UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE des SCIENCES de la VIE et de la SANTÉ

THÈSE présentée par :

Joanna Weronika JACHOWICZ

soutenue le : **17 Décembre 2015**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Aspects moléculaires et cellulaire de la biologie du développement

Molecular mechanisms underlying heterochromatin formation in the mouse embryo

THÈSE dirigée par :

Dr Maria Elena TORRES-PADILLA, Université de Strasbourg

RAPPORTEURS :

Prof Didier TRONO	EPFL, Lausanne
Dr Andrew BANNISTER	The Gurdon Institute, Cambridge

AUTRES MEMBRES DU JURY :

Dr Michael WEBER	ESBS, Illkirch

Table of Contents

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Dr. Maria Elena Torres-Padilla for the continuous support during my studies, for her patience, motivation, and immense knowledge. Her guidance has helped me throughout the research and the thesis writing and without it I would have been probably lost. Thank you for showing me how amazing science is and for proving that one can be a great researcher but still stay a human being. I could not have imagined having a better advisor and mentor for my Ph.D study. Thank you!

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Didier Trono, Dr. Andrew Bannister, and Dr. Michael Weber for accepting to read and comment on my thesis.

I thank my fellow lab mates for the stimulating discussions, enormous patience that they had with me (especially in the beginning!), and all the enthusiasm that they show towards science. In particular, I am grateful to Dr. Yusuke Myianari and Dr. Anas Fadloun for enlightening me the first glance of research. Also I thank my friends at IGBMC for all the fun and crazy scientific and non-scientific conversations we have had in recent years. METP lab members, “get together” team, Rob Schneider’s lab members, I cannot imagine these last 4 years without you! Enormous thanks!

My sincere thanks also go to all the platforms and scientific support at IGBMC, especially to imaging facility. Without their precious support it would not have been possible to conduct this research.

Last but not least, I would like to thank my parents and my friends at home for supporting me spiritually throughout writing this thesis. Thank you for always being there for me!

Abstract

Genomic material within the eukaryotic nucleus can be divided into two functional forms of chromatin: gene rich and actively transcribed *euchromatin* and gene-poor, and often thought as exclusively silenced, *heterochromatin*. Although the DNA of these two states of chromatin is similarly wrapped around core histones forming nucleosomes, they differ in terms of compaction and accessibility. These features are further reinforced given that DNA of heterochromatic regions is methylated and histones display distinctive modifications including global histone hypoacetylation and methylation of H3K9. Another hallmark of heterochromatin is its DNA composition as in general it is composed of repetitive elements (pericentric and centric regions, retrotransposons, endogenous retroviruses), which should remain silenced during the life of the cell. Repressive epigenetic marks and condensed chromatin structure allow the maintenance of the silenced status of heterochromatin and facilitate its inheritance through the cell cycle. Defects in any of above mentioned states often lead to numerous abnormalities, for example improper cell division or abnormal cell cycle progression. All of these events are a potential danger for cell integrity and might generate a significant risk for an organism to develop diseases such as cancer. Moreover, reactivation of these elements is correlated with mutations, deletions and genome instability given their ability to retrotranspose into new genomic regions and/or to affect functionally the neighbouring genes. Most importantly, repetitive elements have arisen as potential major regulators of chromatin state. Therefore, understanding the mechanisms behind the formation and maintenance of heterochromatin has arisen as an important topic in epigenetics and chromatin biology. However, to investigate heterochromatin establishment and its further maintenance throughout cell division, an adequate model system is necessary; preferably one which enables to investigate the complexity of a biological problem at the cellular and organismal level.

Mouse preimplantation embryos are a great candidate as model system since the above-mentioned features of heterochromatin and its epigenetic signatures, which are present in most somatic cells, are erased during development and then acquired de novo. Extensive removal of chromatin marks starts during the formation of the germ cells in which the typical heterochromatin state is altered in order to form functional gametes. After fertilization, intense chromatin remodeling and epigenetic reprogramming continue in order to revert into

a totipotent state, which has all the cellular plasticity that is necessary to start a new developmental program. Thus, mouse preimplantation embryos enable to study the establishment of heterochromatin as naturally occurring phenomena which must take place in the newly formed organism. From an ethical point of view, since this issue cannot be examined in humans, the mouse model provides an ideal alternative where these mechanisms are known to be conserved in humans.

My Ph.D work focused on this precise question, focusing on two different heterochromatic regions, pericentric DNA and L1 elements. Namely what are the mechanisms controlling heterochromatin formation in the mouse embryo, as well as in addressing the impact of manipulating their transcriptional activity on early developmental progression.

Avant propos

Dans le noyau eucaryote, le matériel génétique peut être divisé en deux formes fonctionnelles de chromatine : l'euchromatine riche en gènes et activement transcrite, et l'hétérochromatine souvent considérée comme réprimée. Cette dernière est en général composée d'éléments répétés (régions centriques et péricentriques, rétrotransposons, et virus endogènes), décorée par des marques épigénétiques répressives et des structures chromatiniennes condensées qui permettent le maintien de son statut réprimé et facilitent sa transmission au cours du cycle cellulaire. Si l'une des caractéristiques de l'hétérochromatine mentionnées ci-dessus ne fonctionne pas correctement, de nombreuses anomalies peuvent être trouvées dans les cellules, comme par exemple une division cellulaire incorrecte ou une progression anormale du cycle cellulaire. Tous ces événements présentent un grand danger pour l'intégrité des cellules et peuvent générer un risque important pour un organisme de développer un cancer. De plus, la réactivation des éléments répétés est associée à des mutations, délétions et instabilité du génome, du fait de leur capacité à retrotransposer dans de nouvelles régions génomiques et / ou d'affecter le fonctionnement de gènes voisins. Plus important encore, les événements de rétrotransposition ont été corrélés à la progression tumorale. Par conséquent, la compréhension des mécanismes responsables de la formation et du maintien de l'hétérochromatine devient un sujet important dans la recherche sur le cancer. Étant donné que les caractéristiques de l'hétérochromatine mentionnées ci-dessus et ses signatures épigénétiques sont présentes dans la plupart des cellules somatiques un modèle d'étude différent doit être utilisé. L'embryon préimplantatoire de souris est un excellent candidat du fait que la vaste élimination des marques chromatiniennes ait lieu au cours du développement précoce et que les caractéristiques typiques de l'hétérochromatine sont altérées lors de la formation des gamètes. Cela permet ainsi d'étudier la mise en place de l'hétérochromatine, comme elle se produit naturellement dans un organisme nouvellement formé. Afin d'étudier la formation de l'hétérochromatine dans **l'embryon préimplantatoire de souris**, je me suis concentrée sur **deux régions génétiques différentes**, dans le but notamment de **découvrir les mécanismes qui conduisent à la répression** et le **rôle distinct qu'ils peuvent jouer pendant le processus de développement et la division cellulaire**.

I. L'hétérochromatine et les répétitions péricentriques

L'hétérochromatine centrique et péricentrique se compose respectivement de répétitions de séquences satellites mineurs et majeurs qui ont une organisation spatio-temporelle spécifique au cours du développement précoce. Chez le zygote, les répétitions péricentriques se concentrent autour des NLBs (Nucleolar-like Bodies) - les précurseurs des nucléoles - formant des structures en forme d'anneau dans les noyaux. Après la première division une organisation en "chromocentres" semblable aux cellules somatiques est progressivement initiée de telle sorte qu'au stade 4 cellules les structures en forme d'anneaux ne sont plus présentes. La répression de ces régions pourrait se produire en même temps du fait qu'un pic dans leur transcription est seulement détecté à la fin du stade zygote et dans les embryons au stade 2 cellules. De plus, de précédents travaux dans le laboratoire suggèrent qu'il pourrait y avoir un lien entre la localisation spécifique des répétitions péricentriques dans le noyau et leur répression. En effet, dans les embryons exprimant la mutation H3.3K27R la formation des chromocentres est perturbée, la transcription des satellites majeurs augmente et la protéine d'hétérochromatine HP1β est délocalisée. Ainsi, le **premier objectif principal** de cette partie de mon travail était de **déterminer si dans le noyau la localisation spécifique de la chromatine péricentrique est importante pour sa répression**. Pour y répondre, j'ai réalisé des expériences dans lesquelles j'ai artificiellement perturbé la localisation des régions péricentriques au cours du développement précoce pour forcer leur délocalisation des NLBs vers la membrane nucléaire. Ces expériences montrent que l'organisation spatiale spécifique des domaines péricentriques est essentielle pour leur répression ainsi que pour leur organisation correcte. De plus, mes résultats suggèrent que les défauts d'organisation de l'hétérochromatine conduisent à des défauts de division cellulaire et de prolifération.

II. L'hétérochromatine sur les séquences répétées

Environ la moitié du génome des mammifères est composée d'éléments répétés transposables (Transposable Elements : TE) qui sont regroupés en deux classes : les rétrotransposons et les transposons à ADN. Bien que proportionnellement les TEs représentent une grande partie du génome, seulement une faible proportion de ces éléments est capable de « sauter et se coller » ; ce qui est le cas des LINE-1 sans LTR (Long Interspersed Nuclear Elements L1) qui seraient les TEs les plus actifs chez la souris et pour lesquels les insertions semblent être les plus récentes dans le génome murin. Deux familles de L1, les A et F, ont été confirmés comme étant les éléments L1 les plus jeunes et plus abondants du

génome murin ; des essais de rétrotransposition ayant aussi démontré qu'ils seraient toujours actifs[8]. Cette activité est l'une des raisons principale pour laquelle les TEs, et en particulier les éléments L1, représentent une grande menace pour la stabilité du génome, leur dérégulation étant fréquemment observée dans de nombreux types de cellules cancéreuses. Des études approfondies dans des cellules somatiques et des cellules souches ont révélé certains mécanismes moléculaires les régulant et réprimant leur activité transcriptionnelle, notamment via la méthylation de l'ADN, les modifications d'histone et les voies à ARN. Cependant, la régulation des éléments L1 pendant la période de reprogrammation épigénétique et les divisions cellulaires rapides qui ont lieu au cours du développement reste encore indéterminée. De manière intéressante, des études récentes montrent que les L1 sont réactivés après la fécondation, leur activité transcriptionnelle diminuant ensuite progressivement. Du fait que le niveau de méthylation des L1 reste faible dans les embryons précoces de mammifère[11], leur mode d'expression particulier après la fécondation ne peut s'expliquer par un effet secondaire de l'activation générale du génome mais pourrait plutôt avoir un intérêt fonctionnel. **Quel est cette fonction si il y a, et comment les L1 sont régulés dans les embryons préimplantatoires, restent indéterminés et constituent ainsi le second objectif principal de ma thèse.**

Pour y répondre et étudier l'importance et la possible fonction et régulation des éléments L1, j'ai décidé d'utiliser une approche expérimentale basée sur les TALEs (Transcription Activator–like Effectors) qui sont des protéines liant l'ADN sur des séquences spécifiques[12], récemment découvertes chez *Xanthomonas sp*. Plus précisément, j'ai ciblé les éléments L1 avec des TALEs spécifiques fusionnés à des protéines modifiant l'activité transcriptionnelle. Cette approche m'a permis d'étudier en détail leur processus d'activation et de répression. Ainsi, dans la seconde partie de ma thèse, j'ai conçu et généré les outils adéquats pour répondre à ces questions et vérifier leur expression, localisation, capacité de liaison à l'ADN, et leur habilité à activer les L1 dans les cellules ES de souris. Ces résultats ont ensuite été utilisés pour choisir les meilleurs candidats pour effectuer les expériences dans les embryons de souris. La correcte localisation nucléaire et la capacité à activer les L1 in vivo ont été vérifiés dans les embryons pour les trois TALE-L1 les plus prometteurs. Du fait du mode d'expression connu des L1s dans l'embryon, avec une activation de transcription initiée au stade zygote tardif et une diminution à partir du stade 8 cellules, nous étions particulièrement

intéressés de voir si la répression des L1s est nécessaire pour le développement, et ainsi nous voulions prolonger leur phase d'activation au-delà du stade 2 cellules et observer le phénotype. Ainsi, après micro-injection des activateurs de L1, nous étions capables de détecter un signal plus important des transcrits de L1 dans l'embryon au stade 4 cellules, comparé aux embryons contrôles non-injectés ou injectés avec des TALE-L1 manquant le domaine d'activation. Le groupe expérimental montrait aussi un niveau d'expression plus élevé d'Orf1p codée par les L1, suggérant que nous avions réussi à moduler la traduction des L1 dans l'embryon. De plus, quand les trois groupes ont été cultivés en parallèle, seulement 50% des embryons présentant un niveau plus élevé de L1 ont atteint le stade blastocyste, alors qu'environ 90% des embryons contrôles se sont développés normalement. Ces résultats suggèrent que lors de la perturbation du mode spécifique d'expression des L1, les embryons ne sont plus capables de se développer normalement. Ceci suppose qu'un possible rôle régulateur des éléments L1 pourrait avoir lieu à ce moment, c'est à dire en activant/réprimant la transcription des régions voisines ou de manière plus générale en agissant sur l'organisation du génome, leur surexpression conduisant à une chromatine plus ouverte. En conclusion, ceci représente la première tentative pour élucider la biologie des éléments L1 dans l'embryon précoce de souris par l'utilisation de modificateurs de transcription ciblés spécifiquement, montrant que la sous régulation des éléments L1 est nécessaire pour le développement.

En conclusion, associé à la première partie de mon projet sur l'hétérochromatine péricentrique, j'espère que mon travail contribue à la compréhension des mécanismes responsables du contrôle de l'intégrité du génome.

I. Introduction

1. Characteristics of (hetero)chromatin in mammalian cells

1.1. Brief introduction to chromatin

In eukaryotic cells the genetic material is organized into a complex structure called chromatin which is composed of DNA and proteins and localized in a specialized compartment - the nucleus (Fig.1). The nucleosome is the fundamental unit of chromatin, and is composed of ~147 base pairs of DNA wrapped around an octamer of four core histones: H3, H4, H2A, H2B. The core histones are predominantly globular with an exception of the N-terminal "tail" that can be chemically modified. Posttranslational histone modifications (PMTs) refer to the chemical changes occurring on the specific amino acid residues of histones and include acetylation, methylation, phosphorylation, ubiquitination, SUMOylation, and probably many others, less studied. The information about the patterns of histone modification across the genome offers insights into the regulatory state of promoters, genes, and other regions, as specific modifications can affect gene expression and chromatin compaction depending on the type of the modification, position of the amino acid and the number of modified residues (Kouzarides 2007). Several types of enzymes are known to catalyze addition or removal of histone modifications i.e. methylation of lysines and arginines is performed by histone methyltransferases (HMTs) or acetylation is catalyzed by histone acetyltransferases (HATs). Histones can be not only modified by PMTs by also replaced by their non-canonical variants which display different properties. Although most of the histones are synthesized during S-phase which allows their deposition behind the replication fork, replication-independent replacement of histones can also occur. Both pathways enable exchange of the canonical histones into non-canonical variants by histone chaperons, which may lead to the change of transcriptional state of chromatin. In addition, ATP-dependent chromatin remodeling enzymes mediate rearrangements of the chromatin as they restructure and slide nucleosomes, or eject histones, thereby regulating the dynamic properties of chromatin (Henikoff and Smith 2015).

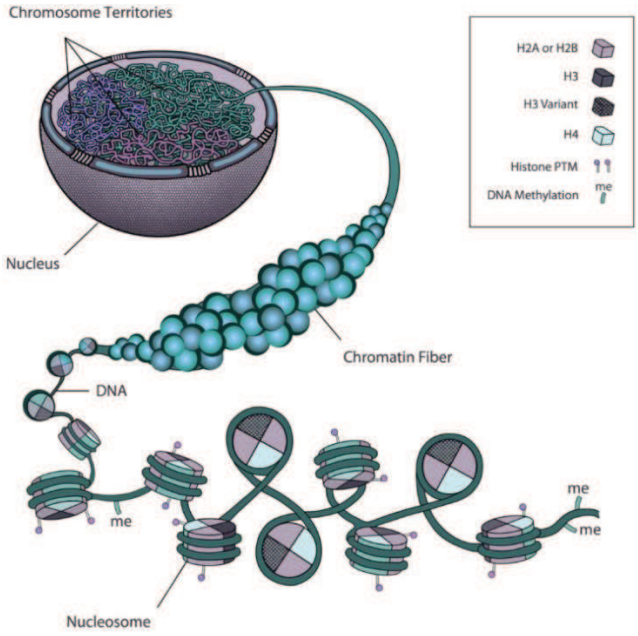


Figure 1. Chromatin organization within the nucleus

Scheme depicting different aspects of chromatin regulation. PTM - post-translational modification. Chromosome territories within the nucleus, shown in different colors, are composed of chromatin fiber, which, in turn, contain packed nucleosomes. (Modified from Rosa and Shaw 2013).

1.2. Euchromatin versus heterochromatin

In mammalian cells, chromatin is organized into two distinct domains known as *euchromatin* and *heterochromatin*. Euchromatin is the gene-rich part of the genome which is more accessible to the transcriptional machinery thanks to the ‘‘open state’’ and “flexibility” which leads to more permissive state and higher probability of expression (Fig. 2). Heterochromatin, on the contrary, is gene-poor and assembles into well compacted domains which have more ‘‘repressive’’ chromatin structure, thus, contain mainly transcriptionally silent regions (Fig. 2) (Jost, Bertulat, and Cardoso 2012). Moreover, heterochromatin shows a distinct pattern of replication, being replicated mostly at the late S-phase (Rhind and Gilbert 2013). Hence, as a general rule these diverse parts of the genome have different chromatin configurations which correspond to a functional state, and can be distinguished by distinct factors present on them i.e. histone modifications or DNA methylation.

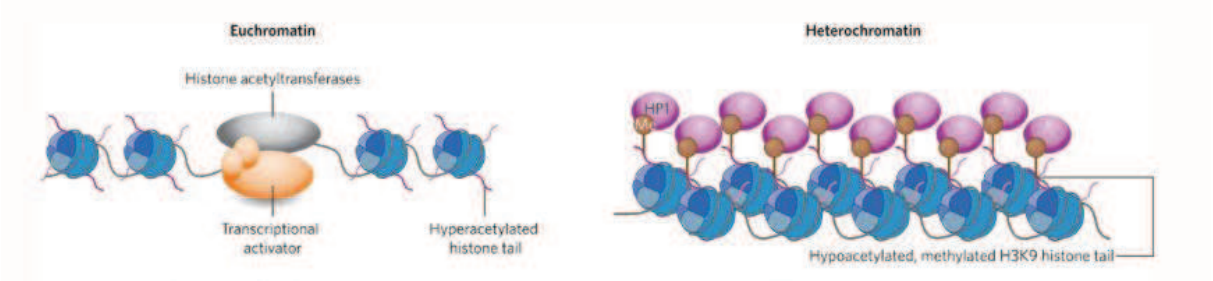


Figure 2. Euchromatin and heterochromatin organization in mammalian cells

Scheme representing differences in heterochromatin and euchromatin organization depicting nucleosome position (in blue), modifications of histones' tails (in purple and brown), transcription factors. (Modified from Grewal and Elgin 2007).

Although the characteristics of euchromatin and heterochromatin seem to be well defined, it is worth mentioning that there are some developmentally regulated loci, where the chromatin state can change in response to cellular signals and gene activity. These regions are referred to as "facultative heterochromatin" and are associated with proteins from Polycomb-group repressive complexes (PRC), distinct histone modifications like H3K27me3, and specific histone variants like macroH2A. "Constitutive heterochromatin" on the other hand, is marked by H3K9me3, H4K20me3, H3K64me3 and HP1β, and contains high density of repetitive sequences and transposable elements. It is mostly found at telomeres and pericentromeric regions - large blocks of chromatin flanking centromeres (Fig. 3). In the nucleus of interphase somatic cells pericentromers from several chromosomes cluster together into the foci called chromocenters that stain intensely by DNA dyes. The centromeres from the same chromosomes can be found around these regions, and are characterized by the presence of the centromere-specific histone H3 variant CENP-A (or Cen-H3) (Guenatri et al. 2004). In mice, centromeric and pericentromeric heterochromatin corresponds to minor and major satellite sequences, respectively (Fig. 3) (Probst and Almouzni 2008). Constitutive heterochromatin from telomeres, centromeres and pericentromeres remains silenced and condensed throughout the cell cycle as it is thought to enable the formation of structures that are essential for chromosomal function. Disruption of the establishment, condensation and/or silencing of pericentromeric chromatin can indeed cause centromere malfunction, incorrect chromosome segregation, and nuclear disassembly (Peters et al. 2001; Bouzinba-Segard, Guais, and Francastel 2006)

An additional feature of heterochromatin is its ability to propagate, and thereby influence gene expression in a region-specific, sequence-independent manner as it happens

during mammalian X-inactivation. This key feature of heterochromatin facilitates the control of the loci that are otherwise incapable of recruiting effectors by themselves. Thus, the heterochromatin can spread in *cis* and is coordinately regulated in *trans* and both these features make heterochromatin indispensable not only for genome organization but also because of its regulatory role (Grewal and Jia 2007).

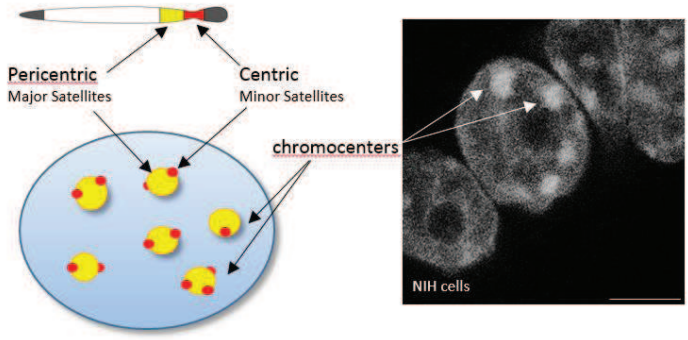


Figure 3. Pericentric and centric repeats in mouse cells

Representation of chromosomal and nuclear location of pericentric (in yellow) and centric (in red) heterochromatin. Scale bar – 10 microns.

1.3. Maintenance and establishment of heterochromatin in somatic and stem cells

Chromatin can be modified and controlled by various epigenetic mechanisms at different levels of its organization, for example at the DNA itself, at the nucleosomes, or even at the higher-order structures which includes nuclear compartmentalization. All these events can lead to more open or closed chromatin configuration and result in on or off state of expression. Thus, while studying mechanisms responsible for chromatin organization, one has always bear in mind a complex picture of interactions whereby active and passive pathways operate in parallel to ensure proper control of chromatin state and gene expression to comply with cellular needs. Importantly, functionally distinct regions have to be regulated in different ways either allowing rapid and dynamic changes like in promoters or enhancers, or enabling complete transcriptional shut downs like in pericentromeric repeats. At the same time, the structure of the chromatin *per se* should stay flexible enough to enable proper progression of replication, which implies existence of tight regulation at many different levels of chromatin organization. In recent years many studies showed that the regulation of the heterochromatic

state is equally dynamic and complex as for gene-rich regions. Surprisingly, the mechanisms that initiate its formation and preserve its distinction from euchromatin, still remain elusive. Nevertheless, our current knowledge shows that maintenance and establishment of these repressive states are largely determined by a complex network that includes: enzymes able to modify DNA and histones tails; complexes with nucleosome remodeling activities; transcription factors (TFs); non-coding RNA; and nuclear organization (Fig.4). Each of the above mentioned pathways will be briefly discussed in this paragraph.

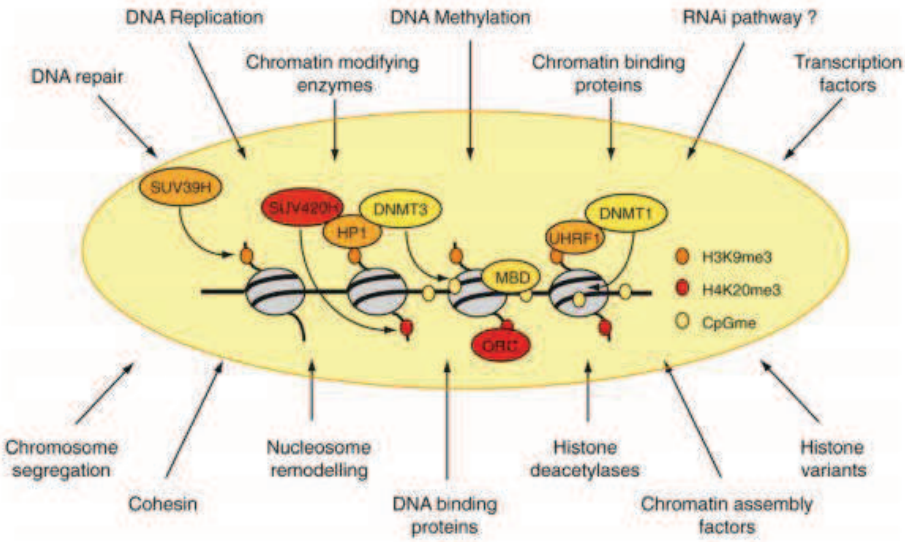


Figure 4. Complex network of factors controlling chromatin state

SUV39H is the responsible HTMase for H3K9me3 on pericentromeres, a histone mark recognized by HP1 proteins. HP1 proteins interact and recruit SUV420H and DNMTs, leading to H4K20me3 and DNAmethylation, respectively. These epigenetic marks function also as docking sites, like H4K20me3 for ORC (origin of replication complex) proteins and CpGme for MBDs (factors with a methyl-binding domain). An alternative for DNMT recruitment might be through UHRF1 that directly interacts with DNMT1 and might read the H3K9me3 mark. (Modified from Nehme Saksouk, Simboeck, and Dejardin 2015).

1.3.1. DNA methylation

DNA methylation refers to the methylation of cytosine residues in CpG dinucleotides, and is catalyzed by DNA methyltransferases (DNMTs): DNMT1 maintains DNA methylation, DNMT3A/3B exhibit both maintenance and *de novo* methylation activities, and DNMT3L which lacks the characteristic N-terminal catalytic domain but acts as a crucial activating cofactor of DNMT3A/B (Ooi, O'Donnell, and Bestor 2009). In general, DNA methylation is viewed as a final state of silencing and maintenance of a transcriptionally inactive state and is

one of the hallmarks of heterochromatin. Thus, hypermethylation of CpG dinucleotides is commonly present at genes-depleted regions including transposable elements or repeats, whereas hypomethylation is found on exones, introns and intergenic regions. There are several mechanism how DNA methylation can impact transcription, for example, through interference with binding of transcriptional factors to imprinting loci like (Bell and Felsenfeld 2000). Another way occurs via 5-methyl cytosine binding proteins (MBP) which interpret methylation levels and recruit chromatin regulators that induce changes in chromatin state as in the case of HMTs recruitment by MBP1 (Sarraf and Stancheva 2004) or chromatin remodeler NuRD via MBP2. Interestingly, even though DNA methylation can be seen as a crucial guardian of gene expression in differentiated cells, its function in heterochromatin establishment during development, remains questionable. Upon depletion of all three DNMT enzymes in mESC (TKO), some heterochromatic regions indeed display transcriptional up-regulation, however, global gene expression is not affected, cells proliferate normally and global pattern of histone marks remains unchanged (Tsumura et al. 2006). Moreover, mice deficient for *de novo* DNMTs display no phenotype prior to implantation (Okano et al. 1999). In fact, alternative factors, described later in this chapter, were shown to compensate lack of DNA methylation implying that it is not a crucial player for heterochromatin establishment in undifferentiated cells.

1.3.2. Histone modifications and related pathways

In fission yeast and higher eukaryotes, histones in heterochromatin are hypoacetylated and selectively methylated at lysine 9 of histone H3. Lysine methylation can exist in three flavors: mono-, di, and tri-methylation, which are catalyzed by different enzymes: Prdm3 and Prdm16 that monomethylate H3K9 in the cytoplasm, which is then converted in the nucleus by the Suv39h1/2 enzymes (also called KMT1A/B) to H3K9me3 (Pinheiro et al. 2012). Other histone methyltransferases are also involved in that conversion i.e. SETDB1 (ESET) that can trimethylate H3K9 (Dodge et al. 2004), and G9a, which acts to mono- and di- methylate H3K9 (Tachibana 2002). However, their activity is mostly restricted to euchromatin, whereas Suv39h1/2 operates on heterochromatin (Martens et al. 2005). Suv39h1/2 role in the maintenance of heterochromatin has been mostly linked to the recruitment of multiple silencing factors i.e. HP1 proteins, Suv4-20h1/h2 enzymes, and DNMT1, which are all lost from pericentromeric chromatin in dn Suv39h1/2 mESC (Schotta et al. 2004). Additionally,

Suv39h1/h2 are required for proper mitosis as AURORA B, a major component of CPC (chromosomal passenger complex) that controls chromosomal segregation, is depleted from pericentromeric regions in the absence of H3K9me3 (Saksouk et al. 2014). Thus, a pathway in which Suv39h enzymes induce H3K9me3 which is then bound by HP1, arises as a hallmark of mammalian heterochromatin (Bannister et al. 2001). Given the ability of HP1 to bind to numerous proteins that are implicated in heterochromatin formation, including HDACs, it has been suggested that HP1, when bound to methylated H3K9, serves as an assembly platform. Indeed, HP1 mediates the recruitment of Suv4-20h HMTs to pericentromers where they catalyze di- and tri-methylation of H4K20 (Schotta et al. 2004). H4K20me3 is highly enriched at pericentric heterochromatin, telomeres, imprinted regions and repetitive elements where it is probably involved in transcriptional silencing. At pericentromeric regions, however, Suv4-20h1/2 activity seems to be engaged also in chromocenters' condensation and proper mitosis as cells deficient for these enzymes display chromosome segregation defects that coincide with reduced sister chromatid cohesion (Hahn et al. 2013). Thus, the recruitment of Suv4-20h1/2 to pericentromers might be more related to heterochromatin organization and global nucleus arrangements than silencing *per se*.

Interestingly, when H3K9me3 activity is impaired, "rescue mechanisms" take over, mainly the Polycomb group proteins (PcG) which are recruited to compensate and maintain heterochromatic, silencing environment. In mouse ES cells, in which Suv39h1/h2 genes have been knocked out this alternative mechanism is switched on and levels of H3K27me3 become elevated (Martens et al. 2005). Similar situation has been observed in Dnmt TKO mESCs where the loss of constitutive heterochromatin marks on pericentromeric repeats was accompanied by the acquisition of factors belonged to Polycomb group, together with the enrichment of H3K27me3 (Saksouk et al. 2014). In general, PcG are hallmarks of facultative heterochromatin and are known to maintain cell fate by repressing hundreds of genes through the activity of two main polycomb repressive complexes (PRC): PRC1 and PRC2. PRC2 contains the H3K27 methyltransferase EZH2 (also called KMT6A), as well as EED, SUZ12, RbAp46/48 (RBBP4/7), and JARID2 (Jumonji/ARID domain-containing protein). The latter, together with Polycomb-like family (Pcl) proteins, have been suggested to be responsible for the recruitment of PRC2 to target genes in mammalian cells. PRC1 contains the E3 ubiquitin ligases RING1A and B (RNF1 and 2) that mediate H2AK119Ub. Canonical PRC1 complexes also contain Cbx - proteins that

recognize and bind H3K27me3 through their chromodomain, leading to coexistence of both PRC2 and PRC1 on the same loci. Non-canonical PRC1 complexes contain Rybp (together with additional proteins, such as L3mbtl2 or Kdm2b) rather than the Cbx proteins, thus, their recruitment to target genes is mostly independent of H3K27me3 (Aloia, Di Stefano, and Di Croce 2013). Lysine demethylase Kdm2b has been proposed as one of the factors that enable binding of the non-canonical PRC1 to unmethylated DNA (Wu, Johansen, and Helin 2013). Surprisingly, this PRC1 complex can also recruit PRC2 in the H2AK119Ub-dependent manner which leads to the deposition of H3K27me3 mark on unmethylated promoters (Blackledge et al. 2014). In addition, PRC1 and PRC2 complexes have been shown to associate and form nuclear foci on their targets which stabilizes gene silencing and suggests another role for Polycomb - in 3D chromatin organization (Cavalli 2015).

1.3.3. Histone variants

Specific histone variants determine the structure of distinct regions of heterochromatin and play a role in their maintenance. For example, when centromere-specific histone H3 variant CENP-A is not present, cells show severe defects in chromosome segregation suggesting its relevance for the integrity and function of kinetochores. In addition to CENP-A, in higher *eukaryotes* there are three other H3 variants: H3.1 and H3.2 which are mainly expressed in S-phase and deposited by the CAF1 complex, and H3.3 expressed throughout the cell cycle with replication-independent deposition by the HIRA and (Tagami et al. 2004) DAXX/ATRX complexes (Lewis et al. 2010; Drané et al. 2010). HIRA is required for localization of H3.3 to actively transcribed regions, while ATRX is essential for H3.3 incorporation at silent regions such as telomeres, and has been shown to play a role in heterochromatin formation and chromosome segregation during mitosis and meiosis (Rabindranath De La Fuente et al. 2004). DAXX is also present on repetitive regions and its deletion leads to disruption of chromocenters, thus, it is thought to be involved in the structural organization of the nucleus (Rapkin et al. 2015). The specific role of DAXX and ATRX chaperons at pericentromeric regions, however, is unclear but it might be linked to transcription of the locus as in Daxx -/- MEFs and in ATRX and H3.3 siRNA depleted cells levels of major satellite transcripts are lower (Drané et al. 2010). ATRX/DAXX complex can also replace the canonical H3.1/H3.2 with the H3.3 variant at specific classes of transposable elements - class I and class II ERVs. It has been previously shown that these TEs are silenced

through H3K9me3 deposited by SETDB1 and its co-repressor complex containing KRAB-associated protein 1 (KAP1) (Rowe et al. 2010; Karimi et al. 2011). Elsässer and collegues suggest a role for H3.3 in this process as well because H3K9me3 and KAP1 occupancy is reduced at some class I and II ERVs upon H3.3 deletion. Thus, the authors draw a link between ERV-associated H3K9me3 and H3.3 deposition, and suggest that the recruitment of KAP1 and H3.3 by DAXX is co-dependent and occurs upstream of the recruitment of SETDB1 (Elsässer et al. 2015).

Exchange of another core histone - H2A - also takes part in the formation of heterochromatin as H2A.Z has been demonstrated to be involved in the recruitment of HP1 at pericentromeric loci, and direct binding to the pericentric heterochromatin-binding protein INCENP (Fan et al. 2004). These interactions facilitate folding of chromatin into high order structures but also play a role in the chromosome segregation (Rangasamy, Greaves, and Tremethick 2004). Although H2A.Z is present on constitutive heterochromatin in some developmental stages, in general it is found on facultative heterochromatin i.e. the inactive X chromosome, where it becomes monoubiquitylated by PRC1 (Sarcinella et al. 2007). Interestingly, its occupancy has been also correlated with the lack of DNA methylation and with H3K4me3/H3K27me3 loci which suggest that H2A.Z brings a dynamic instability to chromatin structure increasing access to chromatin-modifiers. Another variant of H2A - macroH2A is also implicated in heterochromatin regulation as it is mainly present at discrete regions of facultative heterochromatin within the inactive X chromosome that alternate with regions of constitutive heterochromatin (Henikoff and Smith 2015).

1.3.4. DNA binding factors

Heterochromatin function can be also affected by various DNA binding factors which interact with chromatin remodelers and trigger changes in chromatin organization. One of the examples comes from the experiments in mESC where in the absence of H3K9me3 or DNA methylation, (dnSuv30h1/h2 or TKO) the levels of methylation-sensitive DNA binding protein BEND3 arise. It leads to the recruitment of PRC2 via MBD3/NurD chromatin remodeler complex and, only in the case of TKO, PRC1 components, presenting a mechanism which facilitates silencing and operates on unmethylated loci (Saksouk et al. 2014). NoRC is another chromatin remodeler which is known to establish repressive chromatin structure via interaction with DNA binding proteins. The biggest subunit of NoRC - TIP5 - is recruited to

distinct genome loci by specific chromatin-associated proteins, for example, to rDNA by TTF-I, to centromeres by CENP-A and to telomeres by the shelterin complex, and then interacts with HP1α and with histone-modifying enzymes to mediate higher-order chromatin structure. Knockdown of NoRC leads to de-condensation of heterochromatin, abnormalities in mitotic spindle assembly and impaired chromosome segregation (Postepska-Igielska et al. 2013). LSH (lymphoid specific helicase) is another factor involved in the formation of normal heterochromatin structure as it has been shown to maintain nucleosome density leading to more compacted state of the chromatin. Moreover, it plays a specific role in acquisition of *de novo* DNA methylation at transposable elements and pericentromeric repeats (Huang et al. 2004). Its ATP-binding site seems to be crucial for the stable association of Dnmt3 to DNA (Ren et al. 2015) and upon its depletion cells display lower levels of DNA methylation and upregulation of major satellites (Huang et al. 2004).

Recent discoveries in MEF cells have implicated a role for sequence-specific transcription factors in heterochromatin formation/maintenance. These TFs have been shown to use distinct binding sites as targets and their recruitment lead to heterochromatin silencing most likely by RNA dependent pathways (Bulut-Karslioglu et al. 2012). For example Pax3 and Pax9 transcription factors are components of mouse heterochromatin as they localize to DAPI dense regions and can bind to HP1 and KAP1. Upon their depletion accumulation of major satellite transcripts from both strands occurs, and reduction in H3K9me3 and H4K20me3 takes place which, in consequence, leads to genome instability. The putative role of these transcription factor binding sites might be to serve as a base for bidirectional transcription, which is then silenced by heterochromatin formation (Bulut-Karslioglu et al. 2012). This is not such a new concept as the establishment of higher-order chromatin structure at repetitive elements has been already shown to depend on specific noncoding RNAs, including pRNA, TERRA (Postepska-Igielska et al. 2013) and major satellite RNA (Almouzni and Probst). Moreover, the intersection of the RNA-interference (RNAi) pathway and heterochromatin formation has been well documented in *Schizosaccharomyces pombe* where small interfering RNAs generated from flanking outer repeat of a centromere direct histone H3 lysine 9 methyltransferase Clr4 to homologous loci which then recruit Swi6 (homologous of HP1). H3K9me and Swi6 are required to establish CENP-A which further leads to spreading of heterochromatin (Folco et al. 2008). Interestingly, siRNA seem to be important only for the

recruitment of Clr4 as heterochromatin can be still formed without operating iRNA pathway but with artificial tethering of Clr4 to the loci (Kagansky et al. 2009).

1.3.5. Small RNAs and RNAi pathway

Although RNA silencing pathways in mammals work mainly as a post-transcriptional gene silencing (PTGS) mechanism, they can also alter chromatin structure and silence genes at the transcriptional level (Morris et al. 2004). The core machinery includes small RNAs (sRNA) that are complementary to target genomic sequences, which form complexes with the proteins from the Argonaute family (Ago). Depending on the biogenesis of sRNA and the mechanism of silencing, 3 main pathways can be distinguished: RNA interference (siRNA), microRNA (miRNA) and piRNA pathways. siRNA and miRNA require Dicer activity whereas piRNA rely on Piwi proteins: Mili, Miwi1, and Miwi2 (in mouse). Whereas piRNAs are known to function in the male germline to maintain genomic integrity by epigenetic mechanisms including *de novo* DNA methylation of transposable elements (Aravin et al. 2007), siRNAs and miRNAs mainly mediate PTGS through degradation of mRNA and/or inhibition of mRNA translation in the cytoplasm. Nevertheless, they have been also reported to silence heterochromatin and enable acquisition of heterochromatin-specific chromatin modifications such as H3K9me2/3, H3K27me3 or DNA methylation (Li 2014). Moreover, in mouse cells RNA might have a structural role in heterochromatin, as cells treated with RNaseA lose heterochromatic marks (Maison et al. 2002) and RNase treatment of chicken liver cells revealed a role for RNA in the high-order chromatin structure and compaction (Rodríguez-Campos and Azorín 2007). Another evidence coming from human cell lines supports the hypothesis that RNA may have a general structural role in chromatin organization as RNase treatment results in both loss of CENP-C and chromosomal passenger complex (CPC) components from centromeres. This phenotype can be rescued by reintroduction of CENP-C, INCEP and centromeric RNA (Wong et al. 2007). Moreover, loss of Dicer, the main player in the production of siRNA and miRNA, leads to up-regulation of transcription from pericentric heterochromatin and loss of HP1β and H3K9me2/3 from these regions (Kanellopoulou 2005). All of these provide strong argument that RNA is required for silencing and structural integrity of the pericentromeres.

1.3.6. Nuclear organization

The three-dimensional organization of the genome and its importance for the normal functioning of the cell and, in particular, for (hetero)chromatin regulation, attained more attention in recent years. According to the definition by *Nature* nuclear organization refers to *"the spatial distribution of nuclear contents and components in a way that it reflects or facilitates their activities".* Hence, specific regions of the genome tend to localize in the particular nuclear compartments in order to be properly regulated because distinct regions of the nucleus can regulate chromatin in a different way. One of the most striking examples come from the nuclear periphery and nucleolus that, already in the 1950s, have been associated with more dense chromatic regions, most likely heterochromatin.

The nuclear periphery can be separated into two main sub-compartments: nuclear pores and nuclear lamina. Whilst localization to the nuclear pores has been suggested to be correlated with transcriptional activation of genes, the nuclear lamina is most likely associated with heterochromatin and linked to the repression of genes and repetitive sequences (Akhtar and Gasser 2007). Recent advances in genome-wide high-throughput sequencing confirmed these observation and led to the discovery of lamin-associated domains (LADs) which in general are gene-poor and enriched for heterochromatic silencing marks such as H3K9me2 (Kind et al. 2013). Therefore, perinuclear localization appears to regulate chromatin function. Nevertheless, one has to keep in mind that not all genes have the same ability to be modulated by proximity to the nuclear envelope (Finlan et al. 2008) and detachment from the nuclear lamina does not always lead to transcriptional activation as it happens at some regions during ES cells differentiation (Kind et al. 2013). Moreover, even changes in chromatin condensation, without entailing transcriptional activation, are enough to re-localize some genes from the periphery towards the nuclear interior (Therizols et al. 2014). All of the above suggests that in somatic cells, the nuclear periphery functions more as a non-permissive compartment where unused regions are stored, rather than the actual main inhibitor of transcription.

LADs are not the only domains found to be associated with distinct nuclear compartments as the existence of nucleolus-associated domains (NADs) containing repetitive sequences i.e. centromeres, has been also documented (Németh et al. 2010). In general, the nucleolus is an organelle present in most mammalian cells with the major function related to the synthesis of ribosomal RNA and biogenesis of many ribonucleoprotein particles (RNPs).

The nucleolus most likely plays additional role in the establishment of kinetochores and heterochromatin structure as numerous studies have shown accumulation of many centromere-related proteins within its structure i.e. CENP-C or INCENP (Gent and Dawe 2012). Moreover, the alpha-satellite RNA seems to be required for the assembly of centromere-associated nucleoprotein components at the nucleoli and kinetochores in human cells (Wong et al. 2007). The presence of repressive chromatin marks on the NADs together with tethering experiments which showed that proximity to nucleolus reduces gene expression, support the notion that the nucleolus is an important nucleolar compartment involved in silencing and heterochromatin maintenance (Matheson and Kaufman 2015).

Importantly, nuclear architecture is orchestrated by chromatin modifiers similarly to other epigenetic mechanisms as they are all functionally related. Because regions that are in close proximity to the nuclear lamina are irreversibly bound to its components, change in chromatin state or lamina composition often leads to nuclear re-organization. Lamin proteins, but also lamin-associated transmembrane proteins like Emerin or LBR, are the main structural components of lamina that play a role in anchoring of the chromatin. LBR, for example, has been shown to bind HP1 through H3K9me2/3 (Ye and Worman 1996) in HeLa cells. Experiments performed in *C.elegans* where the localization of repetitive gene arrays which recapitulate endogenous heterochromatin behavior were monitored under genome-wide siRNA screen, demonstrate that the peripheral localization of arrays depends also on H3K9 methylation (Towbin et al. 2012). Moreover, the depletion of H3K9 methylation in Prdm3 and Prdm16 deficient mouse embryonic fibroblasts results in disintegration of pericentric heterochromatin and the destabilization of the nuclear lamina (Pinheiro et al. 2012). Another example comes from TKO cells where in the absence of DNA methylation the recruitment of pericentromeric heterochromatin to the nuclear lamina is impaired, as shown by the loss of binding of LBR (Lamin B Receptor), Lamin B1 and B2 to major satellites (Saksouk et al. 2014). The loss of interaction with lamina proteins is sufficient to reorganize the position of major satellites into large aggregates surrounding the nucleoli, which is strikingly similar to the phenotype observed in the LBR/Lamin A/C double-deficient cells (Solovei et al. 2013). Interestingly, the latter work shows that the LBR is the crucial tethering site in early stages of mouse cell differentiation, whereas lamin A/C replaces or supplements its role in terminally

differentiated cells. Mechanisms anchoring repetitive gene arrays to the lamina in *C.elegans* embryos also display a switch in later developmental stages (Towbin et al. 2012).

1.4. Concluding remarks

The above-mentioned studies imply the multilayered regulation of heterochromatin, with various epigenetic pathways shown to play a crucial role. Although heterochromatin formation and maintenance in different organisms is a common theme, the molecular details are difficult to pinpoint because of the complexity and redundancy of the pathways ensuring heterochromatin robustness. Additionally, there is a difficulty in unravelling the biochemical characteristics *in vivo,* and difficulty in combining *in vitro* results with *in vivo* events. Moreover, most of the data derives from models in which heterochromatin is already established, like in somatic cells which already reached their final epigenetic status, or ES cells that, although undifferentiated, have already acquire many of the hallmarks of heterochromatin. As a consequence, it is difficult to elucidate what is the order of events, and which factors are necessary for the formation versus the maintenance of heterochromatin. However, during early development, the epigenetic program has to be erased and reacquired in order to form a new organism. This makes it an ideal time-frame to look into the mechanisms responsible for *de novo* formation of heterochromatin and molecular pathways underlying its establishment.

