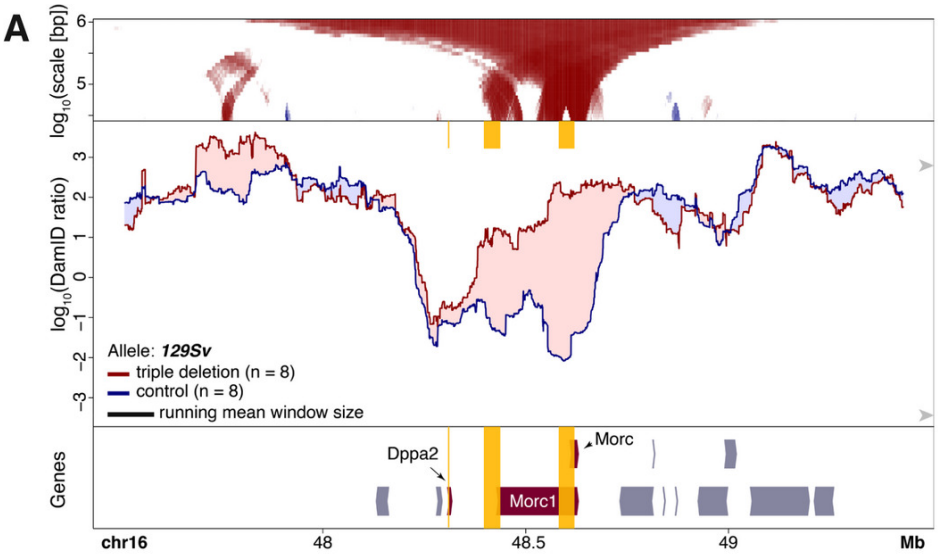
A. DamID profiles of the 129Sv allele of the Morc1 locus with deletions of the promoters of Morc, Morc1, and DppA2 (deletions marked by yellow vertical boxes) and in control cells.

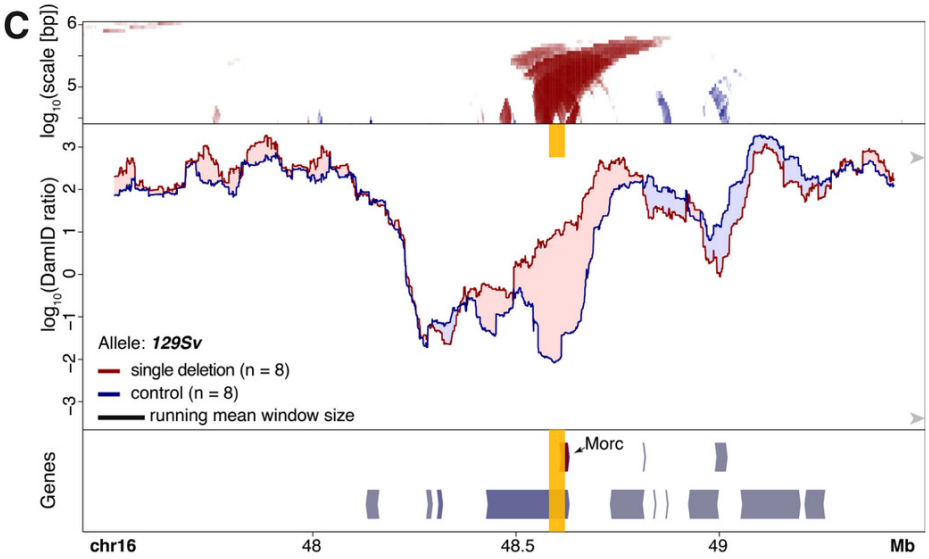
B. Same as (A), but for the non-mutated Cast allele.

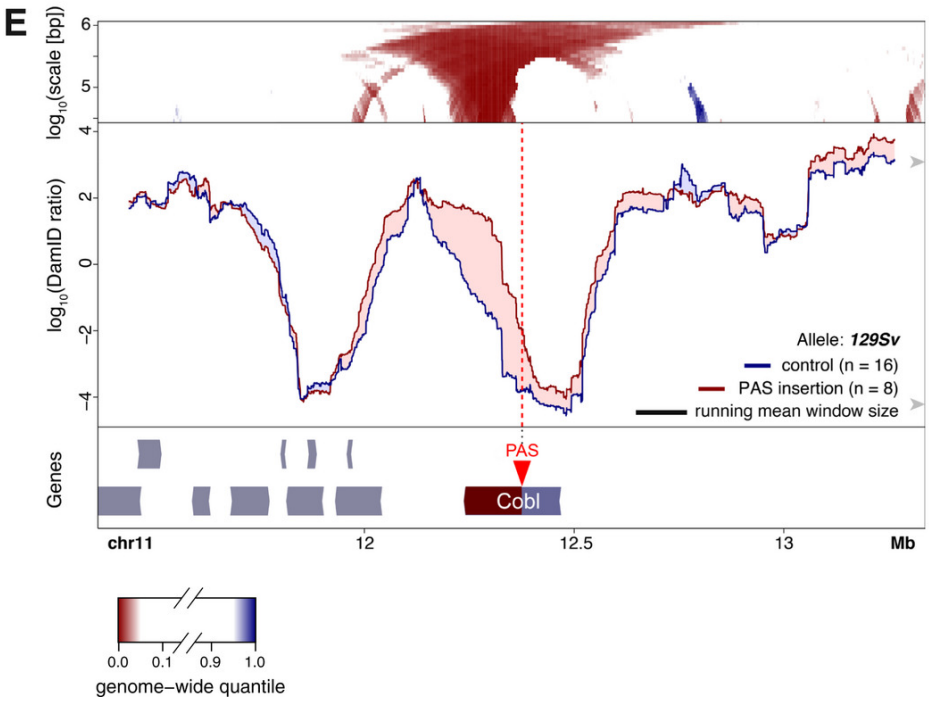
C, D. Same design as (A, B), but with only a single mono-allelic deletion of the Morc promoter (vertical yellow box).

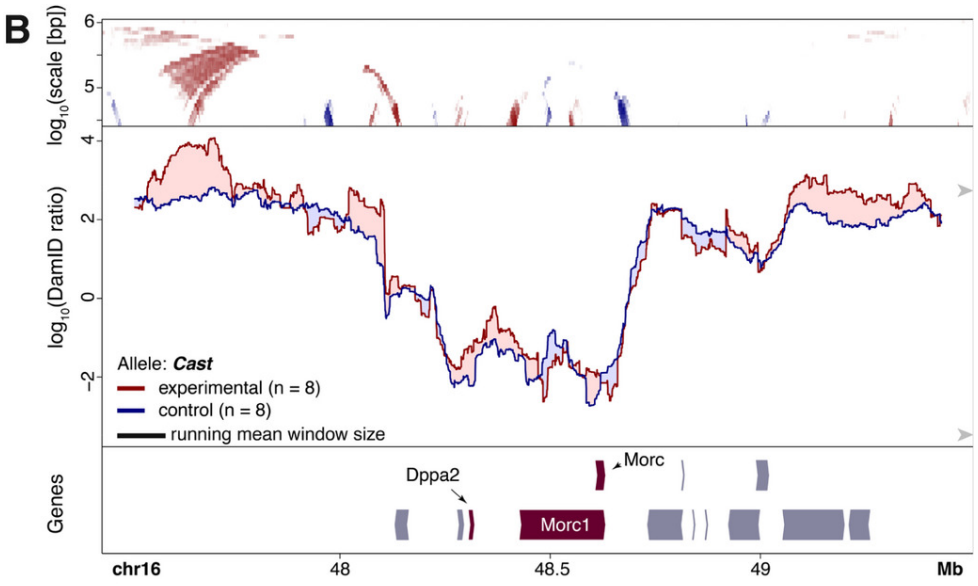
E, F. Effect of PAS insertion on NL interactions of Cobl gene locus. Same design as (A, B) but with insertion of a PAS (located at red triangle and vertical dotted line) that truncates the 129Sv allele of the Cobl transcription unit. 129Sv allele is shown in (E), non-mutated Cast allele in (F).

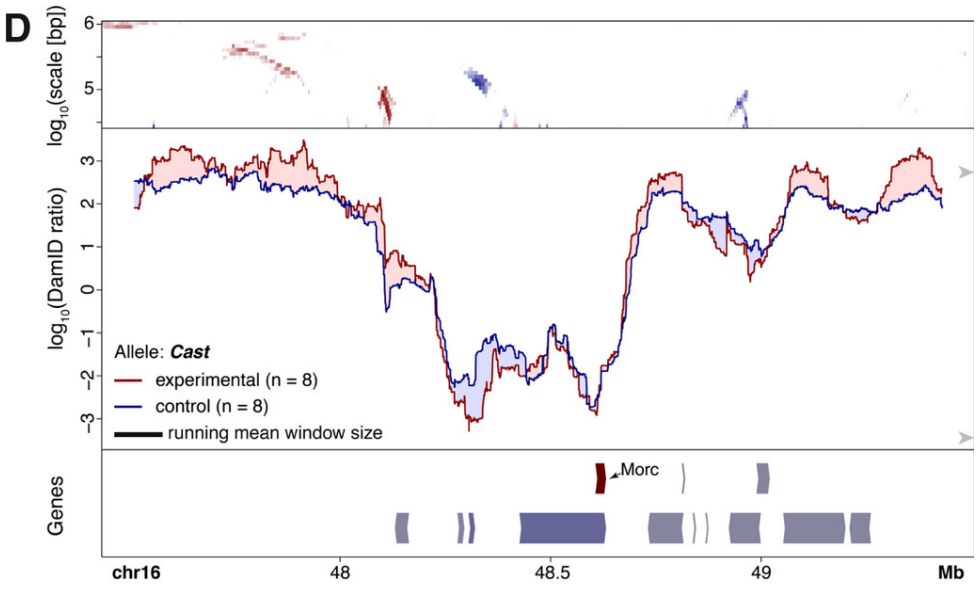
Data information: Clones with Morc1 locus mutations (each assayed in four independent biological replicates) served collectively as control in (E, F), and the clone with the PAS integration (eight independent biological replicates) served as control in (A–D). Visualization of DamID data in all panels is as in Fig 1.

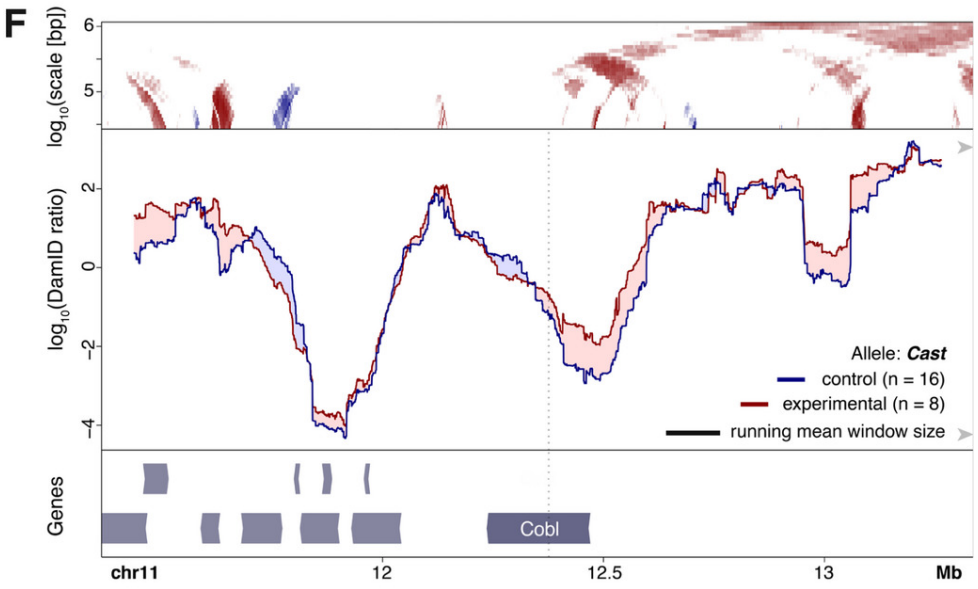

A
log10(scale [bp])
log10(DamID ratio)
Allele: 129Sv
triple deletion (n = 8)
control (n = 8)
running mean window size
Genes
Dppa2
Morc
Morc1
chr16
48
48.5
49
Mb


C
log10(scale [bp])
log10(DamID ratio)
Allele: 129Sv
single deletion (n = 8)
control (n = 8)
running mean window size
Genes
Morc
chr16
48
48.5
49
Mb


E
log10(scale [bp])
log10(DamID ratio)
Allele: 129Sv
control (n = 16)
PAS insertion (n = 8)
running mean window size
Genes
PAS
Cobl
chr11
12
12.5
13
Mb
0.0
0.1
0.9
1.0
genome–wide quantile


B
log10(scale [bp])
log10(DamID ratio)
Allele: Cast
experimental (n = 8)
control (n = 8)
running mean window size
Genes
Dppa2
Morc
Morc1
chr16
48
48.5
49
Mb


D
log10(scale [bp])
log10(DamID ratio)
Allele: Cast
experimental (n = 8)
control (n = 8)
running mean window size
Genes
Morc
chr16
48
48.5
49
Mb


F
log10(scale [bp])
log10(DamID ratio)
Allele: Cast
control (n = 16)
experimental (n = 8)
running mean window size
Genes
Cobl
chr11
12
12.5
13
Mb

Insertion of a small active gene causes moderate detachment from the NL

Finally, complementary to the activation and inactivation of genes in their native context, we tested whether insertion of a highly active transgene into a LAD is sufficient to cause local detachment from the NL. For this purpose, we designed an expression cassette consisting of a transcription unit encoding enhanced green fluorescent protein (eGFP) driven by the strong human PGK promoter, cloned into a PiggyBac transposable element vector (Fig 8A). We integrated this cassette randomly in the genome of F1 hybrid mouse ES cells by co-transfection with PiggyBac transposase. We then isolated clonal cell lines and focused on two with a large number of integrations, reasoning that by random chance several integrations would occur inside LADs. Indeed, by inverse PCR and Tn5 mapping (see Materials and Methods) we found 17 uniquely mappable integrations to be inserted inside LADs, out of a total of 80 in the two cell clones combined (Fig 8B). In comparison with the corresponding wild-type alleles in the same cells, a roughly twofold reduction in average DamID signal was detected around the integration sites, spanning approximately 20 kb on each side (Fig 8C). We conclude that the integrated transcription units tend to detach the directly flanking DNA from the NL, but only partially and within a range of roughly 20 kb.

