Molecular mechanisms underlying the regulation of nuclear organization in early mouse embryos

Vorgelegt von

Mrinmoy Pal

Munich, 2025

Diese Dissertation wurde angefertigt

unter der Leitung von **Prof. Dr. Maria-Elena Torres-Padilla**

am Institut für Epigenetik und Stammzellen

des Helmholtz Zentrum München

Erstgutachter: Prof. Dr. Maria-Elena Torres-Padilla

Zweitgutachter: Prof. Dr. Heinrich Leonhardt

Tag der Einreichung: 13.01.2025

Tag der mündlichen Prüfung: 09.07.2025

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorliegende Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

Ich erkläre weiter, dass ich mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen habe.

Mrinmoy Pal

München, den 13.01.2025

Table of contents

5

Abbreviations

3C	Chromosome conformation capture
3D	Three dimensional
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
BAF	Barrier to autointegration nuclear assembly factor 1
CDK9	Cyclin-dependent kinase 9
CDX2	Caudal-type homeobox protein 2
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
CTCF	CCCTC-binding factor protein
Dam	E. coli DNA adenine methyltransferase protein
DamID	DNA adenine methyltransferase identification
DAPI	4,6-diamidino-2-phenylindole
DPPA3	Developmental pluripotency associated 3 protein
DRB	5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
DUX	Double homeobox protein
E13.5	Embryonic day 13.5
E7.25	Embryonic day 7.25
EED	Embryonic ectoderm development (polycomb) protein
EHMT2	Euchromatic histone lysine methyltransferase 2 protein
ESCs	Embryonic stem cells
EZH2	Enhancer of zeste 2 (polycomb) protein
FGO(s)	Fully-grown oocyte(s)
FRAP	Fluorescence recovery after photobleaching
GO(s)	Growing oocyte(s)
GV	Germinal vesicle
H3.1/3.2	Histone 3.1 or Histone 3.2 (variants of Histone 3)
H3K27ac	acetylation of lysine 27 on the histone H3 tail
H3K27me3	tri-methylation of lysine 27 on the histone H3 tail
H3K36me2	di-methylation of lysine 36 on the histone H3 tail
H3K36me3	tri-methylation of lysine 36 on the histone H3 tail
H3K4me1	mono-methylation of lysine 4 on the histone H3 tail

H3K4me2	di-methylation of lysine 4 on the histone H3 tail
H3K4me3	tri-methylation of lysine 4 on the histone H3 tail
H3K9me2	di-methylation of lysine 9 on the histone H3 tail
H3K9me3	tri-methylation of lysine 9 on the histone H3 tail
HDAC3	Histone deacetylase 3
Hi-C	High-throughput chromosome conformation capture
HP1	Heterochromatin protein 1
ICM	Inner cell mass
IF	Immunofluorescence staining
iLAD(s)	inter lamina-associated domain(s) or inter-LADs
KDM5A/B	Lysine demethylase 5 A/B
LADs	Lamina-associated domain(s)
LAP2β	Lamina-associated polypeptide 2beta
LINE-1	Long interspersed nuclear element-1
MERVL	Murine endogenous retrovirus-L
MZT	Maternal-to-zygotic transition
NAD(s)	Nucleolus-associated domain(s)
NL	Nuclear lamina
NLBs	Nucleolar-like bodies
NPBs	Nucleolar precursor bodies
NSD1	Nuclear receptor binding SET-domain protein 1
NSN	Non-surrounded nucleolus oocytes
OBOX	Oocyte-specific homeobox proteins
OCT4	POU class 5 homeobox 1 protein
P10 or GO-P10	Growing oocytes obtained at postnatal day 10
CBP	CREB binding protein with acetyltransferase activity
PAD(s)	Polycomb-associated domain(s)
PGCs	Primordial germ cells
Pol II	RNA polymerase II
PRC	Polycomb repressive complex
PRDM14	PR/SET domain 14 protein
RIF1	Replication timing regulatory factor 1 protein
RT	Replication timing
SCNT	Somatic cell nuclear transfer

SETD2	SET domain containing 2
SINE-B1	Short interspersed nuclear element B1
SIRT1/6	Sirtuin family deacetylase 1/6
SN	Surrounded nucleolus oocyte
SOX2	SRY (sex determining region Y) box transcription factor 2
SUV39H1/2	Suppressor of variegation 3-9 homolog 1/2
TAD(s)	Topologically associating domain(s)
TCF3/12	Transcription factor 3/12 protein
TET3	Tet methylcytosine Dioxygenase 3 protein
TPR	Translocated promoter region protein (nuclear pore)
ZGA	Zygotic genome activation

List of publications

Published manuscripts

Pal M, Kind J, Torres-Padilla ME. DamID to map genome-protein interactions in preimplantation mouse embryos. Methods Mol Biol. 2021; 2214:265-282.