2. Characteristics of heterochromatin during mouse development

2.1. Mouse development at a glance

The life cycle of an organism is an extremely complex process, which starts with two distinct identities – the oocyte and the sperm - that come together to form a zygote. Within the subsequent 2 days, the zygote divides twice without changing its size and reaches the 8-cell stage with 8 blastomeres of equal volume. Then, individual cells lose their distinctive outlines and maximize intercellular contacts in a process called compaction. After one more day, at the early 32-cell stage, fluid begins to accumulate between cells and creates a cavity – the embryo reaches the blastocyst stage. At one end of the cavity lies a cluster of pluripotent cells known as the inner cell mass (ICM) surrounded by a thin layer of cells that form the polar

trophectoderm. On the opposite side, mural trophoblast cells close the blastocyst cavity (Fig. 5). The trophoblast cells or trophoectoderm (TE) will form the chorion, which is an embryonic part of the placenta, whereas the ICM will give rise to an embryo and its associated yolk sac, allantois, and amnion. These distinct cell types of blastocyst are demarcated by the expression of key tissue-distinctive transcription factors i.e. Cdx2 limited to TE or Oct4 and Nanog associated to the pluripotent ICM. By the 64-cell stage the TE and ICM are completely separate layers becoming the first cell fate specification event in development. The first segregation of ICM cells that happens afterwards gives rise to hypoblast (or primitive endoderm) and the epiblast, thought to contain all the cells that will generate the actual embryo.

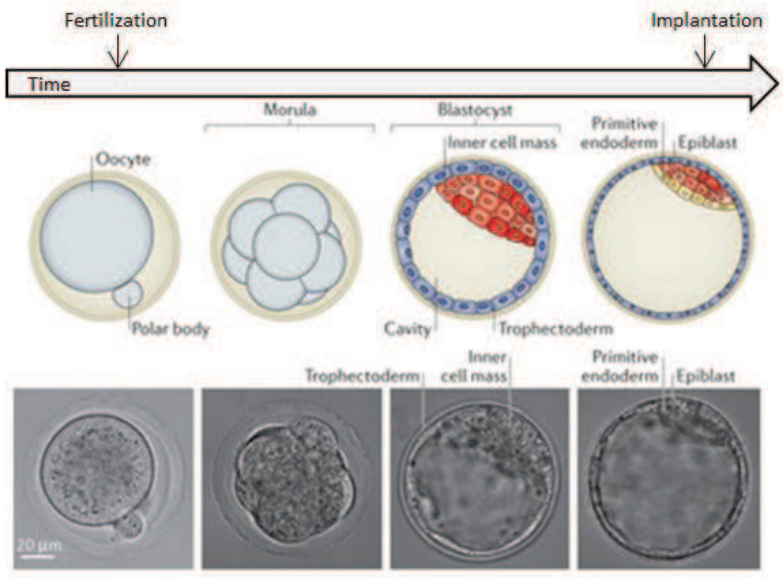


Figure 5. Preimplantation development in mouse: from fertilization to implantation

Scheme depicting different stages of mouse preimplantation development with corresponding bright-field images.(Modified from Wennekamp et al. 2013).

2.2. Formation of gametes

2.2.1. Primordial germ cells specification

Germ cells are highly specialized cells that reached the final step of differentiation after many cell divisions and transcriptional, epigenetic and morphological changes. In the same time, they are the only cell types that, *in vivo*, have the capacity to reset their identity through reprogramming. The long process during which they are formed starts already at the early

gastrulation stage (Fig. 6). Primordial germ cells (PGCs) are first specified in a response to bone morphogenetic protein 4 (BMP4) from the extra-embryonic ectoderm, which induces signaling pathway leading to expression of BLIMP1 (also known as PRDM1)(Vincent et al. 2005) and PRDM14 (Yamaji et al. 2008). As a consequence, a group of ~40 cells with high levels of alkaline phosphatase activity cluster and begin to shut down their somatic transcriptional program. At ~E7.5 PGCs start their migration from the posterior end of the embryo through the hindgut to the developing gonads. Once in gonads, PGCs continue dividing mitotically to accumulate their number before facing the last division – meiosis – in which they reduce their DNA content to 1n.

During migration and mitotic divisions PGCs undergo many changes in order to silence their somatic program (Fig. 6). In early PGCs, for example, levels of DNA methylation (Popp et al. 2010) and H3K9me2 (Seki et al. 2007) decrease, whereas H3K4me2, H3K4me3 and H3K9 acetylation (H3K9ac) together with H3K27me3 accumulate (Seki et al. 2007). Also transcription factors that are characteristics of pluripotent state, such as Sox2, OCT4, Nanog, Stella, start to be expressed. In later PGCs ~11 days after fertilization, the second wave of DNA methylation starts in which erasure of imprints begins (Popp et al. 2010)and further reprogramming takes place: linker histone H1 is removed leading to more loose chromatin configuration and enlargement of nuclei; chromocenters are lost; many repressive histone modifications are removed i.e. H2A/H4 R3 methylation (Ancelin et al. 2006), or heterochromatin specific H3K9me3 and H3K27me3; proteins associated with facultative heterochromatin or constitutive heterochromatin like heterochromatin protein 1α (HP1α), HP1β and HP1γ, ATRX and CBX2 are redistributed and/or disappear (Hemberger, Dean, and Reik 2009). In short words, the epigenome undergoes profound changes. Worth mentioning is the fact that some of these changes are transient, and thought to serve a purpose on facilitating global DNA demethylation. Indeed, soon after reprogramming happens at E12.5, many chromatin marks are reversed i.e. heterochromatic H3K9me3 and H3K27me3 levels go up together with re-clustering of chromocenters (Hajkova et al. 2008).

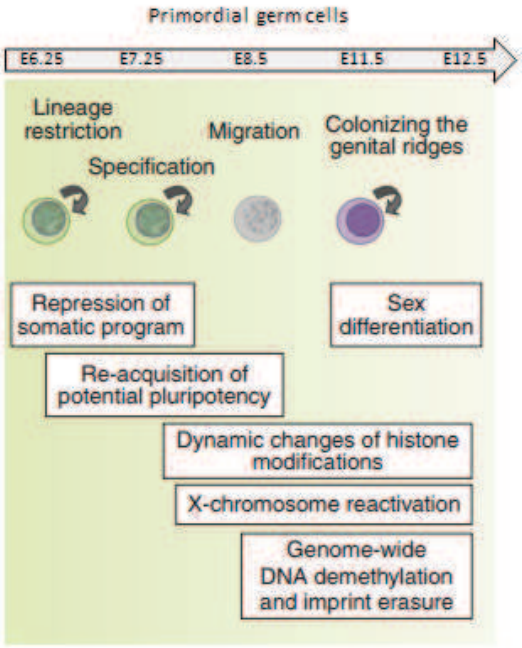


Figure 6. A schematic representation of primordial germ cell formation in mice

A brief outline of germ cell development in mice is shown schematically. Key events associated with each stage of germ cell development are also shown. (Modified from Saitou, Kagiwada & Kurimotu, 2012).

2.2.2. Formation of sex specific gametes

Male gametes

Following epigenetic reprogramming, remarkable differences are observed in the way that male and female gametes are formed after the sex determination has started at ~E12.5 (Fig. 7 & Fig. 8). In males, PGCs enter into mitotic arrest upon entry into the genital ridges, and stay in the G0/G1 phase of the cell cycle for the remaining embryonic period (Western et al. 2008). Only around day 5 postpartum (P5), spermatogenesis begins and many of PGCs resume active proliferation while others are recruited as spermatogonial stem cells (SSCs). Cells that resume mitotic divisions proliferate rapidly until they become primary spermatocyte – product of last mitotic division. They subsequently undergo meiosis I to produce two haploid secondary spermatocytes, which will later divide once more into haploid spermatids and start spermiogenesis in order to become fully mature spermatozoa (Rathke et al. 2014) (Fig. 7).

One of the most prominent changes during this time can be observed on the level of *de novo* DNA methylation. Re-methylation initiates at E14.5 in prospermatogonia and, during G0/G1 quiescence, male PGCs attain global DNA methylation. As a consequence, at birth, male specific methylation pattern is fully established and does not change throughout concomitant mitotic divisions but only when cells enter meiosis (Henckel et al. 2009). The importance of this *de novo* methylation pattern in male gametes has been shown by many studies, i.e. knock-outs of DNA methyltransferases (Bourc'his and Bestor 2004) or disruption of piRNA pathway by knock-outs of PIWI proteins (Aravin et al. 2007), with both suggesting its role in imprinting and meiosis (Sasaki and Matsui 2008). Other factors that seem to be crucial for proper meiosis and sperm formation include the acquisition of some histone modifications (e.g. H3K9me2/3) and an unusually high number of histone variants (e.g. TH2A, TH2B, TH3, H3.3A, H3.3B and HT1) which are incorporated in the nucleosomes of spermatogonia and/or spermatocytes (Bošković and Torres-Padilla 2013). Interestingly, the most dynamic process during specification of male gametes, is the global change in the chromatin structure through histone exchange. This transition starts in post-meiotic spermiogenesis and takes place during elongation and condensation of spermatids. At that phase of maturation, DNA is stripped of most of its nucleosomal packaging and becomes wrapped around so-called transition proteins (TPs) which are thought to facilitate chromatin remodeling (Bošković and Torres-Padilla 2013; Rathke et al. 2014). Moreover, the hyperacetylation of histone H4 takes place in elongating spermatids, just prior to histone removal (Grimes and Henderson 1984; G W van der Heijden et al. 2006), which, most likely, also induce histone displacement by creating more open chromatin structure. Interestingly, transcription does not occur during the elongation process of spermatid maturation, suggesting that core histone acetylation in these cells is not related to gene expression but indeed functions as a signal for the eviction of histones (Bošković and Torres-Padilla 2013). In late spermiogenesis TPs are exchanged for protamines - small, highly basic proteins that bind DNA with high affinity and wrap it in a toroidal structure. This transition is essential for the formation of proper spermatozoa as defects in the process can lead to infertility. The incorporation of protamines into sperm chromatin induces global DNA compaction which is believed to provide safe environment for the genome. Moreover, as a consequence of histones-to-protamines exchange, most of the heterochromatin associated histone modifications (i.e. H3K9 or H4K20) together with HP1, are not present in the fully grown spermatozoa (G W van der Heijden et al. 2006). This leads to extremely rare

phenomena where in *in vivo* system heterochromatin associated marks are not present, thus have to be acquired *de novo*. However, an interesting twist is that mammalian sperm chromatin retains some of the spermatid histones, i.e. H4K8ac or H4K12ac, and some of the nucleosomal proteins, most likely at the constitutive heterochromatin (G W van der Heijden et al. 2006). Because nuclear organization of the sperm genome is well defined, with the telomeres positioned at the outer membrane and centromeric heterochromatin in the center of the nucleus forming the chromocenter (Haaf and Ward 1995), the colocalization of the remaining modified histones with chromocenter suggests that they may play a role in a subsequent reorganization of the male genome.

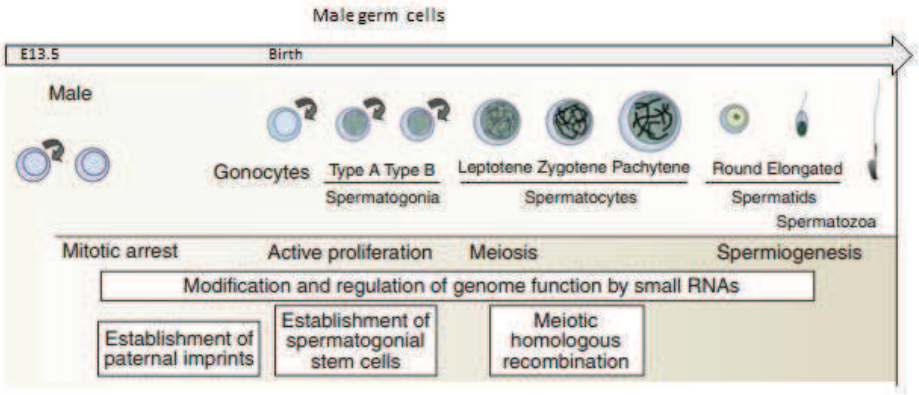


Figure 7. A schematic representation of male germ cell formation in mice

A brief outline of germ cell development in mice is shown schematically. Key events associated with each stage of germ cell development are also shown. (Modified from Saitou, Kagiwada & Kurimotu, 2012).

Female gametes

In females, contrary to males, PGCs continue to proliferate until E13.5. Then, they enter into the prophase I of meiotic division which arrests at the diplotene stage, until female reaches puberty (Speed 1982). This prolonged arrest is maintained by the signaling from somatic follicular cells which surround oocytes creating primordial follicle (Fig. 8). During the storage, oocytes grow and by selectively regulating gene expression, accumulate organelles and macromolecules indispensable for their further fate. In adult females, signal from follicle stimulating hormone (FSH) recruits groups of primary follicles that start to grow and, in consequence, become antral follicles containing fully grown oocytes. Then, upon hormonal

stimulation from luteinizing hormone (LH), groups of fully grown oocytes called germinal vesicles (GV) undergo germinal vesicle break down (GVBD) and complete first meiotic division with concomitant extrusion of the first polar body. The GV oocytes undergo an additional transition in their chromatin organization. Initially, fully grown oocytes contain a 'Non Surrounded Nucleolus' (NSN) which is transcriptionally active and contains less condensed chromatin displaying heterochromatic DAPI-rich regions in several clusters in the nucleoplasm, similarly to somatic cells. At later stages of oogenesis, the so called Surrounded Nucleolus (SN) oocytes form, which are transcriptionally inactive and contain more condensed chromatin of which the majority is wrapped around the nucleolus forming a characteristic ''rim''. Interestingly, only SN oocytes are able to resume maturation whereas others remain unresponsive to LH stimulation (Debey et al. 1993). As a consequence, only SN oocytes undergo the maturation process involving haploidization which is completed upon arrest at metaphase II of meiosis. Termination of meiosis, however, awaits fertilization. Only after fertilization with a haploid spermatozoon, oocytes can complete the second meiotic division, extrudes the second polar body and give rise to a totipotent zygote.

As mentioned before, acquisition of such a potency is possible due to dramatic changes and full reset of somatic program while PGCs are formed, but also afterwards. As a continuation of the reprograming, in female germ cells, DNA methylation levels remain low (at E16.5) which leads to i.e. further erasure of genomic imprinting and reactivation of the inactive X chromosome. The initiation of DNA methylation, leading to female specific imprints, begins in the growing oocytes after birth, and the *de novo* methylation process is complete by the time oocytes are supposed to resume meiosis. Hence, in the fully grown oocyte heterochromatic DNA is hypomethylated, however, some of the specific regions, like major satellites or LINE-1 transposones, remain hypomethylated (Arand et al. 2015). Changes in other epigenetic marks and associated proteins also take place but are not as profound as those observed in the male gametes. For example, some histone modifications change rapidly - histones H3 and H4 are generally acetylated at prophase I of meiosis, but they rapidly become deacetylated at metaphase I by HDACs. On the other hand, levels of other histone modifications stay stable like in case of H3K9me2 which remains dispersed throughout the nucleus, or H3K9me3 that appears concentrated at condensed heterochromatic regions (Meglicki, Zientarski, and Borsuk 2008). Also some heterochromatin associated proteins

display diverge patterns of expression i.e. HP1β protein is present in the primordial oocytes and remains bound to heterochromatin regions of fully-grown oocytes disassociating soon after GVBD. In contrast, HP1α can be detected for the first time in the oocytes at the beginning of their growth phase and dissociates from the chromatin of fully grown oocytes during their transition from NSN to SN type (Meglicki, Zientarski, and Borsuk 2008). Moreover, during meiotic prophase I somatic histone H1 is replaced by an oocyte-specific variant H1FOO, which in consequence leads to a less condensed state of chromatin (Hayakawa et al. 2014). Although H1FOO is indispensable for GV oocyte and early embryogenesis (Furuya et al. 2007), soon after the second cleavage canonical H1 is reacquired.

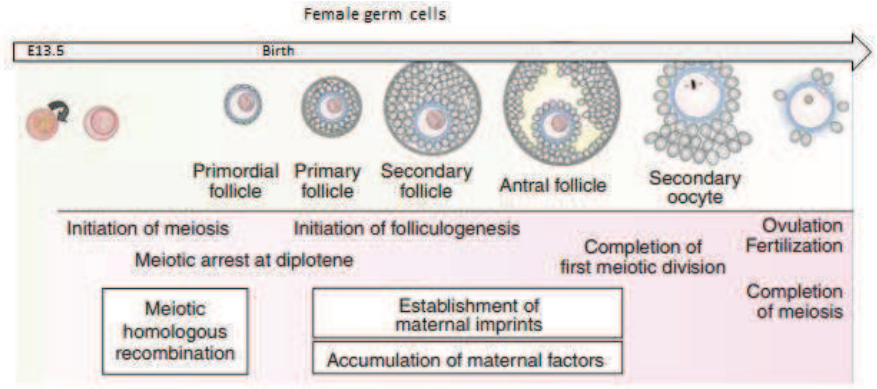


Figure 8. A schematic representation of female germ cell formation in mice

A brief outline of germ cell development in mice is shown schematically. Key events associated with each stage of germ cell development are also shown. (Modified from Saitou, Kagiwada & Kurimotu, 2012).

In conclusion, an oocyte arrested at metaphase II and awaiting fertilization has very distinct epigenetic landscape from a somatic cell with some of the heterochromatic features being already erased (e.g. HP1α), and others still present on the chromatin (such as H4K20me3 and H3K9me3). Also gene expression pattern differs drastically because oocyte is not transcriptionally active during the last phase of maturation, and it relies only on the factors accumulated during its growth and maturation. As mentioned earlier, the sperm undergoes even more drastic changes with its histones-to-protamines transition leading to extreme compaction of DNA. Why all these changes are important, and how exactly they are regulated has been under deep investigation for many years and still leaves many unanswered

questions. What can be concluded, however, is the crucial role of epigenetic reprogramming and global chromatin reorganization in both gametes for the formation of the "first cell of an organism" – the zygote. This cell has to establish *de novo* its entire developmental program, which includes heterochromatinization of some regions and activation of others. How does the newly formed embryo distinguish these specific parts of a genome, acquires the correct histone modifications and reorganizes nuclear architecture, has become one of the main questions in cell biology. In the next chapter, I will provide a more detailed description of our current knowledge of these processes.

2.3. Preimplantation development

2.3.1. Most important events in preimplantation embryos

Development in mammals commences with the fertilization of an oocyte by a sperm, which results in the formation of a zygote that after many divisions give rise to embryonic and extra-embronic tissues (Fig. 5). This capacity is defined as a totipotency, in opposition to pluripotency that leads to the formation of embryonic tissues only i.e. stem cells, or multipotency which enables some differentiation but only according to already taken pathway, for example in neurons. Until now, totipotency has been observed only in 1-cell stage embryos and in each blastomere of 2-cell stage embryos, which after separation can also form a mouse. How such a potency is retained in a zygote and what are the mechanisms of its regulation, remain unknown, however, some of the molecular pathways and events/chromatin changes happening during preimplantation development seem to be indispensable for this process.

Epigenetic reprogramming of parental genomes is a crucial event which results in the creation of a totipotent cell from two differentiated ones. Not surprisingly, if one keeps in mind their morphology at the fertilization time point, maternal and paternal sets of chromatin behave in a distinct way during reprogramming process, showing different chromatin signatures, histone marks, and replication and transcription timings. One of the main reasons lays in the fact that in sperm protamines have to be exchanged for histones which results in a rapid decondensation of the paternal chromatin, whereas female chromatin does not have to undergo such an extensive remodeling. Moreover, the maternal and paternal pronuclei (PN) remain separate nuclear entities throughout the first cell cycle. Based on the

morphological characteristics of the PN, zygotes can be classified into sub-stages from PN0 to PN5 (Fig. 9). Both male and female PN, however, contain nonfunctional nucleoli-like bodies (NLBs) which become mature nucleoli only at the 4/8-cell stage. The function of the NLBs is not known but they may play a role in the organization of pericentric heterochromatin as in the zygote and early 2-cell stage embryos the pericentric repeats form rims surrounding NLBs, and only at the late 2-cell stage they begin to cluster into chromocenter-like structures (A. V. Probst and Almouzni 2008).

Importantly, the differences between male and female chromatin organization should be resolved quite quickly to ensure correct chromosome segregation in the first cleavage, and subsequent embryo formation. Thus, strong asymmetries present in zygote are still maintained at some degree through the second division but are almost undetectable at later stages. Moreover, embryonic genome activation (EGA) has to take place as the newly form zygote can rely on the maternally provided factors only for a limited amount of time. The major wave of transcription activation occurs at the 2-cell stage, but there is also a minor EGA that starts already at the zygote stage (Schultz 2002). All of the above suggests not only how rapidly the changes in early embryos occur but also how important they are for the proper control of gene(ome) expression and global chromatin organization. In the subsequent chapters the most important and dramatic events of reprogramming are discussed, in particular in the context of heterochromatin formation.

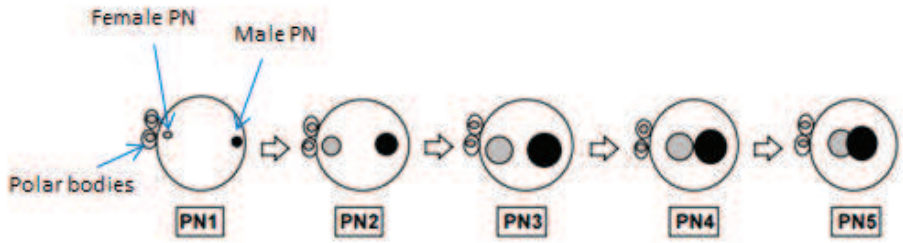


Figure 9. Schematic representation of pronuclear stages in the mouse embryo

Male and female chromatin are represented by dark and grey motifs, respectively. PN1, small pronuclei located at the periphery of the embryo. PN2, pronuclei increased in size and began migration towards the center of the embryo. PN3, large pronuclei migrated towards the center. PN4, large pronuclei were close to each other in the center of the embryo. PN5, large central pronuclei were apposed. (Modified from P. Adenot et al., 1997).

2.3.2. DNA methylation

Upon fusion, gametes contain high level of DNA methylation which is gradually removed after fertilization, reaching the lowest level in blastocyst stage (Messerschmidt, Knowles, and Solter 2014). Interestingly, demethylation of the sperm and oocyte DNA follows different kinetics as it occurs more rapidly in male than in female DNA (Fig. 10). For many years it was thought that the maternal genome is progressively demethylated by dilution of 5mC with each DNA replication, whereas the paternal genome undergoes rapid active DNA demethylation via Tet3-mediated oxidation of 5-methyl-cytosine (5mC) to 5-hydroxymethyl-cytosine (5hmC) (Wossidlo et al. 2011). Recent discoveries have shown, however, that active and passive demethylation occur in both pronuclei (Wang et al. 2014) suggesting the existence of multiple pathways that take part in this process. Moreover, although global, genome-wide demethylation is observed after fertilization, its disruption, for example in the absence of Tet3, produces only a subset of embryos displaying arrest (Gu et al. 2011), leading to confusing interpretations of the role of demethylation. Moreover, worth mentioning is the fact that erasure of DNA methylation is not global, like in PGCs, and is preserved in imprinting control regions (ICRs), enabling parent-of-origin-specific gene expression in tissues at later developmental time-points. Differentially methylated ICRs are protected against demethylation by numerous factors i.e. Stella that binds to chromatin containing H3K9me2 (Nakamura et al. 2007), Dnmt1 which maintains methylation (Hirasawa et al. 2008), or Znf57 which acts through KRAB box-mediated interaction with KAP-1 (X. Li et al. 2008). In the case of Stella, the developmental phenotype is also accompanied by impaired replication (Nakatani et al. 2015). Thus, it seem that histone exchange through replication itself and/or additional epigenetic changes also play a role in this process, although the mechanistic nature of their role is less known.

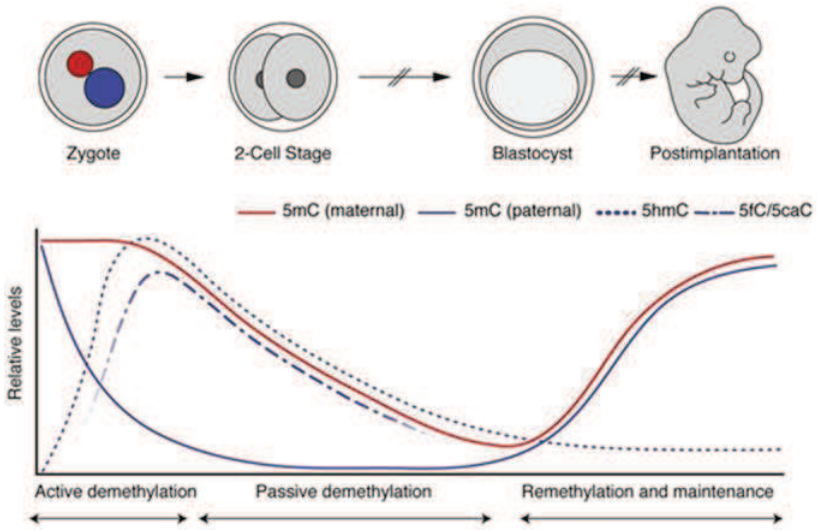


Figure 10. DNA methylation dynamics in preimplantation embryos

Distinct characteristics of maternal and paternal genomes impose an epigenetic asymmetry in the zygote. The maternal genome (red pronucleus; red line) undergoes passive DNA demethylation throughout several rounds of DNA replication. The paternal genome (blue pronucleus; blue lines) undergoes active demethylation before DNA replication in the zygote ensues. Concomitant with global loss of paternal 5mC, 5hmC (blue dotted line) and the further oxidation derivatives (5fC and 5caC; blue dashed line) are enriched. (Modified from Messerschmidt, Knowles, and Solter 2014).

2.3.3. Histone variants

H3 variants

Shortly after gamete fusion, protamines are rapidly exchanged by maternally provided histones that are assembled into nucleosomes. The variant H3.3, initially described as a replacement variant to transcriptionally active loci, is the first and main type of H3 deposited in the paternal genome most likely because its incorporation does not depend on DNA replication (Torres-Padilla et al. 2006). Immunostaining of H3.1/H3.2, on the other hand, have revealed that these variants are initially enriched exclusively in the maternal pronucleus but not in the paternal (Godfried W. van der Heijden et al. 2005). Time-lapse analyses, however, have shown that H3.1 is incorporated into both maternal and paternal pronucleus at around the same time, but a few hours later than H3.3 (Santenard et al. 2010). Thus, H3.3 deposition, which depends on the Hira chaperon, seems to be a crucial event in the formation of the zygote. Although it has been shown that incorporation of H3.3 is not involved in protamine removal, this histone variant is indispensable for pronucleus formation as nucleosomes do not assemble at the paternal genome in maternal Hira mutant zygotes. Moreover, Hira-mediated H3.3 is essential for DNA replication and rDNA transcription in both parental genomes (Lin et

al. 2015) and plays a role in the formation of a functional nuclear envelope through ELYS-mediated assembly of nuclear pore complex in male pronucleus (Inoue and Zhang 2014). Additionally, it has been shown recently that H3.3 replacement facilitates epigenetic reprogramming upon somatic cell nuclear transfer (SCNT) (Wen et al. 2014). Hence, H3.3 is a crucial factor during reprogramming indispensable for chromatin organization and required for reactivation of key pluripotency genes.

H2A variants

Deposition of the core histone H2A also occurs after fertilization together with incorporation of its variants, which seem to be important for early development. For example, absence of histone variant H2A.Z leads to developmental arrest at around implantation stage. H2A.Z displays distinct pattern of incorporation with lowest levels in both pronuclei of the zygote and in 2-cell stage embryos, and high abundancy from the 4-cell stage onwards (Bošković et al. 2012). Although its presence on the constitutive heterochromatin has been previously described together with the putative function it may play in later stages of development in the establishment of facultative heterochromatin (Rangasamy et al. 2003; Fan et al. 2004), in early embryos H2A.Z remains excluded from the DAPI rich regions and becomes targeted to pericentromeric heterochromatin only upon differentiation of the blastocyst (Rangasamy et al. 2003). Interestingly, H2A.Z may be acetylated, which contributes to the overall destabilization of the nucleosome and presumably facilitates transcription (Ishibashi et al. 2009). Its absence in the zygote and 2-cell stage embryos, when EGA occurs, is surprising and suggests that regulation of embryonic transcription might be orchestrated by different chromatin signatures to that in somatic cells. Another H2A variant - H2A.X – has been shown to be involved in the formation of male gametes, as male mice lacking H2A.X are not fertile (Celeste et al. 2002). The physiological role of H2A.X is related to its unique COOH-terminal tail that contains a motif which can be phosphorylated. Phosphorylation of H2AX (γH2AX) is induced by Double-Strand Breaks (DSBs) and recruits the repair factors to damaged DNA, thus, γH2AX is normally present at the sites of recombination (Redon et al. 2002). Interestingly, γH2AX is abundant throughout pre-implantation development starting from the zygote stage where it is enriched in the early paternal pronucleus plausibly in the Tet-dependent manner (Wossidlo et al. 2010). The maternal pronucleus seems to be protected from its presence by the activity of Stella and higher levels of DNA methylation (Nakatani et al. 2015), which is lost

by the 4-cell stage. Nevertheless, the presence of γH2AX is probably independent of DNA damage because another marker of DSB - tumour suppressor p53-binding protein 1 (TP53BP1) - does not colocalize with these foci. Most likely, γH2AX reflects the fact that embryonic nuclei constitute newly assembled chromatin that needs to achieve a proper nucleosomal configuration (Ziegler-Birling et al. 2009). Whether γH2AX plays an important role in this process remains formally unknow.

2.3.4. Pathways responsible for establishment of repressive and active histone marks

The two chromosome sets at the beginning of fertilization are strikingly different, not only at the level of incorporated histone variants and DNA methylation state but also at the accumulation of histone marks. Paternal genome is associated with hyperacetylation already during decondensation, whereas maternal genome has more repressed pattern of histone modifications. Moreover, newly incorporated histones are hyperacetylated and hypomethylated. This implies that the resulting paternal genome is devoid of heterochromatic marks which, in consequence, have to be acquired *de novo*.

H3K9me, H4K20me and HP1 proteins

One of the hallmarks of constitutive heterochromatin - H3K9me2/3 methylation – is present only at the maternal constitutive heterochromatin in early zygotes most likely because of the inheritance from the oocyte, whereas paternal PN is devoid of H3K9 methylation in the early zygote. Very low levels of H3K9me2 start to be detectable at the late zygotes at around PN5, which remain through the later stages. H3K9me3 asymmetry, however, is still detectable at 2-cell stage but it disappears (Santenard et al. 2010; Santos et al. 2005; Godfried W. van der Heijden et al. 2005; Puschendorf et al. 2008) once maternal chromatin loose the signal (Fig. 11A). The establishment of H3K9me3 relays on the activity of Suv39h1/h2 enzymes (Peters et al. 2001) which are not present or detectable at very low levels in these early stages (A. Burton et al. 2013; Puschendorf et al. 2008), hence, the plausible reason why this heterochromatic mark is gone. Another H3K9 methylatransferases - SETDB1 - is expressed in zygotes and might provide for the lack of Suv39h1/h2 in early development, however, these two enzymes have been shown to operate on distinct regions (Dodge et al. 2004).

H4K20me3, which in somatic cells is placed in the same silencing pathway as the Suv39h1/h2 enzymes, is inherited from the oocyte and stays present on the maternal chromatin in zygotes, mostly enriched in the pericentric regions. After the first division it

becomes undetectable until the blastocyst stage when it comes back (Wongtawan et al. 2011). H4K20me1 follow different pattern of localization because it seems to stay stably present on chromatin throughout early development, whereas H4K20me2 starts to be detectable only from the 4-cell stage (Wongtawan et al. 2011) (Fig. 11B). Interestingly, HP1α which is involved in the recruitment of Suv4-20h1/h2 enzymes to heterochromatin (Schotta et al. 2004), is not present during early stages of development (Godfried W. van der Heijden et al. 2005). HP1β, on the other hand, is loaded on the maternal chromatin already after fertilization, with a stronger signal coming from constitutive heterochromatin regions. HP1β is not present at the early male PN and becomes detectable only later, around PN (Santos et al. 2005; Arney et al. 2002). In 2-cell stage embryos it stains the entire nucleus with the enrichments on the pericentromeric regions (Santenard et al. 2010). Interestingly, HP1β acquisition on constitutive heterochromatin in the female PN has been related to the presence of H3K9me3 as in Suv39h2$^{m-/z-}$ knock-out embryos this localization is lost (Puschendorf et al. 2008). Moreover, its recruitment is also disrupted after the RNaseA treatment of early zygotes but this time from pericentric regions of both pronuclei, which suggests a role of RNA-dependent pathway in HP1β recruitment (Santenard et al. 2010). Indeed, in H3.3K27R embryos which display higher levels of major satellite transcription and loss of HP1β from heterochromatic regions, the latter can be rescued by microinjections of ds major satellites RNA (Santenard et al. 2010). Unfortunately, the exact mechanisms of HP1β recruitment is still not know as in the above mentioned experiment it is difficult to point out causality. Moreover, Polycomb group proteins and pathways associated to them are also involved in setting up heterochromatin in the early embryo.

H3K27me3 and Polycomb

In somatic and ES cells H3K27me3 is mostly related to facultative heterochromatin, however, its abundance in the early embryo is not surprising regarding the lack of "conventional" constitutive heterochromatin mentioned earlier. Present from the early stages on the maternal PN, H3K27me3 is gradually acquired on the paternal one, to reach high levels in 2 cell stage embryos (Fig. 11 C). Di-methylation of H3K27 follows a similar pattern of the localization, whereas H3K27me1 becomes enriched in the male PN much earlier and is distributed more equally in both eu- and heterochromatin (Santenard et al. 2010; Puschendorf et al. 2008; Santos et al. 2005; van der Heijden et al. 2005). Depletion of maternal and zygotic

Ezh2 (Ezh2$^{m-/z-}$) results in embryos in which the maternal PN completely lacks H3K27me2 and H3K27me3, whereas in paternal PN both marks are absent from pericentromeric regions, which is otherwise the only accumulation that they display in wild type embryos (Puschendorf et al. 2008). Nevertheless, Ezh2$^{m-/z+}$ as well as Ezh2$^{m+/z-}$ knock-out mice still progress through preimplantation development and display only post-implantation embryonic lethality and defects in gastrulation (Erhardt et al. 2003). Interestingly, also Ezh2$^{m-/z-}$ embryos develop normally to the blastocyst stage and arrest at the similar time-point as Ezh2$^{m-/z+}$ and Ezh2$^{m+/z-}$ knock-out mice (Terranova et al. 2008). These phenotypes would suggest that H3K27me3 is not crucial at the early stages, or that there are compensatory mechanisms in place. On the other hand, experiments in which mutated H3.3K27R was expressed in zygotes show that these embryos cannot develop properly and arrest before blastocyst stage. Moreover, HP1β recruitment to pericentric regions is perturbed and higher level of transcription from these regions is observed together with impaired chromocenter formation (Santenard et al. 2010). This implies that H3.3K27 play an important role in heterochromatin formation and its presence is necessary for proper development. Recently, the BEND3 protein has been shown to be involved in the PcG/H3K27me3 pathway by interacting with MBD3/NuRD and recruiting PRC2 components to pericentric heterochromatin (Saksouk et al. 2014), which may act as a rescue in Ezh2$^{m-/z-}$ embryos. Interestingly, in early development, loading of the cannonical PRC1 components to pericentromers – RING1B, Cbx2, Bm1 or Phc - is independent of Ezh2 and H3K27me3 because in Ezh2$^{m-/z-}$ embryos all of the above mentioned proteins are present on the male PN. Thus, a model has been proposed in which PRC1 is responsible for the silencing and heterochromatin formation on the male PN. Indeed, upon maternal depletion of the enzymatic unit of the PRC1 complex - RING1B - major satellites are up-regulated from paternal PN which supports above hypothesis. PRC1 recruitment relies probably on the AT-hook domain of Cbx2 which can bind to major satellites in H3K27me3-dependent and -independent manner, and subsequently leads to transcriptional silencing (Tardat et al. 2015). In female PN, on the other hand, Suv3-9h1/h2 depended H3K9me3/HP1β establishment in the oocyte has been suggested to be important for the maintenance of heterochromatin on the pericentromers and their subsequent silencing. Moreover, together with DNA methylation, they probably protect female PN from the PRC1 machinery (Puschendorf et al. 2008; Saksouk et al. 2014) as in Suv39h2$^{m-/z+}$ zygotes PRC1 complex starts to accumulate on

the maternal genome. This may explain why no increase in major satellites has been documented in these embryos and why they do not seem to display developmental arrest.

Figure 11. Summary of the histone modifications occurring in germinal vesicle oocyte (GV) and on maternal and paternal chromatin in the zygote

Membrane of the germinal vesicle (GV stage) or the pronucleus (PN stage) and the nucleolar-like bodies (NLBs) are shown as just one inner circle for simplicity. Histone modifications associated with active transcription are colored in green, repressive in red and for those with no clear distinct correlation in blue. The intensity of shading represents the relative intensity of the labelling throughout stages. Those marks that are reported to localize to the perinucleolar (NLB) ring(s) are highlighted on the inner circle in the appropriate color. (Modified from Burton and Torres-Padilla 2010).

Active marks

Although acquisition of repressive marks is extremely important for proper heterochromatin formation and development, active marks are also engaged in the process of reprogramming and may facilitate nucleus reorganization. As mentioned above, the paternal genome is hyperacetylated at the lysines of histone 4 (e.g. H4K8ac or H4K12ac) already at the pre-PN formation. The oocyte, on the other hand, does not contain high levels of H4Kac, which only becomes detectable during early PN formation. H3K9 acetylation, on the other hand, is catalyzed during early PN formation to both sets of chromatin, and remain present and abundant in all stages of mouse preimplantation development (G W van der Heijden et al. 2006; Bošković et al. 2012). The general assumption is that hyperacetylated state of chromatin in early development is linked to more open chromatin necessary for genomes reorganization and subsequent EGA.

H3K4me3 – another hallmark of active chromatin is inherited by the female PN from the oocyte, similarly to H3K4 mono- and di- methylation, and is present in the entire female pronucleus (Lepikhov and Walter 2004). However, there is a dichotomy since the male pronucleus does not have H3K4 methylation marks at the early stages of pronuclear formation and acquire them later (Bošković and Torres-Padilla 2013). As described earlier, H3.3 enrichment on the paternal genome is perhaps overriding the low levels of H3K4me3 and play a role in activation of the genome (Torres-Padilla et al. 2006). Whilst H3K4me3 by 2-cell stage become equally distributed on both sets of chromatin, H3K36me3 - mark of transcriptional elongation – does not follow a similar pattern. H3K36me3 is enriched in early female PN, is lost during PN formation and it is never detected on the male PN. At the 2 cell stage, almost complete lack of that mark is observed, which gradually comes back from the 4-cell stage onwards (Bošković et al. 2012).

2.3.5. Nuclear organization in preimplantation embryos

How the creation of functional nuclear compartments and chromatin organization in the 3D nuclear space is orchestrated during cell cycle progression and development, became one of the key questions in developmental biology. Interestingly, in oocyte to 2-cell stage transition one of the most striking examples of nuclear reorganization has been observed.

Already in oocyte rapid changes of pericentric and centric heterochromatin localization can be observed that, in some cases, can be correlated with the presence of some histone modifications. For example in NSN-type oocytes, the centromeric and pericentromeric heterochromatin is mainly organized in clusters called chromocenters - with the typical pattern of silencing posttranslational histone modifications - enriched in H4K20me3, H3K9me3 and HP1β, and devoid of H3K4me3 and H4K5ac (Bonnet-Garnier et al. 2012; Meglicki, Zientarski, and Borsuk 2008) - as described in somatic mouse cells (Bannister et al., 2001). During the transition to SN-type oocytes, the pericentromeric regions start to condense and expand to form an almost complete ring around NLBs. In parallel, some of the histone marks previously present on these regions, such as H3K9me3 and H4K20me3, seem to follow the repositioning of chromatin within the nucleus, as they can be still found on the rings in SN oocytes. Interestingly, some modifications do not follow this pattern, like H3K27me3, which is not associated with chromocenters in NSN oocytes but becomes present as spots around NLBs in SN oocytes, or H3K4me3 and H4K5ac which co-stain with euchromatic regions in NSN

oocyte but appear as a ring in SN oocytes (Bonnet-Garnier et al. 2012). Therefore, despite the global chromatin compaction that occurs at the SN stage, pericentromeric sequences seem to decondense around the NLB with both repressive and active histone modifications present on them. Moreover, during the transition from NSN to SN, oocytes seem to acquire even higher levels of global histone modifications and have higher expression of enzymes that catalyze these modifications (Kageyama et al. 2007).

A similar situation can be observed in early preimplantation development where pericentromeric repeats of both sets of chromatin are gathered around the NLBs forming ring-like structures despite of an asymmetric organization of heterochromatic marks in the maternal and paternal pronuclei described earlier. After the first division, this particular configuration is still maintained, however, some of the pericentromeric regions start to condense and cluster into structures reminding somatic-like chromocenters. This dramatic reorganization continuous through the entire second cycle, and by the 4-cell stage no ring-like structures are present (Fig. 12) (A. V. Probst and Almouzni 2008; Almouzni and Probst; Aguirre-Lavin et al. 2012). The most striking observation is that a peak in major satellites transcription has been detected only in late zygotes and 2-cell stage embryos (A. V. Probst and Almouzni 2008; A. V. Probst et al. 2010; Puschendorf et al. 2008), hence, silencing of these regions occurs presumably at the same time as their reorganization. Because there is no strict temporal correlation between the global acquisition or removal of any histone modification analyzed and the expression of major satellites, it is possible that transcriptional silencing/activity of pericentric regions occurs independently of changes of histone modifications and that some other mechanisms play a role in that process. Other repeats also display changes in the localization through the early stages of development, however, not so rapid and dramatic. The localization of telomeres, for example, depends if they are distal from centromere part of a chromosome or not (Fig. 3). The half which is in the proximity to the centromere, is present around the NLBs and associated with pericentromeric signals, whereas the other, distal half is located in the nucleoplasm, or close to the nuclear envelope (Aguirre-Lavin et al. 2012). Tandem repetitive clusters of rDNA, on the other hand, are always associated with pericentromeric domains, either at rings surrounding the NLBs, or as a few foci at the nuclear periphery (Aguirre-Lavin et al. 2012). Nevertheless, their transcription

pattern does not change during early development and starts already at the zygotic stage, from those copies that are NLB-associated (Lin et al. 2015).

Worth noting is the emerging role of the NLB and its plausible importance for the nuclear architecture and reprogramming events during early development. For example, when NLBs are micro-surgically removed from the GV oocytes, the maturation process until methaphase II is not disrupted. However, after subsequent fertilization, pronuclei do not form properly and most embryos arrest at the 2-cell stage (S. Ogushi et al. 2008). The above results suggest that zygotes inherit NLBs from the oocyte, and that this event is essential for embryonic development. Moreover, zygotes display heterochromatin disorganization if NLBs are removed. The global pattern of histone modifications (H3K9me3, H3K27me1/2/3, H3K4me3) seems not to be affected (Sugako Ogushi and Saitou 2010), however, it is not clear whether the enrichment of these marks at major satellites becomes disrupted or not. Interestingly, developmental arrest can be rescued when NLBs are re-injected into MII oocytes, and with a much lower ratio, when re-injected into zygotes during PN formation. Hence, all this together suggests that NLBs are important during PN formation, presumably for the correct organization of pericentric heterochromatin, necessary for the subsequent divisions.

In summary, the picture of the nuclear organization in the mouse early embryo appears to differ drastically from the one observed in the somatic and stem cells. Nevertheless, whether nuclear organization plays a crucial role in reprogramming and totipotency remains an open question.

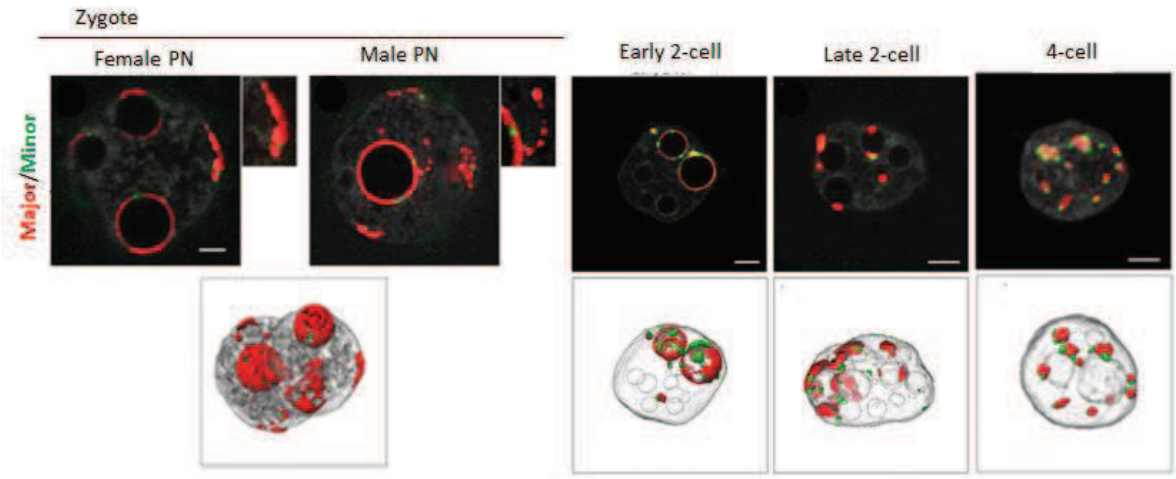


Figure 12. 3D reconstruction of the distribution of the pericentromeres and centromeres signals during early stages of mouse development

Single confocal sections of each preimplantation stage with pericentromeric (major satellite, red), centromeric (minor satellite, green), and DNA (grey) labeling are presented here, as well as the corresponding 3D reconstructions. Scale bar represents 5µm. (Modified from Aguirre-Lavin et al. 2012).