Figure 8

Effects of a highly active integrated transgene on NL interactions in mES cells

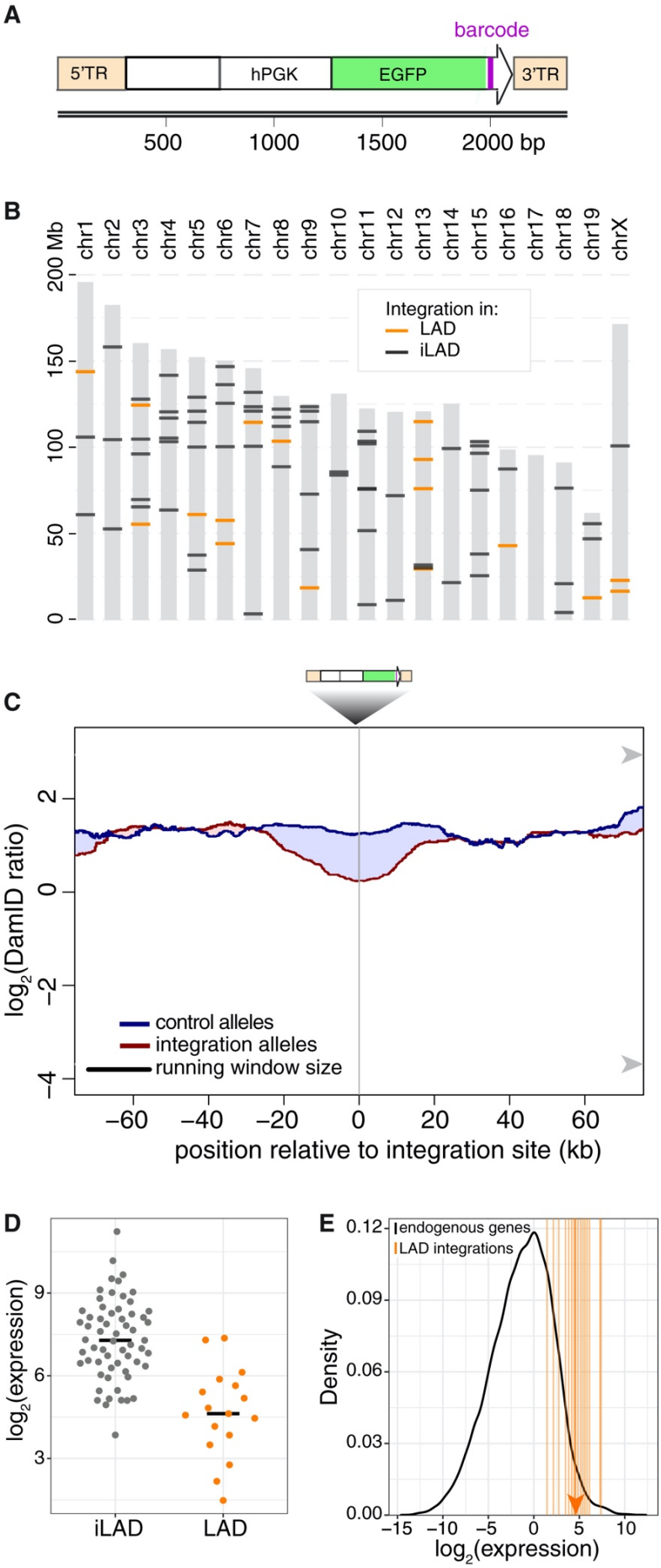
A. Design of the transgene construct, consisting of an enhanced green fluorescent protein (EGFP) transcript, marked at its 3′ end by a random barcode (purple bar) and driven by the human PGK promoter. The construct is flanked by the terminal repeats (5′TR and 3′TR) of the Piggybac transposon that are used for random integration in the genome.

B. Summary of the mapped locations of integrations in the genomes of two F1 hybrid mES cell clonal cell lines. LAD and iLAD integrations are shown in orange and black, respectively.

C. Average DamID profiles across 17 transgene integration sites inside LADs. Blue curve shows alleles without integrations; red curve shows the corresponding alleles with integrations. Shading between the lines shows which curve has the highest value.

D. Relative expression levels of individual barcoded transgenes in LADs and iLADs.

E. Estimated expression levels of integrated transgenes in LADs (vertical orange lines; median value indicated by arrowhead), compared to the distribution of expression levels of all active endogenous genes (black curve).


A
barcode
5'TR
hPGK
EGFP
3'TR
500
1000
1500
2000 bp
B
200 Mb
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chrX
Integration in:
LAD
iLAD
C
log2(DamID ratio)
control alleles
integration alleles
running window size
position relative to integration site (kb)
D
log2(expression)
iLAD
LAD
E
endogenous genes
LAD integrations
Density
log2(expression)

We considered the possibility that our expression cassette was not strong enough to cause more pronounced or extended detachments from the NL. To determine the expression level relative to endogenous genes, we performed RNA-seq and used the barcodes to estimate the expression levels of individual integrations (see Materials and Methods). Transcriptional activity was readily detectable for LAD integrations, although their median expression was about eightfold lower compared to the iLAD integrations (Fig 8D). However, the median expression level of the integrated transgenes inside LADs still ranks approximately in the upper 97th percentile of all active endogenous genes (Fig 8E). Thus, even within LADs, most of the integrated transgenes are expressed at very high levels. These expression levels can be sufficient to reduce NL interactions, but only moderately and locally.

Discussion

Evidence that components of the transcription machinery can affect the spatial organization of the genome is accumulating, but the underlying processes are still poorly understood [37,38]. The data presented here consistently show that activation of genes in LADs leads to detachment from the NL and conversely that inactivation can lead to increased NL contacts. Moreover, the results point to a remarkable flexibility of the chromatin fiber, allowing for the repositioning of individual genes without much effect on flanking DNA.

Several of our results point to a role for transcription elongation in counteracting NL interactions. First, activity-induced detachment from the NL generally extends across the entire activated transcription unit, from the activated promoter until the 3′ end of the gene. We observed this for a wide range of gene sizes. A role for elongation is also strongly supported by premature termination of the active *Cobl* gene by insertion of a PAS, which primarily caused an increase of NL interactions downstream but not upstream of the new termination site. These results are consistent with a study of the *ThymoD* non-coding RNA gene in mouse T-cell progenitors, where insertion of a PAS prevented detachment from the NL as observed by FISH [21]. Conversely, read-through transcription into heterochromatin, elicited by influenza virus NS1 protein, was found to cause relocation from the heterochromatic compartment “B” to the euchromatic compartment “A” [39], which largely correspond to LADs and iLADs, respectively [2]. How transcription elongation may prevent NL interactions remains to be elucidated. It could be a physical effect, for example when a transcribed gene is tethered to a structure in the nuclear interior. It may also be a biochemical effect, such as the removal of particular NL-interacting chromatin proteins by the elongating RNA polymerase complex. We note that our dataset is skewed toward long genes (median length: 366 kb, n = 14) compared to the average genome-wide gene length (which is about 10–15 kb); we cannot rule out that smaller genes often behave differently, although our results for *TRAM1L1* and the transgene insertions (both ~2 kb) suggest that transcription-induced detachment of smaller genes from the NL is also mostly limited to a region of several tens of kb around the transcription unit.

Earlier work found that VP16-induced movement of a LacO repeat toward the nuclear interior could not be blocked by elongation inhibitors [40]. This is not necessarily contradictory to our evidence that supports a role for elongation; it is possible that a transcription activator like VP16 also promotes detachment from the NL independently of transcription elongation. In support of such an additional mechanism, some of the genes that we studied (e.g., Figs 1A and 2A, and 6A) showed the strongest loss of NL interactions near their 5′

end. Similarly, a class of naturally active genes inside LADs exhibits more prominent detachment of their TSS compared to the downstream transcription units [8,41,42]. Furthermore, global tethering of VP64 across all LADs caused virtually no changes in transcription yet triggered loosening of LAD-NL interactions [43], underscoring that VP16 can counteract NL interactions without activating transcription.

That detachment from the NL can involve non-transcribed regions is also suggested by our study of the *PTN* gene. CRISPRa of this gene causes detachment that extends several hundred kb into the neighboring *DGKI* gene, even though the latter gene is not detectably activated. Furthermore, the transcriptionally inactivated Dppa2/4 genes remained mostly dissociated from the NL (Fig 7A). In two earlier reports, relocation from the NL to the nuclear interior was also achieved by tethering of an artificial peptide that induces chromatin decondensation without detectable recruitment of RNA polymerase II [20,40]. It is not understood how this peptide (which is not derived from a naturally occurring protein) exerts this effect, but it suggests that a transcription-independent mechanism of relocation exists in addition to transcription-linked mechanisms. Recent evidence suggests that active chromatin marks such as H3K27ac deposited by p300 may counteract NL interactions [44].

In most cases, the transcriptionally inactive regions adjacent to our activated genes remain relatively unaffected in their NL contacts. Conversely, inhibition of transcription (either by deleting promoters or by insertion of a PAS) leads to increased NL interactions in a very local manner. The latter results were obtained in genomic regions that are facultative LADs, i.e., they may have an intrinsic ability to interact with the NL in the absence of transcription. Genes in constitutive iLADs may lack this ability, either due to spatial constraints or because they lack certain sequence features or chromatin characteristics. It will be of interest to further dissect the molecular mechanisms that underlie the apparent competition between forces that tether chromatin to the NL and forces (such as transcription elongation) that counteract these interactions.

These and previously reported data [1–5] together suggest a balancing act between transcription and LADs: For many genes, LADs pose a repressive environment [8]. This, however, may be overcome by strong transcription activators. Once transcription is active, it causes detachment of the gene from the NL. Possibly this helps to reinforce the active state. It was previously found that a transiently activated gene can remain detached from the NL for several days after the activation signal has subsided [20].

We observed changes from late to early replication timing for all five upregulated genes that we assayed by Repli-Seq. The regions that exhibit shifts in replication timing are roughly 1–2 Mb long, matching in size with the usual span of replication domains observed *in vivo* [22,45]. This suggests that there may be a fundamental minimal size of such domains. This is different from the regions that change NL interactions, which can be smaller. Indeed, the overlap between changes in NL interactions and changes in replication timing, while substantial, was imperfect in most instances. The shifts in replication timing tended to involve a larger region than the shift in NL interactions and were centered around the targeted promoters rather than the entire transcription unit. An exception to this is the PTN locus, where changes in NL interactions and replication timing roughly coincide; this may be due to the putative transcription-independent mechanism discussed above. Together, these results suggest that both NL interactions and replication timing can be modulated by the transcription machinery, but elongation appears to play a more prominent role in counteracting NL interactions, while a signal emanating from activated promoters may evoke a change in replication timing. These distinct but closely linked mechanisms may explain why LADs and late-replicating domains overlap strongly but imperfectly. Together, the large datasets presented here provide a wealth of information on the spatial rewiring of chromosomes in response to transcription activation or inactivation.

Materials and Methods

Cell culture

The RPE-1 cell line stably expressing SunTag-CRISPRa [30] was kindly provided by the R. Medema lab (Netherlands Cancer Institute, Amsterdam) and cultured in DMEM-F12 supplemented with 10% FCS. F121-9 mES cells were kindly provided by J. Gribnau (Erasmus Medical Center, Rotterdam, the Netherlands) and cultured in feeder-free 2i medium according to the 4D Nucleome protocol (https://data.4dnucleome.org/protocols/cb03c0c6-4ba6-4bbe-9210-c430ee4fdb2c/).

TALE-VP64 experiments

TALE-VP64 constructs [20] with Puro resistance marker were kindly provided by Pierre Therizols. The *Sox6* TALE target coordinate is chr7: 116034554–116034570, and the *Nrp1* TALE target coordinate is chr8:128358929–128358945 (mm10 coordinates).
F121-9 cells were transfected with TALE-VP64 constructs targeting *Nrp1* or *Sox2* by electroporation using Lonza Mouse Embryonic Stem Cell Nucleofector™ Kit (VPH-1001) according to the manufacturer's instructions. Cells were selected with Puromycin (1 µg/µl) for 1 week to obtain stable polyclonal cell pools.

Primers

Primer sequences are listed in Tables 2 and 3.