3. Heterochromatin during mouse development – specific regions and open questions about their establishment

3.1. Pericentric repeats

The mammalian centromere is strictly required for accurate chromosome segregation and its disruption may lead to chromosome missegregation and subsequently to aneuploidy and cancer. Its main role is to maintain sister chromatid cohesion until anaphase and to provide the assembly site for large protein complexes and crucial histone modifications which are required for kinetochore assembly and organization (Westhorpe and Straight 2014). Thus, it becomes a main task during the first embryonic divisions to properly establish and maintain pericentric heterochromatin while extensive genome reprogramming is taking place. How this occurs and what mechanisms are responsible for this formation, remains under extensive investigation, however, based on the previously described pathways, some hypothesis can be described.

For example, deposition of the histone variant H3.3 has been suggested as one of the players involved in pericentric chromatin silencing because it localizes to the paternal pericentric chromatin and might facilitate transcription from these regions, an event that is specific to the stages immediately following fertilization. Moreover, the mutation of H3.3K27 — but not of H3.1K27 — leads to developmental arrest (Santenard et al. 2010), which also supports a role for H3.3 in establishing heterochromatin at the pericentromere. Interestingly, H3.3 deposition can be also mediated by ATRX/Daxx chaperon as mentioned before. Whether similar mechanism takes place in embryos, is not clear, however, both factors are present in early embryos and localize to pericentric repeats where they may play a role in heterochromatin formation. Upon depletion of maternal ATRX, Daxx recruitement is impaired and up-regulation of major satellites has been observed from maternal PN. This has been linked with improper function of centromere and subsequent chromosome rearrangement and instability (R. De La Fuente, Baumann, and Viveiros 2015).

Several studies have correlated transcriptional activation of pericentromeric satellites with decondensation of heterochromatin. H3 and H4 acetylation, together with H3S10P might be involved in that process, however, most likely the deposition/reposition of repressive histone marks and associated proteins, is more crucial. In mouse embryos, pericentric regions are marked by constitutive repressive marks – H3K9me3/HP1β - at the maternal PN, and

facultative repressive machinery – H3K27me3 and Polycomb group proteins – on the paternal one. Impairing PRC1 activity leads to the up-regulation of major satellites, similarly to the mutation of H3K27 methylation site. The exact cause of that increase in expression is not clear as there are many players that may contribute to that phenotype i.e. HP1β localization seems to be lost from pericentric regions in Ring1B -/- mutants (Puschendorf et al. 2008). Worth mentioning is the fact that in the above described experiments the reappearing theme is higher transcriptional signal coming from major satellites which, in many cases, seems to be somehow correlated with the improper centromere function and chromosome instability. Unusual levels of pericentromeric transcripts are detected in zygotes but also in the 2-cell stage embryos, depending on the experimental design, and are thought to be derived either from one of the pronuclei, or from both parental sets of chromatin. This suggests the function of RNA-dependent mechanisms in the formation of heterochromatin at pericentromeres.

A burst of transcription from major satellite sequences at the 2-cell stage may, for example, induce global remodeling and association of pericentric domains into the forming chromocenters. In support of this idea, experiments in which LNA probes were injected into zygotes to 'block' major satellite transcripts leads to arrested development at the 2-cell stage, prior to chromocenter formation (A. V. Probst et al. 2010). Nevertheless, the cause of such developmental failure and the exact role of the transcript remain unknown. Whether it is the act of transcription *per se* creating more open chromatin environment or the RNA acting as a structural component recruiting other players, remain among the speculations. On one hand, the recruitment hypothesis is appealing because it is supported by the dsRNA rescue of HP1β recruitment in embryos expressing H3K27R. A similar role could be also envisaged in stabilizing the PRC1 complexes that accumulate at paternal pericentric domains in an Ezh2 and H3K27me3 independent manner. Moreover, RNA-dependent silencing and heterochromatin formation and spreading, has been already documented in cells, especially at the centromeric regions. On the other hand, RING1B and H3K9me3 do not change their localizations upon depletion of major satellite transcripts which is in contradiction with the above idea and would suggest that the transcript does not play a structural role, at least not for these two factors.

Another hypothesis that can be put forward from the studies done so far, is the role of nuclear organization and dramatic rearrangements that take place after fertilization which, as mentioned before, are concomitant with the silencing events. Decondensation and

subsequent clustering into chromocenter may facilitate acquisition of heterochromatic/silenced state and in consequence lead to establishment of somatic-like heterochromatin on the pericentromers. The plausible role of the NLBs as the main organizational compartments which enable proper heterochromatin formation seems to be very likely. First of all, they have been shown to be necessary for the proper development of the early embryo, which does not proceed upon their removal. Secondly, they have been associated with heterochromatic regions in cells and suggested to play a structural role in the kinetochore formation.

Last but not least, is the particular spatio-temporal organization of pericentric regions, first around NLBs, and only afterwards in the late 2-cell stage into clusters, which is concomitant with their silencing and might be linked with the acquisition of heterochromatin marks. Whether silencing of pericentric regions is dependent upon their nuclear localization, remains unknown, and thus became one of main interests during my Ph.D.

To sum up, the emerging picture from the above studies is that the transcriptional regulation of major satellites transcripts underlines a unique regulatory mechanism at a critical time window during development. What the exact time and casual relationship between these events and how this process is regulated, is not yet known.

3.2. L1 transposable elements

Approximately half of the mammalian genome is composed of repetitive transposable elements (TEs) which are grouped into two classes: retrotransposons and DNA transposons. TEs are another example, after pericentric repeats, of a region that in differentiated cells is normally described as heterochromatin and contains most of it hallmarks described in the previous chapters. Reactivation of these elements is correlated with mutations, deletions and genome instability given their ability to retrotranspose into new genomic regions and/or to affect functionally the neighbouring genes. Therefore, understanding the mechanisms behind the formation and maintenance of the heterochromatin on TEs has arisen as an important topic in cell biology. Extensive studies in somatic and stem cells have revealed some of the molecular mechanisms which tightly regulate and repress their transcriptional activity, including DNA methylation, histone modifications and RNA pathways. However, it still remains

unknown how L1 elements are regulated during the period of epigenetic reprogramming and cell division occurring in development. Thus, as a second model to study heterochromatin, I focused on L1 elements, a non-LTR class of retrotransposons. I have recently summarized the main molecular and evolutionary features of L1 elements, as well as the general mechanisms governing their silencing in differentiated and somatic cells in the review below (Publication I).

Chromosoma
DOI 10.1007/s00412-015-0520-2

REVIEW

LINEs in mice: features, families, and potential roles in early development

Joanna W. Jachowicz[1] · **Maria-Elena Torres-Padilla**[1]

Received: 1 December 2014 / Revised: 27 April 2015 / Accepted: 5 May 2015

Abstract Approximately half of the mammalian genome is composed of repetitive elements, including LINE-1 (L1) elements. Because of their potential ability to transpose and integrate into other regions of the genome, their activation represents a threat to genome stability. Molecular pathways have emerged to tightly regulate and repress their transcriptional activity, including DNA methylation, histone modifications, and RNA pathways. It has become evident that Line-L1 elements are evolutionary diverse and dedicated repression pathways have been recently uncovered that discriminate between evolutionary old and young elements, with RNA-directed silencing mechanisms playing a prominent role. During periods of epigenetic reprogramming in development, specific classes of repetitive elements are upregulated, presumably due to the loss of most heterochromatic marks in this process. While we have learnt a lot on the molecular mechanisms that regulate Line-L1 expression over the last years, it is still unclear whether reactivation of Line-L1 after fertilization serves a functional purpose or it is a simple side effect of reprogramming.

Keywords Transposable elements · Line-L1 · Epigenetic reprogramming

✉ Maria-Elena Torres-Padilla
metp@igbmc.fr

[1] Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM U964, Université de Strasbourg, 67404 Illkirch, France

The intrusive genome

In a genomic era when a more global overview on genomes is possible, we can finally appreciate that only a small part of mammalian DNA is composed of genes whereas most of it is composed of different types of repetitive sequences that are either organized into clusters, in blocks, or are interspersed, among which transposable elements (TEs) are the most abundant ones (Lander et al. 2001). TEs—or so called "selfish" (Orgel and Crick 1980) or "junk" DNA (Ohno 1972)—have been thought for many years to be just parasitic relics from a long evolutionary past. Over billions of years, they accumulated using self-propagation mechanisms based on the cellular machinery and on the proteins that they encode. As a consequence, in present times, TEs contribute to ~50 % of the total content of mammalian genomes (Waterston et al. 2002). Naturally, this raises the question of why nature permitted this accumulation when it is thought that evolution tends to dispose off unnecessary and harmful elements. Jeopardy coming from TEs is based on their biology, as they can modify the genome in a variety of ways, for example, through insertional mutagenesis and/or chromosomal rearrangements; they can also affect genome expression by modifying cellular transcription acting as alternative promoters or enhancers, creating or disrupting polyadenylation signals, or splicing sites (Han et al. 2004). Although it might be dangerous and can lead to various diseases, this property of the TEs is also thought to be one of the strongest drivers of genome evolution and an alternative way of controlling gene expression, suggesting that "junk DNA" may not be as selfish and useless after all. Of course, one can argue that the majority of TEs are anyway truncated—that is, they have lost part of their functional sequences due to incomplete insertions in the host genome—and, additionally, many different mechanisms exist that prevent their expression and subsequent jumping. The latter is true in most cases.

Published online: 16 May 2015

However, there are some circumstances when high levels of various TEs transcripts are present in the cell, together with the proteins that they encode. The germ line is a prime example of this (Zamudio and Bourc'his 2010). Activation of TEs is even more striking during pre-implantation development of mouse embryos, when different classes of TEs are reactivated, following different temporal kinetics and different strength in their transcriptional activity (Fadloun et al. 2013; Peaston et al. 2004). Although for most part, it is assumed that their reactivation is a side effect of the large-scale epigenetic reprogramming that the embryo undergoes after fertilization, it is possible that the transcription of TEs at this time point serves a functional purpose that is yet to be established.

Families and features

TEs in mammals are grouped into two classes: class II or DNA transposons that show no current evidence of jumping or being active and class I elements which all share common mechanisms of transposition requiring reverse transcription of an RNA intermediate (Fig. 1a). The latter ones can be further divided into two main groups: long terminal repeat (LTR) retrotransposons, which are primarily derived from endogenous retroviruses (ERVs) and non-LTR retrotransposons that do not contain terminal repeats. Long and short interspersed nuclear elements (LINEs and SINEs) are the most abundant non-LTRs, that along with ERVs comprise, respectively, 19, 8, and 10 % of genomic DNA (Waterston et al. 2002). Although, proportionally, retrotransposons represent a huge part of the genome, only a small portion of these elements can still "jump and paste" among which LINE-1 (L1) insertions seem to be the most recent (Naas et al. 1998) and active ones in the mouse genome (Akagi et al. 2008). The main reason why the majority of TEs is thought to be in quiescence is because most copies are truncated or inverted, and there is no evidence for recent transposition events. Full L1 elements, on the other hand, are still present in the mouse genome and from more than approximately 100,000 L1 fragments around 3000 contain an intact sequence (Goodier et al. 2001). Although estimated copy number of L1 elements is known, the exact number of full-length copies or fragments remains unclear, as it is very difficult to assess it experimentally and the number varies depending on a calculation method. Nevertheless, regardless of their number, full-length L1 elements are potentially capable of activation and retrotransposition. Interestingly, SINEs and even some of truncated L1 are also known to be able to invade the genome. However, they use transposition machinery of intact LINE copies and require their transcription and translation which points out the importance and additional function of full-length L1 elements.

Full-length mouse L1 elements are ~7-kb length and share common features: they have 5′ untranslated region, two open reading frames, and 3′ poly(A) tail (Fig. 1b). The 5′ untranslated region (UTR) functions as a promoter and is composed of monomers—repetitive sequences of ~200 bp organized into tandem repeats that follow single-copy, non-monomeric sequence (Goodier et al. 2001). The ORF1 codes for a nucleic acid chaperone, ORF1p, which co-purifies with L1 RNA as a ribonucleoprotein (RNP) complex (Martin and Bushman 2001) and is required for transposition, but its exact role in this process is not well understood. Likewise, ORF2 encodes for another protein, ORF2p, whose function seems to be better established, as ORF2p contains endonuclease (EN) and reverse transcriptase (RT) activities and is mainly involved in target primed reverse transcription (TPRT). TPRT is a specific propagation mechanism used by all non-LTR elements whereby an EN cleaves the putative target site on the genomic DNA creating a 3′OH, which acts as a primer for reverse transcription. Subsequently, cDNA is generated from the RNA template and its first strand is incorporated within the genome simultaneously. The process of incorporation of the second strand is still largely unknown. LTR retrotransposons also use reverse transcription and RNA for their accumulation in the cell, but the exact process whereby this occurs is more complex and, unlike non-LTR elements, requires virus-like particles.

The above-mentioned characteristics are common to all full L1 elements; however, their sequences are not identical, with some copies displaying a strong deviation from the "consensus" sequences. Furthermore, L1 elements also differ in their transcriptional and retrotransposition activity, which makes it challenging to find one clear way of dividing them into groups, either functionally or based on their sequence identity. Phylogenetic studies brought more insight into L1 diversity and their classification, revealing that these retroelements have a very unusual pattern of expansion as they seem to have evolved as a single lineage. In the mouse genome, one family of L1 is activated at a time, amplifies to thousands of copies, and becomes integrated in the host genome to subsequently start to extinct as it becomes replaced by a new lineage. As a consequence of this phenomenon, older L1 elements accumulate more mutations and are more divergent than young L1 (Adey et al. 1994). Prior to the availability of the complete mouse genome, detailed studies of L1 evolution were not possible. Therefore, the classification of newly identified L1 into families or subfamilies and the corresponding lineages has not been straightforward. Members of each family were put together based mainly on isolated criteria where comparison of their sequences would reveal some similarities, commonly represented by consensus blocks within each element or family of elements. The 5′UTR regions served perfectly for that purpose as monomers differ in their number and sequence among L1 elements, and their promoter activity is proportional to the number of monomers, as indicated by reporter assays performed by DeBerardinis and

A

TRANSPOSABLE ELEMENTS					Recent transposition events	% of genome
CLASS I Retrotransposons	LTRs i.e. ERV-IAP, ERV-L				Yes	10%
	non-LTRs	SINEs			Yes	8%
		LINEs	V type promoter		No	19%
			A type promoter	Af family ~5300 FL*	No	
			F type promoter	Gf family ~1000 FL*	Yes	
				Tf family ~4500 FL*	Yes	
CLASS II DNA transposons					No	0,9%

*FL – full length

B

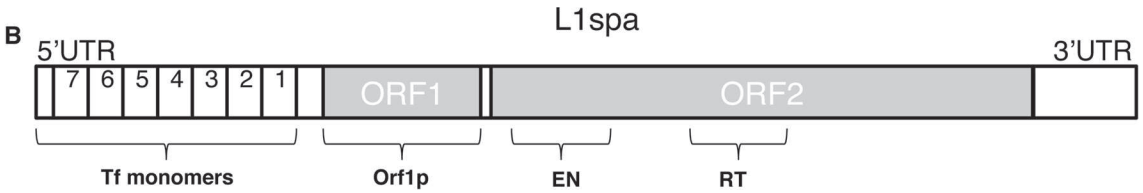


Fig. 1 Overview of the genomic content of transposable elements in the mouse genome and typical structure of Line-L1 elements. **a** Broad classification of transposable elements. The main representatives of each class are mentioned, and the number of full-length (FL) copies of the most recent Line-L1 elements is indicated. **b** Schematic of one of the most studied Line-L1 elements in the mouse, the L1spa, which belongs to the Tf family. The different domains of a representative full-length L1 element are shown. *ORF* open reading frame, *EN* endonuclease, *RT* reverse transcriptase

Kazazian (DeBeradinis and Kazazian 1999). Earlier studies revealed that there are at least three groups of monomers: A, F, and V. The more ancient V type is not very copious, and it lacks the 5′ region, whereas A- and F-type promoters are thought to be equally abundant, and both contain monomer repeats of 208 and 206 bp, respectively (Padgett et al. 1988). Further comparisons between different A, F, and V types identified more commonly conserved characteristics and led to a discovery of complete V, A, and F families of L1 elements with A and F monomers, respectively. In the mouse, V, A, and F belong to the same lineage, but the three types were inserted into the mouse genome during different time periods (Adey et al. 1994). Comprehensive studies of ~20,500 L1 inserts containing full RT sequence confirmed previous results and strengthened a more thorough phylogenetic L1 classification, revealing the existence of additional promoter types and families. Nevertheless, L1 elements that belong to the A and F types were confirmed to be the youngest and most abundant L1 elements in the murine genome (Sookdeo et al. 2013).

Retrotransposition assays have demonstrated that members of the A and F types might be still active (Goodier et al. 2001), although F types are more divergent and were therefore originally thought to be "dead" (Mears and Hutchison 2001). A systematic screening of disease-related insertions identified two recent insertions of L1 elements—$L1_{spa}$ and $L1_{Orl}$—leading to the discovery of new young families of the F lineage: T_F (Naas et al. 1998) and subsequently G_F (Goodier et al. 2001). It was suggested that $L1_{spa}$ and $L1_{Orl}$ probably emerged from an ancient F family because their monomers are similar to F-type promoters. This work also indicated that the L1 of the F type is capable of causing new insertions. Moreover, a reporter system together with retrotransposition assays showed that members of both T_F and G_F families are able to retrotranspose in cultured cells, and their promoters have the ability to drive transcription (Goodier et al. 2001; DeBeradinis and Kazazian 1999; Hardies et al. 2000). Thus, some ~1000 full-length copies of G_F and ~4500 full-length copies of T_F present in the mouse genome may be still active (Sookdeo et al. 2013). Note, however, that there are disparate reports on the exact total number of full-length L1 copies that differ largely based on the approach used to determine the number of elements.

L1 elements are abundant and, as their name indicates, can be found interspersed almost anywhere in the mouse genome.

However, some regions seem to be more prone for their presence. Interestingly, their distribution pattern across genomic regions is similar between rodents and primates, even though most insertions occurred after the divergence of both orders, suggesting that common mechanisms shape or impact new insertional events (Waterston et al. 2002; Lander et al. 2001).

In general, L1 tend to be more abundant within AT-rich and gene-poor regions of the genome (Boyle et al. 1990). This trend might be explained by a lower tolerance to keep L1 insertions in proximity to genes because of their potential deleterious influence including deletions or insertions that could affect regulatory regions of the sequence and structure of the genes themselves. Thus, it seems that these latter events have, in general, not been selected during evolution. When L1 insertions occur close to genes or even within them, the genes in question are often lowly expressed (Han et al. 2004; Muotri et al. 2005). However, L1 localization is also dependent on their age, given that younger L1 elements have been found to be located in closer proximity to genes than the old L1 ones (Medstrand et al. 2002). Moreover, analysis of the most recent insertions reveals that there are genomic regions containing genes that are more likely to be the target of retrotransposition. Whether location of these hotspots is related to the transcriptional activity of these genomic regions remains unclear, as actively transcribed genes are in an open chromatin configuration, potentially making these sites "available" for new insertions. Alternatively, the observations of recent L1 insertions situated next to genes lend support to the hypothesis of a potential role for L1 in inactivation of some genomic regions. Together, all the above indicates that location of L1 is not as random as expected and that in some regions, their presence might be useful for a host, whereas in others, natural selection eliminated them because of the potential danger that they bring.

Regulation of L1 expression

Evolutionary-driven elimination of L1 elements from some genomic regions and "tolerance" for their accumulation in others are not the only means to constrain their activity. Mammalian cells have developed many mechanisms that serve this purpose, and therefore, the control of L1 activity at different steps of their life cycle has been documented: from transcription through translation and posttranslational modifications, up to the insertion event. Possibly, the strongest and most significant mechanism of regulation relates to their chromatin structure, whereby L1 acquires a silent chromatin configuration. The silencing signatures of these regions at the chromatin level are strikingly similar and have particular features that resemble constitutive heterochromatin domains, either within euchromatin or within clusters that are embedded close to heterochromatic regions. In most cases studied, the Tf family has been used as a model for L1, and therefore, what we know on L1 regulation comes mostly from L1 Tf families. However, with the advent of genome wide studies, especially when studying chromatin features, some of the findings described below also apply to other L1 families.

The main silencing signature present on most types of TEs is DNA methylation (Meissner et al. 2008) which, historically linked to transcriptional repression, has been suggested to have evolved for a specific purpose of defending a host genome from TE activation. LTR and L1 elements contain particularly high levels of DNA methylation and hypermethylated canonical promoters, a feature conserved between mice and humans (Meissner et al. 2008; Lander et al. 2001). In mammals, this signature is established primarily by the de novo DNA methyltransferases (DNMTs) DNMT3A and DNMT3B—which display highest activity in the germ line and in early embryogenesis—and is maintained by DNMT1 in somatic tissues. For example, loss of DNA methylation in mouse embryos deficient for DNMT1 leads to activation of intracisternal A particle (IAP)—a type of endogenous retrovirus (ERV)—in somatic tissues (Walsh et al. 1998). Moreover, in mouse ES cells lacking all three DNMTs, reactivation of both IAPs and L1 has been reported (Matsui et al. 2010). IAPs and L1 are also derepressed in perinatal testes of mice lacking DNA methyltransferase-like protein (DNMT3L), a cofactor essential for DNMT3A-mediated methylation in primordial germ cells (PGCs) (Bourc'his and Bestor 2004). Thus, DNA methylation is thought to play an essential role in the silencing of retrotransposons in mammalian cells. Nevertheless, as it is an epigenetic signature, it is not always present on DNA, and there are developmental time windows during which L1 elements are not methylated or not fully methylated. The most prominent differences are observed when genome undergoes global epigenetic reprogramming. The first wave of reprogramming takes place during the formation of the PGCs when almost a complete demethylation of the DNA occurs. Likewise, the second reprogramming phase takes place after fertilization when the newly formed zygote undergoes global DNA demethylation, reaching the lowest point in the early blastocyst, and only then, gradual remethylation occurs. Interestingly, transcriptional activation of L1 and other TEs has been reported to occur at both of these time periods (Peaston et al. 2004; Fadloun et al. 2013). Moreover, retrotransposition events of L1 can occur during embryogenesis and create somatic mosaics (Kano et al. 2009). Thus, germ cells and pre-implantation embryos are two time windows in which retrotransposons have to be tightly regulated, most likely independently of DNA methylation, to guarantee protection from aberrant activation of mobile elements. Furthermore, they are a good model to study how L1 silencing is established and what are the mechanisms involved in this process.

A major safeguard mechanism present in germ cells involves small non-coding (nc) RNAs, which guide

transcriptional and posttranscriptional silencing of TEs and act in an RNAi-related pathway (Fig. 2). P element-induced wimpy testis (Piwi)-interacting RNAs (piRNAs) are germ cell-specific small ncRNA which are bound by the Piwi clade of Argonaute proteins, and—in contrast to other groups of small RNAs, i.e., small interfering RNAs (siRNAs) or microRNAs (miRNAs)—their processing is Dicer-independent. The piRNA pathway keeps TEs in check, at least in male germ cells, as the knockout of either of the two Piwi members—Mili or Miwi2—results in upregulation of L1 and other TEs in testes (Aravin et al. 2007; Carmell et al. 2007). The main function of piRNA pathway can be divided into two main tasks: (i) the establishment of de novo L1 promoter DNA methylation and acquisition of repressive histone marks and (ii) the maintenance of L1 silencing during spermatogenesis when DNA methylation is not yet present. In regard to transcriptional control of L1s, it has been proposed that de novo DNA methylation of TEs in male germ cells is guided by piRNAs (Aravin et al. 2008). Moreover, the loss of DNA methylation on L1 and the sterility phenotypes observed in Mili and Miwi2 mutants are strikingly similar to those observed in the knockout of DNMT3L in mouse male germ cells (Bourc'his and Bestor 2004) which enforces the above hypothesis. Furthermore, more recent studies have shown that piRNA and Piwi proteins are also involved in the establishment of the repressive histone modification H3K9me3 at specific L1 families in the germ line. H3K9me3 is deposited mostly on the promoter 5′UTR of full-length young L1 elements, and the piRNA pathway targets these elements for chromatin-mediated repression in a different manner than the more ancient L1 (Pezic et al. 2014). Interestingly, another repressive mark—H3K9me2—seems to act as an additional mechanism in mitotic spermatogonia until early meiosis because when a functional piRNA pathway and L1 DNA methylation are disrupted experimentally, G9a-mediated H3K9me2 is sufficient to silence L1 elements (Di Giacomo et al. 2014). The second function of piRNAs during spermatogenesis is the maintenance of L1 silencing by posttranscriptional gene silencing pathway (PTGS) that can directly cleave L1 transcripts. Disruption of piRNA biogenesis in mice leads to upregulation of L1 elements in testes and a spermatogenic failure, which implies that PTGS is important during reprogramming in the embryonic male germ line (deFazio et al. 2011). Moreover, it has been shown that during adult spermatogenesis, piRNA-mediated PTGS silences L1 in meiosis (Di Giacomo et al. 2013) and in spermiogenesis (Reuter et al. 2011) where DNA methylation is probably insufficient to maintain transcriptional silencing of L1. All of the above suggests that not only the existence of multiple epigenetic mechanisms involving mainly piRNAs but also other, additional factors are both important for sufficient silencing of L1 elements and proper function of male germ cells at different stages of development and in adulthood.

On the other hand, in female germ cells, the situation differs drastically, as loss of Piwi proteins does not lead to any of the above-mentioned phenotypes (Carmell et al. 2007). One can hypothesize that, perhaps, another group of small RNAs may take over from piRNAs and suppress TE activation at the transcriptional or posttranscriptional level. Indeed, siRNA and miRNA are two other clades of small RNAs present in female germ cells and most of them correspond to transposons (Watanabe et al. 2008) (Fig. 2). Both types are derived from long dsRNA precursors which are processed by Dicer into small RNAs and then loaded into Argonaute effector proteins. The defining difference between miRNAs and siRNAs is an additional processing step, whereby primary miRNAs are converted into pre-miRNAs by the Microprocessor complex (Drosha-DGCR8 complex) and only subsequently become modified through the Dicer-dependent pathway. However, Dicer knockout in germ cells does not lead to upregulation of L1 transcription; on the contrary, L1 expression is enhanced specifically in PGCs (Hayashi et al. 2008). Moreover, L1

	5mC	Small RNAs	Histone modifications/modifiers	
Somatic cells	+	None described	H3K9me3/H4K20me3/H3K27me3	
ES cells	+	Dicer-siRNAs	Young L1:	Suv3-9h1/h2 dependent H3K9me3
			Older L1:	KAP-1
Male germ cells	-	PIWI-piRNAs	Young L1:	Small RNA-dependent H3K9me3
			Older L1:	
Female germ cells	-	?	?	
Preimplantation embryos	-	?	-	

Fig. 2 Summary of the major RNA and chromatin signatures involved in transcriptional silencing of Line-L1 in different cell types in the mouse. While no small RNAs have yet been described in somatic cells, male germ and ES cells have been found to contain different classes of small RNAs that are involved in Line-L1 silencing. The main chromatin modifier pathways associated to Line-L1 silencing are shown on the *right*, and the associated levels of DNA methylation (5mC) are also indicated. *5mC* 5 methyl cytosine

expression in oocytes lacking Dicer is also not severely affected. On the other hand, RLTR10 fragments are upregulated in Dicer −/− oocytes and mouse transcript A (MTA) in Mili −/− oocytes, implying that both piRNA and siRNA are active and can suppress retrotransposons in mouse oocytes, but each pathway has preferred targets (Watanabe et al. 2008). It is noteworthy that oocytes display moderate to high levels of L1 and other TE transcripts, suggesting that either the oocyte has an increased tolerance for TE expression or that TE transcripts carried through the oocyte might play a role in early development. The mechanisms responsible for controlling L1 expression in female germ cells remain still unclear, but the prolonged cell cycle characteristic of oocytes has been suggested to act as a block against the full retrotransposition cycle (Shi et al. 2007).

The situation in pre-implantation embryos differs greatly from the one observed in mouse germ cells, as PIWI proteins are not expressed therein (Aravin et al. 2008), and there is no cell cycle arrest. Recent work has shown that this might vary depending on the species, since unlike in the mouse, PIWI and piRNAs are present in bovine two- and four-cell stage embryos (Roover et al. 2015). As experiments on embryos are challenging and have some limitations, hints of pathways potentially involved in the acquisition of a silent chromatin configuration on L1 elements may come from the studies conducted in ES cells. The role of ncRNAs in this process is very plausible considering that an increasing number of types of small RNAs have been reported in ES cells, i.e., endo-siRNAs, miRNAs, and shRNAs, which differ in function and biogenesis. Indeed, small RNAs derived from L1 elements were found in mES cells and have been shown to map mostly to young L1 families (Chow et al. 2010; Ciaudo et al. 2013). Their regulatory role on L1 activity is also likely because in ES cells lacking Dicer or Ago2, young L1 families are upregulated and retrotransposition events occur more often (Ciaudo et al. 2013) (Fig. 2). Worth mentioning is the fact that Dicer and Ago2 mutant ES cells display severe defects in differentiation both in vitro and in vivo, together with a failure in epigenetic silencing of centromeric repeats (Kanellopoulou et al. 2005). This can be explained by the broad function of Dicer-dependent small RNAs but also indicates the complexity of small RNA pathways and the difficulty in investigating their exact role on L1. Studies conducted in somatic human cells show, for example, the existence of bidirectional L1 transcripts derived from the 5′UTR, which can be processed into endogenous siRNAs (endo-siRNAs) and can potentially prevent retrotransposition events. However, Dicer depletion does not lead to massive upregulation of L1 elements in these cells suggesting different biogenesis of these siRNAs or simply the existence of alternative pathways preventing L1 transcriptional upregulation (Yang and Kazazian 2006). To further investigate the role of small RNAs in the control of L1 elements and to determine which type of small RNAs is mostly involved in L1 regulation, studies in DGCR8 −/− ES cells were conducted. Although most canonical miRNAs are depleted in these cells due to improper function of the Microprocessor complex, no increase in L1 copy number was observed (Ciaudo et al. 2013), in spite of the fact that expression of L1 is elevated (Heras et al. 2013). These observations suggest that other mechanisms, i.e., siRNAs or Microprocessor-independent miRNAs, are involved in preventing the full life cycle of L1. Therefore, the decrease in DNA methylation level rather than miRNAs or non-functional Microprocessor can be speculated as the direct cause of L1 transcriptional activation in DGCR8 −/− ES cells (Ciaudo et al. 2013). Some evidence supporting the latter hypothesis comes from studies conducted in human somatic cells where it has been shown that Microprocessor can directly bind full L1 transcripts and cleave them at 5′ UTR regions, which may interfere with their ability to mobilize. Interestingly, the 5′UTR of the active mouse L1 element—L1spa—has been suggested to be controlled in a similar manner because in DGCR8 −/− mES cells, the expression of luciferase reporter downstream of L1spa 5′UTR is increased. Surprisingly, however, in Dicer −/− cells, the same reporter construct does not show high luciferase activity (Heras et al. 2013), implying that either endogenous mouse L1 is controlled differently than transiently transfected plasmids or that the mechanisms regulating L1 transcriptional activity are not entirely conserved between mouse and human. Caution therefore should be taken when interpreting findings across species, as findings in human cells cannot always be accurately extrapolated on regulation of mouse L1.

The RNA-based mechanisms described above are good candidates for keeping TE in check, and there is enough evidence on their involvement in the regulation of L1 elements at different levels. Nevertheless, unlike DNA methylation and histone modifications, these pathways do not "lock" chromatin state, as most of them operate at the posttranscriptional level. Some histone marks can generate a silent chromatin signature, i.e., trimethylations of H3K9, H4K20, and H3K27, which are well-known repressive marks and are found on TEs (Martens et al. 2005). Studies carried out in mouse ES cells have shown, however, that not all retrotransposons have comparable levels of such histone marks (Day et al. 2010) and that the molecular mechanisms responsible for the addition and the removal of repressive histone modifications can vary among different classes of TEs. Constitutive heterochromatin marks and the corresponding histone methyltransferases (HMT) have been studied most extensively. In particular, H3K9me3, which is thought to be involved in de novo DNA methylation and probably acts upstream of H4K20me3 in silencing TEs. Accordingly, depletion of the HMTs SUV4-20H1 and SUV4-20H2—which leads to loss of H4K20me3—does not lead to the upregulation of TEs (Matsui et al. 2010) (Fig. 2). Depletion of H3K9-specific HMTs, on the other hand, results in more complex

effects that depend on the targeted enzymatic pathway and the class of TEs. For example, lack of G9a in ES cells has no effect on TE activation, even though DNA methylation is reduced (Dong et al. 2008). Double knockout of SUV3-9H1 and SUV3-9H2 (dnSuv39h1/h2) induces a modest derepression of IAPs, but L1 is only mildly activated (Martens et al. 2005; Day et al. 2010). Interestingly, recent studies indicate that L1-associated H3K9me3 mostly decorates the 5′UTR of young and full-length elements across different cell types including ES, male germ cells, or differentiated cells (Bulut-Karslioglu et al. 2014; Pezic et al. 2014). In ES cells, H3K9me3 at young L1 was recently suggested to be dependent on SUV39H1/H2 activity, because SUV39H1/H2 depletion leads to increased transcription of these specific L1 elements exclusively (Bulut-Karslioglu et al. 2014). This implies that Suv39-dependent H3K9me3 may play a role in silencing young, full-length L1 whereas different, alternative pathways are involved in the suppression of other L1. Indeed, depletion of the HMT SETDB1 or its interacting partner KRAB-associated protein 1 (KAP-1) leads to massive derepression of several classes of TEs such as ERV-L elements. However, L1 upregulation in *Setdb1−/−* cells is rather weak (Matsui et al. 2010), which might be potentially because DNA methylation is almost unchanged in *Setdb1−/−* and *Kap1−/−* cells. It will be important to revisit these analyses in the light of the different specificities of various L1 elements toward different repressive machineries, as in the original work from Matsui and colleagues, no analysis to discriminate between distinct classes of L1 was performed. Importantly, more recent studies have shown that KAP-1 occupancy is not uniform across TEs; instead, KAP-1 is present on ancient L1 both in human and mouse ES cells. Accordingly, KAP-1 depletion leads to activation of these specific KAP-1-bound L1 elements only. Younger, full-length elements, on the other hand, are upregulated in cells lacking DNMTs (Castro-Diaz et al. 2014), which implies that separate mechanisms are responsible for suppression of evolutionary distinct L1 in ES cells or at least for their transcriptional repression (Fig. 2).

Naturally, the importance of DNA methylation and the H3K9 methylation pathways in L1 repression is not questioned; however, they are seemingly not the only players. Indeed, in the absence of constitutive heterochromatin marks, facultative heterochromatin can ‘take over” (Peters et al. 2003). This “back-up” mechanism, namely the enrichment for H3K27me3, occurs on L1 and other TEs upon deletion of both SUV39 enzymes (Martens et al. 2005). Distinct enrichment patterns of H3K27me3 and H3K9me3/H4K20me3 have been shown for different groups of LTRs: while the ERV IAP contains high trimethylation levels of H3K9 and H4K20 on average, another ERV class, ERV-L is mostly enriched with H3K27me3 (Day et al. 2010). Depletion of the Polycomb group proteins (PcG), RING1B and EED, leads to reactivation of both ERVs (Leeb et al. 2010), implying that not only H3K9 pathways but also H3K27me3 and/or PcG proteins play a role in ERV silencing. However, there is no evidence linking L1 regulation with Polycomb function.

A remarkable conclusion from the above studies is that regulation of L1 and TEs, in general, can be very redundant, as independent repression pathways seem to converge and to functionally compensate each other. Worth mentioning is the fact that no activation of TEs has been observed when knockout experiments of SETDB1 (Matsui et al. 2010) or SUV39H1/H2 (Bulut-Karslioglu et al. 2014) were performed on MEF cells or differentiated ES cells in which DNA methylation is already fully established. This suggests not only that different mechanisms are responsible for silencing distinct groups of TEs but also that their strength and action depend on cell potency and on chromatin state. Indeed, this is why studies conducted in ES cells, while bringing important new insights into the regulation of TEs, cannot be directly extrapolated to retrotransposon regulation during early development. For example, the above suggested role of H3K9me3 in the initial repression of L1 elements is not very likely to take place in pre-implantation embryos because, after fertilization, the paternal genome is devoid of H3K9me3, and the maternal chromatin progressively looses H3K9me3 with almost no signal observed in four-cell and eight-cell stage embryos (Liu et al. 2004). Moreover, recent ChIP analysis have revealed that the amount of H3K9me3 which is present on L1 during this time period does not seem to change, but their transcriptional activity does, sharply decreasing from the two-cell to the eight-cell stage (Fadloun et al. 2013) (Fig. 3). DNA methylation is also not likely to account for this transcriptional repression because L1 undergoes DNA demethylation after fertilization. Although younger L1 elements are demethylated more drastically than ancient elements in the zygote, all analyzed L1 families reach minimal values at blastocyst stage and, only after that, DNA methylation levels increase (Smith et al. 2012) (Fig. 3). Furthermore, removal of Tet3, which takes part in the active demethylation of DNA in zygotes via oxidation of 5mC to 5hmC, does not lead to changes in L1 transcription (Inoue et al. 2012). Thus, the temporal pattern of L1 transcriptional activation and subsequent silencing cannot be explained through changes in DNA methylation or in H3K9me3 after fertilization and prior to implantation, implying that additional pathways are involved in the process.

Similarly to germ cells and mouse ES cells, RNAi and small RNAs have been proposed to take part in the regulation of TEs in pre-implantation embryos. Indeed, siRNAs, miRNAs, and proteins involved in their processing are present in early mouse embryos. Nevertheless, it is not clear whether the RNAi machinery acts as a protecting mechanism against TE activation and/or mobilization. Inhibition of Dicer by injection of mDicer siRNA into zygotes results in an increased abundance of IAP and MuERV-L transcripts in eight-cell stage

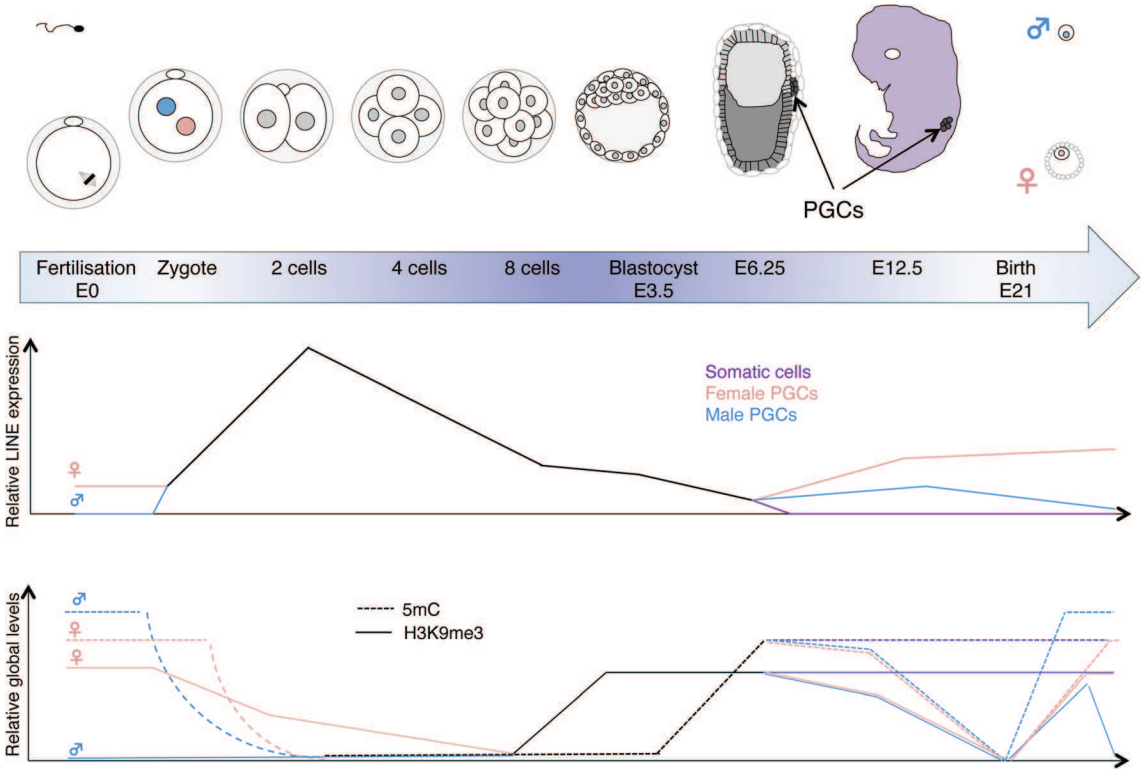


Fig. 3 Temporal dynamics of Line-L1 expression during embryogenesis and correlation with global levels of DNA methylation and H3K9me3. At the *top*, a schematic representation of embryogenesis, starting with fertilization of an oocyte by the sperm and up to the formation of the PGCs, is shown. The *middle graph* shows global levels of Line-L1 elements throughout this period. Note that while, in some cases, the subtypes of Line-L1s expressed have been defined, this has not been addressed in all the stages shown, and therefore, the graph refers to levels of Line-L1 transcription in general. The *bottom graph* shows global levels of DNA methylation (5mC) and H3K9me3 during the corresponding stages of development

embryos, which may support that hypothesis, but preimplantation development is not altered (Svoboda et al. 2004). Perhaps, maternal Dicer provides sufficient stores to support development of these embryos to some extent as Dicer null mice die only after implantation (Bernstein et al. 2003). Moreover, it is difficult to address the role of Dicer and small RNAs in TE regulation because of their additional roles in regulating other cellular processes and important developmental events. For example, Dicer and miRNAs inherited in the oocyte are crucial for the earliest stages of embryonic development, as fertilized oocytes lacking maternal Dicer cannot complete the first cell cycle and display defects in mitotic spindle formation (Tang et al. 2007). That is why, phenotypic defects in Dicer −/− experiments are explained mostly from failure in processing endogenous miRNA and not necessarily linked with TE regulation. Notwithstanding, the involvement of RNAs in the regulation of L1 expression has been addressed in a more targeted approach in which a 17-nt-long single-stranded RNA targeting the L1 ORF1 region was injected into zygotes. Surprisingly, this caused increased L1 transcription in two-cell stage embryos, suggesting that transcriptional regulation of L1 is indeed under control of RNA, but the exact mechanism remains unknown (Fadloun et al. 2013).

Concluding remarks: a role in early development?

It is clear that the period that follows fertilization in mammals represents a great window of opportunity for L1 to be activated, given the chromatin context in the embryo during that time. A current challenge in the field is to determine whether L1 reactivation is just a threatening "side effect" of the open chromatin structure entailing epigenetic reprogramming or whether it has a role in development.

Indeed, a number of hypotheses are plausible; for example, L1 transcripts could serve as scaffolds of chromatin proteins or for the formation of RNPs. Additionally, perhaps, one of the most appealing hypotheses is that L1 helps in nucleating heterochromatin formation. A precedent for this is the suggested role of L1 during inactivation of the X chromosome (XCI) in females. In placental mammals, the nc *Xist* gene is transcribed from the X, and the resulting *Xist* RNA can coat randomly one chromosome, in *cis*, to initiate silencing of the majority of its genes, leading to XCI. However, *Xist* is not abundant and the mechanisms behind spreading and subsequent silencing of the complete chromosome remain unclear. According to the Lyon's hypothesis, L1 elements are the boosters which facilitate spreading of heterochromatin along the entire

chromosome, eventually leading to the repression of all genes (Lyon 1998, 2000). The X contains almost twice as many L1 copies as the mouse autosomes (Waterston et al. 2002; Boyle et al. 1990), and full, young elements are more abundant on the X than on other chromosomes (Abrusan et al. 2008), observations which support Lyon's idea. Also, more recent findings have shown that silent L1 together with Xist RNA takes part in the formation of a silence compartment on the X chromosome (Chow et al. 2010). The link between L1 enrichment and its function in spreading RNA and silencing can be also appreciated from an evolutionary argument because the time of increase in L1 copy number on the X correlates with the emergence of random X inactivation in eutherian (Mikkelsen et al. 2007). The known enrichment of full L1 in the proximity of randomly, monoallelically expressed genes also supports the idea that they can have a role in inactivation of one copy of a gene or even a whole chromosome (Allen et al. 2003). How exactly the induction and/or spreading of a silent state works remains unclear; however, there are at least two plausible ways. The first one depends on L1 genomic sequences, which could act as "attractors" for heterochromatinization factors which subsequently spread on neighboring regions as in the above-mentioned case of *Xist* coating. Moreover, it has been suggested that repressive marks, i.e., H3K9me3 and H4K20me3, can spread from repeats onto proximal unique sequences (Mikkelsen et al. 2007). Although heterochromatin spreading induced by LTRs has been observed, it remains questionable if the same situation occurs with L1 (Rebollo et al. 2014). The second possible way of triggering silencing by L1 may require their transcriptional activity, with transcripts serving as substrates to produce, i.e., small RNAs. XCI can be again a good example, as it was shown that during ES cell differentiation, young, full-length L1 elements are transcribed from regions of inactive X that might be prone to escape inactivation and facilitate the local spreading of silencing via not well-defined siRNA pathways (Chow et al. 2010). This may explain the presence of RNAs derived from L1 in some periods of development and might suggest their role in global chromatin reorganization and heterochromatin formation in the early embryos.

Interestingly, the above-mentioned siRNA pathway includes not only transcripts from L1 but also transcripts from the gene being targeted for silencing, and in this case, the transcriptional activity in question is thought to be driven by the proximal antisense promoter of L1 (Chow et al. 2010). This implies additional function for L1s as alternative promoters that may work in both sense and antisense direction. Thus, during development, L1 might help directing embryonic genome activation (EGA) by providing strong promoters to host genes. Indeed, chimeric transcripts that originate from L1 have been detected in mouse embryos (Peaston et al. 2004; Li et al. 2014). It is also possible that by maintaining a high transcriptional activity, L1 might help promoting an open chromatin packaging of the early embryo to facilitate epigenetic reprogramming and EGA. Evidence in support of this hypothesis comes from experiments in which L1 transcripts have been shown to act as scaffolds that bind euchromatic DNA. After their removal, condensation of chromatin occurs, suggesting the importance of L1 transcripts in keeping a more open chromatin configuration (Hall et al. 2014). An open chromatin state is one of the hallmarks of early embryogenesis and has been associated to totipotency (Boskovic et al. 2014), which would explain the reason for keeping high transcription of L1 elements after fertilization and their subsequent silencing. Moreover, recent discoveries have suggested that L1 elements could potentially play a role in gene regulation in early embryogenesis, as significant associations were found for the genes with intragenic L1 elements and their downregulation from two-cell stage to morula in in silico analyses (Ngamphiw et al. 2014).

While we have learnt a great deal of information about their structural organization, their evolution, and the mechanisms that lead to their silencing, we still do not know whether activation of L1 during the periods of epigenetic reprogramming has a functional purpose and many questions remain open. These questions are currently the subject of intense investigation and promise to provide important findings in the years to follow.

Acknowledgments Work in M.E.T.-P lab is funded through EpiGeneSys NoE, ERC-Stg "NuclearPotency," the EMBO YIP, and the Fondation Schlumberger pour l'Education et la Recherche. JJ is partially supported by a LabEx fellowship (ANR-10-LABX-0030-INRT under the frame programme Investissements d'Avenir).

Conflict of interest The authors declare no conflict of interest.

References

Abrusan G, Giordano J, Warburton PE (2008) Analysis of transposon interruptions suggests selection for L1 elements on the X chromosome. PLoS Genet 4:e1000172

Adey NB, Schichman SA, Graham DK, Peterson SN, Edgell ME, Hutchison CA (1994) Rodent Ll evolution has been driven by a single dominant lineage that has repeatedly acquired new transcriptional regulatory sequences. Mol Biol Evol 11:778–789

Akagi K, Li J, Stephens RM, Volfovsky N, Symer DE (2008) Extensive variation between inbred mouse strains due to endogenous L1 retrotransposition. Genome Res 18:869–880

Allen E, Horvath S, Tong F, Kraft P, Spiteri E, Riggs AD, Marahrens Y (2003) High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. PNAS 100: 9940–9945

Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 316:744–747

Aravin A, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 31:785–799

Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ (2003) Dicer is essential for mouse development. Nature 35:215–217

Boskovic A, Eid A, Pontabry J, Ishiuchi T, Spiegelhalter C, Raghu Ram EV, Meshorer E, Torres-Padilla ME (2014) Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. Genes Dev 28:1042–1047

Bourc'his D, Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 431:96–99

Boyle AL, Ballard SG, Ward DC (1990) Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. Proc Natl Acad Sci 87:7757–7761

Bulut-Karslioglu AB, De La Rosa-Velazquez IA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, Galan C, Winter GE, Engist B, Gerle B, O'Sullivan RJ, Martens JHA, Walter J, Manke T, Lachner M, Jenuwein T (2014) Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. Mol Cell 55:277–290

Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ (2007) MIWI2 is essential for spermatogenesis and repression of transposones in the mouse male germline. Dev Cell 12:503–514

Castro-Diaz N, Ecco N, Collucio A, Kapopaoulou A, Yazdanpanah B, Friedli M, Duc J, Min Jang S, Turelli P, Trono D (2014) Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev 28:1397–1409

Chow JC, Ciaudo C, Fazzari MJ, Mise N, Servant N, Glass JL, Attreed M, Avner P, Wutz A, Barillot E, Greally JM, Voinnet O, Heard E (2010) LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation chromosome inactivation. Cell 141:956–969

Ciaudo C, Jay F, Okamoto I, Chen C-J, Sarazin A, Servant N, Barillot E, Heard E, Voinnet O (2013) RNAi-dependent and independent control of LINE1 accumulation and mobility in mouse embryonic stem cells. PLoS Genet 9:e1003791

Day DD, Luquette LJ, Park PJ, Kharchenko KV (2010) Estimating enrichment of repetitive elements from high-throughput sequence data. Genome Biol 11:R69

DeBeradinis RJ, Kazazian HH Jr (1999) Analysis of the promoter from an expanding mouse. Genomics 56:317–323

deFazio S, Bartonicek N, Di Giacomo M, Abreau-Goodger C, Sankar A, Funaya C, Antony C, Moreira P, Enright AJ, O'Carroll D (2011) The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature 23:259–263

Di Giacomo M, Comazzetto S, De Fazio S, Carrieri C, Morgan M, Vasiliauskaite L, Benes V, Enright A (2013) O'Carroll Donal. Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol Cell 50:601–608

Di Giacomo M, Comazzetto S, Sampath S, O'Carroll D (2014) G9a co-suppresses LINE1 elements in spermatogonia. Epigenetics Chromatin 7:24

Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schubeler D, Tachibana M, Shinkai Y, Lorincz MC (2008) DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. EMBO J 27:2691–2701

Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P, Torres-Padilla ME (2013) Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. Nat Struct Mol Biol 20:332–338

Goodier JL, Ostertag EM, Du K, Kazazian HH Jr (2001) A novel active L1 retrotransposon subfamily in the mouse. Genome Res 11:1677–1685

Hall LL, Carone DM, Gomez AV, Kolpa HJ, Byron M, Mehta N, Fackelmayer FO, Lawrence JB (2014) Stable CoT-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. Cell 156:907–919

Han JS, Szak ST, Boeke JD (2004) Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptiomes. Nature 20:268–274

Hardies SC, Wang L, Zhou L, Zhao Y, Casavant C, Huang S (2000) LINE-1 (L1) lineages in the mouse. Mol Biol Evol 17:616–628

Hayashi K, Chuva de Sousa Lopes SM, Kaneda M, Tang F, Hajkova P, O'Carroll D, Das P, Tarakhovsky A, Miska EA, Surani MA (2008) MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. Plos One 3:e1738

Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyras E, Garcia-Perez JL, Caceres JF (2013) The Microprocessor controls the activity of mammalian retrotransposons. Nat Struct Mol Biol 20(10):1173–1181

Inoue A, Matoba S, Zhang Y (2012) Transcriptional activation of transposable elements in mouse zygotes is independent of Tet3-mediated 5-methylcytosine oxidation. Cell Res 22:1640–1649

Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 15:489–501

Kano H, Godoy I, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH Jr (2009) L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev 23:1303–1312

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHug W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A (2010) Polycomb complexes act redundantly to repress genomic repeats and genes. Gens Dev 24:265–276

Li J, Kannan M, Trivett AL, Liao H, Wu X, Akagi K, Symer DE (2014) An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. Nucleic Acids Res 42:4546–4562

Liu H, Kim JM, Aoki F (2004) Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. Development 131:2269–2280

Lyon MF (1998) X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet 80:133–137

Lyon MF (2000) LINE-1 elements and X chromosome inactivation: a function for "junk" DNA? PNAS 97:6248–6249

Martens JHA, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J 24:800–812

Martin SL, Bushman FD (2001) Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol Cel Biol 21:467–475

Matsui T, Leung D, Miyshita H, Maksakova I, Miyachi H, Kimura H, Tachibana M, Lorincz MC, Shinkai Y (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464:927–931

Mears ML, Hutchison CA (2001) The evolution of modern lineages of mouse L1 elements. J Mol Evol 52:51–62

Medstrand P, van de Lagemaat LN, Mager DL (2002) Retroelements distribution in human genome: variations associated with age and proximity to genes. Genome Res 12:1483–1495

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770

Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, Duke S, Garber M, Gentles AJ, Goodstadt L, Heger A, Jurka J, Kamal M, Mauceli E et al (2007) Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. Nature 10: 167–178

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–562

Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435:903–910

Naas TP, DeBerardinis RJ, Moran JV, Ostertag EM, Kingsmore SF, Seldin MF, Hayashizaki Y, Martin SL, Kazazian HH Jr (1998) An actively retrotransposing, novel subfamily of mouse L1 elements. EMBO Jr 15:590–597

Ngamphiw C, Tongsima S, Mutirangura A (2014) Roles of intragenic and intergenic L1s in mouse and human. Plos One 9:e113434

Ohno S (1972) So much "junk" DNA in our genome. Brookhaven Symp Biol 23:366–370

Orgel LE, Crick FH (1980) Selfish DNA: the ultimate parasite. Nature 284:604–607

Padgett RW, Hutchinson CA III, Edgell MH (1988) The F-type 5′ motif of mouse L1 elements: a major class of L1 termini similar to the A-type in organization but unrelated in sequence. Nucleic Acids Res 16:739–749

Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Dev Cell 7:597–606

Peters A, Kubicek S, Mechtler K, O'Sullican RJ, Derick AAAHA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JHA, Jenuwein T (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12:1577–1589

Pezic D, Manakov SA, Sachidanandam R, Aravin AA (2014) piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes Dev 28:1410–1428

Rebollo R, Karimi MM, Bilenky M, Gagnier L, Miceli-Royer K, Zhang Y, Goyal P, Keane TM, Hirst M, Lorincz MC, Mager DL (2014) Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphism. PLOS Genet 7(9):e1002301

Reuter M, Berninger P, Chuma S, Shah H, Hosokawa M, Funaya C, Antony C, Sachidanandam R, Pillai R (2011) Miwi catalasis is required for piRNA amplification-independent LINE1 transposon silencing. Nature 27:264–267

Roover EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LA, Zischler H, Butter F, Roelen BA, Ketting RF (2015) Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep 10:2069–2082

Shi X, Seluanov A, Gorbunova V (2007) Cell divisions are required for L1 retrotransposition. Mol Cell Biol 27:1264–1270

Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 28:339–344

Sookdeo A, Hepp CM, McClure MA, Boissinot S (2013) Revisiting the evolution of mouse LINE-1 in the genomic era. Mob DNA 4:3

Svoboda P, Stein P, Anger M, Bernstein E, Hannon GJ, Schultz RM (2004) RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. Dev Biol 1:276–285

Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA (2007) Maternal microRNAs are essential for mouse zygotic development. Genes Dev 21:644–648

Walsh CP, Chaillet RJ, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20:116–117

Watanabe T, Totoki Y, Toyoda A, Keneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani AM, Sakaki Y, Sasaki H (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453:539–543

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562

Yang N, Kazazian HH Jr (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol 13:763–771

Zamudio N, Bourc'his D (2010) Transposable elements in the mammalian germline: a comfortable niche or a deadly trap? Heredity 105: 92–104

II. Main questions and objectives

All of the above, led us to formulate several questions in regard to heterochromatin formation during early mouse development. The main, and probably most broad one, is how exactly heterochromatin is assembled, specified and reprogrammed after fertilization. As it is impossible to answer such a general question, I have focused on two different genomic regions that served as model for heterochromatin – pericentric repeats and L1 transposons. These regions are defined by very different characteristics in terms of function, localization, expression, yet both are examples of constitutive heterochromatin, which is established on these elements during development. Thus, during my Ph.D, I aimed to uncover the mechanisms that lead to their silencing and the distinct role they may play during the process of development and cell division.

Part I – Pericentric repeats

The **first main objective** of my work was to investigate whether the spatial nuclear localization of pericentric chromatin within nucleus is important for their silencing. To address it, I asked more specific questions listed below:

- What is the actual spatial organization and the exact transcription pattern of major satellites in zygote and 2-cell stage embryos?
- What is the chronology of the repressive mark acquisition and reorganization of pericentric repeats?
- Does the localization of major satellite around the NLBs play a role in developmental progression?
- Is the particular organization of pericentric regions in the nucleus important for their silencing?
- What might be the epigenetic pathway/mechanism involved in their repression?

Part II – L1 transposable elements

The **second main objective** of my studies was to determine the importance and plausible function of L1 elements during preimplantation development, with a particular focus on the

best characterized Tf family member L1spa. To investigate this, I asked the following specific questions:

- What is the exact pattern of L1spa expression during development with regard to its distinct regions; i.e. whether Tf, Orf1, Orf2 regions are expressed at the same time of development from the same loci?
- Is the silencing of L1s necessary for early development and cell proliferation?
- What is the role that the transcription of L1 plays, if any, during pre-implantation development, and how does it affect development?
- How are L1 regulated in preimplantation embryos?

III. Results & Discussions

Part I - Pericentric repeats

1.1. Summary of Publication 2

Pericentric and centric heterochromatin is composed of major and minor satellite repeats, respectively, which have specific spatio-temporal organization during early development. In the zygote, despite asymmetric organization of heterochromatic marks in pronuclei, pericentric repeats of both sets of chromatin behave in a similar manner and cluster around nucleolar-like bodies (NLBs) forming ring-like structures. After the first division chromocenters' organization is progressively restored and by 4-cell stage no ring-like structures are present. Silencing of these regions occurs presumably at the same time because peak of transcription has been detected only in late zygotes and 2-cell stage embryos. Moreover, previous work in the Torres-Padilla lab has suggested that there must be a link between specific localization of pericentric repeats and their silencing as in the embryos expressing H3.3K27R in which chromocenter formation is disrupted, transcription of major satellites is increased and HP1β is mislocalized. Hence, the main objective of this work is to investigate whether the spatial localization of pericentric chromatin is important for their silencing. To address this question, I performed experiments in which the natural localization of pericentromeric regions during early development was altered experimentally. More precisely, pericentric DNA was mislocalized from NLBs towards the nuclear membrane. Injections of early zygotes with mRNA coding Zinc-finger specific for major satellites and fused with Emerin enabled this manipulation. As a consequence, by use of RNA-FISH technique I could observe higher transcription from these regions present in 2-cell stage embryos and also a change in some of heterochromatin marks i.e. H3K27me3. In addition, embryos with perturbed ring-like pericentromeric heterochromatin localization displayed lagging chromosomes and developmental arrest. Thus, the data that I generated during these experiments indicate that the specific spatial organization of pericentric domains is essential for their silencing and subsequent proper chromatin organization.

RESEARCH COMMUNICATION

Heterochromatin establishment at pericentromeres depends on nuclear position

Joanna W. Jachowicz,[1,3] Angèle Santenard,[1,3] Ambre Bender,[1] Julius Muller,[2] and Maria-Elena Torres-Padilla[1,4]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM U964, Université de Strasbourg, F-67404 Illkirch, France; [2]Laboratory of Chromatin, Epigenetics, and Differentiation, Institute of Molecular and Cell Biology, Agency for Science, Technology, and Research, Singapore 138673, Singapore

Mammalian development begins with fertilization of an oocyte by the sperm followed by genome-wide epigenetic reprogramming. This involves de novo establishment of chromatin domains, including the formation of pericentric heterochromatin. We dissected the spatiotemporal kinetics of the first acquisition of heterochromatic signatures of pericentromeric chromatin and found that the heterochromatic marks follow a temporal order that depends on a specific nuclear localization. We addressed whether nuclear localization of pericentric chromatin is required for silencing by tethering it to the nuclear periphery and show that this results in defective silencing and impaired development. Our results indicate that reprogramming of pericentromeric heterochromatin is functionally linked to its nuclear localization.

Supplemental material is available for this article.

Received June 16, 2013; revised version accepted October 16, 2013.

In mammals, fertilization of an oocyte by the sperm is followed by epigenetic reprogramming, which involves de novo acquisition of chromatin signatures in the two parental genomes, but the molecular determinants underlying such reprogramming are not fully understood. In particular, the formation of heterochromatin de novo is thought to be essential to ensure the subsequent organization of the embryonic epigenome and embryonic development (Probst and Almouzni 2011; Fadloun et al. 2013; Nestorov et al. 2013).

In the zygote, remodeling of the paternal chromatin is particularly extensive, since it is subject to a nearly genome-wide replacement of protamines by maternally supplied histones. A few hours after fertilization, the pericentromeric chromatin must acquire a highly compact chromatin organization for the first time to allow subsequent kinetochore loading and progression through the first mitosis. The pericentromere domain is formed by tandem repeats of major and minor satellite repeats, which constitute the pericentric and centric chromatin, respectively. In the embryo, the initial silencing of pericentromeric chromatin requires the Lys27 of the histone variant H3.3 and is accompanied by progressive acquisition of H3K27 methylation and tethering of HP1β in a mechanism that involves a burst of transcription from the major satellite (MajSat) repeats in the zygote, in analogy to heterochromatin formation in *Schizosaccharomyces pombe* (Grewal and Elgin 2007; Puschendorf et al. 2008; Probst et al. 2010; Santenard et al. 2010). This heterochromatinization in mammals, however, seems to be "atypical," since it is independent of H3K9me3 and H4K20me3, which are absent from the paternal chromatin (Arney et al. 2002; Kourmouli et al. 2004; Santos et al. 2005). In agreement, the expression of an H3.3K27R mutant disturbs heterochromatin silencing at the pericentromeres, while expression of an H3.3K9R mutant has no discernible effect in development (Santenard et al. 2010).

Pericentromeric heterochromatin is typically visualized as DAPI-rich regions in mouse somatic cells in interphase, which are referred to as chromocenters. These form upon clustering of pericentromeric domains from several chromosomes. In the embryo, however, the pericentromeric chromatin does not form chromocenters until the late two-cell/early four-cell stage transition (Martin et al. 2006a; Probst et al. 2007; Aguirre-Lavin et al. 2012). Instead, the mammalian embryo displays a distinctive nuclear organization with a radial arrangement of the chromosome territories with centromeres attached to the nucleoli precursors. The pericentromeric chromatin envelops these precursors, which are referred to as nucleolar-like bodies (NLBs), forming a characteristic ring-like structure.

It is increasingly clear that nuclear organization provides a landmark for gene regulation (Akhtar and Gasser 2007; Kumaran et al. 2008). Although a number of studies have documented changes in the nuclear localization of pericentromeric chromatin after fertilization that correlate with reprogramming efficiency upon somatic cell nuclear transfer (Martin et al. 2006b), a functional role for this process in heterochromatin establishment and/or maintenance has not been addressed, and it is not known whether the spatial organization of the genome can regulate reprogramming. Here, we addressed the temporal dynamics of the acquisition of heterochromatic signatures of the pericentric chromatin in relation to their position in the nuclear three-dimensional (3D) space. We found that the pericentric repeats reach their localization around the nucleoli precursors in the zygote prior to the acquisition of their embryonic heterochromatic signatures (H3K27me1, H3K27me3, and HP1β), suggesting that their spatial configuration is essential for heterochromatic silencing. We tested this hypothesis by artificially

[*Keywords*: pericentromeric chromatin; nuclear organization; epigenetic reprogramming; heterochromatin]
[3]**These authors contributed equally to this work.**
[4]**Corresponding author**
E-mail metp@igbmc.fr
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.224550.113.

tethering pericentromeric chromatin to the nuclear periphery using a designed zinc finger (ZF) fused to emerin, an integral component of the nuclear envelope. Our results show that the spatial localization of the pericentromeric heterochromatin is essential for its silencing and the reprogramming of embryonic chromatin after fertilization.

Results and Discussion

We first addressed the temporal dynamics of the nuclear positioning of pericentromeric chromatin in relation to the acquisition of heterochromatic signatures on the MajSats after fertilization. DNA-FISH was incompatible with the analysis of histone modifications and HP1β, since we noted that the immunostaining profile of several histone modifications and HP1β was perturbed after DNA-FISH as compared with immunostaining alone, presumably due to the denaturing conditions required for FISH with probes other than oligonucleotides (data not shown). Thus, to locate the pericentromeric chromatin in the embryo, we employed a GFP-tagged polydactyl ZF that recognizes specifically the MajSat sequences (see also below) (Lindhout et al. 2007). Importantly, the ZF-GFP showed the same pattern of localization of pericentromeric repeats as that obtained upon DNA-FISH with a MajSat probe, with a progressive decondensation from the center of the male pronucleus toward an organization into rings surrounding the NLBs, in agreement with previous reports (Fig. 1A,B; Martin et al. 2006a; Probst et al. 2007). The female pronucleus also showed the expected pattern of reorganization around the NLBs (Fig. 1A,B). A complete ring-like organization was clearly visible in early two-cell stage embryos (Fig. 1A,B).

While the female chromatin inherits H3K27 methylation from oogenesis, H3K27me3 is undetectable on the paternal pronucleus immediately after fertilization (Santos et al. 2005; Puschendorf et al. 2008; Santenard et al. 2010). Pericentromeric regions become localized in rings around the NLBs by the mid-zygotic stage, as judged from the ZF-GFP signal in pronuclear stage 2–3 (PN2–3) zygotes, without detectable levels of H3K27me3 (Fig. 1C). Remarkably, pericentromeric chromatin acquires H3K27me3 after this stage, and H3K27me3 is clearly and sharply detected around the NLBs at late pronuclear stages, colocalizing with the ZF-GFP signal (Fig. 1C). H3K27me1 and HP1β showed a more dispersed localization pattern in zygotes, but the enrichment of both of these heterochromatic marks with pericentromeric chromatin was only detected at the latest pronuclear stages, after they reach their ring-like destination around the NLBs (Fig. 1D,E). Thus, pericentromeric chromatin becomes localized in sharp rings around NLBs prior to the acquisition of a detectable heterochromatic signature, suggesting that this specific nuclear localization might be a prerequisite for silencing.

To address whether nuclear positioning of pericentromeric chromatin is required for heterochromatin establishment and subsequent silencing, we aimed to manipulate the localization of the MajSat in the nucleus. For this, we designed an approach that would lead to tethering of the pericentromeric chromatin to the nuclear periphery by fusing the above-described ZF to emerin, a constituent of the nuclear membrane (for review, see Bengtsson and Wilson 2004). The targeting fusion protein ZF-Eme is expected to associate with the MajSats via the ZF interaction and with the nuclear lamina via the emerin. We chose the nuclear periphery owing to its known

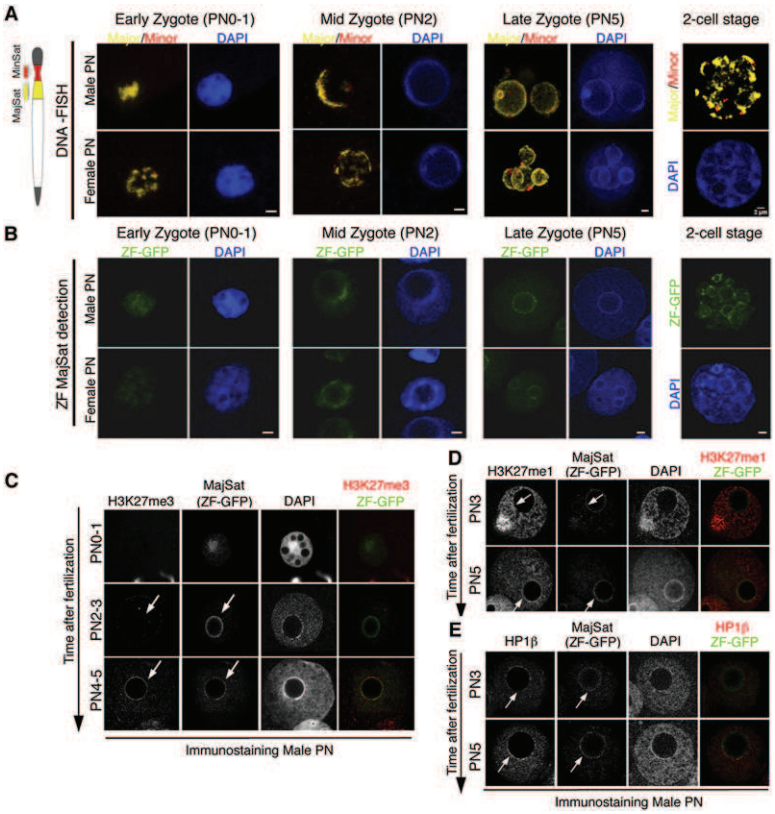


Figure 1. Pericentromeric chromatin localizes around the nucleoli prior to the acquisition of heterochromatic signatures. (*A*) DNA-FISH for major and minor satellite repeats. Representative zygotes immediately after fertilization (early, PN0–1) and at subsequent stages (mid and late) and early two-cell stage embryos are shown. Full projections of confocal sections of the male and female pronuclei are shown. (*Left*) A schematic representation of a mouse chromosome with centromeric regions and the probes color code. (PN0) pronuclear stage 0; (PN1) pronuclear stage 1, etc. Bar, 2 μm. (*B*) Temporal dynamics of the localization of pericentromeric chromatin using ZF design. mRNA coding a ZF targeting MajSats fused to GFP (ZF-GFP) was microinjected at fertilization, and embryos were analyzed at the indicated times. Shown are maximal projections of confocal Z-series as in *A*. Note that the ZF-GFP pattern recapitulates faithfully the pattern obtained by DNA-FISH (yellow label in *A*). Bar, 2 μm. (*C*) Localization of MajSats around the NLBs precedes acquisition of heterochromatic H3K27 methylation signatures on the paternal chromatin. Immunostaining of ZF-GFP-expressing embryos with an H3K27me3 antibody in early (PN0–1), mid (PN2–3), and late (PN4–5) zygotes. Shown are representative single confocal sections of the male pronucleus with the H3K27me3, the GFP, and the DAPI channel in grayscale and the corresponding merge images. Arrows point to places of accumulation of MajSat repeats at the NLBs and the initial absence of H3K27me3 followed by strong accumulation of H3K27me3 in late zygotes. (*D*) Immunostaining analysis as in *C* but with an H3K27me1 antibody. Note that the sharp signal of H3K27me1 colocalizing with the ZF-GFP occurs only after ZF-GFP signal concentrates around the NLBs (arrows) in late (PN5) zygotes. (*E*) Accumulation of HP1β on the pericentromeric chromatin occurs after their relocation around the NLBs at later zygotic stages. Analysis as in *C* and *D* with an HP1β antibody.

potential role in gene regulation and because it is the only other nuclear space that has been clearly identified in the embryo. Because chromatin immunoprecipitation (ChIP) analysis cannot be performed in the embryo, we used transiently transfected NIH3T3 as an additional control to verify binding of the ZF to the MajSats. The ZF-GFP fusion bound specifically to MajSat repetitive sequences—but not to SINE B1 or IAP elements—as demonstrated by ChIP analysis (Supplemental Fig. 1A). We then validated the ZF-Eme construct along with two negative controls—the ZF-GFP construct and a construct expressing emerin alone (HA-Eme)—in NIH3T3 cells. We confirmed by immunostaining that all three constructs were expressed in NIH3T3 cells. The ZF-GFP and HA-Eme controls localized to DAPI-rich regions and the nuclear periphery, respectively, as expected (Supplemental Fig. 1B). In contrast to the ZF-GFP construct, the ZF-Eme construct localized to the nuclear periphery (Supplemental Fig. 1B). Importantly, the ZF-Eme retained the specific ability to bind to MajSat sequences, in contrast to the emerin alone, which did not bind to the MajSat or any of the other repetitive sequences analyzed by ChIP (Supplemental Fig. 1C).

To address whether fusing the ZF to emerin results in efficient tethering of pericentromeric chromatin to the nuclear periphery, we performed 3D DNA-FISH with a MajSat probe in embryos expressing ZF-Eme. We microinjected mRNA for ZF-GFP or ZF-Eme immediately after fertilization and analyzed embryos at the two-cell stage. While the ZF-GFP protein localized as expected around the NLBs and in the one to two chromocenters that are not associated to the NLBs (Probst et al. 2007), the ZF-Eme construct localized to the nuclear periphery in both nuclei of two-cell stage embryos (Fig. 2A). The localization pattern of the pericentromeric chromatin was severely affected in most ZF-Eme embryos and showed a spotty pattern with patches throughout the nuclear membrane rather than the ring and chromocenter configuration typically observed in two-cell stage control embryos (Fig. 2B; Supplemental Movies 1–3). Analysis of combined DNA-FISH for MajSat and immunostaining for endogenous Lamin B through single Z confocal sections revealed that pericentromeric chromatin is largely colocalized with Lamin B in ZF-Eme embryos (Fig. 2C), consistent with the localization of the ZF-Eme construct on the nuclear periphery at this time point (Fig. 2A). While in control embryos, MajSats can be found both around the NLBs and close to the Lamin B domain (Fig. 2C), in ZF-Eme embryos, the proportion of MajSat signal not associated with the nuclear periphery is minimal (Fig. 2C). Importantly, the tethering of MajSats to the nuclear periphery was time-sensitive, since expression of the same ZF-Eme construct at the late two-cell stage did not lead to a significant displacement of the MajSats toward the nuclear periphery at the four-cell stage compared with controls (data not shown). Thus, our ZF-Eme approach in zygotes leads to the efficient tethering of the pericentromeric chromatin to the nuclear periphery.

We next asked whether altering the localization of pericentromeric chromatin results in defective heterochromatic silencing. In the zygote, transcription from the MajSats occurs primarily from the male pronucleus, and these transcripts are believed to be necessary for the subsequent silencing and organization into a somatic-like chromocenter configuration, which occurs progressively from the late two-cell stage (Puschendorf et al. 2008; Probst et al. 2010; Santenard et al. 2010). We therefore analyzed embryos by RNA-FISH with a MajSat probe to detect nascent transcription of pericentromeric chromatin in two-cell stage embryos. Noninjected and ZF-GFP control embryos display a few transcription foci, as expected (Fig. 3A). In contrast, the ZF-Eme embryos displayed a significantly increased RNA-FISH signal, suggesting that the pericentromeric chromatin is transcribed more actively in ZF-Eme embryos as compared with controls (Fig. 3A). Quantification of both the number of MajSat transcription sites (Fig. 3B) and the volume of their transcripts (Fig. 3C) revealed that ZF-Eme embryos have a significantly higher transcriptional output of pericentromeric repeats, suggesting that repositioning of pericentromeric chromatin to the nuclear periphery leads to defective heterochromatin silencing. Moreover, displacing pericentromeric chromatin to the nuclear periphery also resulted in impaired chromocenter formation (Supplemental Fig. 2).

Next, we asked whether recruitment of pericentromeric chromatin to the nuclear periphery affects developmental progression. For this, we injected zygotes as before with ZF-GFP or HA-Eme mRNA as controls and ZF-Eme mRNA. For the two latter constructs, mRNA for GFP was coinjected as a positive control for microinjection. Embryos expressing HA-Eme and ZF-GFP developed at a similar rate and ratios compared with the noninjected control embryos, with between 83% and 88% embryos forming blastocysts, which are the routine values obtained in this type of manipulation (Fig. 3D; Supplemental Fig. 3A). In contrast, expression of ZF-Eme resulted in a significantly

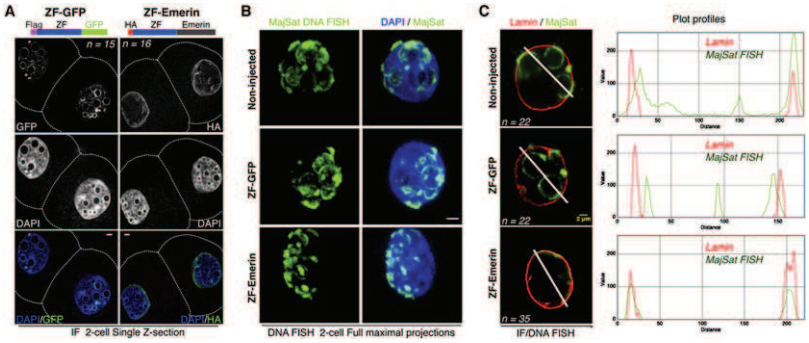


Figure 2. Expression of ZF-Eme results in efficient tethering of pericentromeric chromatin to the nuclear periphery in the early embryo. (*A*) Immunostaining of two-cell stage embryos microinjected with mRNA for ZF-GFP and ZF-Eme using GFP acquisition or an HA antibody as indicated. Dotted lines delineate the cell membrane. Representative single confocal sections are shown. Bar, 2 μm. (*B*) Expression of ZF-Eme in embryos results in displacement of pericentromeric chromatin at the nuclear periphery. Embryos microinjected with the indicated mRNA were processed for 3D DNA-FISH with a MajSat probe (green); DNA was stained with DAPI (blue). Full Z-series projections are shown; serial confocal sections are shown in Supplemental Movies 1–3. Bar, 2 μm. (*C*) Representative single confocal sections of immuno-DNA-FISH with a Lamin B antibody and a MajSat probe of two-cell stage embryos expressing the indicated proteins. The RGB profiles of the corresponding single confocal sections are shown at the *right*. Bar, 2 μm.

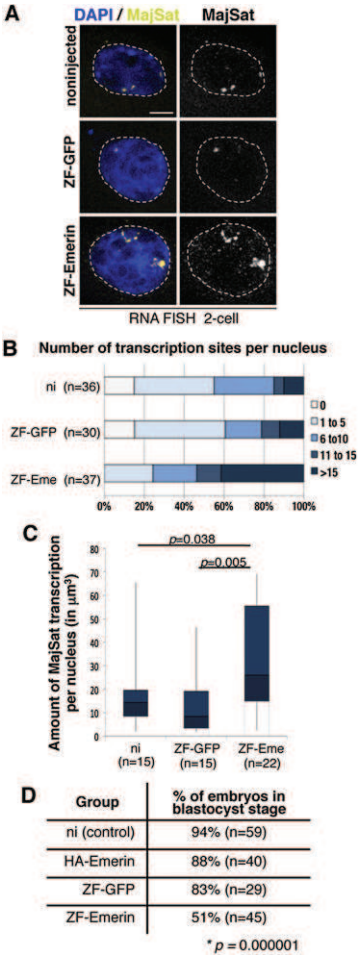


Figure 3. Tethering pericentromeric chromatin to the nuclear periphery impairs silencing and developmental progression. (*A*) RNA-FISH to reveal nascent MajSat transcripts (yellow) in embryos expressing the indicated mRNAs. Representative nuclei at the two-cell stage are shown. The dotted line demarcates the nuclear membrane. Bar, 5 µm. (*B*) Quantification of the number of MajSat transcription sites as determined by RNA-FISH. Noninjected (ni) embryos or embryos expressing ZF-GFP or ZF-Eme were analyzed at the two-cell stage. Embryos were grouped according to the number of transcription sites per nucleus as indicated at the *right* (zero sites, between one and five sites, between six and 10 sites, etc.). The percentage of embryos belonging to each category was plotted for each expression construct. (*C*) Box plot of the quantification of the total volume of MajSats transcribed in noninjected (ni) embryos or embryos expressing ZF-GFP or ZF-Eme. (*D*) Analysis of developmental progression of noninjected embryos or embryos expressing HA-Eme, ZF-GFP, or ZF-Eme. Zygotes were microinjected with the indicated mRNAs. The number of embryos reaching the blastocyst stage for each group was scored after 3 d of development. *n* indicates the total number of embryos analyzed per group.

reduced rate of developmental progression, with only 51% of the embryos reaching the blastocyst stage (P = 0.000001, Kruskal-Wallis test) (Fig. 3D). The remaining 49% stopped development between the two- and eight-cell stages, similarly to embryos with defective heterochromatin silencing (Probst et al. 2010; Santenard et al. 2010). Importantly, the 51% of the embryos that did reach the blastocyst stage displayed delayed development. We also addressed whether embryos expressing ZF-Eme display division errors such as lagging chromatin. In line with our earlier findings of defective heterochromatin formation (Santenard et al. 2010), we observed a high incidence of lagging chromatin in ZF-Eme-expressing embryos (21% compared with 6%, 3%, and 7% for noninjected, ZF-GFP, and HA-Eme, respectively) (Supplemental Fig. 3B). Thus, targeting of endogenous pericentromeric chromatin to the nuclear periphery after fertilization impairs developmental progression. Whether this developmental phenotype is solely due to a silencing defect of pericentromeric chromatin and/or subsequent defects on kinetochore loading and progression to mitosis remains to be determined.

To address whether tethering the pericentromeric chromatin to the nuclear periphery alters gene expression, we analyzed individual embryos expressing ZF-Eme using a microfluidics Biomark approach, which is a robust and quantitative approach amenable to gene expression analysis from low cell number (Supplemental Fig. 4; Guo et al. 2010). We focused specifically on genes that (1) play a role in early development, (2) are activated zygotically between the two-cell and four-cell stages, or (3) are in close proximity to the centromere. The genes analyzed include housekeeping genes, cell cycle-related genes, transcription factors, chromatin modifiers, signaling proteins, developmentally important genes, and genes located at the proximity of the centromere on four different chromosomes (9, 18, 19, and X) (Supplemental Table 1). For the latter group of genes, we performed an in silico search or analyzed genes previously known to be centromere-proximal by cytogenetics: *Suv39h1*, *Suv420h1*, *Yap1*, *Gata6*, and *Rock1*. We analyzed levels of expression of all 41 genes simultaneously in 10 biological replicates and three technical replicates. We found no significant changes in gene expression among the noninjected embryos, embryos expressing ZF-GFP, and embryos expressing ZF-Eme (Supplemental Figs. 4, 5).

Finally, we asked whether tethering the MajSats to the nuclear periphery leads to a defective accumulation of heterochromatic marks, which could potentially explain their increased transcriptional activity. For this, we analyzed H3K27me3 and Ring1b in embryos expressing ZF-Eme (Puschendorf et al. 2008; Santenard et al. 2010). Immunostaining revealed that most ZF-Eme embryos display abnormal localization of Ring1b compared with controls (Fig. 4A; Supplemental Fig. 6). We determined the colocalization of MajSats with H3K27me3 and Ring1b in 3D using DNA-FISH and immunostaining followed by 3D reconstruction (Fig. 4A,B). Quantification of the colocalized volume of H3K27me3/Ring 1b within the MajSats revealed a decreased accumulation of the two heterochromatic marks on MajSats in ZF-Eme embryos compared with the controls (Fig. 4C). These data suggest that tethering the MajSats to the nuclear periphery after fertilization impairs the efficient recruitment of these silencing marks to pericentromeric chromatin.

The defect in heterochromatin formation that we report suggests that embryonic nuclear organization is a key factor of epigenetic reprogramming and that the distinctive organization of embryonic nuclei has a regulatory role. The observations that efficiency of cloning upon nuclear transfer is associated with the acquisition of an NLB-like organization (Martin et al. 2006b) further highlight the uniqueness of such reorganization and point toward a necessary step of reprogramming heterochromatin to restore totipotency.

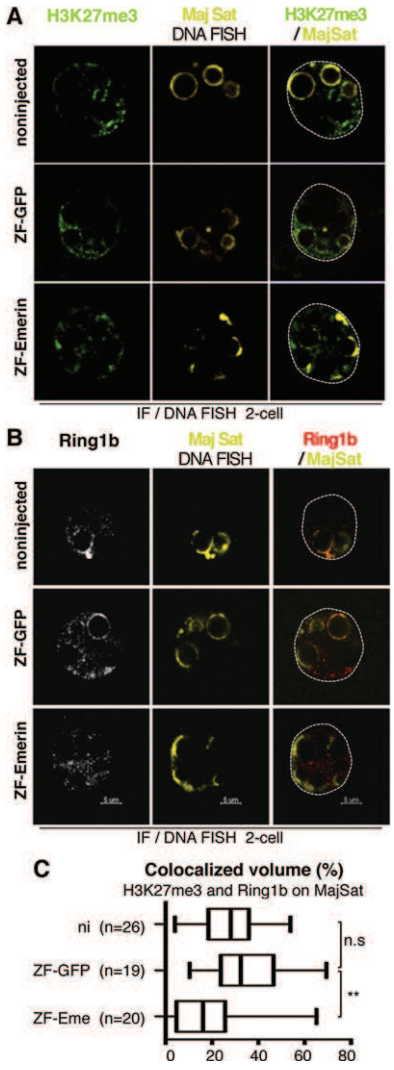


Figure 4. Tethering of pericentromeric chromatin to the nuclear periphery results in decreased accumulation of H3K27me3 and Ring1b. (*A*,*B*) Representative confocal images of immuno-DNA-FISH of early two-cell noninjected, ZF-GFP, and ZF-Eme embryos stained for H3K27me3 (*A*) and Ring1b (*B*). DNA-FISH signal for MajSats is in yellow, and dotted lines delineate the nuclear membrane. Shown is a single confocal section from Z-series acquisitions of representative nuclei (n = 26 embryos for noninjected; n = 19 embryos for ZF-GFP; n = 20 embryos for ZF-Eme). Note that for these experiments, DNA-FISH was performed with oligonucleotide probes. (*C*) Quantification of the volume of MajSats co-occupied by H3K27me3 and Ring1b (in percentage). Whisker plots show the median and the minimum–maximum ranges. (n.s.) Not significant; (**) $P < 0.05$ (ANOVA, Turkey's multiple comparison).

In contrast to the nuclear organization pattern typically found in somatic cells, whereby the most gene-rich chromosomes locate in the center and gene-poor regions are located close to the nuclear periphery (Boyle et al. 2001), studies in bovine embryos revealed that there is no correlation between gene density of chromosome territories and radial positioning prior to the major wave of embryonic genome activation (Koehler et al. 2009). In human cells, relocalization of transgenes to the nuclear periphery alters gene expression (Finlan et al. 2008), but in differentiated cells, the nuclear periphery is believed to be a repressive environment, perhaps through the maintenance of a hypoacetylated chromatin. Our results indicate that the nuclear organization in the early embryo seems to be functionally different from that in differentiated cells.

It remains to be established whether the global dynamics of other chromatin regions within the nuclear space and those of the embryonic chromatin in general further differ from that of somatic cells and whether specific histone modifications would play a role in such regionalization. Our data suggest that the temporal order of events that follow fertilization and the localization of heterochromatin in the 3D nuclear space are tightly regulated and function in parallel to ensure heterochromatic silencing and subsequent development. This adds nuclear reorganization to the molecular cascade of events that dictate establishment of heterochromatin in mammals.

Materials and methods

Embryo collection

For microinjection, zygotes were obtained from superovulated F1 females and microinjected at 17 h post-hCG with in vitro transcribed HA-Eme, ZF-GFP, or Eme-ZF mRNA. Embryos were cultured in KSOM and monitored regularly until fixing for immunostaining, RNA, or DNA-FISH.

FISH

DNA and RNA-FISH were performed with a MajSat probe covering a full repeat (234p) as described (Miyanari and Torres-Padilla 2012), with minor modifications. Embryos were analyzed in drops to preserve 3D information.

Immunostaining

Embryos were fixed ~36–37 h post-hCG as described (Torres-Padilla et al. 2006). Primary antibodies were anti-H3K27me1 (Millipore), anti-H3K27me3 (Millipore), anti-HP1β (IGBMC), anti-Lamin B1 (Abcam), anti-HA (Roche), and anti-RING1b (MBL1).

Acknowledgments

We thank the Institut de Génétique et de Biologie Moléculaire et Cellulaire-Institut Clinique de la Souris imaging facility for support, the sequencing facility for access to the Biomark, B.J. van der Zaal for the PZF:GFO construct, Y. Miyanari for advice on FISH, A. Burton for advice on the Biomark analysis, S. Daujat for advice on ChIP, E. Guccione for support, and A. Bannister for critical reading of the manuscript. M.E.T.-P. acknowledges funding from ANR-09-Blanc-0114, EpiGeneSys Network of Excellence, and European Research Council-Stg "NuclearPotency." J.J. was a recipient of an European Science Foundation Exchange Grant. J.M. acknowledges funding from the Agency for Science, Technology, and Research in E. Guccione's laboratory.

References

Aguirre-Lavin T, Adenot P, Bonnet-Garnier A, Lehmann G, Fleurot R, Boulesteix C, Debey P, Beaujean N. 2012. 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev Biol* **12:** 30.

Akhtar A, Gasser SM. 2007. The nuclear envelope and transcriptional control. *Nat Rev Genet* **8:** 507–517.

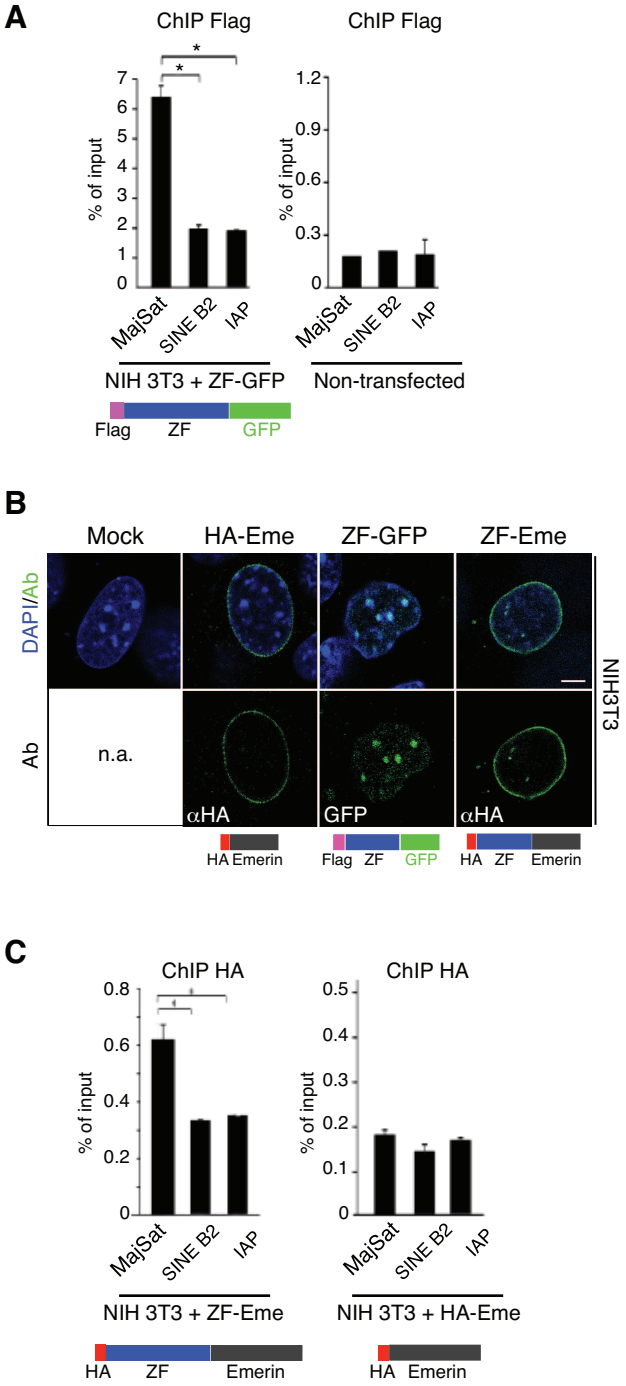
Arney KL, Bao S, Bannister AJ, Kouzarides T, Surani MA. 2002. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int J Dev Biol* **46:** 317–320.