Table 2 Miscellaneous primers used

Oligonucleotide ID	Sequence (5′–3′)
lb877	GACATGGTGCTTGTTGTCCTC
lb982	CTGAGAACGCAGAAGGCTGT
lb991	CTGTTGTCCCACGCATACAG
lb1010	CTGGACCCACCAACTTTGTGG
EB66	CGACAACCACTACCTGAGCA
EB38	CGAACTCCAGCAGGACCATGT
JOYC231	CTCCACTTCCCTCCACCTCT
JOYC232	GAGAGCTTGAACGAAAAACCA
lb563_enrichm_R	CATTGACAAGCACGCCTCAC
lb564_enrichm_F	TAAACCTCGATATACAGACC
lb565_Badapter_ME_5TR_r	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAATTTTACGCAGACTATCTTTCTAG
lb566_Badapter_ME_3TR_f	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTACGTCACAATATGATTATCTTTCTAG

Table 3 Primers for RT–qPCR

Target gene	Oligonucleotide ID	Sequence (5′–3′)
ABCB1	lb667_ABCB1_hs_qPCR_f	CAGTTGAGTGGTGGGCAGAA
	lb668_ABCB1_hs_qPCR_r	GCCTTATCCAGAGCCACCTG
ZNF804B	lb669_ZNF804B_hs_qPCR_f	GCAATCTGAATGTGTTTCTGGA
	lb670_ZNF804B_hs_qPCR_r	ATTCCTTGCTGGAGTTGCT
PTN	lb671_PTN_hs_qPCR_f	CCCAAACCTCAAGAGAAGG
	lb672_PTN_hs_qPCR_r	ACCATCTTCTCAAACTCTCC
SOX6	lb673_SOX6_hs_qPCR_f	TACCAACACTTGTCAGTACCA
	lb674_SOX6_hs_qPCR_r	TCTCTGATTCCATTCTTTGCTG
TRAM1L1	lb675_TRAM1L1_hs_qPCR_f	TCACTGTTGGGTTTCACCT
	lb676_TRAM1L1_hs_qPCR_r	TTTCCAGTAAGGGCATCAG
NLGN1	lb817_qPCR_NLGN1_hs_f	GGTTTCTTGAGTACAGGCG
	lb818_qPCR_NLGN1_hs_r	TGTATGAGATCAAGGAGTCCA
MLK4	lb821_qPCR_KIAA1804_hs_f	GAGGAAGGGCAAGTTTAAGAG
	lb822_qPCR_KIAA1804_hs_r	TTGTGCTGGAAATCTGAAGG
SLC35F3	lb823_qPCR_SLC35F3_hs_f	TTGCCGTTACATATCCCAC
	lb824_qPCR_SLC35F3_hs_r	TGGTGTAGTGATCAATCACTG
ADAM22	lb829_qPCR_ADAM22_hs_f	GTTACTACCAGGGCCATATCC
	lb830_qPCR_ADAM22_hs_r	AGAACATCCCATGAAGTCCG

CRISPRa experiments

sgRNAs were cloned into LentiGuide-Puro vector (Addgene #52963) using restriction enzyme BsmBI, and lentivirus was prepared. RPE-1 cells stably expressing SunTag-CRISPRa were infected with LentiGuide virus and selected with 10 µg/µl puromycin for 1 week to obtain stable polyclonal cell pools.

PAS integration

A PAS was inserted by in-frame integration of a blasticidin resistance (BlastR) cassette followed by the PAS into the *Cobl* gene. For this purpose, sgRNA sequence AGTCATCTGTGCGAAGTGTG was cloned into Blast-TIA vector [46] (kindly supplied by the Brummelkamp lab, Netherlands Cancer Institute) via BbsI restriction digestion. Cells were transfected with the resulting Blast-TIA vector co-transfected with Cas9 expression vector pX330 (Addgene #42230) by nucleofection and subjected to selection by culturing in the presence of 10 µg/µl blasticidin for 1 week. Clones were picked and screened for correct integration of the BlastR cassette by PCR with primers lb877 and lb982. Heterozygosity of the integration was confirmed by PCR using primers lb982 and lb991, and the *129Sv* allele was identified as the targeted allele by PCR using primers lb877 and lb1010, followed by Sanger sequencing with the same primers.

Repli-seq

Repli-seq was performed as described [31]. Sequencing was done on a NovaSeq 6000 system (Illumina), 50-bp read length.

mRNA-seq

As previously described [47], mRNA-seq was performed as follows. Quality and quantity of the total RNA were assessed by the 2100 Bioanalyzer using a Nano chip (Agilent, Santa Clara, CA). Total RNA samples having RIN>8 were subjected to library generation. Strand-specific libraries were generated using the TruSeq Stranded mRNA sample preparation kit (Illumina Inc., San Diego, RS-122-2101/2) according to the manufacturer's instructions (Illumina, Part # 15031047 Rev. E). Briefly, polyadenylated RNA from intact total RNA was purified using oligo-dT beads. Following purification, the RNA was fragmented, random-primed, and reverse-transcribed using SuperScript II Reverse Transcriptase (Invitrogen, part # 18064-014) with the addition of actinomycin D. Second-strand synthesis was performed using polymerase I and RNaseH with replacement of dTTP for dUTP. The generated cDNA fragments were 3′ end adenylated and ligated to Illumina paired-end sequencing adapters and subsequently amplified by 12 cycles of PCR. The libraries were analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent, Santa Clara, CA), diluted and pooled equimolar into a multiplex sequencing pool, and stored at –20°C. The libraries were sequenced with 65 base single reads on a HiSeq2500 using V4 chemistry (Illumina Inc., San Diego). Reads were aligned to hg19 or mm10 using TopHat version 2.1, with Ensembl genome build 75. For *Cobl* PAS integration experiments in the F1 hybrid ES cells, mRNA-seq reads were aligned to mm10 using STAR [48].

RT–qPCR

Cells were collected in TRIsure and total RNA was extracted using PureLink RNA Mini Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. RNA was reverse-transcribed using Tetro Reverse Transcriptase (Bioline) with Oligo(dT)20 primers (Thermo Fisher Scientific) according to the manufacturer's instructions. qPCR was performed using SensiFast no-ROX mix (Bioline) in a 10 µl reaction. Primers are listed in Table 3.

Generation and mapping of random integrations

The hPGK-EGFP cassette was derived from TRIP vector pPTK-Gal4-mPGK-Puro-IRES-eGFP-sNRP-pA [49] by replacing mPGK-Puro-IRES with the human PGK promoter using restriction enzyme cloning with SalI and NcoI. Generation of a barcoded plasmid pool and integration into F121-9 mES cells was performed as described [49]. Clones with high EGFP expression were sorted by FACS and screened for high integration copy number by qPCR with EGFP-specific primers EB66 and EB38, using *Lbr*-specific primers (JOYC231 and JOYC232) for normalization.

Mapping of integrations without linking to barcodes was done by Tagmentation as described [50] with minor modifications: Before PCR for Tn5 adapters, linear amplification of PiggyBac integrations was performed using primers lb565 or lb566 for mapping in reverse or forward orientation respectively. Linear amplification was performed using 0.5 U Phusion polymerase (Bioline) in a 20 µl reaction with Phusion GC-rich buffer, 1 mM dNTPs, 50 nM primer. Reaction was incubated at 98°C for 30 s, then 45 cycles of 98°C for 8 s, 60°C for 5 s and 72°C for 30 s followed by a final step at 72°C for 20 s. For PCR amplification, PiggyBac-specific primers lb565 or 566 were used for mapping in reverse or forward orientation, respectively.

To process the tagmentation mapping reads, the Tn5 adaptor sequence and PiggyBac primer sequence at the ends of the paired-end reads were removed using an adaptation of *cutadapt v1.11*. The genomic part of the sequence was mapped to strain-specific versions of GRCm38 release 68 from Ensembl using *bowtie v2.3.4.1* with mapping set to "*very-sensitive*". To create these strain-specific genomes, SNP information was downloaded from the Mouse Genomes Project [51]; http://www.sanger.ac.uk/science/data/mouse-genomes-project) as VCF files "CAST_EiJ.mgp.v5.snps.dbSNP142.vcf.gz" and "129S1_SvImJ.mgp.v5.snps.dbSNP142.vcf.gz" (version 1 May 2015). *Bcftools* was used to incorporate all SNPs into the GRCm38 reference genome. After mapping to strain-specific genomes, bam files were compared, and for each read, the alignment with the highest alignment score (AS) was used. When the AS was identical, a random choice was made. Read-pairs aligning in opposite orientation and <1,500 bp apart were converted to genomic regions using the *bamToBed* from *bedtools* and *awk,* covering both reads as well as the region in between. *Genomecov* from *bedtools* and *awk* was used to combine regions and calculate coverage. Integration sites were called by combining regions from PCRs from both transposon arms using *closest* from *bedtools* and *awk*. Regions on opposite strands that were at most 5 bp apart were regarded to represent an integration. Next, the allele of the integration was determined by using *mpileup* from *Samtools* [52] *v1.5* with a maximum depth of 50 to count the number of mismatch positions over the complete region compared to

both of the strain-specific GRCm38 modifications. Each position with the allele of the strain-specific genome occurring in a ratio < 0.5 was considered a mismatch position. The allele with the lowest number of mismatch positions was then considered the allele of integration. In case of equal number of mismatch positions, the integration allele was classified as ambiguous. We selected only putative integration sites with at least 1 read having a mapping quality > 10 and > 500 reads mapped on both sides of it.
In addition, allele-specific mapping of the integrations and linking to their barcodes was performed by inverse PCR as described [49], except that the mapping of reads was confined to regions that were initially found by tagmentation mapping. Tagmentation alone identified 50 integrations for clone CM1407 and 56 for clone CM1420, of which 37 and 43 were linked to a single, unique barcode, respectively.

Expression analysis of ES cell clones with random integrations

Clones CM1407 and CM1420 were subjected to mRNA-seq as above. For comparison between eGFP and endogenous mRNA expression, a fasta entry for eGFP was added to the mouse genome version mm10 chromosomes 1-19, X, Y, and M without alternative contigs. The annotation of eGFP transcript was also added to gencode version M18. STAR [48] version 2.6.0c was used to align the cDNA reads to this modified reference genome, and for each transcript, reads were counted. DESeq2 [53] was used to calculate fragments per kilobase million (FPKM) values for each gene, including eGFP.
In addition, barcode-specific expression was determined using sequencing of the barcodes in cDNA similar to the standard TRIP protocol [49]. To discriminate between genuine barcodes and sequencing errors of these barcodes, *starcode* [54] was used. Unlike the standard TRIP protocol, reads were not normalized by gDNA counts. Finally, the eGFP FPKM was scaled by the number of integrations for each clone in order to determine the average eGFP expression per integrated reporter. These numbers were 52 for clone CM1417 and 55 for clone CM1420; barcodes were counted 2 times when found at 2 integration sites.
Lamina-associated domains coordinates in mouse ES cells were obtained from [7] and adjusted to mm10 using the LiftOver tool [55]. To estimate FPKM's for integrations in LADs and iLADs separately, barcodes were used that had a unique location according to a combination of iPCR and tagmentation mapping. In total, 17 barcodes could be confidently linked to LAD locations. To determine LAD and iLAD-specific expression levels, the average eGFP FPKM per integration was scaled by the median LAD and iLAD expression. Finally, the percentile of eGFP FPKM relative to endogenous active genes was calculated by counting the number of genes with higher FPKM than the eGFP estimation, divided by the total number of active genes (defined as genes with FPKM > 0).

DamID-seq

DamID-seq was performed as described [56] with minor modifications. Dam fused to human LMNB1 protein (Dam-LMNB1) or unfused Dam were expressed in cells by lentiviral transduction [57]. Three days after infection, cells were collected for genomic DNA (gDNA) isolation. gDNA was pre-treated with SAP (10 U, New England Biolabs #M0371S) in CutSmart buffer in a total volume of 10 µl at 37°C for 1 h, followed by heat inactivation at 65°C for 20 min to suppress signal from apoptotic fragments. This gDNA was then digested with DpnI (10 U, New England Biolabs #R0176L) in CutSmart buffer in a total volume of 10 µl at 37°C for 8 h followed by heat inactivation at 80°C for 20 min. Fragments were ligated to 12.5 pmol DamID adapters using T4 ligase (2.5 U, Roche #10799009001) in T4 ligase buffer in a total volume of 20 µl incubated at 16°C for 16 h. The reaction was heat-inactivated for 10 min at 65°C. Products were then digested with DpnII to destroy partially methylated fragments. DpnII buffer and DpnII (10 U, New England Biolabs #R0543L) were added in a total volume of 50 µl and incubated at 37°C for 1 h. Next, 8 µl of DpnII-digested products was amplified by PCR with MyTaq Red Mix (Bioline #BIO-25044) and 1.25 µM primers Adr-PCR-Rand1 in a total volume of 40 µl. PCR settings were 8 min at 72°C (1×) followed by 20 s at 94°C, 30 s at 58°C, 20 s at 72°C (24× for Dam, 28× for Dam-LMNB1 samples) and 2 min at 72°C (1×). Remaining steps were performed as previously described. Samples were sequenced on an Illumina HiSeq2500.

Processing of RPE-1 and ES cell DamID data

First, the constant DamID adapter was trimmed from the 65-bp single-end reads using *cutadapt* [58] version 1.11 and custom scripts. The remaining sequence starting with GATC was mapped to hg19 with bowtie2 [59] version 2.2.6. Uniquely mapped reads (filtered for bowtie's XS-tag) were then assigned to individual gDNA sequences between two GATC motifs (referred to as GATC fragments), which are the units of further data processing and analysis because Dam-only methylates GATC motifs. Further processing and analysis were done in R [60] versions 3.4–3.6 using Bioconductor [61], in particular the packages GenomicRanges [62] and Sushi [63].

Replicate experiments were combined by summing the reads for each GATC fragment. Hence, experiments with more reads were weighed proportionally stronger than experiments with fewer reads. Extremely high read counts of individual GATC fragments (> 100 times the genome-wide average) were assumed to be due to PCR artifacts; these read counts were replaced with the genome-wide average read count. Next, smoothing was applied by summing read counts over a running window of 201 consecutive GATC fragments. A pseudocount of 30 was added to each window. The ratio Dam-lamin B1/Dam-

only was calculated for each window and log_2-transformed. Finally, the log_2 ratios were normalized by subtracting the genome-wide average log_2 ratio.
When comparing experimental and control DamID log ratios in genome-wide scatterplots, we noticed modest systematic biases and skews (visible as point clouds that were somewhat banana-shaped rather than cigar-shaped) or differences in the dynamic ranges of the DamID values that are likely to be of technical nature (under the assumption that CRISRPa activation of a single gene is unlikely to cause a genome-wide systematic effect). We estimated such skews empirically by applying a lowess fit (span = 0.5) to the experimental ~ control comparison of a random selection of 50,000 GATC fragments and then used this fit to correct the genome-wide comparison. This effectively removed genome-wide biases, thereby enhancing the sensitivity to detect local changes in NL contacts around the targeted genes.