Bengtsson L, Wilson KL. 2004. Multiple and surprising new functions for emerin, a nuclear membrane protein. *Curr Opin Cell Biol* **16:** 73–79.
Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet* **10:** 211–219.
Fadloun A, Eid A, Torres-Padilla ME. 2013. Mechanisms and dynamics of heterochromatin formation during mammalian development: Closed paths and open questions. *Curr Top Dev Biol* **104:** 1–45.
Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR, Bickmore WA. 2008. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet* **4:** e1000039.
Grewal SI, Elgin SC. 2007. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447:** 399–406.
Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. 2010. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* **18:** 675–685.
Koehler D, Zakhartchenko V, Froenicke L, Stone G, Stanyon R, Wolf E, Cremer T, Brero A. 2009. Changes of higher order chromatin arrangements during major genome activation in bovine preimplantation embryos. *Exp Cell Res* **315:** 2053–2063.
Kourmouli N, Jeppesen P, Mahadevhaiah S, Burgoyne P, Wu R, Gilbert DM, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, et al. 2004. Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J Cell Sci* **117:** 2491–2501.
Kumaran RI, Thakar R, Spector DL. 2008. Chromatin dynamics and gene positioning. *Cell* **132:** 929–934.
Lindhout BI, Fransz P, Tessadori F, Meckel T, Hooykaas PJ, van der Zaal BJ. 2007. Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res* **35:** e107.
Martin C, Beaujean N, Brochard V, Audouard C, Zink D, Debey P. 2006a. Genome restructuring in mouse embryos during reprogramming and early development. *Dev Biol* **292:** 317–332.
Martin C, Brochard V, Migne C, Zink D, Debey P, Beaujean N. 2006b. Architectural reorganization of the nuclei upon transfer into oocytes accompanies genome reprogramming. *Mol Reprod Dev* **73:** 1102–1111.
Miyanari Y, Torres-Padilla ME. 2012. Control of ground-state pluripotency by allelic regulation of Nanog. *Nature* **483:** 470–473.
Nestorov P, Tardat M, Peters AH. 2013. H3K9/HP1 and Polycomb: Two key epigenetic silencing pathways for gene regulation and embryo development. *Curr Top Dev Biol* **104:** 243–291.
Probst AV, Almouzni G. 2011. Heterochromatin establishment in the context of genome-wide epigenetic reprogramming. *Trends Genet* **27:** 177–185.
Probst AV, Santos F, Reik W, Almouzni G, Dean W. 2007. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **116:** 403–415.
Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G. 2010. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* **19:** 625–638.
Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U, Kolb C, Otte AP, Koseki H, Orkin SH, et al. 2008. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* **40:** 411–420.
Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME. 2010. Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol* **12:** 853–862.
Santos F, Peters AH, Otte AP, Reik W, Dean W. 2005. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* **280:** 225–236.
Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M. 2006. Dynamic distribution of the histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* **50:** 455–461.

Heterochromatin establishment at pericentromeres depends on nuclear position

Joanna W. Jachowicz, Angèle Santenard, Ambre Bender, et al.

Genes Dev. 2013 27: 2427-2432
Access the most recent version at doi:10.1101/gad.224550.113



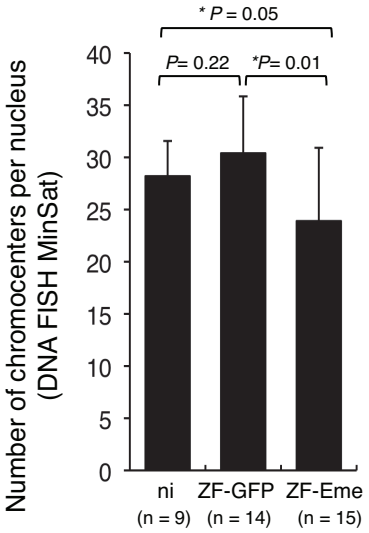
Supplementary Figure S1

Specific binding of the Zn to pericentromeric chromatin is preserved after fusion with Emerin.

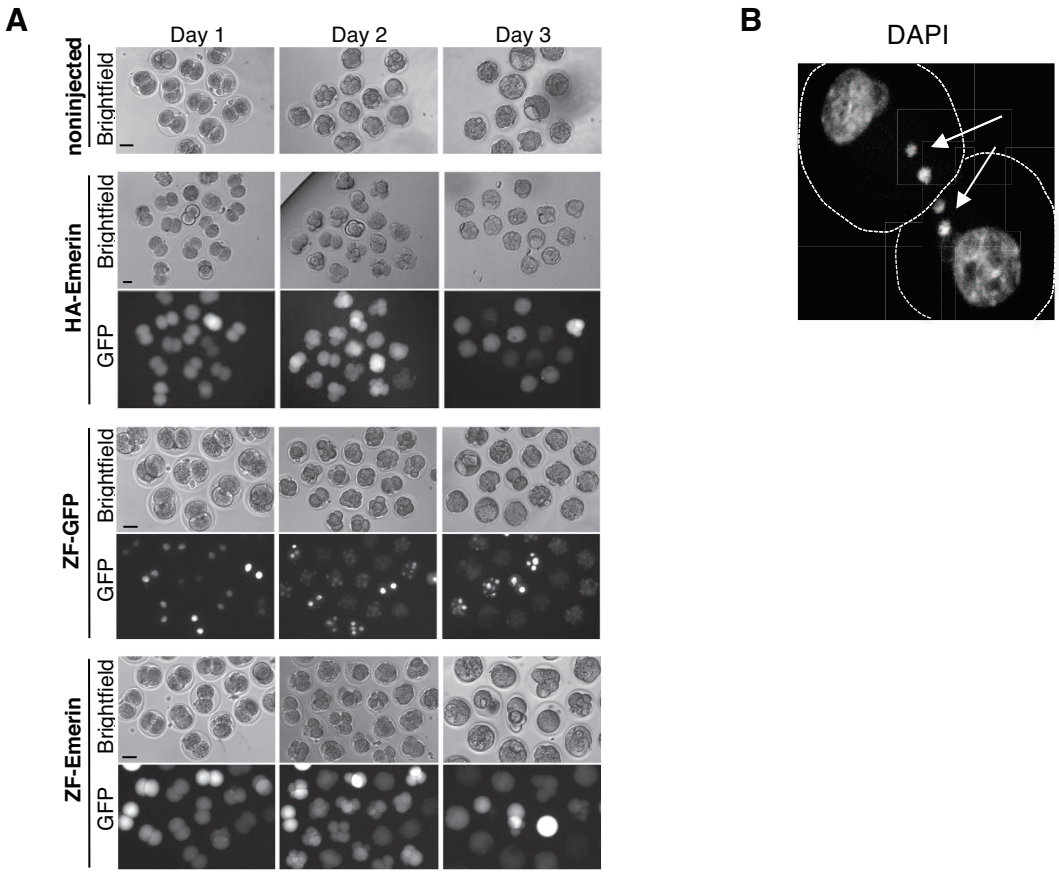
A. ChIP analysis in NIH3T3 cells transiently transfected with the ZF-GFP construct on the indicated genomic sequences. Note that no binding above background levels was detected with the Flag antibody on Sine B2 and IAPs, which were used as negative controls. A Flag immunoprecipitation in non-transfected cells was also included as negative control. A representative experiment of two biological replicates performed in three technical replicates is shown. *$P \leq 0.05$.

B. Immunostaining analysis of NIH3T3 cells expressing the HA-Eme, ZF-GFP and ZF-Eme constructs. NIH3T3 cells were transiently transfected with the indicated constructs and analyzed by immunostaining with an HA antibody or through direct GFP visualization as indicated. Shown are single confocal sections of representative nuclei of >30 cells analyzed for each construct. The tags and domains present in each construct are depicted schematically.

C. Functional binding of the ZF to major satellites upon fusion to Emerin. NIH3T3 cells were transiently transfected with either the ZF-Eme or the HA-Eme construct and analyzed by ChIP using an HA antibody. Note that the emerin alone does not bind to major satellite repeats. A representative experiment of two biological replicates performed in three technical replicates is shown. *$P \leq 0.05$.



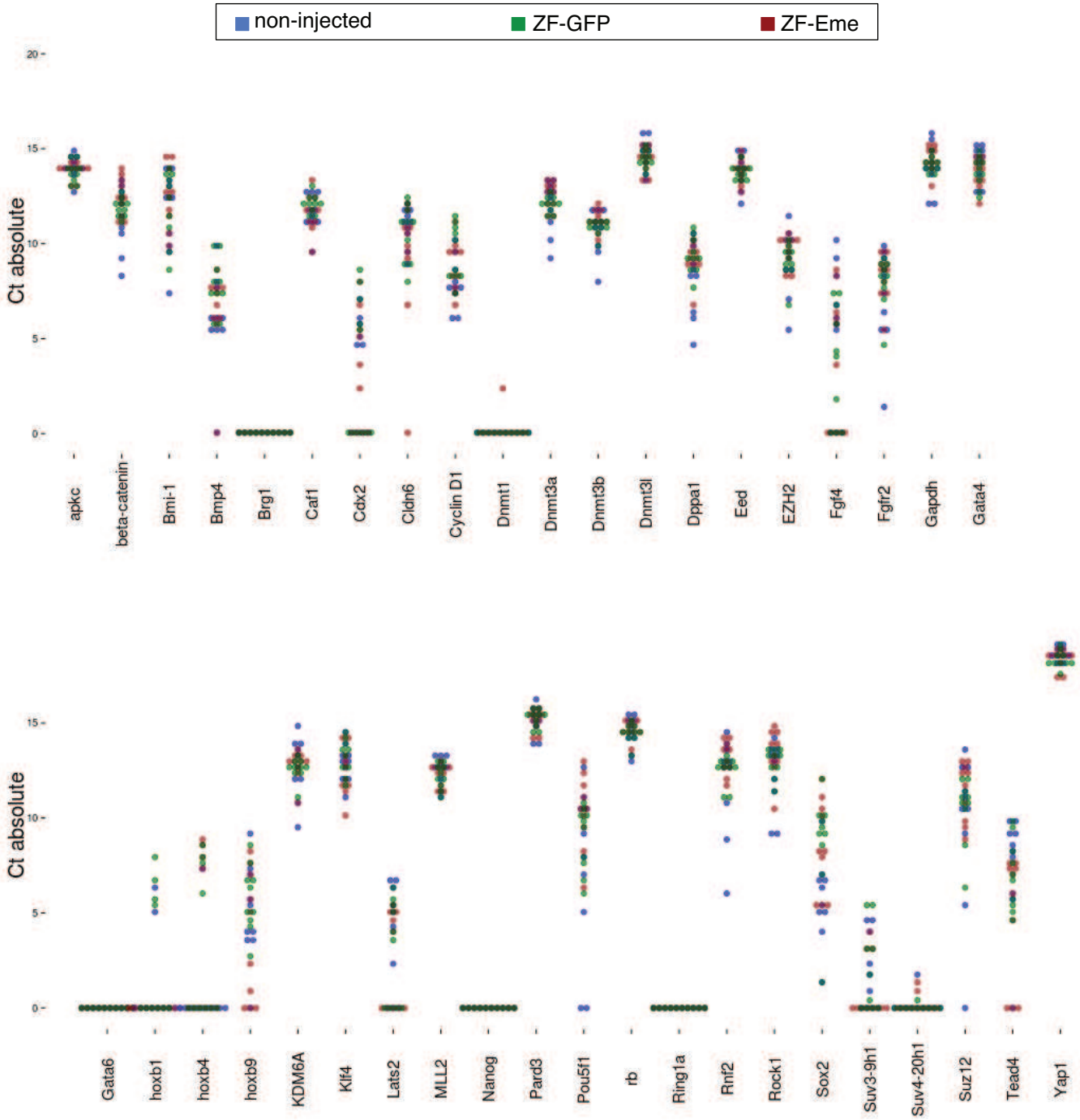
Supplementary Figure S2. Impaired chromocenter formation upon expression of ZF-Emerin. Embryos were microinjected as in Figure 4 and analyzed by DNA-FISH with a minor satellite probe to detect individual chromocenters at the late 2-cell stage. Chromocenters were counted in single nuclei. Note that the ZF-Eme embryos have a reduced number of chromocenters.



Supplementary Figure S3.
Analysis of developmental progression of non-injected embryos or embryos expressing HA-Eme, ZF-GFP or ZF-Eme.

A. Zygotes were microinjected with the indicated mRNAs and monitored daily for developmental progression. Representative brightfield and GFP fluorescence micrographs are shown. Note that co-injection of GFP mRNA was used as a positive control for injection in the HA-Eme and the ZF-Eme groups. Scale bar is 50 μm. The total number of embryos analyzed for each group is shown in Figure 3D.

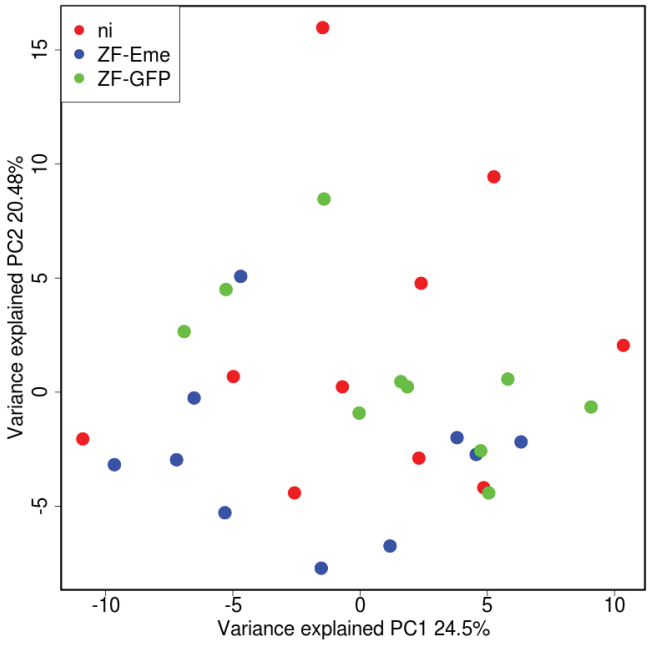
B. Embryos expressing ZF-Eme show a high incidence of chromatin lagging. Representative 2-cell stage embryo expressing ZF-Eme and stained with DAPI. A single confocal section is shown. The dotted line delineates the cell membrane and arrows point to lagging chromatin.



Supplementary Figure S4

Gene expression analysis of non-injected, ZF-GFP and ZF-Eme embryos.

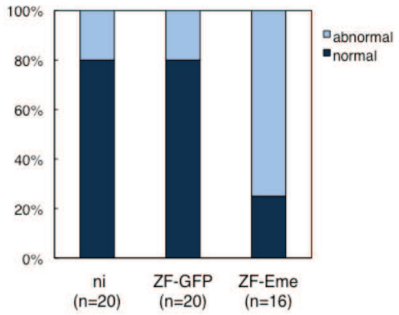
Embryos were microinjected as in Figure 3 and cultured till the early 2-cell stage. Expression levels of the indicated 42 genes was analyzed in single embryos simultaneously using the Biomark Fluidigm system. Each dot is the median of the absolute Ct value of a single embryo obtained from three technical replicates plotted along the y axis. Ten individual embryos were analyzed per experimental group, which are color coded as indicated on the top. A baseline of Ct=28 was used to obtain an absolute expression level. ANOVA analysis revealed no significant difference in gene expression between the three groups of embryos, excepting for *Sox2* and *Cyclin D1* ($p<0.05$), but the subsequent post-hoc tests indicated that these differences were not significant in the paired comparison with the ZF-GFP controls.



Supplementary Figure S5

PCA analysis based on gene expression profiles of non-injected, ZF-GFP and ZF-Eme embryos.

PCA analysis reveals that there is no clustering between the three groups, indicating that there are no global differences in gene expression. Embryos were microinjected as in Figure 3, cultured till the early 2-cell stage and principal component projection of individual cells based upon expression profiles of all 42 analyzed genes was performed. The baseline of Ct=28 was used and the resulting values were normalized to the internal *Gapdh* control by subtracting the corresponding expressing value for each embryo. Each dot represents a single embryo from the non-injected (ni), ZF-GFP or ZF-Eme groups according to the color code indicated.



Supplementary Figure S6

Embryos expressing ZF-Eme display an altered pattern of localization of Ring1b.

Embryos were microinjected as in Figure 3, cultured till the early 2-cell stage, fixed and analyzed by immunostaining with a Ring1b antibody. Shown is the proportion of 2-cell stage embryos showing a normal or abnormal Ring1b localization (qualitatively), where *normal* refers to the typical ring-like distribution of Ring1b with enrichment around the NLBs and *abnormal* refers to a Ring1b pattern that is rather dispersed, of lower intensity and a lack of enrichment around the NLBs (for representative images refer to Figure 4A).

Supplementary Table 1. List of genes analyzed in the Biomark Fluidigm array.

Gene Symbol	ABI Gene Expression Assay ID	Comment
MLL2	Mm02600438_m1	Histone methyltransferase for Histone H3K4
EZH2	Mm00468464_m1	Enhancer of Zest homolog 2, Polycomb Repressive Complex 2
Suz12	Mm01304152_m1	Suppressor of Zeste 12 homolog, Polycomb Repressive Complex 2
Eed	Mm00469651_m1	Embryonic Ectoderm development, Polycomb Repressive Complex 2
Bmi-1	Mm03053308_g1	Polycomb ring finger oncogene, Polycomb Repressive Complex 1
Ring1a	Mm01278940_m1	Histone H2A ubiquitylation, Polycome Repressive Complex 1
Ring1b Rnf2	Mm00803321_m1	Histone H2A ubiquitylation, Polycome Repressive Complex 1
Suv4-20H1	Mm00523065_m1	Centromere proximal location, chr 19
Suv39H1	Mm00468952_m1	Centromere proximal location, chr X
Dnmt1	Mm00599763_m1	Maintenance DNA methyltransferase
Dnmt3a	Mm00432881_m1	De novo DNA methyltransferase
Dnmt3b	Mm01240113_m1	De novo DNA methyltransferase
Dnmt3l	Mm00457635_m1	De novo DNA methylation, cofactor
KDM6A	Mm00801998_m1	Utx, H3K27me2/me3 demethylase
Pou5f1	Mm03053917_g1	Oct4, transcription factors
Nanog	Mm02019550_s1	Transcription factor, zygotically expressed, pluripotency
Sox2	Mm03053810_s1	Transcription factor, ICM maintenance, essential for early development
Gata4	Mm00484689_m1	Transcription factor, specification of primitive endoderm
Fgf4	Mm00438917_m1	Signalling molecule
Cdx2	Mm01212280_m1	Transcription factor, trophectoderm differenciation
Gata6	Mm00802636_m1	Transcription factor, centromere proximal location, chr 18
Bmp4	Mm00432087_m1	TGF signalling, lineage specification
Fgfr2	Mm01269930_m1	Proliferation and differentiation, signalling
Dppa1	Mm00626454_m1	Developmental pluripotency associated 1, zygotically expressed
Tead4	Mm01189836_m1	Hippo pathway (transcription enhancer factor)
klf4	Mm00516104_m1	Transcription factor, pluripotency
pard3	Mm00473929_m1	Partitioning defective 3, cell polarity
apkcλ	Mm00435769_m1	Atypical PKC, cell polarity
caf1	Mm00511230_m1	Chromatin assembly
cldn6	Mm00490040_s1	Claudin 6, tigh junction
beta-catenin	Mm00483039_m1	Signalling, polarity, axis specification, proliferation
hoxb1	Mm00515118_g1	Homeobox b1, transcription factor
hoxb4	Mm00657964_m1	Homeobox b4, transcription factor
hoxb9	Mm01700220_m1	Homeobox b9, transcription factor
yap1	Mm01143263_m1	Centromere proximal location, chr 9
lats2	Mm00497217_m1	Large tumor suppressor 2 kinase, cell division, Hippo pathway
brg1	Mm01151948_m1	ATPase subunit, remodelling complex, required for zygotic genome activation
rock1	Mm00485745_m1	Signalling kinase, Centromere proximal location, chr 18
rb	Mm00485586_m1	Cell cycle, cell proliferation
cyclin D1	Mm00432359_m1	Cell cycle
gapdh	Mm99999915_g1	Housekeeping, internal control

Supplemental Materials and Methods

Embryo collection, culture and microinjection

For microinjection, females F1 (C57BL/6 x CBA/H) were superovulated by I.P. injection of 7.5 international units (IU) pregnant mare's serum gonadotropin (PMSG) followed 46-52h later by I.P injection of 7.5 international units (IU) of human chorionic gonadotropin (hCG) and then crossed with F1 males. Zygotes were collected at the fertilization cone stage about 17h after hCG injection and then microinjected with 1-2pl of 400 ng/μl of in vitro transcribed and capped HA-Eme, ZF-GFP or Eme-ZF mRNA. The emerin constructs spanning N-terminal 256 aminoacids from murine emerin was generated by PCR and fused to the ZF derived from the ZF-GFP construct or cloned directly into the destination vector pRN3P; the ZF-GFP has been described elsewhere (Lindhout et al. 2007). All cDNAs were cloned into the pRN3P plasmid for in vitro transcription with identical 5' and 3'UTRs. mRNAs were transcribed *in vitro* from the pRN3P plasmid with the mMESSAGE mMACHINE Kit (Ambion). After microinjection embryos were cultured in drops of KSOM media under mineral oil at 37˚C under 5% CO_2 and their development was monitored at regular intervals. Embryos were then either cultured until the blastocyst stage (E3.5) or fixed 36-37h post hCG and processed for immunostaing, DNA-FISH or RNA-FISH. For all other experiments, embryos derived from natural matings were used.

Immunostaining

Mouse embryos were fixed approximately 36-37h post hCG as described in (Torres-Padilla et al. 2006). The zona pellucida was removed with acid Tyrode solution (Sigma), followed by three washes in PBS and fixation at 37°C to ensure preservation of nuclear architecture. After permeabilization with 0.5% Triton, embryos were washed three times in PBST, blocked and incubated with the primary antibodies (anti-H3K27me1 (Upstate-Millipore) 1:250; anti-H3K27me3 (Upstate-Millipore) 1:250; anti-HP1β (IGBMC) 1:500; anti-Lamin B1 (Abcam) 1:150; anti-HA (Roche) 1:250); Ring1b (MBL1) 1:200, in 3% BSA 0.1% Tween in PBS. After overnight incubation at 4°C embryos were washed twice in PBS, blocked for 1h and incubated for 2h at RT with the corresponding secondary antibodies in 3% BSA in PBS-T. After washing, embryos were mounted in Vectashield (Vector Laboratories) containing 4'-6-Diamidino-2-phenylindole (DAPI) for visualizing DNA. Confocal microscopy was performed using a 63x oil objective in a Leica TC SP5 inverted microscope. Z-sections were taken every 1 μm. Visualization of the ZF-GFP was done through direct GFP acquisition.

DNA-FISH

Zygotes were microinjected and cultured for around 36h post-hCG after which their zona pellucida was removed with acidic Tyrod's solution (Sigma) and embryos were fixed in 4% paraformaldehyde (PFA) with 0.05% TritonX-100 and 1mg/ml polyvinyl pyrrolidone (PVP) in PBS for 20min at RT. After three washes in PBS, embryos were transferred to permeabilization solution containing 0.5% TritonX-100, 0.02% RNAse A and 1mg/ml PVP in PBS and incubated at 37 ˚C for 1h. Subsequently, embryos were washed three times in PBS and incubated for 2min in 0.1% HCl with 0.7% TritonX-100 and 1mg/ml PVP in water. After washing with PBS three times, embryos were pre-hybridized in 1µl drops of hybridization buffer covered with mineral oil on a 35-mm glass bottom dish at 37 ˚C overnight. The hybridization buffer consisted of 50% formamide, 10% dextran sulphate, 2xSSC, 1 mg/ml polyvinyl pyrrolidone (PVP), 0.05% Triton X-100, 0.5 mg/ml BSA and 50ng of fluorescent probe. Major and minor satellite full-length repeat sequences were used as probes that were firstly cloned to pGEM plasmid and then labeled by PCR reaction with self-made ATTO 647 dATP for major satellites and TAMRA dATPs for minor satellites. Labeled PCR products were then purified with a QIAquick PCR Purification Kit (QIAGEN) and mixed with hybridization buffer. After overnight incubation embryos were kept for 10min at 82 ˚C and then 1h at 37 ˚C after which embryos were transferred to a drop of hybridization mix and incubated overnight at 37 ˚C in humid atmosphere. After hybridization and three washes with 2xSSC, 0.05% Triton-100 and 1mg/ml PVP in PBS at 50˚C for 10 min, embryos were then subjected to a gradient series (20%, 40%, 60%, 80%, 100%) of VECTASHIELD containing DAPI.

Immuno- DNA-FISH

With the exception of the immuno-DNA-FISH with lamin B1, where DNA-FISH was performed prior to immunostaining, in all the other experiments immunostaining was performed first. Immunostaining was performed as described above with an additional post-fixation step after the last wash in PBS. Embryos were then fixed for 45 min, washed three times in PBS and permeabilized for 1h at 37 ˚C in 0.5% TritonX-100, 0.02% RNAse and 1mg/ml PVP in PBS and subsequent DNA-FISH pre-hybridization and washes were performed as above. For the experiments where H3K27me3 and Ring1b co-localization with MajSat was analyzed, oligonucleotide probes were used instead of full-length probes because we noted that otherwise the DNA-FISH procedure with full length probes alters the outcome

of immunostaining (see Results & Discussion). Here, pre-hybridization, hybridization and washes were performed as described above except that instead of full-length MajSat probe, 1 mM short oligonucleotides labeled with Cy5 ([CY5]ttttctcgccatattccaggtc) were used as probes and hybridization mix was prepared with 25% formamide instead of 50%. In all our DNA-FISH experiments there is no amplification step, but probes are directly observed since they are labeled with fluorophores (Cy5 in this case as indicated). To establish the percentage of volume of H3K27me3 and Ring1b colocalized with MajSat we used the Coloc module of Imaris (Bitplane) to quantify the number of voxels in the green and red channels (H3K27me3 and Ring1b, respectively) that localized within the MajSat.

RNA-FISH

RNA FISH and subsequent confocal acquisition were performed as described (Miyanari and Torres-Padilla 2012; Fadloun et al. 2013) and incubation with Cot-1 was omitted. Note that using these RNA-FISH conditions, transcripts analyzed are sensitive to α-amanitin treatment and to RNAse A treatment (Fadloun et al. 2013 and data not shown). Zygotes were microinjected with mRNA, cultured for 36 h post-hCG and then processed for RNA FISH. The Major Satellite probe consisted of the full-length repeat region (234bp) that was labeled with Cy5 and observed directly, without amplification step, allowing for direct quantitation of both signal intensity and volume. Images were acquired on a TCS SP5 confocal microscope with a 63x glycerol immersion objective lens and Z-sections were taken every 1 μm. Embryos were placed in drops for acquisition to preserve 3D. For quantification of the number of transcription sites (TS), images were analyzed with Leica LAF AFlite manually through identification and labelling of each TS throughout all Z-stack. The number of TS per nucleus was determined and five separate experiments were performed. For determining the total volume of transcription output, the IMARIS Bitplane software was used where RNA-FISH signal enabled to build 3D surfaces separately for each nucleus. Three independent experiments were conducted and the same thresholds based on area and intensity were used to reconstruct 3D MajSat transcripts surfaces.

Transfection of NIH3T3 cells and Chromatin immunoprecipitation

For transfection, the mouse PZF:GFP plasmid previously reported was used for the "ZF-GFP" group. The ZF-Eme and Eme-HA plasmids were constructed using the pEGFP-N2

(Clonetech) backbone and replacing the entire GFP coding region with the Eme-ZF-HA or the Eme-HA inserts from the pRN3P plasmids used for mRNA transcription described above, respectively. We calculated the transfection efficiency by FACS analysis (by complementing with the pEGFP-N2 vector where appropriate), that was consistent between the different experimental groups within each experiment. For ChIP, NIH3T3 cells were transfected with the corresponding plasmids and crosslinked for 10 min in 1% formaldehyde in PBS at RT. After washing, scrapping and collection, nuclear pellets were sonicated using a Bioruptor (Diagenode) at maximum intensity with ON/OFF cycles of 30 seconds for 10 minutes. ChIP was performed with equal amounts of DNA-chromatin and chromatin was systematically verified on agarose gel electrophoresis. The equivalent of 25 μg of DNA was used per IP. After washing and elution, samples were incubated in 5 μl of Proteinase K at 20 mg/ml to reverse crosslinking and subjected to phenol/chloroform extraction. PCRs were performed in technical replicates using a Roche Lightcycler as described before (Gazdag et al. 2009) and a representative of at least two biological replicates is shown.

Gene expression analysis using Fluidigm Biomark

Ten individual 2-cell stage embryos from each of the groups analyzed (non-injected, ZF-GFP or ZF-Eme-injected) were washed in PBS at 37 hours post-hCG and flash frozen in liquid nitrogen in 5 μl 2x reaction buffer (CellsDirect™ One-Step qRT-PCR kit, Invitrogen). TaqMan® Gene Expression Assays (20x Applied Biosystems), previously tested using ES cell cDNA for amplification efficiency, were pooled to a final concentration of 0.2x for each of the 43 assays. To each of the single-cell samples in 2x reaction buffer was added 2.5 μl 0.2x assay pool, 0.5 μl RT/Taq enzyme (CellsDirect qRT-PCR kit, Invitrogen) and 2.3 μl of water. Cell lysis and sequence-specific reverse transcription were performed at 50°C for 20 minutes. The reverse transcriptase was inactivated by heating to 95°C for 2 minutes. Sequence-specific pre-amplification was performed by denaturing at 95°C for 15 seconds, then annealing and amplification at 60°C for 4 minutes for 18 cycles. The resulting cDNA was diluted 5-fold before analysis with Universal PCR Master Mix and TaqMan® Gene Expression Assays (Applied Biosystems) in 48:48 Dynamic Arrays on a Biomark™ System (Fluidigm). Ct values were calculated from the system's software (Biomark™ Real-time PCR analysis, Fluidigm). All Raw Ct values were normalized to the assumed detection Ct level of 28 following recommendation from Fluidigm technical support as in Guo et al (2010) and Burton et al (Cell Reports, in press). The list of genes analyzed is depicted in Supplementary

Table 1. All analyses were performed using R with previously established pipelines as described before to generate the dot plot with absolute Ct values and PCA graphical visualization (Guo et al. 2010)(Burton et al., in press). Statistical analyses were performed using ANOVA followed by post-hoc pairwise t-test to assess significant differences in gene expression among the three groups of embryos analyzed.

Supplemental Materials and Methods References

Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P, Torres-Padilla ME. 2013. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol* **20**: 332-338.

Gazdag E, Santenard A, Ziegler-Birling C, Altobelli G, Poch O, Tora L, Torres-Padilla ME. 2009. TBP2 is essential for germ cell development by regulating transcription and chromatin condensation in the oocyte. *Genes Dev* **23**: 2210-2223.

Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. 2010. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* **18**: 675-685.

Lindhout BI, Fransz P, Tessadori F, Meckel T, Hooykaas PJ, van der Zaal BJ. 2007. Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res* **35**: e107.

Miyanari Y, Torres-Padilla ME. 2012. Control of ground-state pluripotency by allelic regulation of Nanog. *Nature* **483**: 470-473.

Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M. 2006. Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* **50**: 455-461.

1.3. Discussion

The general conclusion that can be drawn based on the obtained data, is that nuclear organization is important for the initial formation of heterochromatin on major satellites. Indeed, when their localization is disrupted experimentally, up-regulation of major satellite transcripts is detected, which is accompanied by lower levels of Ring1B and H3K27me3 at these regions.

One of the main questions is, thus, how these phenotypes lead to the developmental arrest. One possibility is that developmental arrests arises as a result of general problems with heterochromatin establishment on major satellites. It is known that in somatic cells pericentric chromatin displays the typical characteristics of heterochromatic regions: presence of repressive marks, silenced state, clustering into chromocenters, late replication. These features are important to enable pericentric regions to serve their main purpose as a platform for kinetochore assembly. Thus, the hypothesis is that during early development pericentric regions have to acquire repressive marks, become silenced, and physically cluster in order to function properly. Most of these characteristics are affected in Em-ZF embryos which suggests that perturbed heterochromatin formation may indeed lead to improper function of kinetochore, potentially explaining the developmental arrest. Results of this part of my Ph.D support this hypothesis, as seen from the presence of increased lagging chromatin in Em-ZF embryos, and provide evidence that nuclear organization is one of the main players in the formation of heterochromatin. In the subsequent chapter I will discuss these ideas in more details providing possible explanations of the phenotypes and how this might be linked with nuclear organization. Moreover, I aim to describe some of the plausible scenarios of how the events leading to heterochromatin formation are orchestrated and what their importance and temporal dynamics are during early stages of development

The presence of major satellite sequences around the NLBs is important for the acquisition of repressive marks, as our data shows that H3K27me3 becomes enriched on major satellites only after they cluster around NLB. Moreover, this mark is decreased on pericentric regions when they are pulled towards the membrane. It is, thus, possible that more open structure of major satellites, when they are spread on the surface of the NLBs, facilitates acquisition of histone modifications and chromatin modifiers. Although, upon removal of the NLBs, global changes of histone modifications have not been observed (S. Ogushi et al. 2008), it was never

assessed if the levels of H3K9me3 or H3K27me3, which should be associated with major satellites, decreased upon such manipulations. It might be worth testing this experimentally to investigate the above hypothesis on NLB function. Lower level of enzymatic unit of PRC1 - Ring1B - in Em-ZF embryos is in support of the idea that the particular localization of major satellites around NLBs plays a role in changes in their chromatin state. Nevertheless, decrease in H3K27me3 and Ring1B is not enough to explain the developmental arrest of Em-ZF embryos because Ezh2$^{m-/z-}$ and Ring1B $^{m-/z-}$ embryos (Terranova et al. 2008) seem to develop properly until blastocyst stage, at least until some extent. It is possible, thus, that other pathways involved in the repression of major satellites, which might compensate Ezh2$^{m-/z-}$ and Ring1B $^{m-/z-}$ phenotypes, might be affected in Em-ZF embryos. For example, binding of BEND3, which is able to recruit PRC2 and PRC1 to pericentromeric regions (Saksouk et al. 2014), or maternally provided Ring1a (Posfai et al. 2012). It has been shown recently that a portion of paternal major satellites remains DNA-methylated until the 2-cell stage (Arand et al. 2015). The function of this methylation is not known but it might be associated with heterochromatin formation on these regions, as lack of methylation in PGCs leads to improper pericentric chromatin organization (Messerschmidt, Knowles, and Solter 2014). Thus, a change in DNA methylation levels may contribute to the phenotype of Em-ZF embryos, if delocalization of pericentric regions from NLBs made them more prone to active demethylation, which remains to be addressed. A crucial role of the NLBs in controlling the state of the major satellite chromatin by acquisition of repressive marks and facilitating chromatin remodeling is also supported by the experiments in which removal of NLBs led to loss of Daxx and up-regulation of major satellites in 2-cell stage embryos (Fulka and Langerova 2014). Although we did not verify the presence and localization of Daxx in Em-ZF embryos, it is likely that Daxx localization was affected. Moreover, in the absence of NLBs, improper decondensation of paternal chromatin has been observed, phenotype which can be rescued by injections of nucleophosmin 2 (Nmp2) – component of NLBs (Inoue et al. 2011). Thus, the NLBs seem to play an important role in the chromatin remodeling after fertilization which is crucial for proper silencing of major satellites.

Another reason why nuclear localization of major satellites around NLBs might be important, concerns the regulation of their transcription. For example, in mES cells Pax3 and Pax9 transcription factors can drive strand specific transcription from pericentric regions,

which is involved in heterochromatin formation, and their depletion leads to up-regulation of major satellite transcripts (Bulut-Karslioglu et al. 2012). The association of transcription factors with major satellites in 1-cell and 2-cell stage embryos has never been assessed but it is possible that binding of transcription factors to major satellites could be affected in Em-ZF embryos and the transcriptional balance between both strands is changed. As the reverse strand is known to play a role during pre-implantation development, but not the forward, and its depletion leads to embryonic arrest (Casanova et al. 2013), it will be important to determine which strand is affected in Em-ZF embryos. Another possibility is that the parent-specific transcription is changed in Em-ZF embryos. Because it has been shown that paternal sets of chromatin transcribes mainly forward transcripts and maternal chromatin transcribes mainly reverse strands (Casanova et al. 2013), this would also lead to the disruption of the global transcriptional pattern of major satellites. Moreover, developmental arrest together with the up-regulation of major satellites from maternal PN has been documented in ATRX mutant embryos (R. De La Fuente, Baumann, and Viveiros 2015) but the up-regulation of major satellites in Ring1B $^{m-/z+}$ embryos (Puschendorf et al. 2008) has no effect on early development, implying that parent-specific transcripts may have different influence on developmental progression. We did not analyze whether the higher levels of major satellites transcripts in Em-ZF are of paternal or maternal origin, but both scenarios are plausible. Because in the absence of NLBs it is the paternal chromatin that displays decreased transcription (Fulka and Langerova 2014), it is possible that the up-regulation observed in Em-ZF embryos is from the maternal part of the genome. On the other hand, Ring1B has been used as a marker of paternal chromatin in several studies (A. V. Probst et al. 2010; Casanova et al. 2013; Puschendorf et al. 2008) suggesting that it preferentially binds to paternal set of chromatin at zygote and 2-cell stage. Thus, its loss in Em-ZF might cause the up-regulation of paternal major satellites, as shown in Ring1B deficient embryos (Puschendorf et al. 2008).

Several studies suggest that major satellite transcripts or transcription *per se* are involved in the early events of development (A. V. Probst et al. 2010; Casanova et al. 2013). Disruption of this function by pulling pericentric regions towards the membrane might be a cause of the arrest of Em-ZF embryos. However, it is not clear what the exact function of these transcripts is and whether and how they can affect chromatin and development. One of the possible scenarios is through RNAi related pathways which may play a role in heterochromatinization

as it is known for yeast (Grewal and Elgin 2007) and vertebrate cells (Fukagawa et al. 2004; Kanellopoulou 2005). The main arguments supporting such a hypothesis is that disruption of RNAi pathway by Dicer mutation in mES cells leads to accumulation of major satellite transcripts, abnormal mitotic cells and defects in the acquisition of heterochromatic marks (Kanellopoulou 2005), situation strikingly similar to the one observed in Em-ZF embryos. Moreover, rescue experiments by dsRNA in H3.3K27R mutant embryos (Santenard et al. 2010) are also along these lines. On the other hand, the potential presence of small RNAs of the size of 19-21 nt, typical of the RNAi pathway, was never reported, and depletion of only reverse but not forward strand cause developmental arrest (Casanova et al. 2013), suggesting that major satellite transcripts have some other function than RNAi, during this time. For example, it is well documented that an 'RNA component' is necessary for binding of some factors to pericentric regions (Maison et al. 2002), and the existence of long-non coding transcripts from major satellites, which can interact with SUMOylated HP1alpha, has been reported (Maison et al. 2011). These lncRNA might, thus, be involved in the formation of a proper heterochromatin structure on pericentromers during preimplantation development serving as a scaffolding or recruiting platform. Because the dsRNA injections mentioned above that rescued the developmental phenotype of the H3.3K27R embryos, were done with the full-length repeat of major satellite (Santenard et al. 2010), it is possible that one of the strands was responsible for the re-deposition of HP1β in the H3.3K27R embryos.

As discussed in the above paragraphs, removal of major satellites from around NLBs has many putative consequences on the global regulation of pericentromeric regions, which can be the cause of the arrest of Em-ZF embryos. The most likely hypothesis is, however, that the described deregulation of the formation of the heterochromatin on major satellites affects kinetochore assembly, which leads to abnormal mitosis and arrest. Em-ZF embryos show disruption of the chromocenter formation or clustering and lagging chromosomes, which supports indeed the above idea. Previous studies have already shown that upon NLB removal, the centromeric heterochromatin is disorganized together with a delay in mitosis (Sugako Ogushi and Saitou 2010). Moreover, similar mitotic defects including cleavage delay and subsequent arrest at 2-cell stage, were observed in Nmp2 -/- embryos lacking NLBs (Burns et al. 2003). Thus, the presence of the pericentric regions around NLBs seems to be related to kinetochore function and cell division. How exactly localization of major satellites around NLBs

and kinetochore assembly are linked, remains an open question. It has been already shown that centromere transcripts are important for mitotic kinetochore function during mitosis. The evidence of centromere RNA association with CENP-C and non-enzymatic partners of the CPC (Wong et al. 2007; Erhardt et al. 2008) suggests that this RNA could act as a molecular scaffold in the recruitment and organization of key centromere proteins. Thus, the pericentric RNA could function to recruit centromere/kinetochore proteins, and to stabilize the overall kinetochore structure during the G2/M-phase of the first divisions when normal organization of pericentromers is still not present. A change in the level of transcripts could potentially affect the structural integrity of the kinetochore, which in turn could further disturb the spatial regulation of CPC, including Aurora B and/or other mitotic checkpoint signaling protein activities at the kinetochore–microtubule interface and led to developmental delay and arrest.

Although the localization of major satellites around NLBs arises as a crucial regulator of heterochromatin formation in preimplantation embryos and its disruption seems to be a cause of developmental failure, it is worth mentioning that the functionality of the nuclear periphery might be also involved in that arrest. Because in the performed experiments major satellites were pulled towards the nuclear membrane, it is possible that this nuclear compartment imposes a change in the chromatin state of pericentromeres. Gene expression analysis after the delocalization towards the nuclear membrane revealed no significant changes in the expression of the regions in proximity to major satellites. It is possible, however, that single copy genes are regulated in a different manner than repeats, which would explain the up-regulation of major satellites with no change in expression of neighbouring genes. To answer this issue, it will be important to address whether the nuclear periphery itself is a compartment permissive for transcriptional activity in the early embryo in general, or whether this is a specificity of major satellites. More thorough experiments need to be performed including relocalization of other repetitive regions and single copy gene towards the membrane, or even pulling specific regions towards the NLB and monitoring levels of their expression.

In conclusion, the work I presented in this part shed a light on the mechanisms responsible for the formation of heterochromatin of pericentric regions during development and revealed that nuclear organization is a crucial factor in this process.

Part II - L1 transposable elements

2.1. L1 biology during early mouse development

2.1.1. Description of the transcription pattern of L1spa elements from fertilization until morula stage

Previous work in the Torres-Padilla lab revealed the particular pattern of nascent L1 transcription in preimplantation embryos, with a peak in expression at the 2-cell stage and decrease at 8-cell stage (Fadloun et al. 2013). To verify if I can observe the same up- and down-regulation, I performed RNA-FISH experiments in which nascent transcripts of L1 were visualized by the use of Tf member L1spa-specific probes (Fig. 13A). I confirmed that there is low transcription at zygotic stage, up-regulation at 2-cell and subsequent down-regulation at 8-cell stage (Fig. 13B).

Then, I asked the question whether the visualized transcripts are mostly full-length or maybe distinct parts of the L1 element produce their own, shorter transcripts, and whether this may change during the 2- to 8-cell stage transition. As the RNA-FISH was performed with probes specific for three distinct regions: Tf monomers, Orf1, and Orf2, I could visualize each of them separately (Fig. 13A). Experiments revealed that, indeed, some of the transcripts are most likely full length as the signal from Tf, Orf1, and Orf2 probes colocalized (Fig 13B). However, most of the RNA-FISH signal derives from Orf1 and Orf2, with the slight bias towards Orf2, suggesting that there are sites of Orf2 transcription independent of Tf or Orf1 expression. This is not surprising considering the existence of truncated versions of L1 within the genome. Interestingly, all three regions followed a similar pattern of decrease in their expression during developmental time, suggesting that full and short transcripts are not specific just to one developmental stage. To study in more detail the transcription of the monomer, RT-qPCR was performed and expression levels of three different L1 families: Tf, Af, and Gf, were compared between 2- and 8-cell stage embryos. The specificity of the PCR reaction was verified by sequencing the amplified products. Because primers were specific for monomer repeats only, I assumed that they targeted full length-elements mostly, not the truncated version. The presence of the transcripts derived from all three families was confirmed, together with the down-regulation from 2- to 8-cell stage for the Tf family (Fig. 13C).

2.1.2. Protein expression

Next, I asked the question if the observed nascent transcription is also represented at the protein level. As it is known that L1 contains two open reading frames which encode: Orf1p from the first and endonuclease (EN) and reverse transcriptase (RT) from the second, I decided to visualize Orf1p as there is no available mouse antibody specific for L1-EN or L1-RT. Immunostaining revealed that Orf1p is abundant during preimplantation mouse development and it localizes within cytoplasm in what resembles ribonucleoprotein particles (RNPs) (Fig. 13D) described in mES cells. The highest signal for Orf1p was observed in 2-cell stage embryos with a gradual decrease by the 8-cell concomitant with lower number of RNPs. Nevertheless, to properly assess the levels of Orf1p expression I am in the process of collecting embryos for Western blot, as quantification of the fluorescent intensity of a rather broad signal from the cytoplasm is often inconclusive.

All the above experiments indicate that L1 elements are highly transcribed and expressed at 2-cell stage embryos with a down-regulation at both transcriptional and, most likely, translational level when development progresses. Previous experiments in which L1 activity was depleted by morpholino injections into zygotes, suggested that L1 transcripts indeed play an important role as these embryos arrest before blastocyst formation (Beraldi et al. 2006). What is the function of this activity is still unclear. Also, whether this phenotype is due to translational repression, transcript degradation, or to a potential function of L1s on the chromatin in cis, has not been addressed. Moreover, whether the observed decrease in L1 transcriptional activity in development is important for developmental progression has never been addressed experimentally. Likewise, whether the expression of L1 elements during this short time window plays any role for the proper development, heterochromatin formation, global genome activation, cell fate, or gene expression, remains unknown. To address this questions, I have decided to use an experimental approach to directly manipulate the expression of L1 in early embryos. The approach I undertook is based on the Transcription Activator–like Effectors (TALEs) from *Xanthomonas sp.* which are site-specific DNA-binding proteins (Zhang et al. 2011). More precisely, I aimed to target L1 elements with specific TALEs fused with the VP64 activating domain (Sadowski et al. 1988) and analyzed the consequences that their activation may have on embryo development.

2.2. Establishment of tools and experimental conditions

2.2.1. TALEs design and verification in human and mouse cells

TALEs are natural bacterial effector proteins used by *Xanthomonas sp.* to modulate gene transcription in host plants to facilitate bacterial colonization. The central region of the protein contains highly conserved tandem repeats (monomers) which are necessary for DNA binding and recognition. Thanks to their repeat variable di-residues (RDV) which are at the 12th and 13th sites, each monomer can bind to the specific nucleotide within target DNA. Thus, each RDV-containing repeat targets one nucleotide and the linear sequence of monomers in a TALE specifies the target DNA sequence in the 5' to 3' orientation. As a consequence, TALEs can be designed in a way that they can be specific for almost any target site within a genome (Sanjana et al. 2012). During my Ph.D, I took an advantage of this feature and decided to use TALEs approach to modulate L1 expression. First, based on the RNA-Cage data available in the lab (Fadloun et al. 2013) and guidelines from the TAL effector Nucleotide Targeter (TALE-NT) 2.0 website (https://tale-nt.cac.cornell.edu/about), I chose regions within the L1 element that would be the best candidates for the targeting (Fig. 14B). As 5' monomers are known to play a promoter role, they were a first obvious choice. Moreover, the most conserved regions within Orf1 and Orf2, which aligned in the proximity to TSS from the CAGE analysis, where also taken into consideration. The list of TALEs and their respective target sequence was then checked for the number of target sites within a mouse genome using the Galaxy platform tools to assess the specificity for L1 elements binding *in silico* (Fig. 14A). Only TALEs that passed that *in silico* screen were constructed using a modified Golden Gate cloning system described in (Miyanari 2014). To verify if the remaining nine TALEs indeed had high affinity to their target sites, luciferase assays were performed (Fig. 14C). Each TALE was fused with VP64 activator and then co-transfected with Firefly reporter plasmids containing the target sequences of TALEs and minimal CMV promoter (Miyanari 2014). All the transfections were performed in human HEK293T to avoid off-target effects and TALEs binding to the host genome. Comparison of luciferase expression between cells expressing TALEs containing activating domains and controls without VP64 domain revealed that eight out of nine designs enabled up-regulation of Firefly expression, thus, presented satisfying binding affinity to their targets. Next, the localization and expression of the constructs was assessed in mouse ES cells by immunostaining with FLAG antibody specifically recognizing FLAG-tag presented on the 3' of TALEs constructs. Images revealing the expression and localization of TALEs proteins (Fig. 14E)

were compared globally with the nuclear localization of L1 visualized by DNA-FISH with an L1spa probe (Fig. 14D). All constructs were found to be expressed efficiently, however, only four out of eight TALEs showed clear nuclear localization. Others were present also within the nucleolus, which in our experience is a sign of either too high expression or not proper binding to genomic targets. Thus, TALEs Tf1, Orf1.1, Orf1.3, and Orf2.4, were used for the subsequent experiments in the mouse embryos.

2.2.2. Establishment of experimental conditions for TALE expression in mouse embryos

The main aim of the experimental design was to up-regulate or prolong the high activity from L1 elements during early stages of development, specifically beyond the 2-cell stage, by use of TALE effectors fused with VP64 activating domain. To reach that point, however, conditions for the expression of all TALE constructs in mouse embryos had to be established, to ensure that TALE proteins were properly expressed and localized at the specific stages of development. Moreover, I had to verify that microinjection of these DNA-binding proteins into the embryo *per se* had no global effect on developmental progression. In addition, I aimed to optimize conditions to obtain the maximum transcriptional effect possible and target as many L1 elements as possible. Thus, I spent a significant amount of my Ph.D establishing the most suitable experimental design. As a first step, four TALEs sequences were cloned into pRN3P vectors commonly used for the *in vitro* transcription to produce mRNA for mouse embryo microinjections. Two separate backbone vectors were used – one containing VP64 domain, and the other one with multi-cloning site instead, as a negative control. Thus, I could make mRNA for each of the four TALEs in two flavors: TALE-VP64 and TALE-Ctrl. For the following experiments only TALE-Ctrls vectors were used as I intended to test that conditions for these experiments are not detrimental for developmental progression prior to addressing the scientific question (Fig. 15).

Firstly, I was interested in addressing what is the best stage of development to start TALEs expression in a way that the protein would be efficiently expressed and localize in the nucleus. Because the aim of the project was to look into the expression of L1 elements after the peak in transcription at 2-cell stage, two time-points of microinjections were tested - at 1-cell stage and at 2-cell stage - after which I performed immunostainings with HA antibody recognizing HA-tag on the 5' of the pRN3P vector. These experiments revealed that all four tested TALEs were present in the 4-cell stage embryos after microinjection into the zygotes or 2-cell stage

embryos, however, strong signal at the 8-cell stage was observed only after microinjections into 2-cell stage embryos (Fig. 15A&B). Next, I intended to see if embryos expressing DNA-binding protein – TALE – which targets specific regions of L1 element, progress through early development properly. To address this, I again performed microinjections into 1-cell and 2-cell stage embryos, but this time testing two mRNA concentrations and scoring for proper blastocyst formation after four days of culture (Fig. 15C). The use of single target site for TALE binding (TALE Orf1.1 or Orf2.4) gave satisfactory results with most embryos reaching blastocyst stage regardless of the injection time and mRNA concentration. Because I wanted to manipulate both ORF1 and ORF2 at the same time and to increase the number of target sites, I decided to validate developmental progression when more than one TALE was expressed. Thus, the combinations of two TALEs (Orf1.1 and Orf2.4) or three TALEs (Tf1, Orf1.3 and Orf2.4) were used. The first variant seemed to have no undesired effect on development as most embryos reached the blastocyst stage when injected at the 2-cell stage with 200 ng/µl of mRNA of each TALEs. The second group, however, did not develop properly when injected at the zygote stage but proceeded through development normally with the 2-cell stage injections of 200 ng/µl mRNA. Based on these observations, the combination of two or three mRNA TALEs at the concentration of 200 ng/µl each and microinjected into blastomeres of 2-cell embryos, was chosen for further experimental conditions. As the number of target sites was much higher in the second variant, a mixture of TALEs binding to Tf1, Orf1.3 and Orf2.4 was given priority in subsequent experiments. Nevertheless, before addressing the main question of the project, the last step of validation was performed in which localization of three TALEs proteins was compared with the nuclear localization of L1 genomic sequences visualized by DNA-FISH (Fig. 15D&E). To be able to look into the TALEs DNA-bound fraction only and to exclude the free proteins present in the nucleoplasm, pre-extraction prior to fixation was done (Hajkova et al. 2010). Indeed, lower intensity of the signal from TALEs proteins was observed in the chromatin-bound fraction that in the total fraction (Fig. 15E) but they both localized in the nucleus and presented similar pattern to the one observed with DNA-FISH probes (Fig. 15D). Interestingly, TALEs localization within a nucleus seemed to be more similar to the monomers than ORF2 regions which may suggest that not all L1 elements were targeted. However, it is likely that a single event of TALE binding to its target cannot be visualized by the immunostaining and only several TALEs together can be easily identified in the images, hence, repetitive monomer regions appear more visible. Regardless, these experiments suggest that

the global distribution of TALEs in 2-cell embryos resembles that of the endogenous L1s as determined by DNA-FISH.

2.3. Role of the repression of L1 elements during preimplantation development of mouse embryos

2.3.1. Artificial tethering of the L1 TALE-VP64 activators and their effect on the level of L1 transcript at 4- and 8-cell stage mouse embryos

To address the main objective of my Ph.D I used the TALE-VP64 constructs described above and assessed their ability to up-regulate L1 elements in the mouse embryos. Based on the established conditions (Fig 15), the combination of three TALE-VP64 was microinjected into 2-cell stage embryos in order to prolong L1 expression. Additionally, three control groups were used - non-injected embryos, embryos injected with mRNA coding for GFP, and embryos injected with mRNA coding for TALE-Ctrl (Fig. 16A). Each of these controls serves a different purpose: non-injected embryos give the overall information about every set of isolated embryos together with the culture conditions of the experiment; GFP embryos are a control of the microinjection conditions and are of high importance because microinjection tends to lead to the slight developmental delay (~2-4h) which has no consequence on proper development but may be crucial when transcription data is compared; TALE-Ctrl embryos are the main control for the TALEs binding and their influence on the targeted regions, thus, providing necessary information if any conclusion based on the VP64 activity wants to be drawn.

After microinjections the effect of the TALE-VP64 on the levels of transcription was analyzed in late 4-cell stage (Fig. 16) and mid 8-cell stage (Fig. 17) embryos (~64h and ~72h after hCG). First, to visualize nascent transcripts from L1 elements, RNA-FISH was performed in 4-cell stage embryos (Fig. 16B) with the probes that I described above (Fig. 13A). All three control groups showed moderate level of L1 expression with the highest signal coming from the probe binding to Orf2 and lowest from Tf, as shown before (Fig. 13B). TALE-VP64 injected embryos, on the other hand, displayed higher levels of signal from all three probes in comparison with controls. The unspecific binding of RNA-FISH probes to i.e. DNA, was ruled out by performing similar RNA-FISH experiments on the RNaseA treated embryos (Fig 16B). No signal from the FISH probes was observed, suggesting that signal derived from the FISH procedure reflect RNA transcripts. Then, the FISH signal was quantified for Tf, Orf1 and Orf2

separately, normalized to the nuclear volume, and plotted for each of the microinjected group (Fig 16C). Because, as explained above, the microinjection procedure generates a slight developmental delay, we expected overall L1 transcripts from non-injected embryos to be lower than in both injected controls. Therefore, GFP and TALE-Ctrl were chosen as more appropriate controls for the analysis of expression levels. Quantifications revealed an up-regulation of L1 element in TALE-VP64 group in comparison to controls and regardless of the probe used. Next, to verify the effect of TALE-VP64 injection on the level of L1 transcripts, RT-qPCR was performed with Taqman probes specific for three L1 families, the most conserved part of the Orf2 region, and Gapdh, as a negative control (Fig. 16D). Expression analysis showed that Tf family which was the main target of three used TALE designs, was indeed up-regulated. Similarly, transcripts containing a conserved stretch of the Orf2 region, were more highly expressed in TALE-VP64 than in controls. Although Gf monomers, which share more similarities with Tf than Af, also seemed to be slightly up-regulated, they expression is in general much lower at this stages (Fig. 13C) which may explain such a small difference between the control groups. Af monomers, on the other hand, which are highly transcribed throughout the entire early development (Fig. 13C), are also affected by the TALE-VP64 injections. Lack of the difference between experimental and non-injected group in case of Af and Gf might be explained again by the slight delay of microinjected embryos but also by the fact that they were not the main target of the three used TALEs. Moreover, there was high variability across the control groups for the Af analyses, suggesting that rather than a specific effect of our TALE manipulation, changes observed for this family are a result of biological variability. Thus, in conclusion our experimental approach using TALE-VP64 positively regulates L1 transcription in 4-cell stage embryos.

To verify if the up-regulation of L1 elements can be prolonged beyond the 4-cell stage, RNA-FISH experiments were performed in 8-cell stage embryos (Fig. 17). Similar patterns of expression to previously shown 4-cell stage embryos, were observed with higher signal coming from the probes targeting Tf, Orf1 and Orf2 regions of L1 transcripts (Fig. 17B). Image analysis also revealed that TALE-VP64 injected embryos displayed up-regulation of L1 with the clear difference at the Tf and Orf2 regions and less pronounced at the Orf1 region (Fig. 17C). Thus, the effect of the TALE-VP64 targeting on the up-regulation of L1 transcription is sustained at the 8-cell stage, albeit to a lesser extent than in 4-cell stage embryos. Thus, these

experimental conditions allowed us to address the consequences for the sustained activation of L1s on early development.

2.3.2. The consequences of the L1 activation on the preimplantation development of mouse embryos

Because the RNA-FISH data (Fig. 16 & Fig. 17) showed that TALE-VP64 injection in 2-cell stage embryos led to the up-regulation of L1 transcription, the next step was to test the effect of that manipulation on mouse development. To address it, three TALE-VP64 proteins were expressed in mouse embryos from the 2-cell stage and embryos were cultured for three additional days after which blastocyst ratio was scored. Similarly to previous experiments, three control groups were used (Fig. 18A). The percentage of embryos which reached the blastocyst stage within the three control groups showed no difference (as compared by Z test for the distribution) varying between 80 % and 90 % (Fig. 18B). Interestingly, only 50 % of TALE-VP64 injected embryos developed timely and formed blastocysts (Fig. 18B&C), which indicates that the up-regulation of L1 elements has a detrimental effect for developmental progression. To monitor in more detail the phenotype of TALE-VP64 embryos, all four groups were fixed at E4 and stained with an antibody for Cdx2, a transcription factor which is a marker of trophoectoderm (Fig. 18D). The control blastocysts showed normal morphology with correct pattern of Cdx2 expression in trophectoderm cells, whereas TALE-VP64 arrested embryos had various defects including lower cell number, absence of blastocaelic cavity, or unequal divisions between blastomeres. Albeit lower, Cdx2 expression was observed in most of them, suggesting that they do initiate the expression of lineage specific markers.

Next, to address whether the above described developmental arrest is not due to the specificity of this particular TALE-VP64 combination but, indeed, results from L1 up-regulation, similar experiments were performed with the second combination of TALEs mixture (Fig. 15C). A combination of two TALE-VP64 constructs (Orf1.1 and Orf2.4) was expressed in mouse embryos from 2-cell stage using non-injected and TALE-Ctrl (Orf1.1 and Orf2.4) injected embryos as controls (Fig. 19A). First, transcription levels were assessed at 4- and 8-cell stage embryos by quantification of images acquired after RNA-FISH was done with the probe specific for the entire L1spa element (Fig. 19B). Up-regulation of L1 was observed in both stages, with a clear difference in comparison to controls at the 4-cell, and smaller but reproducible difference at the 8-cell stage (only when compared with TALE-Ctrl group). Next, experiments

to assess developmental progression were performed (Fig. 19C&D) which revealed that most of the TALE-VP64 embryos arrested before reaching blastocyst stage whereas both controls progressed normally through development. Altogether, these data suggest that the up-regulation of L1 elements by TALE-VP64 binding leads to developmental arrest before the blastocyst formation.

2.3.3. Dissecting the plausible cause of the developmental arrest

To address potential endogenous molecular mechanisms engaged in the prolongation of the L1 expression at TALE-VP64, I decided to check the level of H3K4me3 which has been previously associated with L1 elements during preimplantation development (Fadloun et al. 2013). I reasoned that this active mark might be abundant on L1 for a longer time than in endogenous situation which, as a consequence, may lead to higher expression of L1 or even spread on neighbouring genes that might result in global higher levels of H3K4me3. Images of 4- and 8-cell embryos stained for H3K4me3 showed no difference in the abundance of H3K4me3 between TALE-VP64 and control embryos globally (Fig. 20A&B), suggesting that L1 upregulation might not be linked to higher levels of H3K4me3. This implies that either putative changes in H3K4me3 on L1s is not pronounced enough to be visualized by immunostaining or that some other active mark may accompany up regulation of L1s, i.e. H3K36me3. I am currently in the process of addressing whether alternative histone modifications are changed upon TALE-VP64 expression.

Next, because L1 elements are translated in early mouse embryos with high expression of Orf1p (Fig. 13D), I asked the question if the up-regulation of L1 leads to change in their proteins level. Immunostaining with the Orf1p antibody revealed that embryos injected with TALE-VP64 at 2-cell stage and fixed at either 4- or 8-cell stage display higher expression from Orf1p (Fig. 20A&C). However, one has to keep in mind that the localization of Orf1p is cytoplasmic, thus, the quantification of its total amount based on images is not very accurate. Therefore, ideally, Orf1p levels should also be analyzed by western blot.

Nevertheless, based on the intensity of the signal from Orf1p antibody, we conclude that the protein levels are higher in TALE-VP64 embryos. Although we do not have an antibody to test this, the observations that Orf1p translation is increased suggests that also the two other L1 proteins – RT and EN - might be expressed in higher degree, potentially leading to the increase of double strand breaks. To answer this, I analyzed the levels of γH2A.X, a well-known

marker of DSBs, and verified its presence in the nuclei of TALE-VP64 injected embryos and two controls (Fig. 20A&D). Immunostaining of 4- and 8-cell stage embryos did not reveal any pronounced difference in the γH2A.X abundance between the three groups suggesting that either DSBs do not occur more often or that EN is not expressed at the higher level when L1s are up-regulated. On the other hand, γH2A.X is very abundant during this time-point of development, which is not always correlated with DSBs (Ziegler-Birling et al. 2009), thus, the developmental arrest might be still a result of higher genome instability caused by L1 overexpression

To address whether higher L1 expression might result in higher genome stability I aimed to verify hypothesis that up-regulation of L1 and higher expression of Orf1p, which is required for the retrotransposition (S. L. Martin 2006), leads to increased number of "jumping" events and this is the cause of the developmental arrest. Previous reports showed that azidothymidine (AZT or Zidovudine) is a nucleoside analog inhibitor of RT which can block retrotransposition of L1 elements *in vitro* and *in vivo* (Dai, Huang, and Boeke 2011; Malki et al. 2014). Although it has been reported that AZT treatment in high concentration has lethal effect on mouse embryo development and affect pre-implantation stages, experiments when 2-cell embryos were exposed to low-concentration (1 µM) AZT indicate that these embryos could proceed to the blastocyst stage (Toltzis, Mourton, and Magnuson 1993). Thus, I first verified if the presence of 1 µM AZT in the culture media has an effect on mouse pre-implantation development (Fig. 21A). Embryos cultured in the AZT-free medium from 1-cell or late 2-cell stage reached the blastocyst stage in an expected ratio, whereas embryos exposed to 1 µM AZT from the zygotic stage arrested before blastocyst formation. Interestingly, 2-cell stage embryos cultured under similar conditions of 1 µM AZT exposure, did not display a delay in development. Hence, I concluded that low concentrations of the RT inhibitor AZT had no morphological effect of preimplantation embryo development when applied from the late 2-cell stage. Next, I used this information to test if developmental arrest of TALE-VP64 injected embryos can be rescued by inhibiting L1-RT activity, thus, blocking potential retrotransposition (Fig. 21B). Non-injected embryos cultured in AZT-free conditions or with the addition of 1 µM AZT from the late 2-cell stage, developed normally, similarly to TALE-Ctrl injected embryos (Fig. 21C&D) whereas a significant percent of TALE-VP64 embryos arrested before reaching blastocyst stage as observed in previous experiments (Fig. 18B). This suggests

that the developmental phenotype elicited upon L1 upregulation using TALE-VP64, is not due to higher activity of RT or retrotransposition.

2.4. Figures

List of figures – Part II

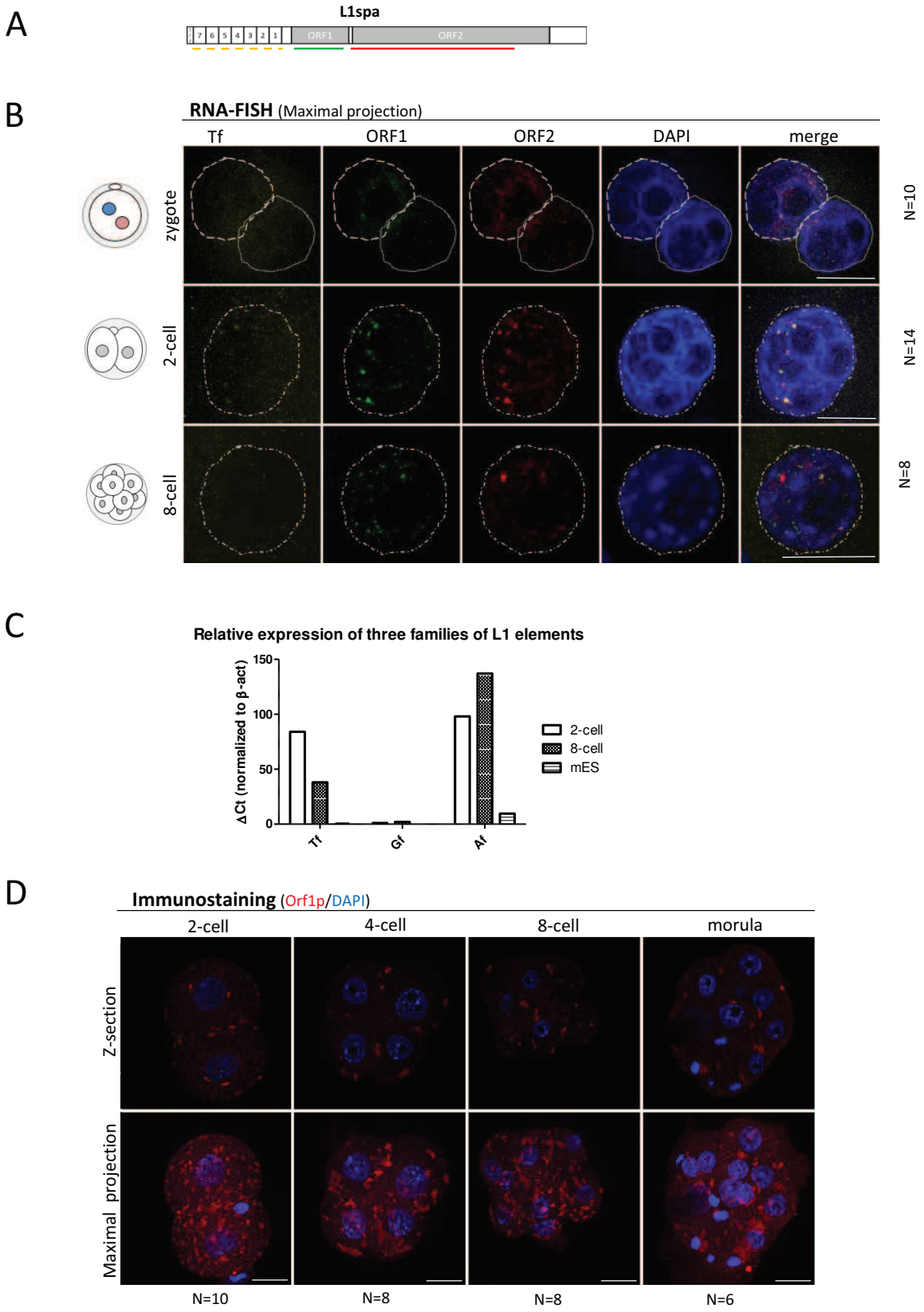


Figure 13. Expression of L1 elements in preimplantation embryos

A) Graphical representation of L1spa element and RNA-FISH probes visualizing transcripts from different regions; B) Images of RNA-FISH detecting L1 transcripts in zygote, 2-cell stage and 8-cell stage embryos; n=1 experiment, 5 embryos/group; C) Relative expression of three different families of L1 elements compared between mES cells, 2-cell and 8-cell stage embryos determined by RT-qPCR using Taqman probes; Plotted are Ct values normalized to ß-actine levels and transformed into linear scale; D) Immunostaining representing Orf1 protein during early mouse development

Scale bars – 10 microns;

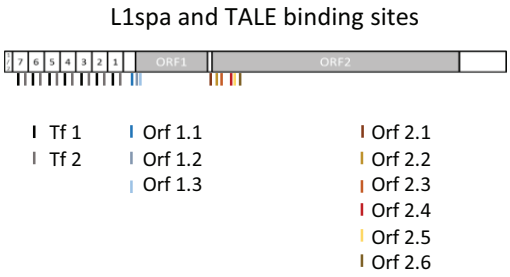
A

In silico analysis of TALEs in the mouse genome

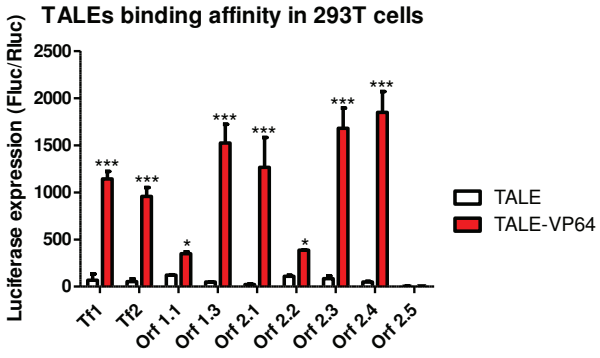
Target within L1 body	TALE construct name	Length of binding site (bp)	Number of target sites	Number of off-targets (non-LINE-1)	Full length targets (>6000 bp)
Tf monomer	Tf1	18	760	✓	718
Tf monomer	Tf2	18	1975	(5) *	1047
ORF 1	1.1	18	1870	✓	1841
ORF 1	1.2	18	3321	(29)	-
ORF 1	1.3	18	3911	✓	
ORF 2	2.1	18	2013	✓	
ORF 2	2.2	18	3886	✓	
ORF 2	2.3	17	5629	(1)*	
ORF 2	2.4	18	4630	✓	1999
ORF 2	2.5	18	11719	✓	
ORF 2	2.6	17	-	-	

* Truncated TE

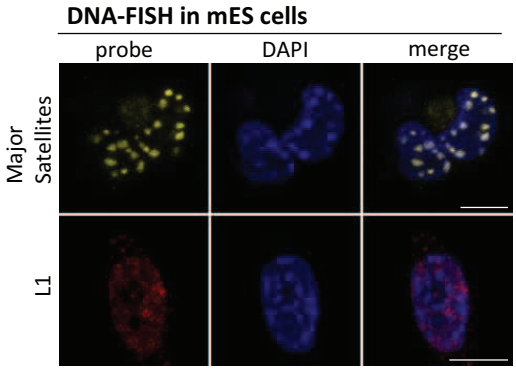
B



C



D



E

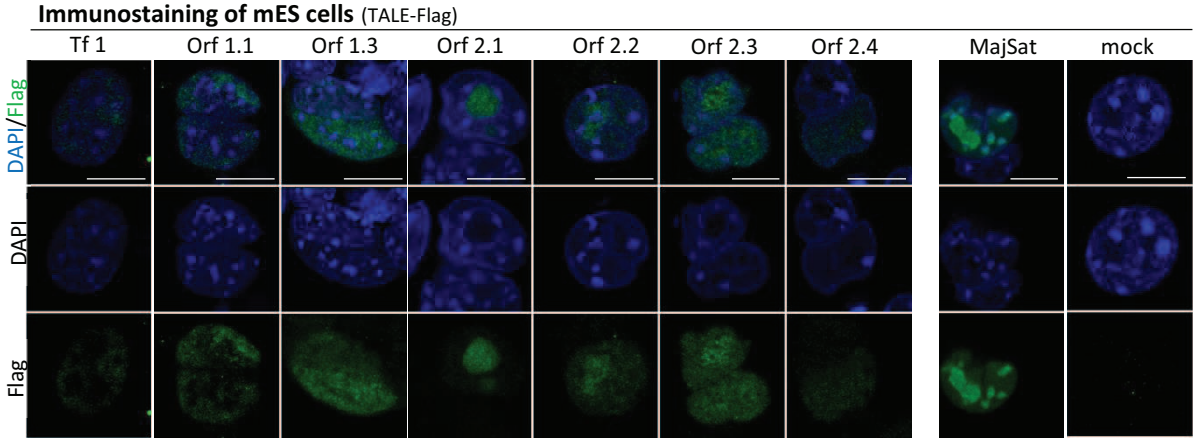
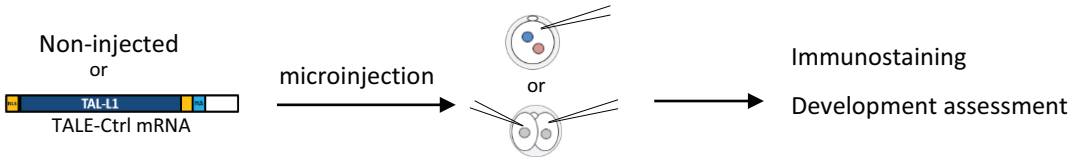


Figure 14. Verification of TALE constructs

A) Summary of in silico analysis of target sites of TALE designs; B) Graphical representation of binding regions of TALE-L1 constructs into the body of L1 element; C) Luciferase assay performed in 293T cells testing binding affinity of TALE-L1 constructs to their target sites; Relative expression of Firefly luciferase normalized to expression of Renilla luciferase plotted at the Y axis; n=3 (biological replicates); D) DNA-FISH visualizing different repetitive regions in mES cells; Maximal projection presented; E) Expression of TALE-Flag-L1 constructs in mES cells targeting different regions of LINE-1 elements according to Table A; anti-Flag antibody represents localization and expression of 7 different TALE-L1 constructs, TALE-major satellites represents positive control, mock represents negative control; Maximal projection presented;

*Scale bars – 10 microns; * p<0.05, ** p<0.01, *** p<0.001*

A



B

Summary of TALE-Ctrl expression in embryos (HA tag)							
TALE construct	Number of target sites	Microinjection into zygotes			Microinjection into 2-cell		
		2-cell	4-cell	8-cell	4-cell	8-cell	16-cell
Tf 1	760	✓	✓/-	nt	✓	✓	nt
ORF 1.1	1870	✓	✓/-	nt	✓	✓	✓/-
ORF 1.3	3911	✓	✓/-	nt	✓	✓	✓/-
ORF 2.4	4630	✓	✓/-	nt	✓	✓	✓/-

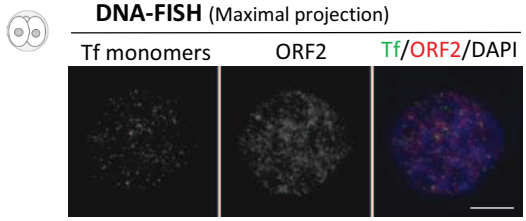
nt – not tested; ✓ - represents positive signal for expression; ✓/- - represents not clear signal for expression

C

Summary of developmental progression after TALE-Ctrl microinjections					
TALE construct	Number of target sites	Microinjection into zygotes		Microinjection into 2cell	
		100ng/ul	200ng/ul	100ng/ul	200ng/ul
Tf 1	760	nt	nt	nt	nt
ORF 1.1	1870	70% (n=59)	100% (n=20)	88% (n=16)	nt
ORF 1.3	3911	nt	nt	nt	nt
ORF 2.4	4630	nt	nt	75% (n=20)	nt
1.1 and 2.4	6500	nt	nt	nt	95% (n=39)
Tf 1, 1.3 and 2.4	9301	80% (n=42)	40% (n=32)	nt	81% (n=42)

nt – not tested; ✓ - represents positive signal for expression; ✓/- - represents not clear signal for expression

D



E

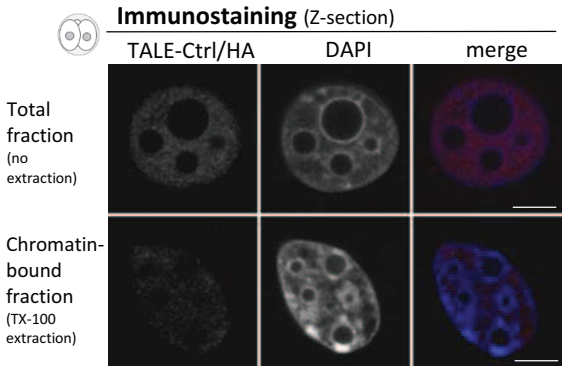
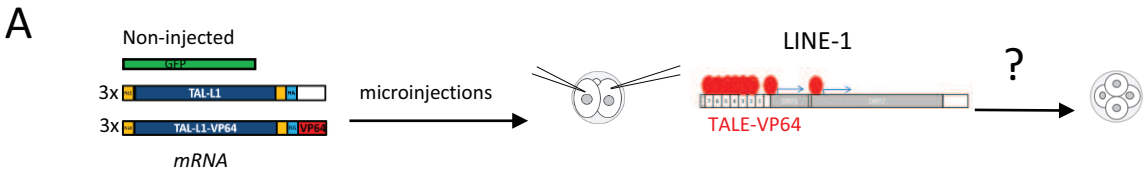
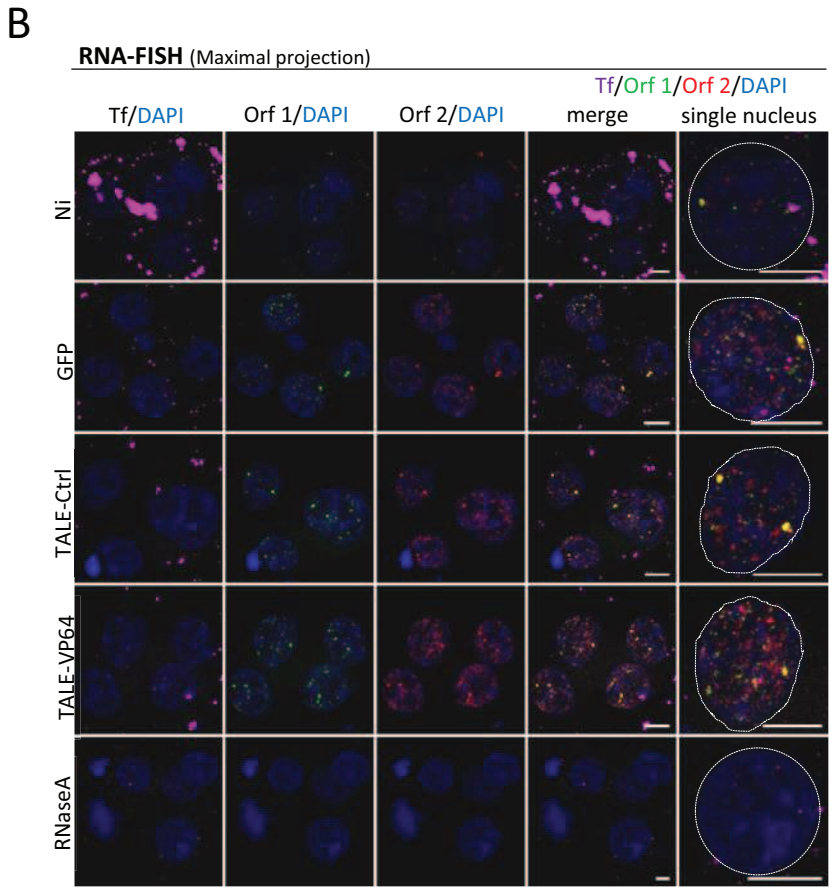


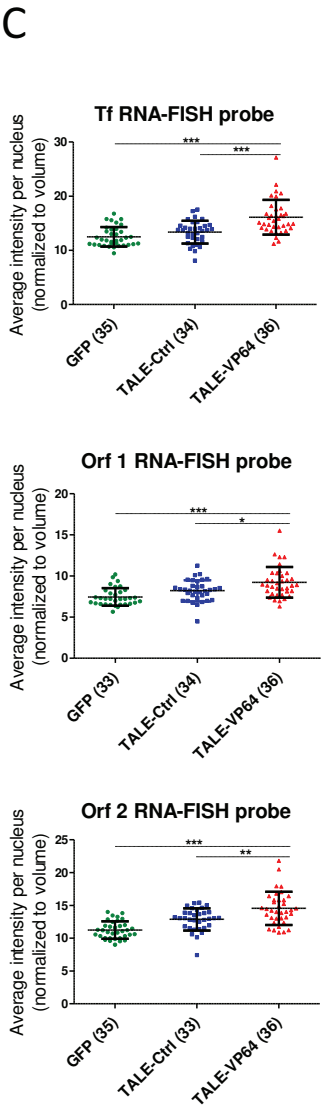
Figure 15. Establishment of experimental conditions for TALE expression in mouse embryos

A) Graphical representation of experimental design: zygotes or 2-cell stage embryos were injected with 100 ng/µl or 200 ng/µl of each of four TALE-Ctrl mRNA and then stained using an anti-HA antibody to verify expression and localization at different stages (Table B) or cultured for five days (Table C) to monitor development; B) Summary of conditions tested to assess proper expression and nuclear localization of TALE-Ctrl mRNAs after microinjections; C) Summary of conditions used to determine injection time and mRNA concentration for proper embryonic development; D) DNA-FISH performed to detect nuclear localization of Tf monomers and ORF2 regions of LINE-1 elements in 2-cell stage embryo shown as maximal projection of single nucleus; E) Immunostaining of an HA-tag visualizing TALE-Ctrl localization after microinjections of mRNA coding for TALEs binding to Tf1, 1.3 and 2.4 (100 ng/µl of each): total fraction vs chromatin bound fraction after TritonX-100 (TX-100) extraction

Scale bars – 10 microns;


A
Non-injected
GFP
3x
TAL-L1
3x
TAL-L1-VP64
VP64
mRNA
microinjections
LINE-1
TALE-VP64
?


B
RNA-FISH (Maximal projection)
Tf/DAPI
Orf 1/DAPI
Orf 2/DAPI
merge
Tf/Orf 1/Orf 2/DAPI
single nucleus
Ni
GFP
TALE-Ctrl
TALE-VP64
RNaseA


C
Tf RNA-FISH probe
Average intensity per nucleus (normalized to volume)
GFP (35)
TALE-Ctrl (34)
TALE-VP64 (36)
Orf 1 RNA-FISH probe
GFP (33)
TALE-Ctrl (34)
TALE-VP64 (36)
Orf 2 RNA-FISH probe
GFP (35)
TALE-Ctrl (33)
TALE-VP64 (36)

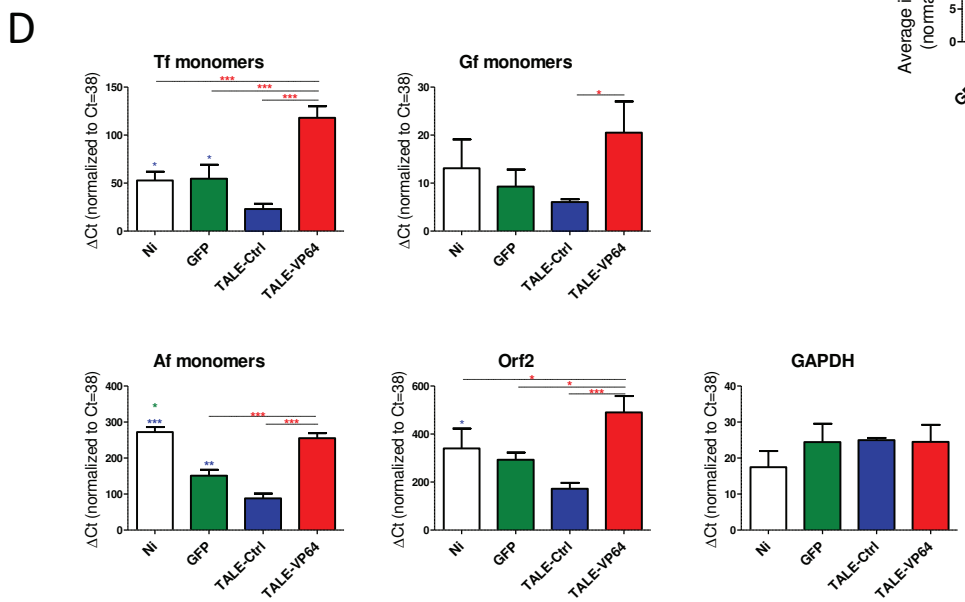

D
Tf monomers
Gf monomers
Af monomers
Orf2
GAPDH
ΔCt (normalized to Ct=38)
Ni
GFP
TALE-Ctrl
TALE-VP64

Figure 16. L1 transcription at 4-cell stage embryos – comparison between control and TALE-VP64 injected mouse embryos

A) Graphical representation of experimental approach – three groups of 2-cell stage embryos were injected with the same amount of total RNA (200 ng/µl of each of three TALE constructs) and cultured until late 4-cell stage; B) Representative images of RNA-FISH of 4-cell stage embryos – non-injected group (Ni), GFP injected, TALE-Ctrl injected, TALE-VP64 injected, and non-injected treated with RNaseA prior to RNA-FISH procedure; Region-specific RNA-FISH probes labeled with different fluorophores are presented in the first three columns, merge images and enlarged single nuclei are presented in the last two columns; C) Quantification of the fluorescent intensity of RNA-FISH probes visualizing transcripts from specific regions (Tf monomers, Orf1, Orf2) in control groups of embryos (GFP, TALE-Ctrl) and experimental group (TALE-VP64); Total intensity normalized to the nuclear volume plotted for each group of embryos; Horizontal line represents mean value, vertical lines represent standard deviation, each dot represents single nucleus; Shown is one representative experiment out of three independent repeats; D) Analysis of the expression of L1 elements belonging to Tf, Af, Gf families and global L1 expression (Orf2) among four groups of embryos; GAPDH presented as a control; Ct values are normalized to the baseline of 38 and inverted into the linear scale; n=1 experiment (5 embryos per group);
*Scale bars – 10 microns; * p<0.05, ** p<0.01, *** p<0.001*

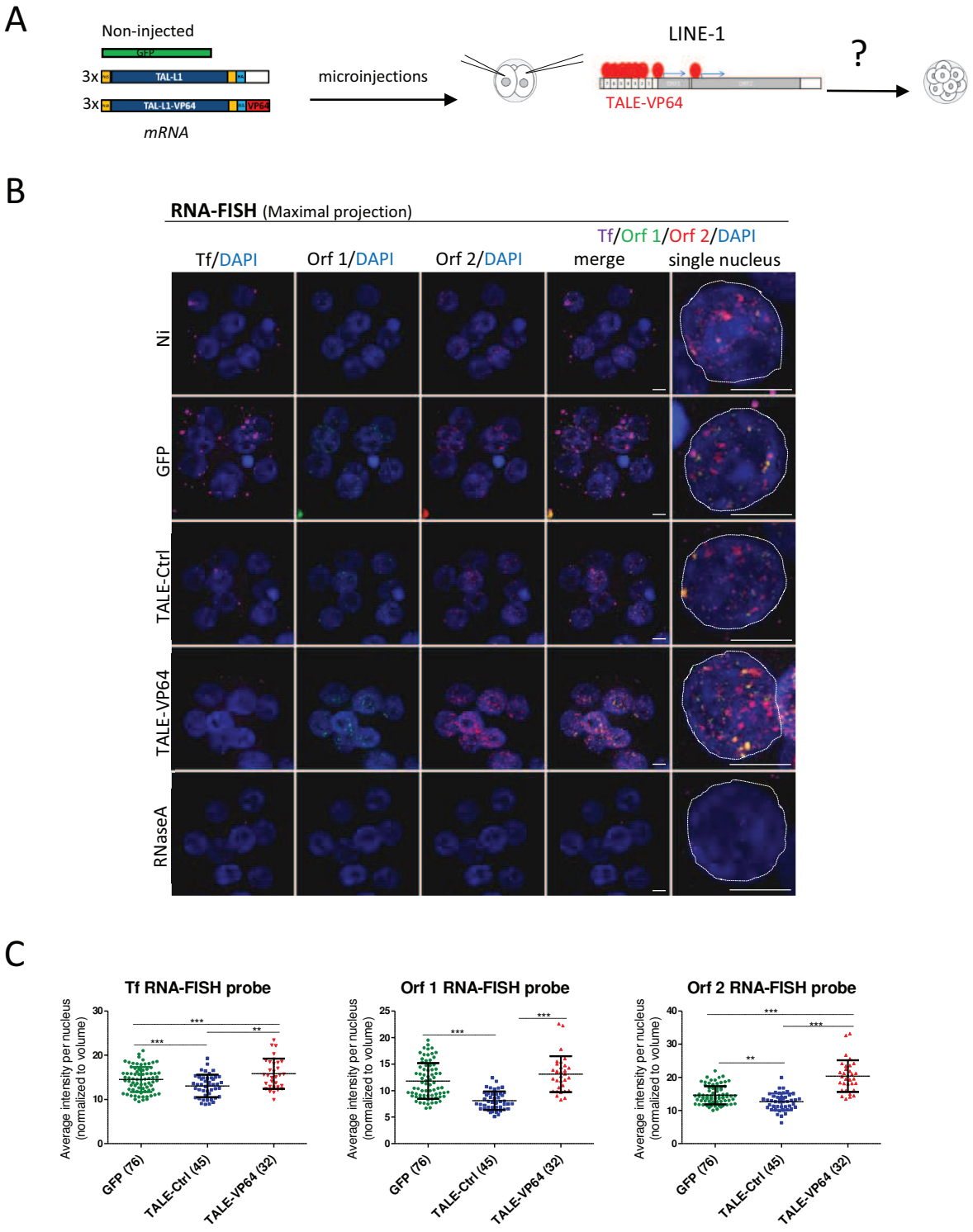
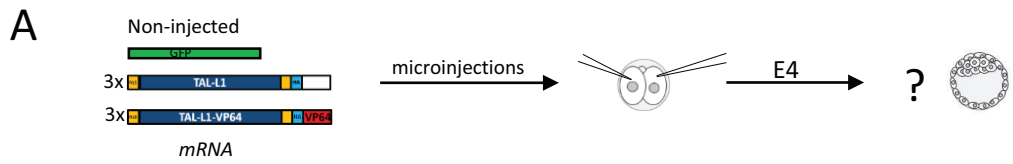
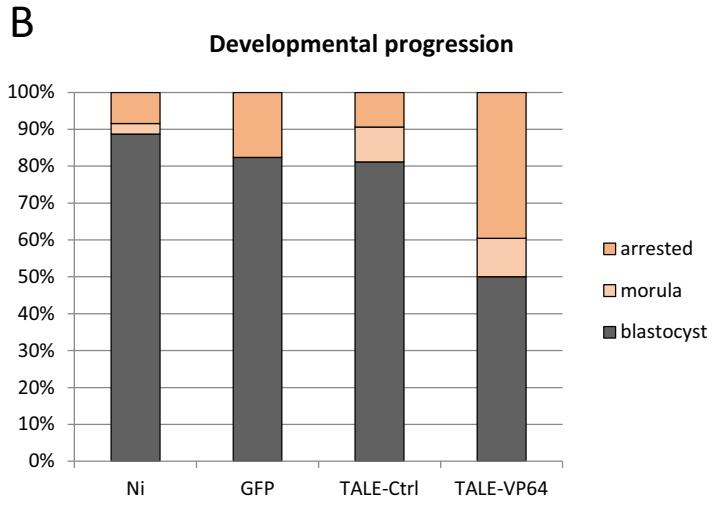