Processing of F1 hybrid mouse ES cell DamID data

For DamID on F1 hybrid *129/Cast* mouse ES cells, 150 or 200 nt single reads were trimmed to remove the DamID adaptor sequence and then strain-specifically mapped to mm10 with WASP [64] using bowtie2 and VCF files from the mouse genomes project ([51]; https://www.sanger.ac.uk/science/data/mouse-genomes-project; version 5). Data were further processed as described for RPE-1 cells, except that a smoothing window size of 301 instead of 201 GATC fragments was applied.

Domainograms

The domainograms in this study are related to those reported previously [25,26], but do not show estimated *P*-values, which are not easily calculated for our experimental design. Rather, for a given window size, they show the ranking of changes in DamID log ratios (experimental minus control) relative to windows of the same size genome-wide. Briefly, in a window of *w* neighboring GATC fragments, the difference in mean DamID log-ratio is calculated between the experimental and control samples. This is done for all possible windows of size *w* genome-wide. Next, windows in which both experimental and control sample showed only baseline DamID signals (i.e., both log ratios are in the respective lower 0.3 quantiles genome-wide) are discarded. This is done because baseline fluctuations can appear strong on a logarithmic scale but are generally of minor amplitude on a linear scale and therefore unlikely to be of biological relevance. The remaining windows are ranked by their log-ratio differences; ranks < 5% or > 95% are visualized by blue or red color scales, respectively. This is repeated for 28 different window sizes *w* that are logarithmically ranging from 67 and 2,917 GATC fragments, i.e., from ~30 kb to ~1 Mb.

Data analysis of Repli-seq samples

Repli-seq reads from early and late-replicating fractions were mapped and processed in the same way as DamID reads, using the same smoothing window size. Instead of Dam-lamin B1/Dam-only, the ratio late/early replication was calculated.

Hi-C data analysis

Hi-C data from wild-type RPE-1 cells are from [34,35] and visualized using Juicebox 1.8.8 [65].

Data availability

The datasets (and computer code) produced in this study are available in the following databases:

Sequencing reads and processed data of DamID, Repli-seq and RNA-seq experiments: Gene Expression Omnibus GSE133275 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133275).

Data analysis code: https://github.com/vansteensellab/LAD_rewiring.

Author contributions

LB: conceived and designed study, conducted majority of experiments, initial data analysis, and wrote manuscript. PAZ: performed Repli-seq experiments and initial Repli-seq data analysis. DP-H: performed experiments and data analysis. TS: processing of DamID data. CL: performed data analysis. JS: *Morc1* locus deletions and initial data analysis. DMG: supervised Repli-seq experiments, generation of *Morc1* locus deletions and initial data analysis. BS: designed study, performed coding and data analysis, wrote manuscript, and supervised project.

Acknowledgments

We thank the NKI Genomics, Flow Cytometry, and RHPC core facilities, as well as Tom Rieuwerts and Ludo Pagie for technical assistance and Lorenzo Bombardelli for Tn5 protein. We thank Andrew Belmont, Jian Ma, and other members of the 4DN Center for Nuclear Cytomics for helpful discussions. Supported by NIH Common Fund “4D Nucleome” Program grant U54DK107965 (BvS and DMG) and European Research Council Advanced Grant 694466 (BvS). The Oncode Institute is partly supported by KWF Dutch Cancer Society.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Gonzalez-Sandoval A, Gasser SM. On TADs and LADs: spatial control over gene expression. *Trends Genet*. 2016;32:485-495.

2. Steensel B, Belmont A. Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell*. 2017;169:780-791.

3. Leeuw R, Gruenbaum Y, Medalia O. *Nuclear Lamins: Thin Filaments with Major Functions.Trends Cell Biol 28:34–45*.; 2018.

4. Kim Y, Zheng X, Zheng Y. Role of lamins in3D genome organization and global gene expression. *Nucleus*. 2019;10:33-41.

5. Lochs SJA, Kefalopoulou S, Kind J. Lamina associated domains and gene regulation in development and cancer. *Cells*. 2019;8:E271.

6. Guelen L, Pagie L, Brasset E, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*. 2008;453(7197):948-951. doi:10.1038/nature06947

7. Peric-Hupkes D, Meuleman W, Pagie L, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell*. 2010;38(4):603-613. doi:10.1016/j.molcel.2010.03.016

8. Leemans C, van der Zwalm MCH, Brueckner L, et al. Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs. *Cell*. 2019;177(4):852-864.e14. doi:10.1016/j.cell.2019.03.009

9. Lund E, Oldenburg AR, Delbarre E, et al. Lamin A/C-promoter interactions specify chromatin state-dependent transcription outcomes. *Genome Res*. 2013;23:1580-1589.

10. Robson MI, Las Heras JI, Czapiewski R, et al. Tissue-specific gene repositioning by muscle nuclear membrane proteins enhances repression of critical developmental genes during myogenesis. *Mol Cell*. 2016;62:834-847.

11. Robson MI, Las Heras JI, Czapiewski R, Sivakumar A, Kerr ARW, Schirmer EC. Constrained release of lamina-associated enhancers and genes from the nuclear envelope during T-cell activation facilitates their association in chromosome compartments. *Genome Res*. 2017;27:1126-1138.

12. Shevelyov YY, Lavrov SA, Mikhaylova LM, et al. The B-type lamin is required for somatic repression of testis-specific gene clusters. *Proc Natl Acad Sci*. 2009;A106:3282-3287.

13. Kohwi M, Lupton L JR, SL M, MR D, C.Q. Developmentally regulated subnuclear genome reorganization restricts neural progenitor competence in Drosophila. *Cell*. 2013;152:97-108.

14. Chen H, Zheng X, Zheng Y. Age-associated loss of lamin-B leads to systemic inflammation and gut hyperplasia. *Cell*. 2014;159:829-843.
15. Finlan LE, Sproul D, Thomson I, et al. Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. *PLOS Genet*. 2008;4(3):e1000039. doi:10.1371/journal.pgen.1000039
16. Kumaran RI, Spector DL. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol*. 2008;180(1):51-65. doi:10.1083/jcb.200706060
17. Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. 2008;452(7184):243-247. doi:10.1038/nature06727
18. Dialynas G, Speese S, Budnik V, Geyer PK, Wallrath LL. The role of Drosophila Lamin C in muscle function and gene expression. *Dev Camb Engl*. 2010;137(18):3067-3077. doi:10.1242/dev.048231
19. Tumbar T, Belmont AS. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat Cell Biol*. 2001;3(2):134-139. doi:10.1038/35055033
20. Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science*. 2014;346(6214):1238-1242. doi:10.1126/science.1259587
21. Isoda T, Moore AJ, He Z, et al. Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell*. 2017;171:103-119.
22. Pope BD, Ryba T, Dileep V, et al. Topologically associating domains are stable units of replication-timing regulation. *Nature*. 2014;515:402-405.
23. Harr JC, Luperchio TR, Wong X, Cohen E, Wheelan SJ, Reddy KL. Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biol*. 2015;208:33-52.
24. Kind J, Pagie L, de Vries SS, et al. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell*. 2015;163(1):134-147. doi:10.1016/j.cell.2015.08.040
25. de Wit E, Braunschweig U, Greil F, Bussemaker H, Steensel B. Global chromatin domain organization of the Drosophila genome. *PLoSGenet*. 2008;4:e1000045.
26. Tolhuis B, Blom M, Kerkhoven RM, et al. *Interactions among Polycomb Domains Are Guided by Chromosome Architecture*.; 2011.

27. Greil F, Moorman C, Steensel B. DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. *Methods Enzym*. 2006;410:342-359.

28. Filion GJ, van Bemmel JG, Braunschweig U, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*. 2010;143(2):212-224. doi:10.1016/j.cell.2010.09.009

29. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014;159(3):635-646. doi:10.1016/j.cell.2014.09.039

30. Tame MA, Manjon AG, Belokhvostova D, Raaijmakers JA, Medema RH. TUBB3 overexpression has a negligible effect on the sensitivity to taxol in cultured cell lines. *Oncotarget*. 2017;8:71536-71547.

31. Marchal C, Sasaki T, Vera D, et al. Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. *NatProtoc*. 2018;13:819-839.

32. Dixon JR, Selvaraj S, Yue F, et al. Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature*. 2012;485(7398):376-380. doi:10.1038/nature11082

33. Nora EP, Lajoie BR, Schulz EG, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485:381-385.

34. Darrow EM, Huntley MH, Dudchenko O, et al. Deletion of DXZ4on the human inactive X chromosome alters higher-order genome architecture. *Proc Natl Acad Sci USA*. 2016;113:E4504-E4512.

35. Darrow EM, Huntley MH, Lieberman Aiden E. *Gene Expression Omnibus GSE71831(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE71831*. DATASET; 2016.

36. Sima J, Chakraborty A, Dileep V, et al. Identifying cis elements for spatiotemporal control of mammalian DNA replication. *Cell*. 2019;176:816-830.

37. Steensel B, Furlong E. The role of transcription in shaping the spatial organization of the genome. *Nat Rev Mol Cell Biol*. 2019;20:327-337.

38. Vermunt MW, Zhang D, Blobel GA. *The Interdependence of Gene-Regulatory Elements and The3D Genome.J Cell Biol 218:12–26*.; 2019.

39. Heinz S, Texari L, Hayes MGB, et al. Transcription elongation can affect genome3D structure. *Cell*. 2018;174:1522-1536.

40. Chuang C-H, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS. Long-range directional movement of an interphase chromosome site. *Curr Biol CB*. 2006;16(8):825-831. doi:10.1016/j.cub.2006.03.059

41. Luperchio TR, Sauria MEG, Wong X, et al. *Chromosome conformation paints reveal the role of lamina association in genome organization andregulation.bioRxivhttps://doi.org/10.1101/122226[PREPRINT*.; 2017.
42. Wu F, Yao J. Identifying novel transcriptional and epigenetic features of nuclear lamina-associated genes. *Sci Rep*. 2017;7(100).
43. Kind J, Pagie L, Ortabozkoyun H, et al. Single-cell dynamics of genome-nuclear lamina interactions. *Cell*. 2013;153(1):178-192. doi:10.1016/j.cell.2013.02.028
44. Cabianca DS, Munoz-Jimenez C, Kalck V, et al. Active chromatin marks drive spatial sequestration of heterochromatin in C. elegans nuclei. *Nature*. 2019;569:734-739.
45. Hiratani I, Ryba T, Itoh M, et al. *Global Reorganization of Replication Domains during Embryonic Stem Cell Differentiation. PLoS Biol 6:E245*.; 2008.
46. Lackner DH, Carre A, Guzzardo PM, et al. A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nat Commu*. 2015;6(10237).
47. Gogola E, Duarte AA, Ruiter W JR, et al. Selective lossof PARG restores PARylation and counteracts PARP inhibitor-mediated synthetic lethality. *Cancer Cell*. 2018;33:1078-1093.
48. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15-21. doi:10.1093/bioinformatics/bts635
49. Akhtar W, Pindyurin AV, Jong J, et al. Using TRIP for genome-wide position effect analysis in cultured cells. *Nat Protoc*. 2014;9:1255-1281.
50. Stern DL. *Tagmentation-Based Mapping (TagMap) of Mobile DNA Genomic Insertion Sites.BioRxivhttps://Doi.Org/10.1101/037762[PREPRINT*.; 2017.
51. Keane TM, Goodstadt L, Danecek P, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. 2011;477:289-294.
52. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078-2079.
53. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15. doi:10.1186/s13059-014-0550-8
54. Zorita E, Cusco P, Filion GJ. Starcode: sequence clustering based on all-pairs search. *Bioinformatics*. 2015;31:1913-1919.
55. Hinrichs AS, Karolchik D, Baertsch R, et al. The UCSC genome browser database: update2006. *Nucleic Acids Res*. 2006;34:D590-D598.
56. Brueckner L, van Arensbergen J, Akhtar W, Pagie L, van Steensel B. High-throughput assessment of context-dependent effects of chromatin proteins. *Epigenetics Chromatin*. 2016;9(1):43. doi:10.1186/s13072-016-0096-y

57. Vogel MJ, Peric-Hupkes D, Steensel B. Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nat Protoc*. 2007;2:1467-1478.

58. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal*. 2011;17:10-12.

59. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie2. *Nat Methods*. 2012;9:357-359.

60. R Core Team. *R: a language and environment for statistical computing.R Core Team*.; 2017. https://www.R-projectorg

61. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12:115-121.

62. Lawrence M, Huber W, Pages H, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol*. 2013;9:e1003118.

63. Phanstiel DH, Boyle AP, Araya CL, Snyder MP. *Sushi.R: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures.Bioinformatics30:2808–2810*.; 2014.

64. van de Geijn B, McVicker G, Gilad Y, Pritchard J. WASP: allele-specific software for robust molecular quantitative trait locus discovery. *NatMethods*. 2015;12:1061-1063.

65. Durand NC, Robinson JT, Shamim MS, et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst*. 2016;3:99-101.

Appendix

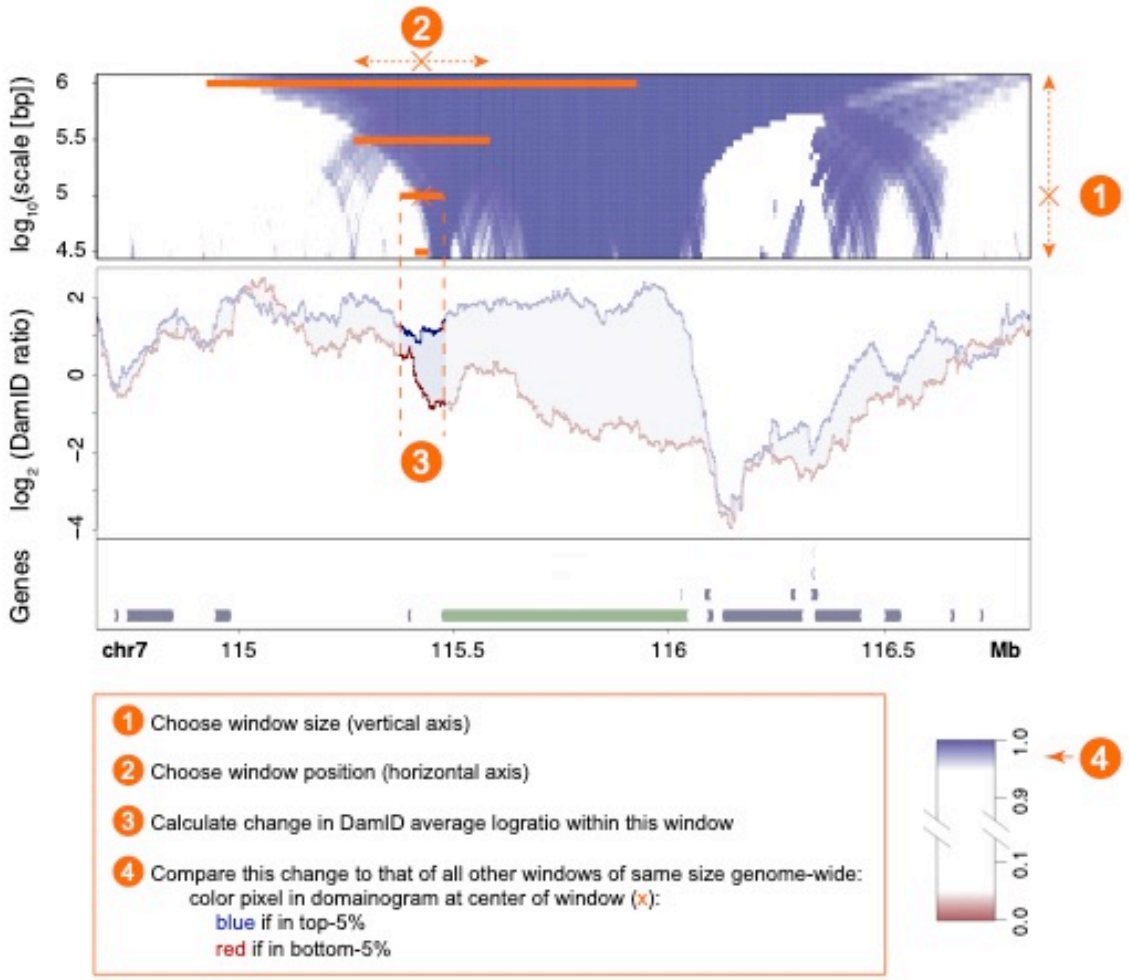


Figure S 1

Explanation of domainogram principle. Text box at the bottom describes the major steps followed to calculate and plot a domainogram.

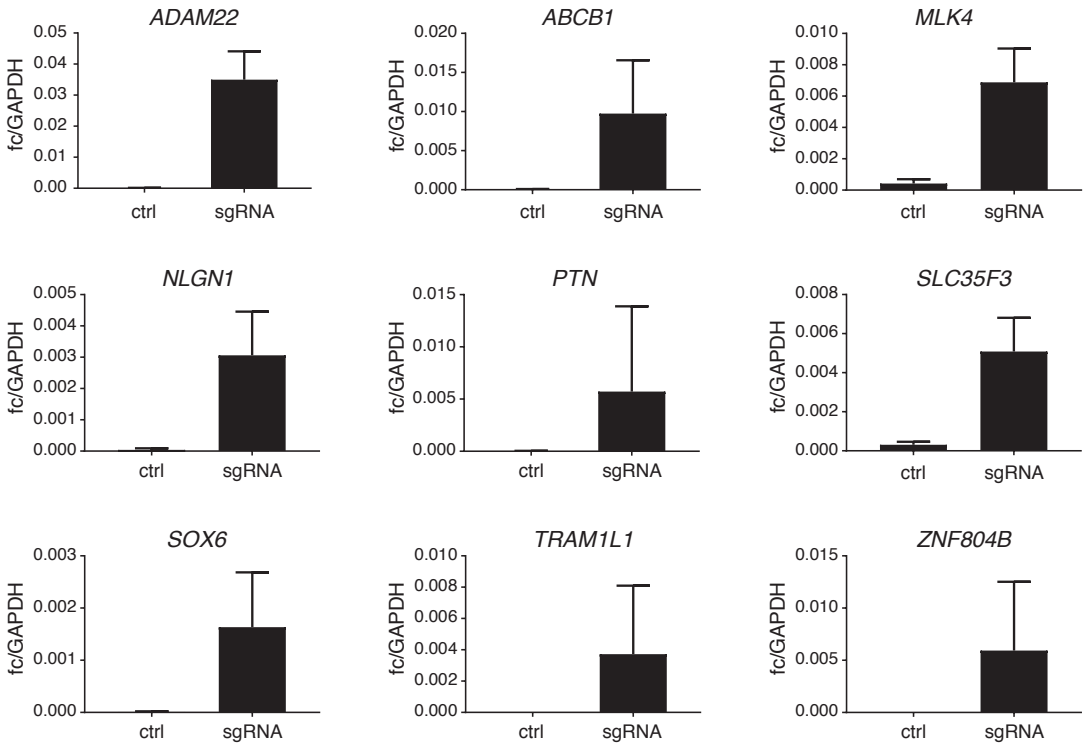


Figure S 2

Expression of genes in RPE-1 cells targeted by CRISPRa. Expression levels were determined by RT-qPCR. Average of three technical replicates, error bar indicates standard deviation. fc/GAPDH: expression level normalized to GAPDH gene. sgRNA: transfection with sgRNA targeting the indicated gene; ctrl: untransfected control

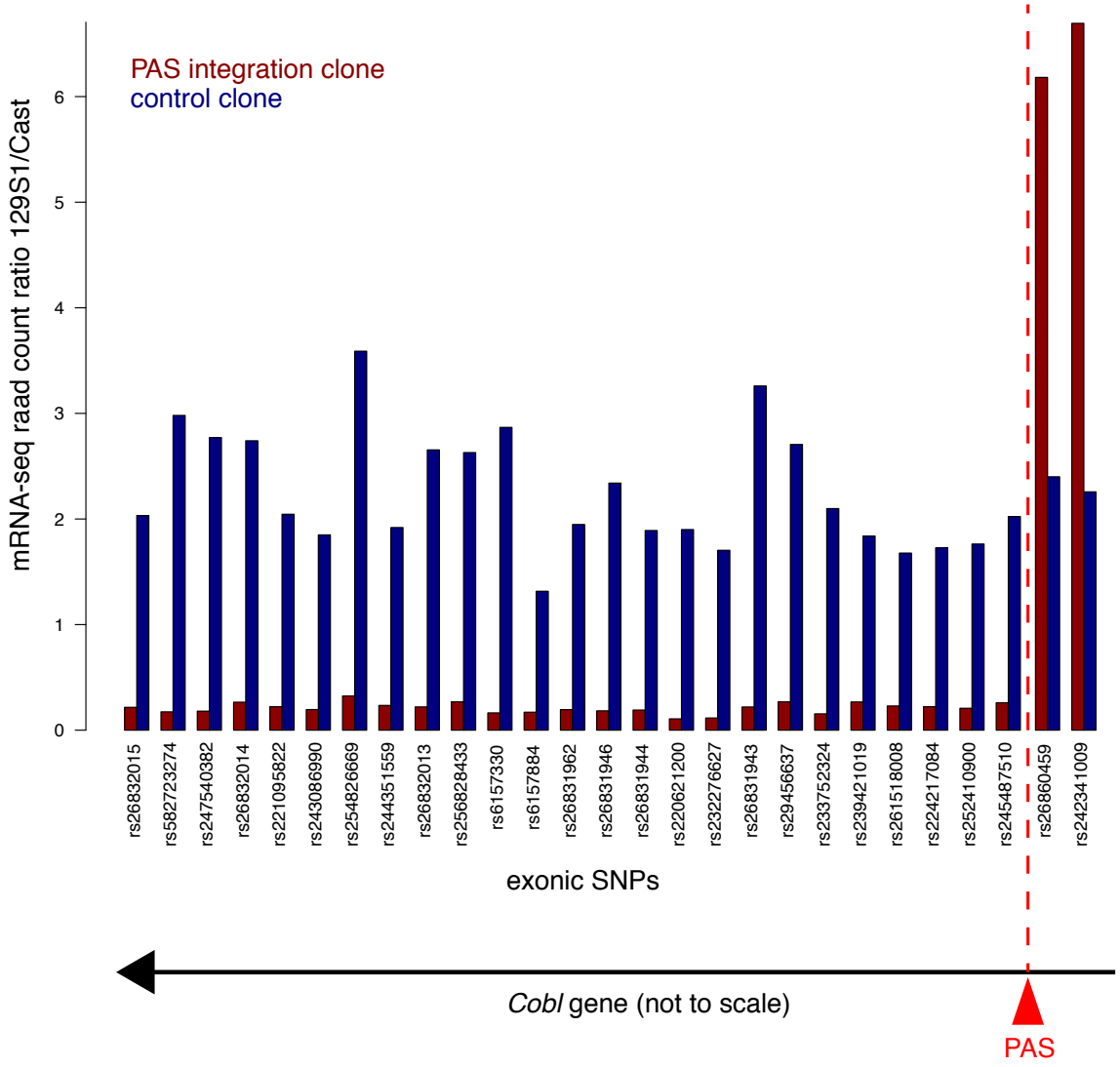


Figure S 3

Effect of PAS integration on Cobl gene transcription. Ratio of allele-specific mRNA-seq read counts of 129Sv and Cast SNP variants located in the Cobl gene. Two SNPs are located upstream of the PAS integration site; all other SNPs are downstream. Note that the 5' to 3' orientation of the gene is from right to left. In a control clone without the PAS, the Sv129 allele is consistently expressed ~2-fold higher than the Cast allele. In the clone with the PAS integration in the 129Sv allele, expression of the 129Sv variants downstream of the PAS is strongly reduced. The increased level of the 129Sv variants upstream may be due to greater stability of the mRNA due to its altered 3' end. Data are average of two independent biological replicates

Chapter 5

Discussion

Discussion

In this study, a range of experimental manipulations was performed in order to perturb chromatin state. These approaches include random, genome-wide reporter integration, tethering of chromatin modifiers and artificial upregulation of endogenous genes using CRISPRa and TALE-VP64 technology. We made use of powerful genome-wide mapping techniques to study chromatin organization on a systems biology level.

Strengths and limitations of the experimental systems used in this study

Strengths and limitations of the TRIP system

The second and third chapter of this thesis are based on the TRIP system, specifically in this case of around 1000 barcoded reporters integrated randomly in *Drosophila* cells. These reporters are present in a heterogeneous cell pool. Classic reporter assays usually work with a single gene either expressed in a plasmid or integrated in a single genomic locus. However, by only looking at a single genomic locus at a time, it is difficult to make generalizable statements, because the following variables are fixed:

- level of transcription
- genomic location
- chromatin context

The strength of the TRIP approach lies in the high statistical power provided by the large number of integrations. Essentially 1000 experiments can be performed at once and at the same time a plethora of genomic contexts can be screened for their effect on gene expression.

One limitation of the TRIP approach, however, are the options to examine the individual integrations. When using TRIP reporters as recruitment platform for chromatin modifiers, apart from looking at the effect on reporter transcription, it would for example be interesting to know if this targeting affects the 3D position of integrations or if their chromatin state has changed. Upon HP1 recruitment, the strong repression of all reporters indicates that even integrations in euchromatic sites had been converted to heterochromatin. This hypothesis could so far not be tested. It would also be worthwhile to find out how far these putative heterochromatin regions might have spread and if the extent of spreading varies by chromatin state. In order to gain information on the chromatin state of TRIP integrations, their barcodes would have to be assayed by ChIP or DamID. However, these are genome-

wide techniques that are not suitable to deliver information about barcoded fragments coming from a heterogeneous cell pool.

The TRIP experiments in this thesis used the original readout of the approach which is the quantification of individual reporter expression by barcode counts [1,2]. Apart from this readout, other studies have been performed in the meantime which looked at double-strand break repair pathways working on TRIP reporters [3] (the following are published as pre-print) [4,5]. Other TRIP reporters could be developed to assess for example alternative splicing or replication timing in combination with chromatin state.

Using TRIP as a platform for recruitment of proteins

We used the TRIP reporters as platform to target a protein of choice to multiple genomic environments. Chromatin proteins are not distributed randomly all over the genome but co-occur in distinct "states"[6]. By combining the TRIP assay with targeted recruitment, it is now possible to probe the function of a chromatin protein of choice in a multitude of genomic contexts.

Performing high-throughput tethering of HP1 and ACF1 revealed context-depending effects which contribute to understanding the function of these proteins. It would be interesting to extend this assay to other candidates, for example proteins involved in chromatin 3D organization such as CTCF. Performing a full TRIP-tethering assay is costly and cumbersome, therefore it would be desirable to have a quick and cheap method to test if a candidate protein actually has context-dependent effects. One possibility would be to isolate representative clones for each chromatin state. Upon targeting of a candidate protein, a quick fluorescence- or qPCR-based screen of reporter GFP activity could indicate if the candidate protein has an effect of reporter activity, and if the effect differs between chromatin states.