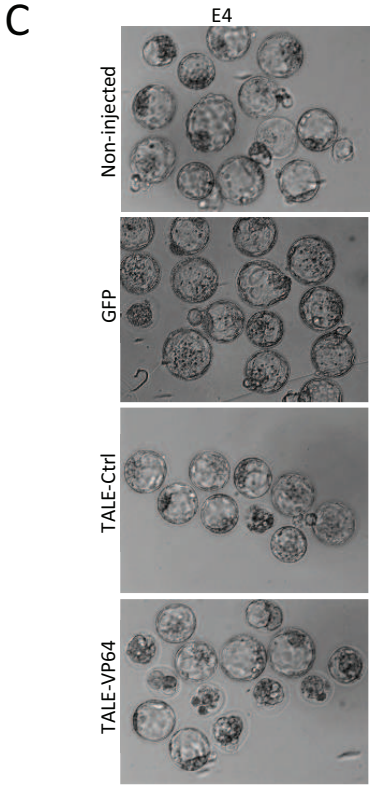
Figure 17. L1 transcription at 8-cell stage embryos – comparison between control and TALE-VP64 injected mouse embryos

*A) Graphical representation of experimental approach – three groups of 2-cell stage embryos were injected with the same amount of total RNA (200 ng/µl of each of three TALE constructs) and cultured until early 8-cell stage; B) Representative images of RNA-FISH of 8-cell stage embryos – non-injected group (Ni), GFP injected, TALE-Ctrl injected, TALE-VP64 injected, and non-injected treated with RNaseA prior to the RNA-FISH procedure; C) Quantification of the fluorescent intensity of RNA-FISH probes visualizing transcripts from specific regions (Tf monomers, Orf1, Orf2) in control groups of embryos (GFP, TALE-Ctrl) and experimental group (TALE-VP64); Total intensity normalized to the nuclear volume plotted for each group of embryos; Horizontal line represents mean value, vertical lines represent standard deviation, each dot represents single nucleus; Shown is one representative experiment out of three independent repeats; Scale bars – 10 microns; * p<0.05, ** p<0.01, *** p<0.001*





Group	% of embryos at blastocyst stage	Difference in blastocyst distribution compare to TALE-VP64 group
Ni	89% (n=71)	***
GFP	82% (n=32)	***
TALE-Ctrl	81% (n=53)	**
TALE-VP64	50% (n=48)	-



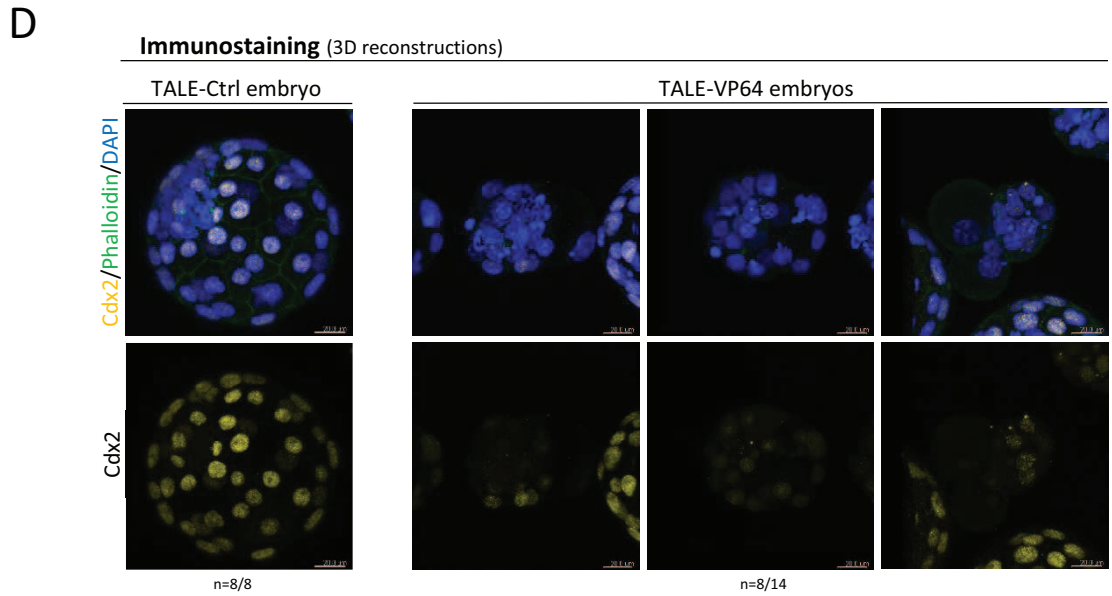


Figure 18. Embryonic development - comparison between control and TALE-VP64 injected mouse embryos
A) Graphical representation of an experimental procedure - three groups of 2-cell stage embryos were injected with the same amount of total RNA (200 ng/µl of each of three TALE constructs) and cultured for five days; B) Developmental progression after 4 days of culture of non-injected embryos (Ni), control groups of embryos injected with mRNA GFP (GFP) and mRNA TALE-Ctrl, and embryos injected with TALE-VP64; Table presents p values of a Z test comparing distributions between groups, no significance between controls observed; C) Images of embryos after 4 days of culture representing one of five independent experiments; D) 3D reconstruction of control embryo and arrested embryos after 4 days of culture showing representative phenotypes;
*Scale bars – 20 microns; * p<0.05, ** p<0.01, *** p<0.001*

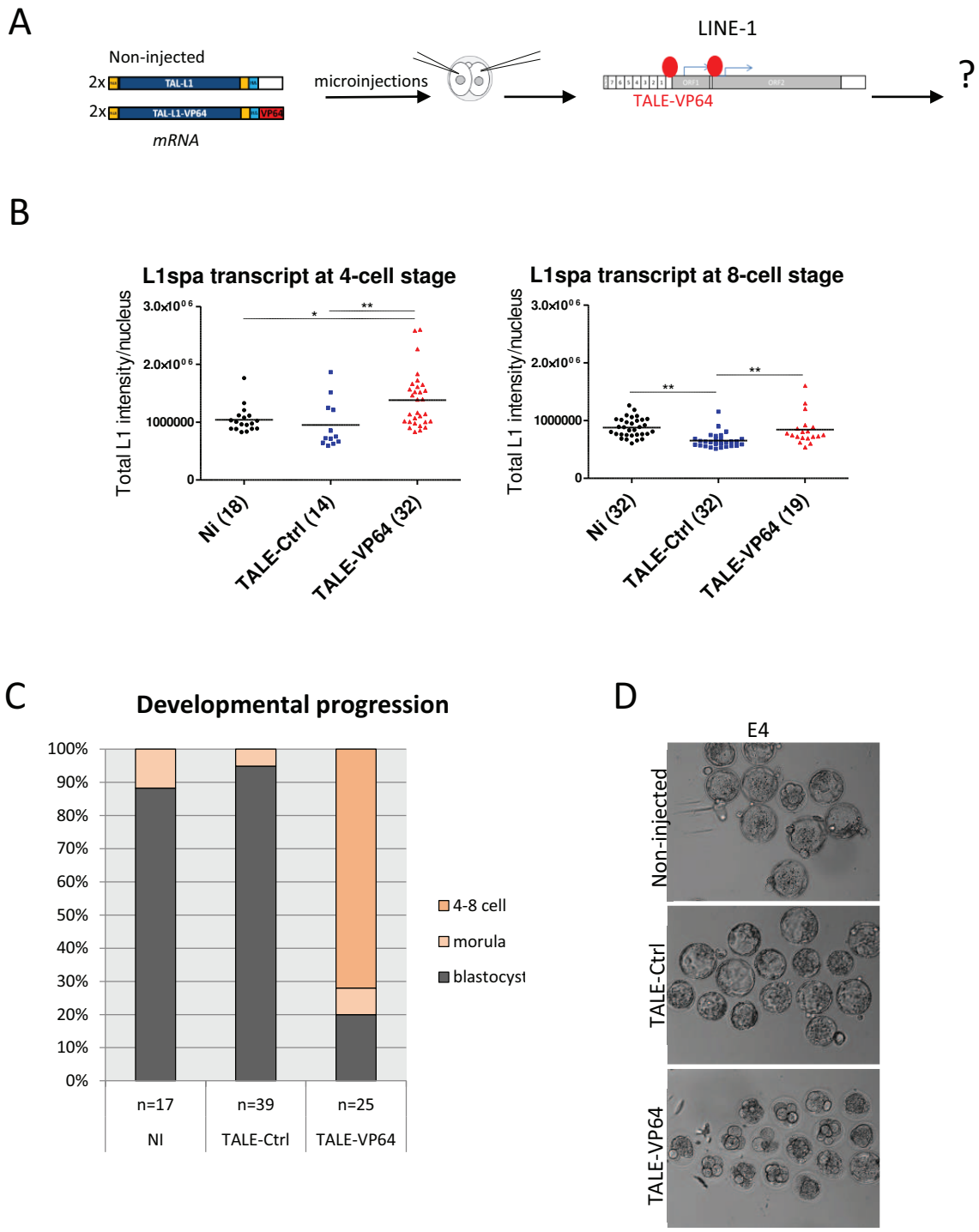


Figure 19. L1 transcription at 8-cell stage embryos – comparison between control and TALE-VP64 injected mouse embryos

A) Graphical representation of experimental approach – three groups of 2-cell stage embryos were injected with the same amount of total RNA (200 ng/µl of two TALE constructs targeting Orf1 and Orf2 regions of L1 elements) and cultured until late 4-cell stage or early 8-cell stage; B) Quantification of the fluorescent intensity of RNA-FISH probes visualizing transcripts from L1spa element in control groups of embryos (Ni – non-injected, TALE-Ctrl) and experimental group (TALE-VP64); Total intensity per nucleus plotted for each group of embryos; Horizontal line represents mean value, each dot represents single nucleus; Shown is one representative experiment out of three independent repeats; C) Developmental progression after 4 days of culture of non-injected embryos (Ni), control group of embryos injected with mRNA TALE-Ctrl, and embryos injected with TALE-VP64; D) Images of embryos after 4 days of culture representing one of three independent experiments;

*Scale bars – 10 microns; * p<0.05, ** p<0.01, *** p<0.001*

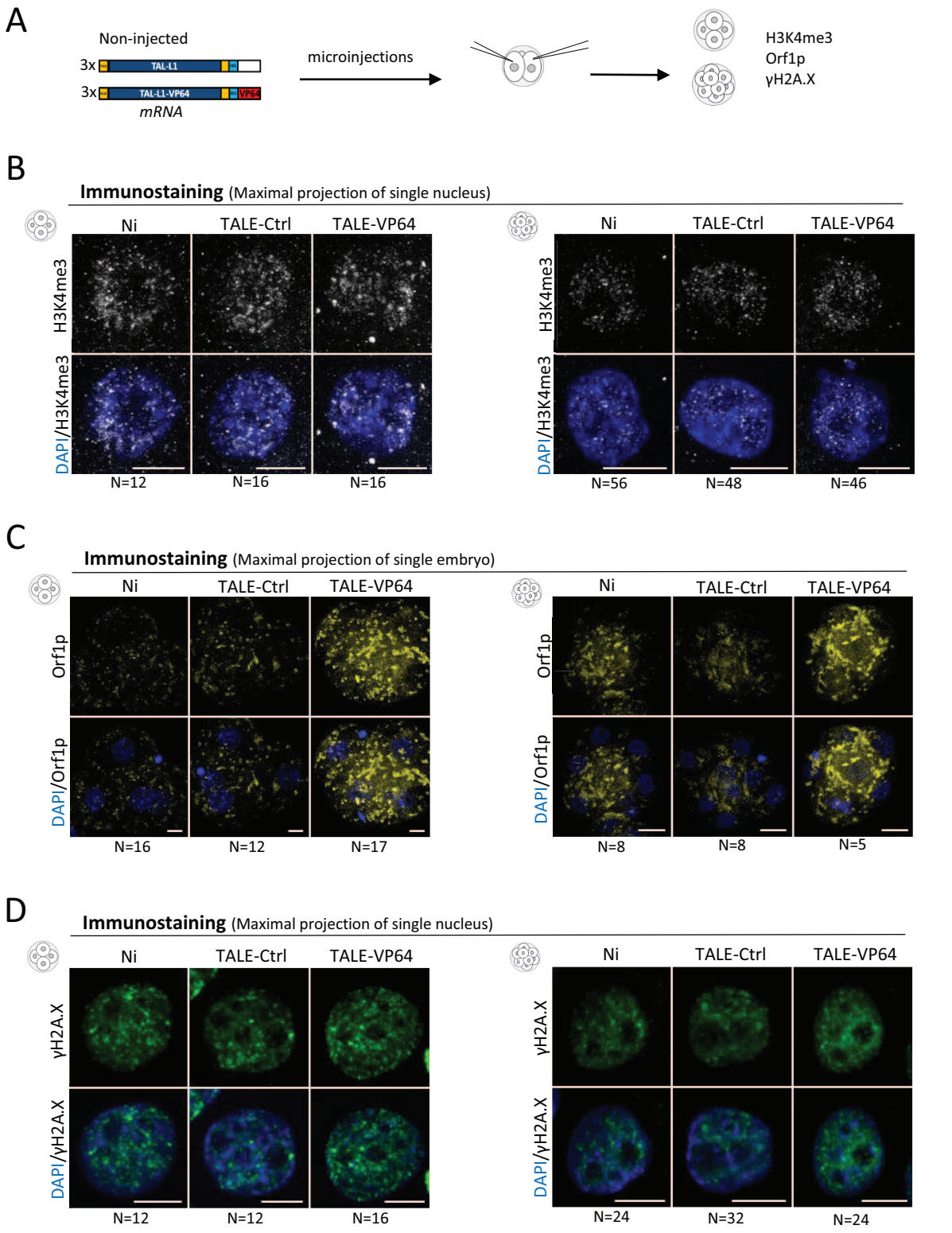


Figure 20. Expression of H3K4me3, Orf1p and γH2A.X in 4-cell and 8-cell stage embryos after microinjections

A) Graphical representation of an experimental design: non-injected (Ni), TALE-Ctrl, and TALE-VP64 microinjected embryos (200 ng/µl of each of three TALEs) were fixed at 4-cell and 8-cell stage and stained with B) anti-H3K4me3 antibody C) anti-Orf1p antibody D) anti-γH2A.X antibody;

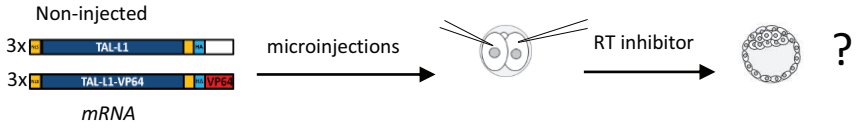
Scale bars – 10 microns;

A

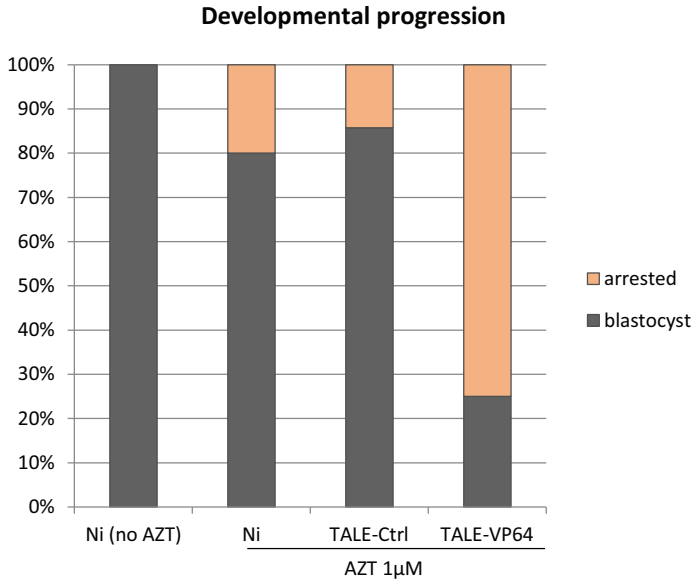
% of embryos at the blastocyst stage

Group	Cultured from 1-cell	Cultured from 2-cell
Ctrl	91% (n=22)	100% (n=19)
AZT 1µM	24% (n=29)	86% (n=14)

B



C



Group	% of embryos at blastocyst stage	Difference in blastocyst distribution compare to TALE-VP64 group
Ni (no AZT)	100% (n=18)	***
Ni	80% (n=15)	***
TALE-Ctrl	86% (n=21)	***
TALE-VP64	25% (n=28)	-

D

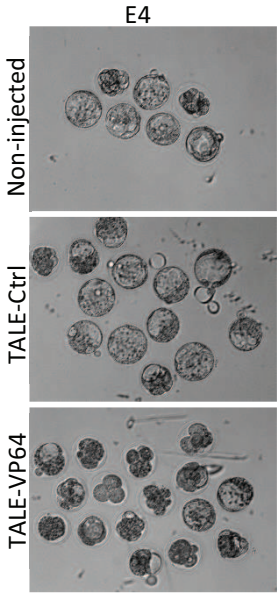


Figure 21. Effect of the RT inhibitor AZT on embryos development

A) Table representing developmental progression of embryos cultured with or without 1 µM AZT from 1-cell or 2-cell stage; B) Graphical representation of experimental design: 2-cell stage embryos were injected with either control TALE-Ctrl of TALE-VP64 (200 ng/µl of each of three TALE constructs) and cultured in the presence of 1 µM RT inhibitor AZT, non-injected embryos cultured as a control with or without AZT; C) Comparison of developmental progression of non-injected embryos, injected with TALE-Ctrl, or TALE-VP64 mRNA, after four days of culture in 1 µM AZT , non-injected embryos cultured in standard medium presented as a developmental control; Table presents p values of a Z test comparing distributions between groups, no significance between controls observed except from p=0,046 for Ni no AZT vs Ni; D) Images of embryos after 4 days of culture in 1 µM AZT representing one of two independent experiments;
**** p<0.001*

2.5. Materials & Methods

Embryo collection and culture

Embryos were collected from ~6 week-old F_1 (C57BL/6J × CBA/H) superovulated females crossed with F_1 males. Embryos were collected at the following times after human chorionic gonadotrophin injection: the 2-cell stage (46 h), 4-cell stage (54 h), 8-cell stage (72h), 16-cell stage (80 h), 32-cell stage (90 h), early blastocyst stage (98 h) and late blastocyst stage (114 h). All experiments were performed in accordance with the current legislation in France and the approval of the Regional Ethics Committee (ComEth's).

ES-cell culture

Mouse ES-cell lines, E14, were cultured in DMEM with GlutaMAX (Invitrogen) containing 15% FCS, LIF, 1 mM sodium pyruvate, penicillin/streptomycin and 0.1 mM 2-mercaptoethanol. The treatment of ES cells with inhibitors was performed using 3 µM CHIR99021 (a GSK3β inhibitor), 1 µM PD0325901 (a MEK inhibitor), 5 µM PD173074 (an FGF receptor tyrosine kinase activity inhibitor) and 3 µM PD184352 (a MEK inhibitor).

TALE construction

TALEs design, construction and luciferase assays were performed according to the instructions provided at (Miyanari 2014).

Immunostaining

Embryos or ES cells were fixed with 4% paraformaldehyde in PBS for 20 min at room temperature. After washing with PBS, embryos were permeabilized with 0.5% Triton X-100 in PBS for 10 min and then incubated in blocking solution (0.2% BSA in PBS) for at least 1h. Primary antibodies were anti-Orf1p (gift from Dr D. O'Carroll), anti-Flag (Wako), anti-HA (Milipore 07-442), anti-H3K4me3 (Abcam 1012), anti- γH2A.X (Millipore 05-636). After over-night (4°C) incubation in blocking solution containing primary antibodies, cells or embryos were washed three times with 0.01% Triton X-100 in PBS for 5 min each and then incubated in blocking solution containing corresponding secondary antibodies. After washing with PBS, mounting was done in VECTASHIELD (Vector Labs). Images were collected on a TCS SP8 confocal microscope (Leica) using a 63x oil objective. Z-sections were taken every 0.3 µm. Image analysis was performed using the software ImageJ and Imaris (Bitplane).

Whole mount RNA-FISH and DNA-FISH

RNA FISH and DNA-FISH were performed as described (Miyanari and Torres-Padilla 2012), except that incubation with Cot-1 was omitted. LINE-1 probes consisted of the full-length Tf element, Tf monomer fragment, Orf1 coding region, or Orf2 coding region from L1spa conserved sequence (Naas et al., 1998) cloned into pBluescript (Chow et al., 2010) and major satellite probe was as described (Myianari et al, 2012).

RT-qPCR

TaqMan® Gene Expression Assays (Applied Biosystems) and Taqman custom design assays (TIB MolBiol) were pooled to a final concentration of 0.2x for each of the 5 assays. Cell lysis and sequence-specific reverse transcription were performed at 50°C for 20 minutes, followed by sequence-specific pre-amplification as describe (A. Burton et al. 2013). The resulting cDNA was diluted 5-fold before analysis with Universal PCR Master Mix and TaqMan® Gene Expression Assays (Applied Biosystems) in light cycler 480 (Roche).

AZT administration

To inhibit L1 reverse transcriptase during preimplantation development embryos were cultured in standard KSOM medium with addition of 1 µM AZT (Sigma Aldrich, Cat#. A2169). The route and doses of AZT used in this study were based on previous reports (Toltzis et al. 1993; Dai et al. 2011; Malki et al. 2014)

2.6. Discussion

The main conclusion that can be drawn from the experiments on L1 elements described earlier is that the tethering of an artificial activator to L1 elements in mouse embryos leads to their up-regulation and subsequent developmental arrest. How these events are linked, remains unknown, however, understanding why L1 up-regulation causes developmental arrest may help us to dissect the function of L1 elements during early stages of development and the mechanisms that control their activity. There are several plausible explanations of how higher levels of L1 transcription may affect developmental progression and each of them indicates a slightly different role for the L1 elements during this period. In this part, I aim to discuss these hypotheses and propose some experiments that may help us distinguish the relevance of the ideas that I put forward.

Genome instability is not a cause of developmental arrest of embryos displaying L1 up-regulation

Firstly, and probably the simplest explanation of the developmental phenotype is that the higher ratio of retrotransposition events led to genome instability and arrest. It has been shown that integration of L1 elements into the genome can occur during early stages of development, prior to the blastocyst stage (Kano et al. 2009). Thus, a similar situation might arise in embryos overexpressing L1, facilitated by their higher RNA levels and protein abundance. However, experiments in which RT activity was blocked by the nucleoside analog inhibitor AZT, showed that the TALE-VP64 arrest was not rescued. This result likely rules out retrotransposition events as a cause for the developmental arrest because functional RT is indispensable for L1 retrotransposition. Moreover, levels of γH2A.X - a marker of DSBs and genome instability - did not seem to be increased in 4- or 8-cell stage embryos, suggesting that developmental problems are not related to increased genome instability caused by potential retrotransposition events. Our failure to detect Orf1p in the nuclei of TALE-VP64 embryos also supports this interpretation. Additionally, it has not been established whether *de novo* insertions can actually happen at such an early stage of development because in the only previously described retrotransposition assays performed in early embryos (Kano et al. 2009), the authors verified the presence of the L1 insertion cassette in blastocysts, but not earlier. They also suggest that it is the carry over L1-RNA from gametes that contribute mostly to the observed retrotransposition events in blastocyst, whereas it is known that the levels of the embryonic L1 transcription are very high before blastocyst formation (Fadloun et al. 2013). It is, thus likely that embryos have a mechanisms that inhibit the jeopardy of transposons at the

time when their expression is the highest. One of such mechanisms might be through post translational regulation of Orf1p by phosphorylation, which has been shown to be important for the transposition of L1 elements in human cells without affecting nuclease acids chaperon activity of the protein (Cook, Jones, and Furano 2015). Nevertheless, no proof for existence of this kind of mechanism has been reported in a mouse model so far.

Plausible deleterious effect of overexpression of L1-derived proteins on mouse preimplantation embryos

Another explanation of developmental arrest might be related to higher levels of expression of L1 proteins that we observed in TALE-VP64 embryos, at least for Or1p. L1 elements comprise two open reading frames encoding Orf1 protein and endonuclease/reverse transcriptase which are thought to be transcribed as bicistronic RNA and translated sequentially through an unconventional mechanism (Alisch et al. 2006). The three proteins are indispensable for retrotransposition, however, it has been suggested that they may play additional roles. Excessive level of L1 proteins in the cell, may elicit toxicity or adverse effects in cell division.

Despite the crucial role of Orf1p in retrotransposition (S. L. Martin 2006), the exact function of this protein is not known, and it may therefore play additional, not-yet identified roles during development. For example, recent work showed that ORF1p is associated with SMAD4 - a core factor of the TGF-β signal pathway- which upon TGF-β/pathway activation translocates from the cytoplasm to the nucleus and can regulate gene expression or even trigger apoptosis (Zhu et al. 2013). In oocytes depleted for Orf1p by injections of dsRNA, Smad4 displayed a different pattern of localization and lower levels of Cyclin B and Cdc2 were detected (Luo et al. 2015), suggesting that Orf1p-Smad4 interaction may operate during development. Hence, it is possible that overexpression of Orf1p in TALE-VP64 embryos leads to dysregulation of Smad4 pathway and, in consequence, affects gene expression. Orf1p has been also shown to be involved in the chromosome dynamics in mouse oocytes where its overexpression leads to oocytes with abnormal chromosome alignment and spindle organization, and subsequent meiotic arrest (Luo et al. 2015). The exact role of the protein in that process is not known but authors speculate that it might be related to the higher ration of retrotransposition events as they observed accumulation of γH2A.X foci. It is possible that TALE-VP64 embryos have abnormal mitosis which leads to the arrest, however, it would not be due to genome instability as discussed earlier. Moreover, I did not observe signs of lagging

chromosomes or micronuclei, thus, the interplay between Orf1p and Smad4, or another function of ORF1p is more likely to be involved.

Additionally to Orf1p, changes in L1 RT activity might be involved in the developmental phenotype of TALE-VP64 embryos as RTs are thought to play an important role in preimplantation development and embryos lacking RT activity, as measured by pharmacological inhibition of RT, depletion of L1 transcripts by morpholino, and anti-RT antibodies injections, show developmental arrest (Pittoggi et al. 2003; Beraldi et al. 2006). Interestingly, exposure of the L1 up-regulated embryos and control embryos to the RT inhibitor – AZT – had no effect on the ratio of blastocyst after 4 days of culture. Several explanations are plausible in regard to this experiment and one of them is a simple conclusion that RT is not involved in the TALE-VP64 embryos arrest. This would also mean that RT plays no crucial role during this time of development, or at least that the short exposure to low levels of the inhibitor from 2-cell stage do not interfere with blastocyst formation. As shown before, AZT is deleterious for embryos in high dosages but also low levels, even though AZT does not block divisions completely, affect the number of cells within blastocyst (Toltzis, Mourton, and Magnuson 1993). Although I have not quantified the number of cells in AZT treated embryos, if the RT had rescued TALE-VP64 arrest but then had caused the RT-deficiency related phenotype, these embryos should have still reached blastocyst, similarly to the controls, which was the case. Noteworthy, control zygotes incubated in the medium containing AZT inhibitor displayed severe developmental defects concomitantly with the previous reports (Toltzis, Mourton, and Magnuson 1993) implying that the AZT treatment that we used here is efficient. This also indicates that the levels of RT activity are necessary for development from the zygote stage, but not from the 2-cell stage onwards. The most likely scenario is that, even if L1-upregulation led to higher expression of RT, the RT activity does not contribute to the developmental arrest observed. A cautionary note, however, must be raised, which is that RT inhibitors may have distinct effects on mouse embryos which are not exclusively related to the function of L1 RT. Anti-RT antibody injection, however, seem to be more specific. It will be interesting to use anti-RT antibody injections at low concentrations on the TALE-VP64 embryos to ask whether by lowering levels of RT their development is rescued.

Recent discovery of the Orf0 protein encoded as an antisense transcript from human LINE-1 (Denli et al. 2015) might be also in relevance for the mouse L1. It has been already

reported that mouse L1 contains antisense promoter (J. Li et al. 2014) which drives transcription leading to the formation of chimeric transcripts. Although the existence of additional open reading frames in the mouse has not been documented, it remains possible that they are present and contain undiscovered L1 protein of unknown function.

Higher levels of L1 transcription and transcripts may affect numerous mechanism during early embryo formation

The way in which L1 could affect events in early embryogenesis, and how, potentially, their up-regulation causes developmental arrest might be also related to the L1 transcripts. Some of the plausible functions were already discussed in the final chapter of the Review, and include: L1 lncRNA functioning as a scaffold for the recruitment of chromatin-regulating factors i.e. spreading of silencing (Chow et al. 2010); L1-derived small RNAs playing a role in heterochromatin formation on L1 elements but also spreading on the neighbouring regions; L1 transcripts and abundant transcription promoting an open chromatin packaging (Hall et al. 2014) of the early embryo to facilitate epigenetic reprogramming and EGA; last but not least, L1 serving as alternative promoters driving expression of the host genes (J. Li et al. 2014; Peaston et al. 2004). All of these may be affected when the artificial activators are brought to the endogenous sequences of L1 elements.

If the scaffold hypothesis is correct, long sense and anti-sense RNAs derived from L1 might bind to genomic regions and recruit yet-to-be defined silencing factors. Antisense transcripts from L1 have been already linked to the changes in gene expression by chromatin modifications, like recruitment of H3K9me3 or H4K20me3 in ES cells (Cruickshanks et al. 2013). Thus, a similar situation may occur in the mouse embryos where in the excess of the transcripts, or with the transcription from elements that should have been already repressed, binding of L1 RNA might lead to the silencing of i.e. some of the developmentally important genes. To address this question, analysis of gene expression, together with more detailed overview on L1 transcription i.e. which strand of L1 is up-regulated or if the effect is strand specific, can be performed.

L1 transcripts might be involved in dysregulation of the global gene expression but also may disable appropriate control of their own repression by i.e. small RNA pathways. Several reports have already shown that L1 transcripts give rise to piRNAs (Aravin et al. 2007; Pezic et al. 2014), siRNAs and microRNAs (Ciaudo et al. 2013; Heras et al. 2013), which control L1

activity at transcriptional and post-transcriptional levels (Pezic et al. 2014; Ciaudo et al. 2013; Chow et al. 2010; DiGiacomo et al. 2013). Moreover, existence of such RNAs has been documented during early development where they are believed to down-regulate levels of L1 RNA (Ohnishi et al. 2009). It is likely that this small RNAs act in an endogenous situation to diminish L1 from 2- to 8-cell stage via depletion of L1 transcripts or establishment silencing environment on the genome. Thus, in case of TALE-VP64 embryos, the balance between levels of the transcripts necessary for the proper L1 functioning and regulation, might be affected with i.e. one strand being more transcribed than the other. This may lead to dysregulation of L1 as it has been shown for L1 derived short antisense RNA injected into zygotes which modulated L1 transcription at 2-cell stage (Fadloun et al. 2013). However, there is no experimental evidence that unambiguously shows that small RNAs pathways indeed operate in preimplantation embryos and play a role in the control of L1 elements. The result of a rescue of the TAL-VP64 phenotype with injections of dsRNA might shed a light on this event.

Although L1-dependent regulation of gene expression through silencing and heterochromatin spreading is an interesting concept, another explanation of the arrest of TALE-VP64 embryos is also possible. L1 elements are known to encode strong antisense promoters which drive a production of chimeric transcripts (Peaston et al. 2004; J. Li et al. 2014). Because in our current study, we targeted ~10.000 elements from which some of them should most likely be silenced by the 8-cell stage, such a high level of transcriptionally active loci might affect expression of regions which are in proximity to activated L1. Global analysis of transcription in embryos displaying up-regulation of L1 at 4-cell stage would help to address this issue. It is worth mentioning that L1 elements might have an influence also on chromatin state by creating more open chromatin configuration and leading to i.e. more permissive transcriptional state in TALE-VP64 embryos. For example, such a high level of transcription *per se* or presence of long non-coding L1 transcripts stably bound to chromatin, might lead to the above described phenomena as it has been suggested for X chromosome (Hall et al. 2014). As mouse L1 elements are thought to be transcribed as bicistronic RNA (Alisch et al. 2006), and the presence of full-length L1 transcripts in mouse embryos is also very likely (Peaston et al. 2004), these lncRNA derived from L1 may indeed take a part in regulation of chromatin state. Thus, an important question arises whether the arrest of TALE-VP64 embryos is related to the change in their transcriptional activity, more open chromatin state, or higher levels of L1

transcripts. To clarify it, additional experiments are necessary which would distinguish between these events, i.e. injections of the full-length transcript into the embryos or use of the DEL domain which has been shown to decondense chromatin (Carpenter et al. 2005). Comparison of the phenotypes of these embryos with the TALE-VP64 might help to, firstly, find a cause of their arrest, secondly, elucidate the function of high L1 transcription in development. Moreover, elimination of L1 transcripts by repression of L1 or by depleting L1 transcripts would also add information necessary for addressing the role of L1 in the early embryo.

To sum up, the data presented indicate that tight control of expression of L1 elements is important for proper development of mouse embryos. As discussed above, there are several layers of how L1s can be controlled and how they can influence events during early development which include changes on chromatin, transcription, transcripts, or proteins level. The use of site-specific transcriptional effectors, together with targeted depletion or overexpression of L1 RNAs, and injections of antibodies blocking L1 protein function, may help to decipher their regulatory pathways and levels at which they operate.

IV. Concluding remarks

The aim of my Ph.D project was to deepen into the mechanisms of heterochromatin formation during preimplantation embryos. To assess how it is established I focused on pericentric repeats and L1 retrotransposes as they represent two distinct types of heterochromatin regions.

The first part of my work revealed that nuclear organization plays an important role in regulating heterochromatin establishment and pointed towards the crucial function of localization of major satellites around NLBs in that process. Experiments that I have performed suggest that NLBs could serve as scaffold for a more open chromatin state which leads to chromatin remodeling, acquisition or maybe removal of specific proteins and epigenetic modifications and regulate transcription pattern of major satellites. The importance of such a distinct nuclear organization after fertilization is also highlighted by SCNT experiments in which this particular localization of pericentric regions around NLBs is restored, suggesting its involvement in the reprogramming (C. Martin et al. 2006). Thus, nuclear organization arises as a crucial player in the formation of heterochromatin in early development.

The second part focused on L1 retrotransposones and role of their silencing during early development. By prolonging the phase of their very abundant transcription we counteract formation of heterochromatin on these regions, which leads to developmental failure. Although we cannot link directly these events, these manipulations indicate that the tight regulation of L1 transcription is crucial for early development. Now, the important question to address is how this regulation works and what mechanisms drive heterochromatin formation on L1 elements in preimplantation embryos. Much work has been already done in ES cells which pinpoints some of the pathways that are important for that process i.e. small RNAs, Suv39h1/h2 mediated H3K9me3, or KRAB-ZNF proteins. In preimplantation embryos, however, which differ from ES cells in terms of molecular pathways, gene expression, and global chromatin layout, there is little evidence which would indicate how L1 are regulated. Lower levels of H3K4me3 present on L1s but not H3K9me3 has been shown to be concomitant with the down-regulation of their expression (Fadloun et al. 2013) suggesting that removal of active histone modifications rather than acquisition of repressive ones may play a role in that process. However, a role for RNA-driven regulation of L1 is also plausible (Fadloun et al. 2013).

To sum up, my experiments suggest that establishment of heterochromatin during early development is crucial event for the formation of functional embryo. Although distinct regions, undergo different changes and temporal dynamics in order to form the final heterochromatic structure (which may include i.e. change in nuclear localization or removal of active marks) there are some aspects that seem to be similar. The involvement of RNA component in the process of heterochromatin formation is one of the striking examples.

V. Bibliography

Aguirre-Lavin, Tiphaine, Pierre Adenot, Amélie Bonnet-Garnier, Gaétan Lehmann, Renaud Fleurot, Claire Boulesteix, Pascale Debey, and Nathalie Beaujean. 2012. "3D-FISH Analysis of Embryonic Nuclei in Mouse Highlights Several Abrupt Changes of Nuclear Organization during Preimplantation Development." *BMC Developmental Biology* 12 (1): 30. doi:10.1186/1471-213X-12-30.

Akhtar, Asifa, and Susan M Gasser. 2007. "The Nuclear Envelope and Transcriptional Control." *Nature Reviews. Genetics* 8 (7). Nature Publishing Group: 507–17. doi:10.1038/nrg2122.

Alisch, Reid S., Jose L. Garcia-Perez, Alysson R. Muotri, Fred H. Gage, and John V. Moran. 2006. "Unconventional Translation of Mammalian LINE-1 Retrotransposons." *Genes and Development* 20 (2): 210–24. doi:10.1101/gad.1380406.

Almouzni, Geneviève, and Aline V Probst. "Heterochromatin Maintenance and Establishment: Lessons from the Mouse Pericentromere." *Nucleus (Austin, Tex.)* 2 (5): 332–38. doi:10.4161/nucl.2.5.17707.

Aloia, Luigi, Bruno Di Stefano, and Luciano Di Croce. 2013. "Polycomb Complexes in Stem Cells and Embryonic Development." *Development (Cambridge, England)* 140 (12): 2525–34. doi:10.1242/dev.091553.

Ancelin, Katia, Ulrike C Lange, Petra Hajkova, Robert Schneider, Andrew J Bannister, Tony Kouzarides, and M Azim Surani. 2006. "Blimp1 Associates with Prmt5 and Directs Histone Arginine Methylation in Mouse Germ Cells." *Nature Cell Biology* 8 (6): 623–30. doi:10.1038/ncb1413.

Arand, Julia, Mark Wossidlo, Konstantin Lepikhov, Julian R Peat, Wolf Reik, and Jörn Walter. 2015. "Selective Impairment of Methylation Maintenance Is the Major Cause of DNA Methylation Reprogramming in the Early Embryo." *Epigenetics & Chromatin* 8 (1): 1. doi:10.1186/1756-8935-8-1.

Aravin, Alexei A, Ravi Sachidanandam, Angelique Girard, Katalin Fejes-Toth, and Gregory J Hannon. 2007. "Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control." *Science (New York, N.Y.)* 316 (5825): 744–47. doi:10.1126/science.1142612.

Arney, Katharine L, Siqin Bao, Andrew J Bannister, Tony Kouzarides, and M Azim Surani. 2002. "Histone Methylation Defines Epigenetic Asymmetry in the Mouse Zygote." *The International Journal of Developmental Biology* 46 (3): 317–20. http://www.ncbi.nlm.nih.gov/pubmed/12068953.

Bannister, A J, P Zegerman, J F Partridge, E A Miska, J O Thomas, R C Allshire, and T Kouzarides. 2001. "Selective Recognition of Methylated Lysine 9 on Histone H3 by the HP1 Chromo Domain." *Nature* 410 (6824): 120–24. doi:10.1038/35065138.

Bell, A C, and G Felsenfeld. 2000. "Methylation of a CTCF-Dependent Boundary Controls Imprinted Expression of the Igf2 Gene." *Nature* 405 (6785): 482–85. doi:10.1038/35013100.

Beraldi, Rosanna, Carmine Pittoggi, Ilaria Sciamanna, Elisabetta Mattei, and Corrado Spadafora. 2006. "Expression of LINE-1 Retroposons Is Essential for Murine Preimplantation Development." *Molecular Reproduction and Development* 73 (3): 279–87. doi:10.1002/mrd.20423.

Blackledge, Neil P, Anca M Farcas, Takashi Kondo, Hamish W King, Joanna F McGouran, Lars L P Hanssen, Shinsuke Ito, et al. 2014. "Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation." *Cell* 157 (6): 1445–59. doi:10.1016/j.cell.2014.05.004.

Bonnet-Garnier, Amelie, Prisca Feuerstein, Martine Chebrout, Renaud Fleurot, Habib Ullah Jan, Pascale Debey, and Nathalie Beaujean. 2012. "Genome Organization and Epigenetic Marks in Mouse Germinal Vesicle Oocytes." *International Journal of Developmental Biology* 56 (10-12): 877–87.

doi:10.1387/ijdb.120149ab.

Bošković, Ana, Ambre Bender, Laurence Gall, Celine Ziegler-Birling, Nathalie Beaujean, and Maria-Elena Torres-Padilla. 2012. “Analysis of Active Chromatin Modifications in Early Mammalian Embryos Reveals Uncoupling of H2A.Z Acetylation and H3K36 Trimethylation from Embryonic Genome Activation.” *Epigenetics* 7 (7): 747–57. doi:10.4161/epi.20584.

Bošković, Ana, and Maria Elena Torres-Padilla. 2013. “How Mammals Pack Their Sperm: A Variant Matter.” *Genes and Development* 27 (15): 1635–39. doi:10.1101/gad.226167.113.

Bourc’his, Déborah, and Timothy H Bestor. 2004. “Meiotic Catastrophe and Retrotransposon Reactivation in Male Germ Cells Lacking Dnmt3L.” *Nature* 431 (7004): 96–99. doi:10.1038/nature02886.

Bouzinba-Segard, Haniaa, Adeline Guais, and Claire Francastel. 2006. “Accumulation of Small Murine Minor Satellite Transcripts Leads to Impaired Centromeric Architecture and Function.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (23): 8709–14. doi:10.1073/pnas.0508006103.

Bulut-Karslioglu, Aydan, Valentina Perrera, Manuela Scaranaro, Inti Alberto de la Rosa-Velazquez, Suzanne van de Nobelen, Nicholas Shukeir, Johannes Popow, et al. 2012. “A Transcription Factor–based Mechanism for Mouse Heterochromatin Formation.” *Nature Structural & Molecular Biology* 19 (10): 1023–30. doi:10.1038/nsmb.2382.

Burns, Kathleen H, Maria M Viveiros, Yongsheng Ren, Pei Wang, Francesco J DeMayo, Donald E Frail, John J Eppig, and Martin M Matzuk. 2003. “Roles of NPM2 in Chromatin and Nucleolar Organization in Oocytes and Embryos.” *Science (New York, N.Y.)* 300 (5619): 633–36. doi:10.1126/science.1081813.

Burton, a., and M.-E. Torres-Padilla. 2010. “Epigenetic Reprogramming and Development: A Unique Heterochromatin Organization in the Preimplantation Mouse Embryo.” *Briefings in Functional Genomics* 9 (5-6): 444–54. doi:10.1093/bfgp/elq027.

Burton, Adam, Julius Muller, Shengjiang Tu, Pablo Padilla-Longoria, Ernesto Guccione, and Maria Elena Torres-Padilla. 2013. “Single-Cell Profiling of Epigenetic Modifiers Identifies PRDM14 as an Inducer of Cell Fate in the Mammalian Embryo.” *Cell Reports* 5 (3). The Authors: 687–701. doi:10.1016/j.celrep.2013.09.044.

Carpenter, Anne E, Sevinci Memedula, Matthew J Plutz, and Andrew S Belmont. 2005. “Common Effects of Acidic Activators on Large-Scale Chromatin Structure and Transcription.” *Molecular and Cellular Biology* 25 (3): 958–68. doi:10.1128/MCB.25.3.958-968.2005.

Casanova, Miguel, Michał Pasternak, Fatima El Marjou, Patricia Le Baccon, Aline V. Probst, and Geneviève Almouzni. 2013. “Heterochromatin Reorganization during Early Mouse Development Requires a Single-Stranded Noncoding Transcript.” *Cell Reports* 4 (6): 1156–67. doi:10.1016/j.celrep.2013.08.015.

Cavalli, Giacomo. 2015. “PRC1 Proteins Orchestrate Three-Dimensional Genome Architecture.” *Nature Genetics* 47 (10). Nature Publishing Group: 1105–6. doi:10.1038/ng.3411.

Celeste, Arkady, Simone Petersen, Peter J Romanienko, Oscar Fernandez-Capetillo, Hua Tang Chen, Olga A Sedelnikova, Bernardo Reina-San-Martin, et al. 2002. “Genomic Instability in Mice Lacking Histone H2AX.” *Science (New York, N.Y.)* 296 (5569): 922–27. doi:10.1126/science.1069398.

Chow, Jennifer C., Constance Ciaudo, Melissa J. Fazzari, Nathan Mise, Nicolas Servant, Jacob L. Glass, Matthew Attreed, et al. 2010. “LINE-1 Activity in Facultative Heterochromatin Formation during X Chromosome Inactivation.” *Cell* 141 (6). Elsevier Ltd: 956–69. doi:10.1016/j.cell.2010.04.042.

Ciaudo, Constance, Florence Jay, Ikuhiro Okamoto, Chong-Jian Chen, Alexis Sarazin, Nicolas Servant, Emmanuel

Barillot, Edith Heard, and Olivier Voinnet. 2013. “RNAi-Dependent and Independent Control of LINE1 Accumulation and Mobility in Mouse Embryonic Stem Cells.” *PLoS Genetics* 9 (11): e1003791. doi:10.1371/journal.pgen.1003791.

Cook, Pamela R., Charles E. Jones, and Anthony V. Furano. 2015. “Phosphorylation of ORF1p Is Required for L1 Retrotransposition.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (14): 4298–4303. doi:10.1073/pnas.1416869112.

Cruickshanks, Hazel A, Natasha Vafadar-Isfahani, Donncha S Dunican, Andy Lee, Duncan Sproul, Jonathan N Lund, Richard R Meehan, and Cristina Tufarelli. 2013. “Expression of a Large LINE-1-Driven Antisense RNA Is Linked to Epigenetic Silencing of the Metastasis Suppressor Gene TFPI-2 in Cancer.” *Nucleic Acids Research* 41 (14): 6857–69. doi:10.1093/nar/gkt438.

Dai, Lixin, Qing Huang, and Jef D Boeke. 2011. “Effect of Reverse Transcriptase Inhibitors on LINE-1 and Ty1 Reverse Transcriptase Activities and on LINE-1 Retrotransposition.” *BMC Biochemistry* 12 (1). BioMed Central Ltd: 18. doi:10.1186/1471-2091-12-18.

De La Fuente, R., C. Baumann, and M. M. Viveiros. 2015. “ATRX Contributes to Epigenetic Asymmetry and Silencing of Major Satellite Transcripts in the Maternal Genome of the Mouse Embryo.” *Development*, 1806–17. doi:10.1242/dev.118927.

De La Fuente, Rabindranath, Maria M Viveiros, Karen Wigglesworth, and John J Eppig. 2004. “ATRX, a Member of the SNF2 Family of helicase/ATPases, Is Required for Chromosome Alignment and Meiotic Spindle Organization in Metaphase II Stage Mouse Oocytes.” *Developmental Biology* 272 (1): 1–14. doi:10.1016/j.ydbio.2003.12.012.