It has to be kept in mind that candidate proteins need to be fused to the Gal4 domain for this assay which could directly abolish their function or their interactions with secondary proteins. This already limits the number of possible candidates, for example proteins which work as dimers or proteins which directly bind to certain DNA motifs would probably not be suitable for Gal4-fusion. However, this limitation could be circumvented by directly introducing the respective binding site into the TRIP reporter cassette.

Advantages of generating clonal cell lines from TRIP pools

The original TRIP assay is performed by amplification of barcode RNA from a heterogeneous pool of cells. This way, the only information gained from the individual reporters is their expression. What if a researcher wants to look at the nuclear position of these integrations? As mentioned before, due to the heterogeneous nature of these cell pools, DamID is not suitable to deliver information on the nuclear position of individual barcodes. Could DamID be performed on barcodes directly? In the TRIP cassettes used in this assay, there are GATC sites flanking the barcode which could be used. However, the DamID signal is not reliable for single fragments which is why DamID sequencing reads are usually pooled over several kb. Information could be averaged for large groups of barcodes, for example to show that integration in LADs still retain their lamina association.

But DamID can be performed on clonal cell lines derived from TRIP pools. In chapter 4 we employed a medium-throughput approach by generating TRIP clones with a high number of integrations. One limitation still remained: in diploid cell lines, the DamID signal around the integration site would be diluted by the signal coming from the homologous allele without integration. Therefore, we generated TRIP clones in F1 hybrid mouse ES cells. This made it possible to look at nuclear position around individual integrations and indicated that integration of a transcription unit leads to local detachment from the nuclear lamina. Consequently, we could overcome one limitation of the maximum-throughput TRIP system. Still, some limitations remained: we were unable to directly map the 3D position of each integrated reporter gene itself and could only look at the flanking genomic regions. Furthermore, DamID resolution was not high enough to see if individual integrations were detached or not, and we could only see average detachment by pooling all 25 LAD integrations.

Gaining higher resolution by using the CRISPRa system

In order to observe detachment of single transcription units we had to switch to targeting of large endogenous genes by CRISPRa. We could now beautifully see single genes reposition upon induction at unprecedented resolution. We observed detachment from the nuclear lamina along the transcription unit for each upregulated gene. There was a strong correlation between expression level and lamina association. The cost of this focused study however was the loss of statistical power. The number of genes assayed in chapter 4 is high enough not to speak of anecdotal evidence. However, the throughput is still too low to deduce general rules for gene detachment from the nuclear lamina. Therefore the “perfect”

methodology that combines both high throughput and high resolution for assaying genome architecture is still missing.

It should also be noted that the DamID data obtained from CRISPRa-targeted cells represent an average from a pool of cells. Therefore, it cannot be stated what gene detachment looks like in individual cells. Do all genes detach in the same manner? Is detachment gradual or more in an all-or-nothing fashion? How does the extent of detachment correspond to the level of transcription in individual cells? Do genes first detach and then start to transcribe or the other way around?

Over the last decade it has become possible to use transcriptome analysis as well as chromatin architecture assays such as DamID or Hi-C in single cells [7,8]. However, using these genome-wide techniques in single cells comes at the cost of low resolution and are therefore not optimal when looking for example at the 3D position of single genes. At the same time, microscopy techniques are advancing rapidly. Chromosome-painting techniques using tiled oligonucleotides together with super-resolution microscopy have made it possible to trace the architecture of large structural domains in single cells at a resolution of several kilobases [9–11]. Imaging-based techniques also enable researchers to observe genome 3D structure and gene activity in the same cell. This way it can be determined in which temporal order changes in transcription and nuclear position are occurring or how transcription level corresponds to displacement.

These advances have made it possible to move beyond experiments in cultured cells and study the role of chromatin architecture in whole tissues or during development [12–14]. Sequencing-based techniques such as Hi-C require high numbers of cells to achieve high resolution but this forces the researcher to work with specific cell lines. Looking at tissues enables researches to study multiple cell types at the same time, addressing questions such as which features of genome organization are static during development and which are flexible. It would for example be possible to observe if regions become “primed” for activation or inactivation by moving to a different chromatin compartment during differentiation. Of course, in comparison to sequencing-based techniques, imaging-based methods require the researcher to focus on a defined region of interest and can thereby limit the possibility for new, unbiased discoveries.

TRIP expression range in comparison to activation by CRISPRa

In mouse cells, expression of the integrated reporters varied up to 1000-fold, showing how strongly gene expression is influenced by genomic context [1]. We observed a similar expression range when barcoded GFP reporters were integrated in *Drosophila* cells in chapter 2.

The SunCas CRISPRa system used in chapter 4 was initially developed for imaging in living cells. The technique was successfully repurposed to generate a strong transcriptional activator by using the Cas9 protein to recruit up to 24 copies of VP64. Previous techniques using only a single copy of VP64 fused to Cas9 only achieved up to 2-fold upregulation of target genes [15]. In chapter 4, the SunCas system was used for activation of genes in LADs which are unfavorable for transcription. Candidate genes were moved to the nuclear interior as could be clearly seen by DamID.

This raises the question: Is the SunCas-based activation so powerful because it is able to push genes out of the “inactive” zone at the nuclear periphery, thereby overcoming chromatin position effects? And if yes, how is this “push” achieved? Are the multiple copies of VP16 more likely to stochastically capture the pre-initiation complex even at a site far away from transcriptional activity? Or any other chromatin factors that could bring the target locus to a position or chromatin state which is more favorable to transcription?

In a study looking at lamina association in single cells, lamina-associated regions were targeted with VP16 which led to dissociation from the nuclear envelope [16]. However, this dissociation seemingly did not affect all targeted loci. Also, some loci moved much further towards the interior than others. It would be interesting to perform DamID on these cells to see if some loci are more susceptible to detachment from the nuclear lamina upon VP16 targeting than others. However, this would require the use of a methyltransferase with a target motif other than GATC for the second round of LAD-labelling or another technique for determining position.

Notably, the CRISPRa and VP16 systems are highly artificial. This raises the question how comparable the observed repositioning is to endogenous genes. It would be interesting to pick a more natural system for gene activation and dissect the events which lead to relocalisation.

HP1 TRIP results and implications for 3D architecture

In the TRIP-tethering experiment described in chapter 2, HP1 was initially chosen as a positive control when setting up the assay. This protein is known as a strong transcriptional repressor. Therefore, it was not surprising that HP1, in fact, silenced every single reporter, regardless of expression level. The native function of HP1 is to shut off transposons which are genomic "parasites", therefore it is desirable that HP1-induced repression is as efficient as possible. Upon further inspection, we noticed something interesting: HP1 seems to repress more efficiently in its native environment which is pericentric heterochromatin. This has a functional implication for chromatin states – by concentrating proteins such as HP1 in certain genomic compartments, their efficiency can be enhanced.

The initial hypothesis when starting the HP1 TRIP experiments was, that some loci would retain HP1-silencing even after loss of the initial targeting event and that we could thereby find chromatin states that would support epigenetic memory. We therefore examined if any of the reporters retained HP1-induced repression after loss of HP1 tethering. However, all integrations completely recovered expression two weeks after targeting. It has been shown in mouse cells that repression after HP1 targeting can be maintained but this effect is dependent on DNA methylation [17]. We therefore attributed the lack of "memory" in our *Drosophila* system to the fact that fly cells barely have any DNA methylation. We subsequently moved the HP1 high-throughput tethering system to human K562 cells. We found that some reporters in LADs showed memory of silencing (unpublished data) but the results showed insufficient reproducibility to warrant further follow-up. It is therefore still unclear if there are genomic environments that support stable gene silencing.

HP1-occupied chromatin is found at the nuclear periphery or in the "chromocenters" seen in some cell types. This raises the question of what happens to the reporters that are targeted by HP1: Are they changing spatial position and move to the nuclear periphery or another HP1-dense region? When looking at endogenous loci, inactive genes tend to be associated with the nuclear lamina but a large number of silent genes also resides in the nuclear interior. High-throughput targeting of HP1 combined with a method to determine the 3D position of each integration could help to unravel why some inactive regions move to the nuclear periphery while others do not. However, the Kc167 cell pools are not ideal for imaging-based experiment due to the small size of their nuclei. Also, for visualizing single integrations, a signal-amplifying method such as the SunCas system would have to be used with is readily available in mammalian cells [15].

As previously mentioned, a domain which binds methylated GATC motifs was used to target VP16 to lamina-associated loci. The same system could be repurposed by first specifically methylating regions in the nuclear interior and subsequently targeting these regions with HP1 to see if and where they will relocate.

Chromatin factors for spatial repositioning

It has been shown that decondensation of chromatin is sufficient to relocalize peripheral genes to the nuclear interior [18]. This implies that there are two modes for relocalization of genomic loci: 1) transcription-dependent and 2) transcription-independent. Since in CRISPRa experiments genes detached as single units along the transcription unit, one could hypothesize that any events where larger or intergenic regions move to the nuclear interior require the help of chromatin remodeling or - modifying proteins. But not much is known about the chromatin factors involved in transcription-independent repositioning.

Unbiased screens have been performed in *C. elegans* which used a reporter gene to screen for factors involved in gene position [19,20] Another screen in human cells used a FISH-based imaging readout and found components of the replication machinery to be involved in repositioning [21]. For the human cystic fibrosis gene, apart from A-type lamins, CTCF and a histone deacetylase were found to be involved in peripheral positioning [22]. Early studies suggested that nuclear actin and myosin are players in gene relocalization [23]. Regarding heterochromatin, for H3K9 methyltransferase SUV39H1 it has been shown that targeting this protein to an A-compartment (iLAD) locus caused this locus to switch to a B-compartment (LAD). Interestingly, the chromodomain of SUV39H1 was sufficient for this relocalization irrespective of gene expression [24]. This result indicates how recruitment of certain chromatin proteins can induce repositioning. This recruitment might be the first step in preparation for later downregulation as has been observed for mouse genes during differentiation which first moved to the nuclear periphery and later became downregulated [25].

It remains challenging to perform a screen for proteins that mediate gene repositioning, especially when aiming at factors that are independent from transcription. This experiment would require three components:
1) a reporter locus which can inducibly be relocalised to the nuclear interior. One possibility would be the use of a decondensation peptide fused to a TALE protein as used in previous studies to target an endogenous (non-gene) locus of choice [18]

2) a method to detect the position of the reporter locus. Ideally the detection method would be cheap and suitable for high-throughput screening.
3) a method to perform loss-of function screens, for example a CRISPR sgRNA library to perform genome-wide knockout screens.

Notably, cartoon representations of genome nuclear lamina interactions usually depict a single, extended string of DNA meandering to and from the nuclear envelope (Fig 1A). However, this representation neglects the condensation state of DNA. Instead of linear detachment, it might rather look like a gene "spilling its guts" upon activation (Fig 1B). However, high-resolution and time-resolved microscopy techniques would be needed to trace the exact way and mode in which a gene moves away from the nuclear periphery upon activation.

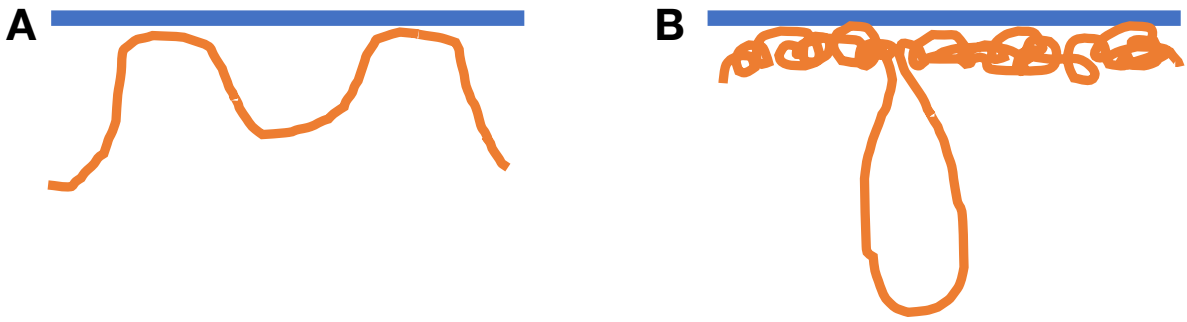


Figure 1: Gene position at the nuclear lamina depicted as (A) linear detachment or (B) decondensation

Is genome organization efficient?

It remains the question, how much of genome organization has evolved, giving the cell a functional advantage either due to more efficient activation of important genes or due to more efficient silencing of genes that need to stay inactive
Is the current organization in a cell really the most optimal one, and how much of it is just random and without any major importance? Or asked differently: If one could be a genome architect, would it be possible to arrange the genome in an even more efficient way? Recently, "minimal yeast genomes" were created by reducing the s. cerevisiae genome to two chromosomes [26] or one chromosome [27] by chromosome end-to-end fusions. The fused chromosomes in the one-chromosome strain adopted a more globular shape compared to the stretched conformation of chromosomes in the 16-chromosome wild type strain [27]. Surprisingly, transcriptome analysis revealed that gene expression was nearly unchanged [26,27]. However, 3D architecture analysis was not detailed enough to reveal to what extent nuclear envelope association had changed. It would be interesting to study the role of

genome architecture by large-scale perturbation studies, for example by use of chromothripsis cell lines [28] or with scrambled yeast genomes [29].

LADs cover 30-50% of the genome which is disproportionally large in relation to the space available at the nuclear periphery as compared to the interior. Also, LADs typically cover large stretches of DNA. This poses the question if the generation of large gene deserts sequestered at the nuclear periphery is somehow favorable to the cell instead of a more random arrangement of active and inactive regions. Alternatively, inactive regions could have gradually extended themselves in a form of “gene desertification” for example by heterochromatin encroaching into active regions and gradually silencing adjacent genes. For example, transplanting LADs into nuclear interior regions and vice versa could help to elucidate the importance of this large-scale domain structure. Also, what would happen if all LADs be genetically deleted. Would this kind of cell still be viable or able to differentiate? CRISPR-based tools for genome engineering allow for targeted rearrangements but can also be employed to induce DNA looping without genetic modifications [30]. To dissect the effects of genome architecture from gene expression it would be desirable to have more options available to directly manipulate genome 3D structure that do not rely on the rather large Cas9 protein.

Nonetheless these and future precision tools for genomic manipulation in combination with imaging techniques steadily increasing in resolution should allow for exciting new insights into genome regulation.

Bibliography

1. Akhtar W, de Jong J, Pindyurin AV, et al. Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell*. 2013;154(4):914-927. doi:10.1016/j.cell.2013.07.018
2. Vansant G, Chen H-C, Zorita E, et al. The chromatin landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res*. 2020;48(14):7801-7817. doi:10.1093/nar/gkaa536
3. Gisler S, Gonçalves JP, Akhtar W, et al. Multiplexed Cas9 targeting reveals genomic location effects and gRNA-based staggered breaks influencing mutation efficiency. *Nat Commun*. 2019;10(1):1598. doi:10.1038/s41467-019-09551-w
4. Pokusaeva VO, Diez AR, Espinar L, Filion GJ. Strand asymmetry influences mismatch repair during single-strand annealing. *bioRxiv*. Published online November 20, 2019:847160. doi:10.1101/847160
5. Schep R, Brinkman EK, Leemans C, et al. Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance. *bioRxiv*. Published online May 5, 2020:2020.05.05.078436. doi:10.1101/2020.05.05.078436
6. Filion GJ, van Bemmel JG, Braunschweig U, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*. 2010;143(2):212-224. doi:10.1016/j.cell.2010.09.009
7. Kind J, Pagie L, Vries SS, et al. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell*. 2015;163:134-147.
8. Nagano T, Lubling Y, Stevens TJ, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*. 2013;502(7469):59-64. doi:10.1038/nature12593
9. Bintu B, Mateo LJ, Su J-H, et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*. 2018;362(6413). doi:10.1126/science.aau1783
10. Nir G, Farabella I, Pérez Estrada C, et al. Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLoS Genet*. 2018;14(12):e1007872. doi:10.1371/journal.pgen.1007872
11. Wang S, Su J-H, Beliveau BJ, et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science*. 2016;353(6299):598-602. doi:10.1126/science.aaf8084
12. Liu M, Lu Y, Yang B, et al. Multiplexed imaging of nucleome architectures in single cells of mammalian tissue. *Nature Communications*. 2020;11(1):2907. doi:10.1038/s41467-020-16732-5

13. Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature*. 2019;568(7750):49-54. doi:10.1038/s41586-019-1035-4
14. Cardozo Gizzi AM, Cattoni DI, Fiche J-B, et al. Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. *Mol Cell*. 2019;74(1):212-222.e5. doi:10.1016/j.molcel.2019.01.011
15. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014;159(3):635-646. doi:10.1016/j.cell.2014.09.039
16. Kind J, Pagie L, Ortabozkoyun H, et al. Single-cell dynamics of genome–nuclear lamina interactions. *Cell*. 2013;153:178-192.
17. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. Dynamics and memory of heterochromatin in living cells. *Cell*. 2012;149(7):1447-1460. doi:10.1016/j.cell.2012.03.052
18. Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science*. 2014;346(6214):1238-1242. doi:10.1126/science.1259587
19. Towbin BD, González-Aguilera C, Sack R, et al. Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell*. 2012;150(5):934-947. doi:10.1016/j.cell.2012.06.051
20. Gonzalez-Sandoval A, Towbin BD, Kalck V, et al. Perinuclear Anchoring of H3K9-Methylated Chromatin Stabilizes Induced Cell Fate in C. elegans Embryos. *Cell*. 2015;163(6):1333-1347. doi:10.1016/j.cell.2015.10.066
21. Shachar S, Voss TC, Pegoraro G, Sciascia N, Misteli T. Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping. *Cell*. 2015;162(4):911-923. doi:10.1016/j.cell.2015.07.035
22. Muck JS, Kandasamy K, Englmann A, Günther M, Zink D. Perinuclear positioning of the inactive human cystic fibrosis gene depends on CTCF, A-type lamins and an active histone deacetylase. *J Cell Biochem*. 2012;113(8):2607-2621. doi:10.1002/jcb.24136
23. Chuang C-H, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS. Long-range directional movement of an interphase chromosome site. *Curr Biol*. 2006;16(8):825-831. doi:10.1016/j.cub.2006.03.059
24. Wijchers PJ, Krijger PHL, Geeven G, et al. Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments. *Mol Cell*. 2016;61(3):461-473. doi:10.1016/j.molcel.2016.01.001

25. Peric-Hupkes D, Meuleman W, Pagie L, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell*. 2010;38(4):603-613. doi:10.1016/j.molcel.2010.03.016
26. Luo J, Sun X, Cormack BP, Boeke JD. Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. *Nature*. 2018;560(7718):392-396. doi:10.1038/s41586-018-0374-x
27. Shao Y, Lu N, Wu Z, et al. Creating a functional single-chromosome yeast. *Nature*. 2018;560(7718):331-335. doi:10.1038/s41586-018-0382-x
28. Stephens PJ, Greenman CD, Fu B, et al. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell*. 2011;144(1):27-40. doi:10.1016/j.cell.2010.11.055
29. Richardson SM, Mitchell LA, Stracquadanio G, et al. Design of a synthetic yeast genome. *Science*. 2017;355(6329):1040-1044. doi:10.1126/science.aaf4557
30. Morgan SL, Mariano NC, Bermudez A, et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nature Communications*. 2017;8(1):15993. doi:10.1038/ncomms15993

Addendum

Addendum

Summary

Human cells contain around 2 meters of DNA which in a “beads-on-a-string”-like fashion is wrapped around histone proteins into higher structures called nucleosomes. These proteins help to package and compact the DNA, building up a condensed state called heterochromatin. Condensation not only serves to organize DNA but can also prevent certain genes from becoming activated. Therefore, the packaging state of the genome provides a higher level of information, determining which parts of the genome are expressed and which are silent. DNA is not only organized in terms of compaction but also in 3D space. Taking the nuclear envelope as a landmark, there are certain genomic regions that preferentially locate to the nuclear periphery. These regions can be identified by mapping their association with the nuclear lamina (a protein layer that covers the inside of the nuclear envelope) and are therefore called “lamina-associated domains” or LADs. They tend to be heterochromatic and therefore low in transcriptional activity. Consequentially, transcriptionally active regions are usually found in the interior of the nucleus. In addition, a multitude of proteins associate with DNA making up a structure called chromatin. Some of these proteins can directly modify chromatin structure. The distribution of these chromatin proteins is not random. Instead, certain proteins tend to co-localize, making up distinct “chromatin states”. Chapter 1 introduces various aspects of genome organization as well as the major techniques used in this study.

These chromatin states are still poorly understood. For example, one question is what happens if a chromatin protein is taken out of its native state. Does its function become diminished? We addressed this question in chapter 2 of this thesis by taking heterochromatin protein 1 (HP1) and recruiting it to a reporter gene randomly integrated in approximately 1000 sites in *Drosophila* cells by using the TRIP (thousands of reporters integrated in parallel) technique. Depending on the integration site, the expression of the reporter gene varied up to 1000-fold. HP1 silenced highly transcribed reporters just as well as less active ones. The study revealed that HP1 is a powerful repressor that works in any chromatin state surveyed. However, a closer look showed that reporters integrated in HP1’s native chromatin state were silenced more potently. The accumulation of proteins into distinct states might therefore increase their efficiency. Since it has been shown in mouse cells that HP1-induced silencing can be inherited throughout mitotic cell divisions, we also examined if we could observe this kind of memory in our *Drosophila* system. However, all

reporters regained their original expression level after two weeks, showing no indication of memory.

In chapter 3, we used the same system to elucidate the function of nucleosome remodeling complex CHRAC/ACF in a collaborative study. Tethering the subunit ACF1 to around 1000 integrated reporters revealed that ACF1 has a context-dependent effect on gene expression. It preferentially downregulates poorly transcribed reporters in repressive chromatin states.

The results mentioned so far mainly describe changes in transcription. As mentioned previously, there is a connection between transcriptional activity and 3D genome organization. Inactive regions tend to localize to the nuclear lamina and vice versa. In chapter 4 we were looking at changes in genome 3D organization as a consequence of transcription. Specifically, we generated clonal cell lines with multiple integrations of a short transcription unit and then mapped the genome-wide nuclear lamina association. We were wondering if these integrations could maybe lead a whole LAD to detach from the nuclear lamina. However, we only observed very modest and local detachment. Since this system did not allow us to look at the gene body, we moved to activating large endogenous genes in LADs with CRISPRa. Indeed, we observed that all upregulated genes hat detached from the nuclear lamina and moved to the interior. However, detachment only covered the transcription unit and did not extend far towards surrounding regions. Even when neighboring genes were activated in combination, the intermediate region was not affected. To confirm that transcription was indeed the delimiting factor for detachment, we truncated an endogenous gene and observed increased lamina association after the newly inserted transcription stop. We also observed increased lamina interaction upon abolishing transcription of three genes in the nuclear interior. We also looked at replication timing and found a switch from late to early replication for CRISPRa-activated genes. In summary, these results show how transcription can shape chromatin architecture.

In chapter 5 these results are discussed and put in relation to recent findings in the field of 3D genome architecture. Open questions are addressed and potential research strategies are presented to further advance the understanding of genome organization.

Nederlandse Samenvatting

Menselijke cellen bevatten ongeveer 2 meter DNA dat op een "kralen-aan-een-touwtje" - achtige manier om histoneiwitten wordt gewikkeld in hogere-orde structuren die nucleosomen worden genoemd. Deze eiwitten helpen het DNA te verpakken en samen te pakken, waardoor een compacte toestand wordt opgebouwd die heterochromatine wordt genoemd. Deze compactie dient niet alleen om het DNA te organiseren, maar kan ook voorkomen dat bepaalde genen geactiveerd worden. Daarom biedt de verpakkingstoestand van het genoom een hoger niveau van informatie, waardoor wordt bepaald welke delen van het genoom tot expressie worden gebracht en welke inactief zijn. DNA is niet alleen georganiseerd in termen van compactie, maar ook in 3D-ruimte. Als we de nucleaire envelop als een referentiepunt nemen, zijn er bepaalde genomische regio's die zich bij voorkeur in de nucleaire periferie bevinden. Deze regio's kunnen worden geïdentificeerd door hun associatie met de nucleaire lamina (een eiwitlaag die de nucleaire envelop aan de binnenzijde van de celker bedekt) in kaart te brengen. Ze worden daarom "lamina-geassocieerde domeinen" of LAD's genoemd. Ze zijn meestal heterochromatisch en daarom laag in transcriptionele activiteit. Daarentegen worden transcriptioneel actieve gebieden meestal in het binnenste van de celkern aangetroffen. Bovendien associëren een groot aantal eiwitten zich met DNA en vormen ze een structuur die chromatine wordt genoemd. Sommige van deze eiwitten kunnen de chromatinestructuur direct wijzigen. De verdeling van deze chromatine-eiwitten is niet willekeurig. In plaats daarvan hebben bepaalde eiwitten de neiging om te co-lokaliseren, waardoor ze verschillende types chromatine vormen. Hoofdstuk 1 introduceert verschillende aspecten van genoomorganisatie, evenals de belangrijkste technieken die in deze studie worden gebruikt.

Er is nog veel onbekend over het samenspel tussen chromatine-eiwitten. Een vraag is bijvoorbeeld wat er gebeurt als een chromatine-eiwit uit zijn oorspronkelijke chromatine-type wordt gehaald. Wordt de functie ervan verminderd? We hebben deze vraag beantwoord in hoofdstuk 2 van dit proefschrift door heterochromatine proteïne 1 (HP1) te recruteren naar een reportergen dat willekeurig is geïntegreerd in ongeveer 1000 plaatsen in het genoom van *Drosophila*-cellen. Hiervoor hebben we gebruik gemaakt van de TRIP-techniek *(thousands of reporters integrated in parallel*). Afhankelijk van de integratieplaats varieerde de expressie van het reportergen tot 1000-voudig. We vonden dat HP1 zowel sterk actieve als minder actieve reporters onderdrukte. De studie onthulde dat HP1 een krachtige repressor is die werkt in elk onderzocht type chromatine. Uit nadere analyse bleek dat reporters die in het oorspronkelijke chromatine-type van HP1 waren geïntegreerd,

krachtiger geinactiveerd werden door de extra HP1. De accumulatie van eiwitten in verschillende chromatine-types zou daarom hun efficiëntie kunnen verhogen. Omdat in cellen van de muis is aangetoond dat HP1-geïnduceerde inactivatie kan worden overgeërfd tijdens mitotische celdelingen, hebben we ook onderzocht of we dit soort geheugen konden waarnemen in ons Drosophila-systeem. Alle reporters kregen echter na twee weken hun oorspronkelijke expressieniveau terug en lieten geen indicatie van een geheugen zien.

In hoofdstuk 3 hebben we hetzelfde systeem gebruikt om in een collaboratieve studie de functie van nucleosoomremodelleringscomplex CHRAC / ACF op te helderen. Door het ewit ACF1 vast te maken aan ongeveer 1000 geïntegreerde reporters, bleek dat ACF1 een contextafhankelijk effect heeft op genexpressie. Het reguleert bij voorkeur zwak getranscribeerde reporters in repressieve chromatine-types.