Debey, P, M S Szöllösi, D Szöllösi, D Vautier, A Girousse, and D Besombes. 1993. “Competent Mouse Oocytes Isolated from Antral Follicles Exhibit Different Chromatin Organization and Follow Different Maturation Dynamics.” *Molecular Reproduction and Development* 36 (1): 59–74. doi:10.1002/mrd.1080360110.

Denli, Ahmet M., Iñigo Narvaiza, Bilal E. Kerman, Monique Pena, Christopher Benner, Maria C.N. Marchetto, Jolene K. Diedrich, et al. 2015. “Primate-Specific ORF0 Contributes to Retrotransposon-Mediated Diversity.” *Cell* 163 (3): 583–93. doi:10.1016/j.cell.2015.09.025.

DiGiacomo, Monica, Stefano Comazzetto, Harpreet Saini, Serena DeFazio, Claudia Carrieri, Marcos Morgan, Lina Vasiliauskaite, Vladimir Benes, AntonJ Enright, and Dónal O’Carroll. 2013. “Multiple Epigenetic Mechanisms and the piRNA Pathway Enforce LINE1 Silencing during Adult Spermatogenesis.” *Molecular Cell* 50 (4): 601–8. doi:10.1016/j.molcel.2013.04.026.

Dodge, Jonathan E, Yong-Kook Kang, Hideyuki Beppu, Hong Lei, and En Li. 2004. “Histone H3-K9 Methyltransferase ESET Is Essential for Early Development.” *Molecular and Cellular Biology* 24 (6): 2478–86. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=355869&tool=pmcentrez&rendertype=abstract.

Drané, Pascal, Khalid Ouararhni, Arnaud Depaux, Muhammad Shuaib, and Ali Hamiche. 2010. “The Death-Associated Protein DAXX Is a Novel Histone Chaperone Involved in the Replication-Independent Deposition of H3.3.” *Genes & Development* 24 (12): 1253–65. doi:10.1101/gad.566910.

Elsässer, Simon J., Kyung-Min Noh, Nichole Diaz, C. David Allis, and Laura a. Banaszynski. 2015. “Histone H3.3 Is Required for Endogenous Retroviral Element Silencing in Embryonic Stem Cells.” *Nature* 522 (7555): 240–44. doi:10.1038/nature14345.

Erhardt, Sylvia, Barbara G Mellone, Craig M Betts, Weiguo Zhang, Gary H Karpen, and Aaron F Straight. 2008.

"Genome-Wide Analysis Reveals a Cell Cycle-Dependent Mechanism Controlling Centromere Propagation." *The Journal of Cell Biology* 183 (5): 805–18. doi:10.1083/jcb.200806038.

Erhardt, Sylvia, I-Hsin Su, Robert Schneider, Sheila Barton, Andrew J Bannister, Laura Perez-Burgos, Thomas Jenuwein, Tony Kouzarides, Alexander Tarakhovsky, and M Azim Surani. 2003. "Consequences of the Depletion of Zygotic and Embryonic Enhancer of Zeste 2 during Preimplantation Mouse Development." *Development (Cambridge, England)* 130 (18): 4235–48. doi:10.1242/dev.00625.

Fadloun, A, S Le Gras, B Jost, C Ziegler-Birling, H Takahashi, E Gorab, P Carninci, and M E Torres-Padilla. 2013. "Chromatin Signatures and Retrotransposon Profiling in Mouse Embryos Reveal Regulation of LINE-1 by RNA." *Nat Struct Mol Biol* 20 (3). Nature Publishing Group: 332–38. doi:10.1038/nsmb.2495.

Fan, Jun Y., Danny Rangasamy, Karolin Luger, and David J. Tremethick. 2004. "H2A.Z Alters the Nucleosome Surface to Promote HP1α-Mediated Chromatin Fiber Folding." *Molecular Cell* 16 (4): 655–61. doi:10.1016/j.molcel.2004.10.023.

Finlan, Lee E., Duncan Sproul, Inga Thomson, Shelagh Boyle, Elizabeth Kerr, Paul Perry, Bauke Ylstra, Jonathan R. Chubb, and Wendy a. Bickmore. 2008. "Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells." *PLoS Genetics* 4 (3): e1000039. doi:10.1371/journal.pgen.1000039.

Folco, H. D., A. L. Pidoux, T. Urano, and R. C. Allshire. 2008. "Heterochromatin and RNAi Are Required to Establish CENP-A Chromatin at Centromeres." *Science* 319 (5859): 94–97. doi:10.1126/science.1150944.

Fukagawa, Tatsuo, Masahiro Nogami, Mitsuko Yoshikawa, Masashi Ikeno, Tuneko Okazaki, Yasunari Takami, Tatsuo Nakayama, and Mitsuo Oshimura. 2004. "Dicer Is Essential for Formation of the Heterochromatin Structure in Vertebrate Cells." *Nature Cell Biology* 6 (8). Nature Publishing Group: 784–91. doi:10.1038/ncb1155.

Fulka, Helena, and Alena Langerova. 2014. "The Maternal Nucleolus Plays a Key Role in Centromere Satellite Maintenance during the Oocyte to Embryo Transition." *Development (Cambridge, England)* 141 (8): 1694–1704. doi:10.1242/dev.105940.

Furuya, Masataka, Mamoru Tanaka, Takahide Teranishi, Kazuya Matsumoto, Yoshihiko Hosoi, Kazuhiro Saeki, Hitoshi Ishimoto, Kazuhiro Minegishi, Akira Iritani, and Yasunori Yoshimura. 2007. "H1foo Is Indispensable for Meiotic Maturation of the Mouse Oocyte." *The Journal of Reproduction and Development* 53 (4): 895–902. doi:10.1262/jrd.19008.

Gent, Jonathan I., and R. Kelly Dawe. 2012. "RNA as a Structural and Regulatory Component of the Centromere." *Annual Review of Genetics* 46 (1): 443–53. doi:10.1146/annurev-genet-110711-155419.

Grewal, Shiv I S, and Sarah C R Elgin. 2007. "Transcription and RNA Interference in the Formation of Heterochromatin." *Nature* 447 (7143). Nature Publishing Group: 399–406. doi:10.1038/nature05914.

Grewal, Shiv I S, and Songtao Jia. 2007. "Heterochromatin Revisited." *Nature Reviews. Genetics* 8 (1): 35–46. doi:10.1038/nrg2008.

Grimes, S R, and N Henderson. 1984. "Hyperacetylation of Histone H4 in Rat Testis Spermatids." *Experimental Cell Research* 152 (1): 91–97. http://www.ncbi.nlm.nih.gov/pubmed/6714327.

Guenatri, Mounia, Delphine Bailly, Christèle Maison, and Geneviève Almouzni. 2004. "Mouse Centric and Pericentric Satellite Repeats Form Distinct Functional Heterochromatin." *The Journal of Cell Biology* 166 (4): 493–505. doi:10.1083/jcb.200403109.

Haaf, T, and D C Ward. 1995. "Higher Order Nuclear Structure in Mammalian Sperm Revealed by in Situ Hybridization and Extended Chromatin Fibers." *Experimental Cell Research* 219 (2): 604–11.

doi:10.1006/excr.1995.1270.

Hahn, Matthias, Silvia Dambacher, Stanimir Dulev, Anastasia Yurievna Kuznetsova, Simon Eck, Stefan Wörz, Dennis Sadic, et al. 2013. “Suv4-20h2 Mediates Chromatin Compaction and Is Important for Cohesin Recruitment to Heterochromatin.” *Genes & Development* 27 (8): 859–72. doi:10.1101/gad.210377.112.

Hajkova, Petra, Katia Ancelin, Tanja Waldmann, Nicolas Lacoste, Ulrike C Lange, Francesca Cesari, Caroline Lee, Genevieve Almouzni, Robert Schneider, and M Azim Surani. 2008. “Chromatin Dynamics during Epigenetic Reprogramming in the Mouse Germ Line.” *Nature* 452 (7189). Nature Publishing Group: 877–81. doi:10.1038/nature06714.

Hajkova, Petra, Sean J Jeffries, Caroline Lee, Nigel Miller, Stephen P Jackson, and M Azim Surani. 2010. “Genome-Wide Reprogramming in the Mouse Germ Line Entails the Base Excision Repair Pathway.” *Science (New York, N.Y.)* 329 (5987): 78–82. doi:10.1126/science.1187945.

Hall, Lisa L, Dawn M Carone, Alvin V Gomez, Heather J Kolpa, Meg Byron, Nitish Mehta, Frank O Fackelmayer, and Jeanne B Lawrence. 2014. “Stable C0T-1 Repeat RNA Is Abundant and Is Associated with Euchromatic Interphase Chromosomes.” *Cell* 156 (5): 907–19. doi:10.1016/j.cell.2014.01.042.

Hayakawa, Koji, Jun Ohgane, Satoshi Tanaka, Shintaro Yagi, and Kunio Shiota. 2014. “Oocyte-Specific Linker Histone H1foo Is an Epigenomic Modulator That Decondenses Chromatin and Impairs Pluripotency.” *Epigenetics* 7 (9): 1029–36. doi:10.4161/epi.21492.

Hemberger, Myriam, Wendy Dean, and Wolf Reik. 2009. “Epigenetic Dynamics of Stem Cells and Cell Lineage Commitment: Digging Waddington’s Canal.” *Nature Reviews. Molecular Cell Biology* 10 (8): 526–37.

Henckel, Amandine, Kazuhiko Nakabayashi, Lionel a. Sanz, Robert Feil, Kenichiro Hata, and Philippe Arnaud. 2009. “Histone Methylation Is Mechanistically Linked to DNA Methylation at Imprinting Control Regions in Mammals.” *Human Molecular Genetics* 18 (18): 3375–83. doi:10.1093/hmg/ddp277.

Henikoff, Steven, and M Mitchell Smith. 2015. “Histone Variants and Epigenetics.” In *Cancer Research*, 7:1–16. doi:10.1101/cshperspect.a019364.

Heras, Sara R, Sara Macias, Mireya Plass, Noemí Fernandez, David Cano, Eduardo Eyras, José L Garcia-Perez, and Javier F Cáceres. 2013. “The Microprocessor Controls the Activity of Mammalian Retrotransposons.” *Nature Structural & Molecular Biology* 20 (10): 1173–81. doi:10.1038/nsmb.2658.

Hirasawa, Ryutaro, Hatsune Chiba, Masahiro Kaneda, Shoji Tajima, En Li, Rudolf Jaenisch, and Hiroyuki Sasaki. 2008. “Maternal and Zygotic Dnmt1 Are Necessary and Sufficient for the Maintenance of DNA Methylation Imprints during Preimplantation Development.” *Genes & Development* 22 (12): 1607–16. doi:10.1101/gad.1667008.

Huang, Jiaqiang, Tao Fan, Qingsheng Yan, Heming Zhu, Stephen Fox, Haleem J Issaq, Lionel Best, Lisa Gangi, David Munroe, and Kathrin Muegge. 2004. “Lsh, an Epigenetic Guardian of Repetitive Elements.” *Nucleic Acids Research* 32 (17): 5019–28. doi:10.1093/nar/gkh821.

Inoue, Azusa, Sugako Ogushi, Mitinori Saitou, Masataka G Suzuki, and Fugaku Aoki. 2011. “Involvement of Mouse Nucleoplasmin 2 in the Decondensation of Sperm Chromatin after Fertilization.” *Biology of Reproduction* 85 (1): 70–77. doi:10.1095/biolreprod.110.089342.

Inoue, Azusa, and Yi Zhang. 2014. “Nucleosome Assembly Is Required for Nuclear Pore Complex Assembly in Mouse Zygotes.” *Nature Structural & Molecular Biology* 21 (7). Nature Publishing Group: 609–16. doi:10.1038/nsmb.2839.

Ishibashi, Toyotaka, Deanna Dryhurst, Kristie L Rose, Jeffrey Shabanowitz, Donald F Hunt, and Juan Ausió. 2009.

"Acetylation of Vertebrate H2A.Z and Its Effect on the Structure of the Nucleosome." *Biochemistry* 48 (22): 5007–17. doi:10.1021/bi900196c.

Jost, K. Laurence, Bianca Bertulat, and M. Cristina Cardoso. 2012. "Heterochromatin and Gene Positioning: Inside, Outside, Any Side?" *Chromosoma* 121 (6): 555–63. doi:10.1007/s00412-012-0389-2.

Kagansky, Alexander, Hernan Diego Folco, Ricardo Almeida, Alison L Pidoux, Abdelhalim Boukaba, Femke Simmer, Takeshi Urano, Georgina L Hamilton, and Robin C Allshire. 2009. "Synthetic Heterochromatin Bypasses RNAi and Centromeric Repeats to Establish Functional Centromeres." *Science (New York, N.Y.)* 324 (5935): 1716–19. doi:10.1126/science.1172026.

Kageyama, Sun Ichiro, Honglin Liu, Naoto Kaneko, Masatoshi Ooga, Masao Nagata, and Fugaku Aoki. 2007. "Alterations in Epigenetic Modifications during Oocyte Growth in Mice." *Reproduction* 133 (1): 85–94. doi:10.1530/REP-06-0025.

Kanellopoulou, C. 2005. "Dicer-Deficient Mouse Embryonic Stem Cells Are Defective in Differentiation and Centromeric Silencing." *Genes & Development* 19 (4): 489–501. doi:10.1101/gad.1248505.

Kano, Hiroki, Irene Godoy, Christine Courtney, Melissa R. Vetter, George L. Gerton, Eric M. Ostertag, and Haig H. Kazazian. 2009. "L1 Retrotransposition Occurs Mainly in Embryogenesis and Creates Somatic Mosaicism." *Genes and Development* 23 (11): 1303–12. doi:10.1101/gad.1803909.

Karimi, Mohammad M., Preeti Goyal, Irina a. Maksakova, Misha Bilenky, Danny Leung, Jie Xin Tang, Yoichi Shinkai, et al. 2011. "DNA Methylation and SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements, and Chimeric Transcripts in Mescs." *Cell Stem Cell* 8 (6). Elsevier Inc.: 676–87. doi:10.1016/j.stem.2011.04.004.

Kind, Jop, Ludo Pagie, Havva Ortabozkoyun, Shelagh Boyle, Sandra S de Vries, Hans Janssen, Mario Amendola, Leisha D Nolen, Wendy A Bickmore, and Bas van Steensel. 2013. "Single-Cell Dynamics of Genome-Nuclear Lamina Interactions." *Cell* 153 (1): 178–92. doi:10.1016/j.cell.2013.02.028.

Kouzarides, Tony. 2007. "Chromatin Modifications and Their Function." *Cell* 128 (4): 693–705. doi:10.1016/j.cell.2007.02.005.

Lepikhov, Konstantin, and Jörn Walter. 2004. "Differential Dynamics of Histone H3 Methylation at Positions K4 and K9 in the Mouse Zygote." *BMC Developmental Biology* 4 (1): 12. doi:10.1186/1471-213X-4-12.

Lewis, Peter W, Simon J Elsaesser, Kyung-Min Noh, Sonja C Stadler, and C David Allis. 2010. "Daxx Is an H3.3-Specific Histone Chaperone and Cooperates with ATRX in Replication-Independent Chromatin Assembly at Telomeres." *Proceedings Of The National Academy Of Sciences (Of The United States Of America)* 107 (32): 14075–80. doi:10.1073/pnas.1008850107.

Li, Jingfeng, Manoj Kannan, Anna L. Trivett, Hongling Liao, Xiaolin Wu, Keiko Akagi, and David E. Symer. 2014. "An Antisense Promoter in Mouse L1 Retrotransposon Open Reading Frame-1 Initiates Expression of Diverse Fusion Transcripts and Limits Retrotransposition." *Nucleic Acids Research* 42 (7): 4546–62. doi:10.1093/nar/gku091.

Li, Long-Cheng. 2014. "Chromatin Remodeling by the Small RNA Machinery in Mammalian Cells." *Epigenetics* 9 (1). Landes Bioscience: 45–52. doi:10.4161/epi.26830.

Li, Xiajun, Mitsuteru Ito, Fen Zhou, Neil Youngson, Xiaopan Zuo, Philip Leder, and Anne C. Ferguson-Smith. 2008. "A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints." *Developmental Cell* 15 (4): 547–57. doi:10.1016/j.devcel.2008.08.014.

Lin, Chih-Jen, Fong Ming Koh, Priscilla Wong, Conti Marco, and Miguel Ramalho-Santos. 2015. "Hira-Mediated

H3.3 Incorporation Is Required for DNA Replication and Ribosomal RNA Transcription in the Mouse Zygote." *Dev Cell.* 29 (6): 997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted.

Luo, Yi-Bo, Li Zhang, Zi-Li Lin, Jun-Yu Ma, Jialin Jia, Suk Namgoong, and Qing-Yuan Sun. 2015. "Distinct Subcellular Localization and Potential Role of LINE1-ORF1P in Meiotic Oocytes." *Histochemistry and Cell Biology* 1 (October). Springer Berlin Heidelberg: 1–12. doi:10.1017/CBO9781107415324.004.

Maison, Christèle, Delphine Bailly, Antoine H F M Peters, Jean-Pierre Quivy, Danièle Roche, Angela Taddei, Monika Lachner, Thomas Jenuwein, and Geneviève Almouzni. 2002. "Higher-Order Structure in Pericentric Heterochromatin Involves a Distinct Pattern of Histone Modification and an RNA Component." *Nature Genetics* 30 (3). Nature Publishing Group: 329–34. doi:10.1038/ng843.

Maison, Christèle, Delphine Bailly, Danièle Roche, Rocio Montes de Oca, Aline V Probst, Isabelle Vassias, Florent Dingli, et al. 2011. "SUMOylation Promotes de Novo Targeting of HP1α to Pericentric Heterochromatin." *Nature Genetics* 43 (3). Nature Publishing Group: 220–27. doi:10.1038/ng.765.

Malki, Safia, Godfried W. vanderHeijden, Kathryn a. O'Donnell, Sandra L. Martin, and Alex Bortvin. 2014. "A Role for Retrotransposon LINE-1 in Fetal Oocyte Attrition in Mice." *Developmental Cell* 29 (5). Elsevier Inc.: 521–33. doi:10.1016/j.devcel.2014.04.027.

Martens, Joost H A, Roderick J O'Sullivan, Ulrich Braunschweig, Susanne Opravil, Martin Radolf, Peter Steinlein, and Thomas Jenuwein. 2005. "The Profile of Repeat-Associated Histone Lysine Methylation States in the Mouse Epigenome." *The EMBO Journal* 24 (4): 800–812. doi:10.1038/sj.emboj.7600545.

Martin, Catherine, Vincent Brochard, Carole Migné, Daniele Zink, Pascale Debey, and Nathalie Beaujean. 2006. "Architectural Reorganization of the Nuclei upon Transfer into Oocytes Accompanies Genome Reprogramming." *Molecular Reproduction and Development* 73 (9): 1102–11. doi:10.1002/mrd.20506.

Martin, Sandra L. 2006. "The ORF1 Protein Encoded by LINE-1: Structure and Function during L1 Retrotransposition." *Journal of Biomedicine and Biotechnology* 2006: 1–6. doi:10.1155/JBB/2006/45621.

Matheson, Timothy D., and Paul D. Kaufman. 2015. "Grabbing the Genome by the NADs." *Chromosoma* 1. doi:10.1007/s00412-015-0527-8.

Meglicki, Maciej, Maciej Zientarski, and Ewa Borsuk. 2008. "Constitutive Heterochromatin during Mouse Oogenesis: The Pattern of Histone H3 Modifications and Localization of HP1?? And HP1?? Proteins." *Molecular Reproduction and Development* 75 (2): 414–28. doi:10.1002/mrd.20772.

Messerschmidt, Daniel M., Barbara B. Knowles, and Davor Solter. 2014. "DNA Methylation Dynamics during Epigenetic Reprogramming in the Germline and Preimplantation Embryos." *Genes and Development* 28 (8): 812–28. doi:10.1101/gad.234294.113.

Miyanari, Yusuke. 2014. "TAL Effector-Mediated Genome Visualization (TGV)." *Methods (San Diego, Calif.)* 69 (2). Elsevier Inc.: 1–7. doi:10.1016/j.ymeth.2014.03.030.

Morris, Kevin V, Simon W-L Chan, Steven E Jacobsen, and David J Looney. 2004. "Small Interfering RNA-Induced Transcriptional Gene Silencing in Human Cells." *Science (New York, N.Y.)* 305 (5688): 1289–92. doi:10.1126/science.1101372.

Nakamura, Toshinobu, Yoshikazu Arai, Hiroki Umehara, Masaaki Masuhara, Tohru Kimura, Hisaaki Taniguchi, Toshihiro Sekimoto, et al. 2007. "PGC7/Stella Protects against DNA Demethylation in Early Embryogenesis." *Nature Cell Biology* 9 (1): 64–71. doi:10.1038/ncb1519.

Nakatani, Tsunetoshi, Kazuo Yamagata, Tohru Kimura, Masaaki Oda, Hiroyuki Nakashima, Mayuko Hori, Yoichi Sekita, Tatsuhiko Arakawa, Toshinobu Nakamura, and Toru Nakano. 2015. "Stella Preserves Maternal

Chromosome Integrity by Inhibiting 5hmC-Induced γH2AX Accumulation." *EMBO Reports* 16 (5): 582–89. doi:10.15252/embr.201439427.

Németh, Attila, Ana Conesa, Javier Santoyo-Lopez, Ignacio Medina, David Montaner, Bálint Péterfia, Irina Solovei, Thomas Cremer, Joaquin Dopazo, and Gernot Längst. 2010. "Initial Genomics of the Human Nucleolus." *PLoS Genetics* 6 (3). Public Library of Science: e1000889. doi:10.1371/journal.pgen.1000889.

Ogushi, S., C. Palmieri, H. Fulka, M. Saitou, T. Miyano, and J. Fulka. 2008. "The Maternal Nucleolus Is Essential for Early Embryonic Development in Mammals." *Science* 319 (5863): 613–16. doi:10.1126/science.1151276.

Ogushi, Sugako, and Mitinori Saitou. 2010. "The Nucleolus in the Mouse Oocyte Is Required for the Early Step of Both Female and Male Pronucleus Organization." *The Journal of Reproduction and Development* 56 (5): 495–501. doi:10.1262/jrd.09-184H.

Ohnishi, Yusuke, Yasushi Totoki, Atsushi Toyoda, Toshiaki Watanabe, Yasuhiro Yamamoto, Katsushi Tokunaga, Yoshiyuki Sakaki, Hiroyuki Sasaki, Hirohiko Hohjoh, and Mortonarb. 2009. "Small RNA Class Transition from siRNA/piRNA to miRNA during Pre-Implantation Mouse Development." *Nucleic Acids Research* 38 (15): 1–16. doi:10.1038/nature06814.

Okano, Masaki, Daphne W Bell, Daniel A Haber, and En Li. 1999. "DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development." *Cell* 99 (3): 247–57. doi:10.1016/S0092-8674(00)81656-6.

Ooi, Steen K T, Anne H O'Donnell, and Timothy H Bestor. 2009. "Mammalian Cytosine Methylation at a Glance." *Journal of Cell Science* 122 (Pt 16): 2787–91. doi:10.1242/jcs.015123.

Peaston, Anne E., Alexei V. Evsikov, Joel H. Graber, Wilhelmine N. de Vries, Andrea E. Holbrook, Davor Solter, and Barbara B. Knowles. 2004. "Retrotransposons Regulate Host Genes in Mouse Oocytes and Preimplantation Embryos." *Developmental Cell* 7 (4): 597–606. doi:10.1016/j.devcel.2004.09.004.

Peters, Antoine H F M, Dónal O'Carroll, Harry Scherthan, Karl Mechtler, Stephan Sauer, Christian Schöfer, Klara Weipoltshammer, et al. 2001. "Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability." *Cell* 107 (3): 323–37. doi:10.1016/S0092-8674(01)00542-6.

Pezic, Dubravka, Sergei A Manakov, Ravi Sachidanandam, and Alexei A. Aravin. 2014. "piRNA Pathway Targets Active LINE1 Elements to Establish the Repressive H3K9me3 Mark in Germ Cells." *Genes & Development* 28 (13): 1410–28. doi:10.1101/gad.240895.114.

Pinheiro, Inês, Raphaël Margueron, Nicholas Shukeir, Michael Eisold, Christoph Fritzsch, Florian M. Richter, Gerhard Mittler, et al. 2012. "Prdm3 and Prdm16 Are H3K9me1 Methyltransferases Required for Mammalian Heterochromatin Integrity." *Cell* 150 (5): 948–60. doi:10.1016/j.cell.2012.06.048.

Pittoggi, Carmine, Ilaria Sciamanna, Elisabetta Mattei, Rosanna Beraldi, Anna Maria Lobascio, Antonello Mai, Maria Giovanna Quaglia, Rodolfo Lorenzini, and Corrado Spadafora. 2003. "Role of Endogenous Reverse Transcriptase in Murine Early Embryo Development." *Molecular Reproduction and Development* 66 (3): 225–36. doi:10.1002/mrd.10349.

Popp, Christian, Wendy Dean, Suhua Feng, Shawn J Cokus, Simon Andrews, Matteo Pellegrini, Steven E Jacobsen, and Wolf Reik. 2010. "Genome-Wide Erasure of DNA Methylation in Mouse Primordial Germ Cells Is Affected by AID Deficiency." *Nature* 463 (7284). Macmillan Publishers Limited. All rights reserved: 1101–5. doi:10.1038/nature08829.

Posfai, Eszter, Rico Kunzmann, Vincent Brochard, Juliette Salvaing, Erik Cabuy, Tim C. Roloff, Zichuan Liu, et al.

2012. “Polycomb Function during Oogenesis Is Required for Mouse Embryonic Development.” *Genes and Development* 26 (9): 920–32. doi:10.1101/aad.188094.112.

Postepska-Igielska, Anna, Damir Krunic, Nina Schmitt, Karin M Greulich-Bode, Petra Boukamp, and Ingrid Grummt. 2013. “The Chromatin Remodelling Complex NoRC Safeguards Genome Stability by Heterochromatin Formation at Telomeres and Centromeres.” *EMBO Reports* 14 (8). Nature Publishing Group: 704–10. doi:10.1038/embor.2013.87.

Probst, Aline.V., Ikuhiro Okamoto, Miguel Casanova, Fatima El Marjou, Patricia Le Baccon, and Geneviève Almouzni. 2010. “A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development.” *Developmental Cell* 19 (4): 625–38. doi:10.1016/j.devcel.2010.09.002.

Probst, Aline V., and Geneviève Almouzni. 2008. “Pericentric Heterochromatin: Dynamic Organization during Early Development in Mammals.” *Differentiation* 76 (1): 15–23. doi:10.1111/j.1432-0436.2007.00220.x.

Puschendorf, Mareike, Rémi Terranova, Erwin Boutsma, Xiaohong Mao, Kyo-ichi Isono, Urszula Brykczynska, Carolin Kolb, et al. 2008. “PRC1 and Suv39h Specify Parental Asymmetry at Constitutive Heterochromatin in Early Mouse Embryos.” *Nature Genetics* 40 (4): 411–20. doi:10.1038/ng.99.

Rangasamy, Danny, Leise Berven, Patricia Ridgway, and David John Tremethick. 2003. “Pericentric Heterochromatin Becomes Enriched with H2A.Z during Early Mammalian Development.” *The EMBO Journal* 22 (7): 1599–1607. doi:10.1093/emboj/cdg160.

Rangasamy, Danny, Ian Greaves, and David J Tremethick. 2004. “RNA Interference Demonstrates a Novel Role for H2A.Z in Chromosome Segregation.” *Nature Structural & Molecular Biology* 11 (7): 650–55. doi:10.1038/nsmb786.

Rapkin, Lindsy M., Kashif Ahmed, Stanimir Dulev, Ren Li, Hiroshi Kimura, Alexander M. Ishov, and David P. Bazett-Jones. 2015. “The Histone Chaperone DAXX Maintains the Structural Organization of Heterochromatin Domains.” *Epigenetics & Chromatin* 8 (1). BioMed Central: 44. doi:10.1186/s13072-015-0036-2.

Rathke, Christina, Willy M Baarends, Stephan Awe, and Renate Renkawitz-Pohl. 2014. “Chromatin Dynamics during Spermiogenesis.” *Biochimica et Biophysica Acta* 1839 (3): 155–68. doi:10.1016/j.bbagrm.2013.08.004.

Redon, Christophe, Duane Pilch, Emmy Rogakou, Olga Sedelnikova, Kenneth Newrock, and William Bonner. 2002. “Histone H2A Variants H2AX and H2AZ.” *Current Opinion in Genetics and Development* 12 (2): 162–69. doi:10.1016/S0959-437X(02)00282-4.

Ren, J., V. Briones, S. Barbour, W. Yu, Y. Han, M. Terashima, and K. Muegge. 2015. “The ATP Binding Site of the Chromatin Remodeling Homolog Lsh Is Required for Nucleosome Density and de Novo DNA Methylation at Repeat Sequences.” *Nucleic Acids Research* 43 (3): 1444–55. doi:10.1093/nar/gku1371.

Rhind, Nicholas, and David M Gilbert. 2013. “DNA Replication Timing.” *Cold Spring Harbor Perspectives in Biology* 5 (8): a010132. doi:10.1101/cshperspect.a010132.

Rodríguez-Campos, Antonio, and Fernando Azorín. 2007. “RNA Is an Integral Component of Chromatin That Contributes to Its Structural Organization.” Edited by Giacomo Cavalli. *PLoS ONE* 2 (11). Public Library of Science: e1182. doi:10.1371/journal.pone.0001182.

Rosa, Stefanie, and Peter Shaw. 2013. “Insights into Chromatin Structure and Dynamics in Plants.” *Biology* 2 (4). Multidisciplinary Digital Publishing Institute: 1378–1410. doi:10.3390/biology2041378.

Rowe, Helen M, Johan Jakobsson, Daniel Mesnard, Jacques Rougemont, Séverine Reynard, Tugce Aktas, Pierre V Maillard, et al. 2010. “KAP1 Controls Endogenous Retroviruses in Embryonic Stem Cells.” *Nature* 463 (7278): 237–40. doi:10.1038/nature08674.

Sadowski, I, J Ma, S Triezenberg, and M Ptashne. 1988. “GAL4-VP16 Is an Unusually Potent Transcriptional Activator.” *Nature*. doi:10.1038/335563a0.

Saksouk, N, T K Barth, C Ziegler-Birling, N Olova, a Nowak, E Rey, J Mateos-Langerak, et al. 2014. “Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation.” *Mol Cell* 56 (4): 580–94. doi:10.1016/j.molcel.2014.10.001.

Sanjana, Neville E, Le Cong, Yang Zhou, Margaret M Cunniff, Guoping Feng, and Feng Zhang. 2012. “A Transcription Activator-like Effector Toolbox for Genome Engineering.” *Nature Protocols* 7 (1): 171–92. doi:10.1038/nprot.2011.431.

Santenard, Angèle, Céline Ziegler-Birling, Marc Koch, Làszlò Tora, Andrew J Bannister, and Maria-Elena Torres-Padilla. 2010. “Heterochromatin Formation in the Mouse Embryo Requires Critical Residues of the Histone Variant H3.3.” *Nature Cell Biology* 12 (9). Nature Publishing Group: 853–62. doi:10.1038/ncb2089.

Santos, Fátima, Antoine H. Peters, Arie P. Otte, Wolf Reik, and Wendy Dean. 2005. “Dynamic Chromatin Modifications Characterise the First Cell Cycle in Mouse Embryos.” *Developmental Biology* 280 (1): 225–36. doi:10.1016/j.ydbio.2005.01.025.

Sarcinella, Elizabeth, Philip C Zuzarte, Priscilla N I Lau, Ryan Draker, and Peter Cheung. 2007. “Monoubiquitylation of H2A.Z Distinguishes Its Association with Euchromatin or Facultative Heterochromatin.” *Molecular and Cellular Biology* 27 (18): 6457–68. doi:10.1128/MCB.00241-07.

Sarraf, Shireen A, and Irina Stancheva. 2004. “Methyl-CpG Binding Protein MBD1 Couples Histone H3 Methylation at Lysine 9 by SETDB1 to DNA Replication and Chromatin Assembly.” *Molecular Cell* 15 (4): 595–605. doi:10.1016/j.molcel.2004.06.043.

Sasaki, Hiroyuki, and Yasuhisa Matsui. 2008. “Epigenetic Events in Mammalian Germ-Cell Development: Reprogramming and beyond.” *Nature Reviews Genetics* 2008 (2). Nature Publishing Group: 129–40. doi:10.1038/nrg2295.

Schotta, Gunnar, Monika Lachner, Kavitha Sarma, Anja Ebert, Roopsha Sengupta, Gunter Reuter, Danny Reinberg, and Thomas Jenuwein. 2004. “A Silencing Pathway to Induce H3-K9 and H4-K20 Trimethylation at Constitutive Heterochromatin.” *Genes & Development* 18 (11): 1251–62. doi:10.1101/gad.300704.

Schultz, R. M. 2002. “The Molecular Foundations of the Maternal to Zygotic Transition in the Preimplantation Embryo.” *Human Reproduction Update* 8 (4): 323–31. doi:10.1093/humupd/8.4.323.

Seki, Yoshiyuki, Masashi Yamaji, Yukihiro Yabuta, Mitsue Sano, Mayo Shigeta, Yasuhisa Matsui, Yumiko Saga, Makoto Tachibana, Yoichi Shinkai, and Mitinori Saitou. 2007. “Cellular Dynamics Associated with the Genome-Wide Epigenetic Reprogramming in Migrating Primordial Germ Cells in Mice.” *Development (Cambridge, England)* 134 (14): 2627–38. doi:10.1242/dev.005611.

Solovei, Irina, Audrey S. Wang, Katharina Thanisch, Christine S. Schmidt, Stefan Krebs, Monika Zwerger, Tatiana V. Cohen, et al. 2013. “LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation.” *Cell* 152 (3). Elsevier Inc.: 584–98. doi:10.1016/j.cell.2013.01.009.

Speed, R M. 1982. “Meiosis in the Foetal Mouse Ovary. I. An Analysis at the Light Microscope Level Using Surface-Spreading.” *Chromosoma* 85 (3): 427–37. http://www.ncbi.nlm.nih.gov/pubmed/6180868.

Tachibana, M. 2002. “G9a Histone Methyltransferase Plays a Dominant Role in Euchromatic Histone H3 Lysine 9 Methylation and Is Essential for Early Embryogenesis.” *Genes & Development* 16 (14): 1779–91. doi:10.1101/gad.989402.

Tagami, Hideaki, Dominique Ray-Gallet, Geneviève Almouzni, and Yoshihiro Nakatani. 2004. “Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis.” *Cell* 116 (1): 51–61. doi:10.1016/S0092-8674(03)01064-X.

Tardat, Mathieu, Mareike Albert, Rico Kunzmann, Zichuan Liu, Lilia Kaustov, Raphael Thierry, Shili Duan, Urszula Brykczynska, Cheryl H. Arrowsmith, and Antoine H.F.M. Peters. 2015. “Cbx2 Targets PRC1 to Constitutive Heterochromatin in Mouse Zygotes in a Parent-of-Origin-Dependent Manner.” *Molecular Cell* 58 (1). Elsevier Inc.: 157–71. doi:10.1016/j.molcel.2015.02.013.

Terranova, Rémi, Shihori Yokobayashi, Michael B. Stadler, Arie P. Otte, Maarten van Lohuizen, Stuart H. Orkin, and Antoine H F M Peters. 2008. “Polycomb Group Proteins Ezh2 and Rnf2 Direct Genomic Contraction and Imprinted Repression in Early Mouse Embryos.” *Developmental Cell* 15 (5): 668–79. doi:10.1016/j.devcel.2008.08.015.

Therizols, Pierre, Robert S Illingworth, Celine Courilleau, Shelagh Boyle, Andrew J Wood, and Wendy A Bickmore. 2014. “Chromatin Decondensation Is Sufficient to Alter Nuclear Organization in Embryonic Stem Cells.” *Science (New York, N.Y.)* 346 (6214). American Association for the Advancement of Science: 1238–42. doi:10.1126/science.1259587.

Toltzis, P., T. Mourton, and T. Magnuson. 1993. “Effect of Zidovudine on Preimplantation Murine Embryos.” *Antimicrobial Agents and Chemotherapy* 37 (8): 1610–13. doi:10.1128/AAC.37.8.1610.

Torres-Padilla, Maria-Elena, Andrew J. Bannister, Paul J. Hurd, Tony Kouzarides, and Magdalena Zernicka-Goetz. 2006. “Dynamic Distribution of the Replacement Histone Variant H3.3 in the Mouse Oocyte and Preimplantation Embryos.” *The International Journal of Developmental Biology* 50 (Next): 455–61. doi:10.1387/ijdb.052073mt.

Towbin, Benjamin D., Cristina González-Aguilera, Ragna Sack, Dimos Gaidatzis, Véronique Kalck, Peter Meister, Peter Askjaer, and Susan M. Gasser. 2012. “Step-Wise Methylation of Histone H3K9 Positions Heterochromatin at the Nuclear Periphery.” *Cell* 150 (5): 934–47. doi:10.1016/j.cell.2012.06.051.

Tsumura, Akiko, Tomohiro Hayakawa, Yuichi Kumaki, Shin Ichiro Takebayashi, Morito Sakaue, Chisa Matsuoka, Kunitada Shimotohno, et al. 2006. “Maintenance of Self-Renewal Ability of Mouse Embryonic Stem Cells in the Absence of DNA Methyltransferases Dnmt1, Dnmt3a and Dnmt3b.” *Genes to Cells* 11 (7): 805–14. doi:10.1111/j.1365-2443.2006.00984.x.

van der Heijden, G W, A A H A Derijck, L Ramos, M Giele, J van der Vlag, and P de Boer. 2006. “Transmission of Modified Nucleosomes from the Mouse Male Germline to the Zygote and Subsequent Remodeling of Paternal Chromatin.” *Developmental Biology* 298 (2): 458–69. doi:10.1016/j.ydbio.2006.06.051.

van der Heijden, Godfried W., Jürgen W. Dieker, Alwin a.H.a. Derijck, Sylviane Muller, Jo H.M. Berden, Didi D.M. Braat, Johan van der Vlag, and Peter de Boer. 2005. “Asymmetry in Histone H3 Variants and Lysine Methylation between Paternal and Maternal Chromatin of the Early Mouse Zygote.” *Mechanisms of Development* 122 (9): 1008–22. doi:10.1016/j.mod.2005.04.009.

Vincent, Stéphane D, N Ray Dunn, Roger Sciammas, Miriam Shapiro-Shalef, Mark M Davis, Kathryn Calame, Elizabeth K Bikoff, and Elizabeth J Robertson. 2005. “The Zinc Finger Transcriptional Repressor Blimp1/Prdm1 Is Dispensable for Early Axis Formation but Is Required for Specification of Primordial Germ Cells in the Mouse.” *Development (Cambridge, England)* 132 (6): 1315–25. doi:10.1242/dev.01711.

Wang, Lu, Jun Zhang, Jialei Duan, Xinxing Gao, Wei Zhu, Xingyu Lu, Lu Yang, et al. 2014. “Programming and Inheritance of Parental DNA Methylomes in Mammals.” *Cell* 157 (4). Elsevier Inc.: 979–91. doi:10.1016/j.cell.2014.04.017.

Wen, Duancheng, Laura A Banaszynski, Zev Rosenwaks, C David Allis, and Shahin Rafii. 2014. “H3.3 Replacement Facilitates Epigenetic Reprogramming of Donor Nuclei in Somatic Cell Nuclear Transfer Embryos.” *Nucleus* 5 (5): 369–75. doi:10.4161/nucl.36231.

Wennekamp, Sebastian, Sven Mesecke, François Nédélec, and Takashi Hiiragi. 2013. “A Self-Organization Framework for Symmetry Breaking in the Mammalian Embryo.” *Nature Reviews. Molecular Cell Biology* 14 (7). Nature Publishing Group: 452–59. doi:10.1038/nrm3602.

Western, Patrick S, Denise C Miles, Jocelyn A van den Bergen, Matt Burton, and Andrew H Sinclair. 2008. “Dynamic Regulation of Mitotic Arrest in Fetal Male Germ Cells.” *Stem Cells (Dayton, Ohio)* 26 (2): 339–47. doi:10.1634/stemcells.2007-0622.

Westhorpe, Fg, and Af Straight. 2014. “The Centromere: Epigenetic Control of Chromosome Segregation during Mitosis.” *Cold Spring Harb Perspect Biol.*, 1–26. doi:10.1101/cshperspect.a015818.

Wong, Lee H, Kate H Brettingham-Moore, Lyn Chan, Julie M Quach, Melissa A Anderson, Emma L Northrop, Ross Hannan, et al. 2007. “Centromere RNA Is a Key Component for the Assembly of Nucleoproteins at the Nucleolus and Centromere.” *Genome Research* 17 (8): 1146–60. doi:10.1101/gr.6022807.

Wongtawan, Tuempong, Jane E Taylor, Kirstie A Lawson, Ian Wilmut, and Sari Pennings. 2011. “Histone H4K20me3 and HP1α Are Late Heterochromatin Markers in Development, but Present in Undifferentiated Embryonic Stem Cells.” *Journal of Cell Science* 124 (Pt 11): 1878–90. doi:10.1242/jcs.080721.

Wossidlo, Mark, Julia Arand, Vittorio Sebastiano, Konstantin Lepikhov, Michele Boiani, Richard Reinhardt, Hans Schöler, and Jörn Walter. 2010. “Dynamic Link of DNA Demethylation, DNA Strand Breaks and Repair in Mouse Zygotes.” *The EMBO Journal* 29 (11): 1877–88. doi:10.1038/emboj.2010.80.

Wossidlo, Mark, Toshinobu Nakamura, Konstantin Lepikhov, C. Joana Marques, Valeri Zakhartchenko, Michele Boiani, Julia Arand, Toru Nakano, Wolf Reik, and Jörn Walter. 2011. “5-Hydroxymethylcytosine in the Mammalian Zygote Is Linked with Epigenetic Reprogramming.” *Nature Communications* 2: 241. doi:10.1038/ncomms1240.

Wu, Xudong, Jens Vilstrup Johansen, and Kristian Helin. 2013. “Fbxl10/Kdm2b Recruits Polycomb Repressive Complex 1 to CpG Islands and Regulates H2A Ubiquitylation.” *Molecular Cell* 49 (6): 1134–46. doi:10.1016/j.molcel.2013.01.016.

Yamaji, Masashi, Yoshiyuki Seki, Kazuki Kurimoto, Yukihiro Yabuta, Mihoko Yuasa, Mayo Shigeta, Kaori Yamanaka, Yasuhide Ohinata, and Mitinori Saitou. 2008. “Critical Function of Prdm14 for the Establishment of the Germ Cell Lineage in Mice.” *Nature Genetics* 40 (8): 1016–22. doi:10.1038/ng.186.

Ye, Q, and H J Worman. 1996. “Interaction between an Integral Protein of the Nuclear Envelope Inner Membrane and Human Chromodomain Proteins Homologous to Drosophila HP1.” *The Journal of Biological Chemistry* 271 (25): 14653–56. http://www.ncbi.nlm.nih.gov/pubmed/8663349.

Zhang, Feng, Le Cong, Simona Lodato, Sriram Kosuri, George M Church, and Paola Arlotta. 2011. “Efficient Construction of Sequence-Specific TAL Effectors for Modulating Mammalian Transcription.” *Nature Biotechnology* 29 (2): 149–53. doi:10.1038/nbt.1775.

Zhu, Yunfeng, Fan Feng, Jiyun Yu, Bin Song, Mingming Hu, Xudong Gao, Yue Wang, and Qing Zhang. 2013. “L1-

ORF1p, a Smad4 Interaction Protein, Promotes Proliferation of HepG2 Cells and Tumorigenesis in Mice." *DNA and Cell Biology* 32 (9): 531–40. doi:10.1089/dna.2013.2097.

Ziegler-Birling, Celine, Anne Helmrich, Lszl Tora, and Maria-Elena Torres-Padilla. 2009. "Distribution of p53 Binding Protein 1 (53BP1) and Phosphorylated H2A.X during Mouse Preimplantation Development in the Absence of DNA Damage." *The International Journal of Developmental Biology* 53 (7): 1003–11. doi:10.1387/ijdb.082707cz.

Joanna Weronika JACHOWICZ

Molecular mechanisms underlying heterochromatin formation in the mouse embryo

Résumé

heterochromatin, répétitions péricentriques, L1 retrotransposon, epigenetics, pluripotence, embryon de souris

Afin d'étudier la formation de l'hétérochromatine dans l'embryon préimplantatoire de souris, je me suis concentrée sur deux régions génétiques différentes - répétitions péricentriques et L1 éléments transposables - dans le but notamment de découvrir les mécanismes qui conduisent à la répression et le rôle distinct qu'ils peuvent jouer pendant le processus de développement et la division cellulaire. Mes expériences montrent que l'organisation spatiale spécifique des domaines péricentriques est essentielle pour leur répression ainsi que pour leur organisation correcte. De plus, mes résultats suggèrent que les défauts d'organisation de l'hétérochromatine conduisent à des défauts de division cellulaire et de prolifération. La seconde partie de ma thèse montre que la réglementation stricte de L1 éléments transposables est nécessaire pour le développement préimplantatoire d'embryons de souris. En outre, représente la première tentative pour élucider la biologie des éléments L1 dans l'embryon précoce de souris par l'utilisation de modificateurs de transcription ciblés spécifiquement.

Résumé en anglais

heterochromatin, pericentric repeats, L1 retrotransposones, epigenetics, pluripotency, mouse preimplantation embryos

To study the formation of heterochromatin in mouse preimplantation embryo, I focused on two different genetic regions – pericentric repeats and L1 transposable elements - in order to investigate the mechanisms that lead to their repression and the distinct role that these regions can play during the process of development and cell division. My experiments show that the specific spatial organization of pericentric domains is essential for their repression and for their correct organization. Moreover, my findings suggest that defects in organization of heterochromatin lead to improper cell division and proliferation. The second part of my thesis shows that the tight regulation of L1 transposable elements is required for the preimplantation development of mouse embryos. Additionally, it is the first attempt to elucidate the biology of L1 elements in the early mouse embryo through the use of targeted transcription modifiers.