Bovenstaande resultaten beschrijven voornamelijk veranderingen in transcriptie. Zoals eerder vermeld, is er een verband tussen transcriptionele activiteit en 3D-genoomorganisatie. Inactieve gebieden hebben de neiging zich te lokaliseren bij de nucleaire lamina en vice versa. In hoofdstuk 4 bestuderen we de veranderingen in de genoom 3D-organisatie als gevolg van transcriptie. Hiertoe hebben we klonale cellijnen gegenereerd met geïntegreerde activeerbare genen en vervolgens de genoombrede nucleaire lamina-associatie in kaart gebracht. We vroegen ons af of deze integraties er misschien toe zouden kunnen leiden dat een hele LAD zich losmaakt van de lamina. Echter, we vonden slechts een subtiele en heel lokale verplaatsing. Omdat dit systeem ons niet goed in staat stelde om naar het gen zelf te kijken, zijn we overgestapt op het activeren van grote endogene genen in LADs door middel van de CRISPRa technologie. We vonden dat alle geactiveerde genen zich losmaakten van de nucleaire lamina. Deze onthechting was meestal beperkt tot de transcriptie-eenheid en strekte zich niet ver uit naar het flankerende DNA. Zelfs wanneer naburige genen in combinatie werden geactiveerd, werd het tussenliggende gebied niet beïnvloed. Om te bevestigen dat transcriptie inderdaad de bepalende factor was voor loslating, hebben we de transcriptie van een actief gen in het midden geblokkeerd. Dit resulteerde in een verhoogde lamina-associatie na de nieuw ingevoegde transcriptiestop. We hebben ook een verhoogde lamina-interactie waargenomen na het inactiveren van drie genen die oorspronkelijk geen contact met de lamina hadden. Tot slot vonden we dat de activatie van genen lokaal leidde tot een vroegere replicatie van het DNA. Deze resultaten laten zien hoe transcriptie de ruimtelijke structuur van chromatine kan beinvloeden.

In hoofdstuk 5 worden deze resultaten besproken en in relatie gebracht met recente bevindingen op het gebied van 3D genoomorganisatie. Een aantal openstaande vragen worden besproken, alsmede mogelijke onderzoeks-strategiën die het begrip van genoom-organisatie verder kunnen vergroten.

Abbreviations

ATP	adenosine triphosphate
ChIP	chromatin immuno-precipitation
CRC	chromatin remodeling complex
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
DBD	DNA binding domain
DNA	desoxyribonucleic acid
e.g.	exempli gratia (for example)
ES cells	embryonic stem cells
FACS	fluorescence-activated cell sorting
FISH	fluorescent in-situ hybridization
gDNA	genomic DNA
GFP	green fluorescent protein
HP1	heterochromatin protein 1
i.e.	id est (that is)
iLAD	inter-LAD
kb	kilobase
LAD	lamina-associated domain
Mb	megabase
mRNA	messenger RNA
NGS	next-generation sequencing
NL	nuclear lamina
PCA	principal component analysis
qPCR	quantitative polymerase chain reaction
RNA	ribonucleic acid
sgRNA	short guide RNA
TAD	topologically associated domain
TALE	transcription activator-like effector
TRIP	thousands of reporters integrated in parallel
UAS	upstream activating sequence
Wt	wild type

Curriculum Vitae

Laura Brueckner was born on January 30th in Frankfurt am Main, Germany. In 2006, she first started to study Psychology at Vienna University but then after one year changed to Biology.

For her Bachelor's project, she worked in the laboratory of Prof. Michael Kiebler at the Center for Brain Research in Vienna on the role of Staufen proteins in synaptic plasticity. She completed her Bachelor's degree in 2010 and continued with a Master's program in Molecular Biology.

During her Master's studies, Laura worked as a tutor at the University of Vienna, supervising undergraduate lab courses in Microbiology and Genetics. She performed a research internship in the laboratory of Prof. Denise Barlow at the Center for Molecular Medicine (CeMM) in Vienna on the role of macro non-coding RNAs in genomic imprinting. She also worked in the laboratory of Prof. Christian Seiser at the Max F. Perutz Laboratories (MFPL) in Vienna on the effects of HDAC1 and HDAC2 knockouts in mouse neurons and epidermis. She then went to Japan part of the "Frontier Lab" foreign student internship program and worked for 9 months in the laboratory of Prof. Shoji Tajima at the Institute for Protein Research of Osaka University where she developed a method for the synthesis of long hemimethylated DNA for nucleosome reconstitution. She also supervised and trained a Master student. The resulting protocol majorly contributed to a publication on RFTS-dependent negative regulation of Dnmt1 by nucleosome structure and histone tails. This internship was sponsored by a JASSO scholarship. Laura then went back to Vienna and joined the laboratory of Dr. Stefan Kubicek at the CeMM to work on her Master's thesis titled "Inhibition of JmjC-domain histone demethylases with small molecules". She graduated in 2013 as a Master of Science with distinction.

In 2013 she joined the group of Prof. Bas van Steensel at the Netherlands Cancer Institute Antoni van Leuwenhoek (NKI-AVL) in Amsterdam where she worked on her PhD thesis as part of the Oncology Graduate School. Her research contributed to the 4D Nucleome Consortium to understand the principles behind genome 3D organization.

List of Publications

1. **Brueckner L**, Zhao PA, van Schaik T, Leemans C, Sima J, Peric-Hupkes D, Gilbert DM, van Steensel B. Local rewiring of genome-nuclear lamina interactions by transcription. EMBO J. 2020 Mar 16;39(6):e103159. doi: 10.15252/embj.2019103159. Epub 2020 Feb 21. PMID: 32080885; PMCID: PMC7073462.

2. Leemans C, van der Zwalm MCH, **Brueckner L,** Comoglio F, van Schaik T, Pagie L, van Arensbergen J, van Steensel B. Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs. Cell. 2019 May 2;177(4):852-864.e14. doi: 10.1016/j.cell.2019.03.009. Epub 2019 Apr 11. PMID: 30982597; PMCID: PMC6506275.

3. Scacchetti A, **Brueckner L,** Jain D, Schauer T, Zhang X, Schnorrer F, van Steensel B, Straub T, Becker PB. CHRAC/ACF contribute to the repressive ground state of chromatin. Life Sci Alliance. 2018 Feb 9;1(1):e201800024. doi: 10.26508/lsa.201800024. PMID: 30456345; PMCID: PMC6238394.

4. Mishima Y, **Brueckner L**, Takahashi S, Kawakami T, Arita K, Oka S, Otani J, Hojo H, Shirakawa M, Shinohara A, Watanabe M, Suetake I. RFTS-dependent negative regulation of Dnmt1 by nucleosome structure and histone tails. FEBS J. 2017 Oct;284(20):3455-3469. doi: 10.1111/febs.14205. Epub 2017 Sep 11. PMID: 28834260.

5. **Brueckner L**, van Arensbergen J, Akhtar W, Pagie L, van Steensel B. High-throughput assessment of context-dependent effects of chromatin proteins. Epigenetics Chromatin. 2016 Oct 18;9:43. doi: 10.1186/s13072-016-0096-y. PMID: 27777628; PMCID: PMC5069885.

PhD Portfolio

Name of the PhD student:	Laura Brueckner
Erasmus MC department:	Cell Biology
Graduate School:	Onderzoekschool Oncologie Amsterdam, Netherlands Cancer Institute
PhD period:	2013-2018
Promotors:	Prof.dr. Bas van Steensel, Prof.dr. Fred van Leeuwen

PhD Training

Courses

2014	OOA Course "Histopathology of human tumors", Amsterdam
2015	OOA Course "Epigenetics and non-coding RNA: from mechanism to disease", Amsterdam
2016	6th Winter School of the Collaborative Research Centre TRR81 "Chromatin Changes in Differentiation and Malignancies", Kleinwalsertal
2016	MGC "Epigenetic Regulation in Health and Disease", Rotterdam
2016	EMBO/FEBS Lecture Course "Chromatin and the Environment", Spetses
2017	7th Winter School of the Collaborative Research Centre TRR81 "Chromatin Changes in Differentiation and Malignancies", Kleinwalsertal

Seminars and Workshops

2014-2015	OOA PhD student retreat (poster/oral presentation)
2013-2018	NKI research club (oral presentation)
2013-2018	Weekly department meetings (oral presentation)
2013-2018	International NKI seminar, lunch meetings with speaker
2013-2018	Annual Supervisory Committee Meeting

(Inter-)national Conferences

2014-2018	Cancer Genomics Annual Meeting, Amsterdam
2014-2018	Dutch Chromatin Day, the Netherlands
2015	21st Century Genetics: Genes at Work, Cold Spring Harbor (poster presentation)
2016	Dynamics of Genome Structure – 4DGenome ERC Synergy Project, Barcelona (poster presentation)
2016	4D Nucleome Annual Meeting

2017 France-Japan Epigenetics Workshop, Paris (oral presentation)

2018 Nuclear Organization & Function, Cold Spring Harbor (oral presentation)

Teaching

2017 Supervision of a Master student (6 months of guidance in lab work, presenting, writing)

Acknowledgements

First of all, to my supervisor and promoter Bas van Steensel, thank you for giving me the opportunity to work in your lab. You made sure that I was never limited in my research in terms of necessary materials and you were always available when I needed guidance or feedback. You have high standards for quality which you enforce without compromise. This is something that left a strong impression on me. Paying attention to detail, proper quantification and presentation style are among the many things I learned from you. Also, thank you for all the conferences and meetings you let me join, these were invaluable experiences.

Fred van Leeuwen, I would never have finished this thesis without you giving me a push. Thank you so much for the support and encouragement that I needed and also for agreeing to be co-promoter.

The other group leaders in the division of Gene Regulation: Reuven Agami, Elzo de Wit, Benjamin Rowland, Tineke Lenstra – thank you for your feedback during our B4 meetings, I always enjoyed the presentations and discussions.

My lab colleagues:

Eva, you agreed to start a rowing course with me which led to many fun trips on the Amstel. We even made it around NEMO that one time! Your creative spirit and can-do attitude are amazing. I always enjoyed your drawings and I will never forget the “Bioline” cake by Simon Baker. Thank you for helping me out during difficult times and also thank you for getting me my Amsterdam apartment of five years. I hope you will expand your skiing skills in Sweden.

Carolyn, as my almost neighbor you made it much easier feeling welcome in Amsterdam. I was so happy having an original Aussie in the lab. Also, thanks for the amazing food!

Omar, you were always encouraging me when I hit a rough patch. Thank you so much for your support and your friendly attitude. I hope I can visit Shina and you one day in Edinburgh or we can meet up in Vienna. Sorry for all the Sombreros and Burritos, I learned a lot about Mexico thanks to you. Maybe you can spot what the title image was inspired by?

Marcel, I still have photos of your lab decorations and I laugh every time I look at them. Thank you for always playing good music, every Monday should be Metal Monday.

Christ, if you ever get tired of science you can still become a master chef without problem. Or a professional dancer? Thank you for always helping me out with my bioinformatics problems.

Tom, you were my other go-to person for informatics stuff. You are extremely smart but always patiently helped me with the stupidest of problems. I actually started doing bike tours after making fun of you for doing so.
Ludo, also thanks to you for helping me with anything informatics-related and the nice lab dinners we had in your garden.
Tao, you are full of interesting ideas and conversations with you never got boring. I am glad you shared my enthusiasm for Karaoke.
对了，我认真地开始学中文！谢谢你教我中国的文化的东西。
Ezequiel, the tower of Sauron, Italian hands plant and lightning buzzer auditorium are things that still make me smile when I think of them.
Ruben, Miguel, Federico and Stefano, the lab retreat in Mieussy is in my collection of happy memories. I wish you all the best for your research.
Suzanne, I enjoyed chatting with you whenever I came to your office for a conference trip.

My collaborators from the Becker lab, Peter and Alessandro, it was great publishing a paper with you. Thank you for letting me include this publication in my thesis.
From the Gilbert lab, Dave, Jiao and Peiyao and Juan Carlos, it was wonderful working with you! Also, thanks to Andy Belmont and Jian Ma for your helpful feedback and discussions.

Other people I met at the NKI:
My PABs Tessy and Chelsea, my time at the NKI would not have been the same without our coffee breaks or days/nights out in Amsterdam. Tessy, thank you for introducing me to important parts of Dutch culture, such as that nice song about a girl that got burned or Salsa Tequila. Thank you for making me feel at home in Amsterdam and taking me fun places. Chelsea, I'm so glad I got to name your cats. I have many happy memories thanks to hanging out with you and our shared love for sushi and Sublime. I'm still amazed by how extremely strong you are! Lauren, LYLT! Thank you for your hospitality when I came to Toronto.

Gödze and Rui, you were always brightening up my lunch breaks. Alba, unfortunately we did not get to see each other much after you left the NKI but I still hope to meet your little family one day. Robin, your bad taste jokes always made my day. Santiago and Lorenzo, thanks for helping me out with TRIP or TagMap.

Waseem, the reason why I joined the van Steensel lab is because of your TRIP technique. You were so welcoming and patient whenever I needed help. I wish I would have had more time to talk to you.

My rowing team, Sophie and Min, I was looking forward to our Amstel trips every week. Sam, Nabil and Liam. I enjoyed hanging out with you and talking about something other than science for a change. Thanks for discovering the holy chicken place.
Tatiana, I will forever be grateful for every DnB night you went with me, even when you had to drag me out of my cave. Cayci and Selamina, I'm glad we met up at the Meetup!
Asami, you are my idol and I hope you will achieve whatever you want in your career. Many bowls of ramen were enjoyed together with you, Adina and Signe.

Mama und Papa, Danke dass ihr mich immer unterstützt habt. Diese Arbeit ist euch gewidmet.