Liu B, He Y, Wu X, Lin Z, Ma J, Qiu Y, Xiang Y, Kong F, Lai F, **Pal M**, Wang P, Ming J, Zhang B, Wang Q, Wu J, Xia W, Shen W, Na J, Torres-Padilla ME, Li J, Xie W. Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators. Nat Cell Biol. 2024 Jun; 26(6):962-974.

Pal M, Altamirano-Pacheco L, Schauer T, Torres-Padilla ME. Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity. Genes Dev. 2023 Oct 1; 37(19-20):901-912.

Nakatani T, Schauer T, Altamirano-Pacheco L, Klein KN, Ettinger A, **Pal M**, Gilbert DM, Torres-Padilla ME. Emergence of replication timing during early mammalian development. Nature. 2024 Jan; 625(7994):401-409.

List of unpublished manuscripts

Nakatani T, Schauer T, **Pal M**, Ettinger A, Altamirano-Pacheco L, Zorn J, Gilbert DM, Torres-Padilla ME. (*Unpublished manuscript, under review*). RIF1 regulates the consolidation of replication timing in early mouse embryos independently of changes in nuclear organization towards the nuclear lamina.

Pal M, Schauer T, Burton A, Nakatani T, Pecori F, Hernandez-Gimenez A, Nadelson I, Marti-Renom MA, Torres-Padilla ME. (*Unpublished manuscript, under review*). The establishment of nuclear organization in mouse embryos is orchestrated by multiple epigenetic pathways.

Statement of contribution as co-author

I hereby state that my contribution to the publication:

Liu B, He Y, Wu X, Lin Z, Ma J, Qiu Y, Xiang Y, Kong F, Lai F, **Pal M**, Wang P, Ming J, Zhang B, Wang Q, Wu J, Xia W, Shen W, Na J, Torres-Padilla ME, Li J, Xie W. Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators. Nat Cell Biol. 2024 Jun; 26(6):962-974.

consisted in performing LaminB1 DamID in P10 growing oocytes, analyzing the DamID-seq data and helping in writing the corresponding section of the manuscript.

Mrinmoy Pal

Mrinmoy Pal
Munich, 05.12.2024

Confirmation of contribution

I hereby confirm the statement of contribution reproduced above is both truthful and accurate.

Prof. Dr. Maria-Elena Torres-Padilla
Munich, 05.12.2024

Statement of contribution as co-author

I hereby state that my contribution to the publication:

Nakatani T, Schauer T, Altamirano-Pacheco L, Klein KN, Ettinger A, **Pal M**, Gilbert DM, Torres-Padilla ME. Emergence of replication timing during early mammalian development. Nature. 2024 Jan; 625(7994):401-409.

consisted in performing KDM5B overexpression in embryo for H3K4me3 immunostaining and single-cell Repli-seq sample collection at the 2-cell stage.

Mrinmoy Pal

Mrinmoy Pal

Munich, 05.12.2024

Confirmation of contribution

I hereby confirm the statement of contribution reproduced above is both truthful and accurate.

Prof. Dr. Maria-Elena Torres-Padilla

Munich, 05.12.2024

Statement of contribution as co-author

I hereby state that my contribution to the publication:

Nakatani T, Schauer T, **Pal M**, Ettinger A, Altamirano-Pacheco L, Zorn J, Gilbert DM, Torres-Padilla ME. (Unpublished manuscript, under review). RIF1 regulates the consolidation of replication timing in early mouse embryos independently of changes in nuclear organization towards the nuclear lamina.

consisted in performing LaminB1 DamID in 4-cell and 8-cell stage embryos (siControl and siRif1), contributing to the bioinformatic analysis of DamID-seq data and to the writing of the corresponding section of the manuscript.
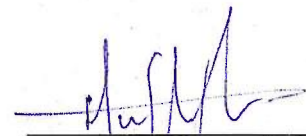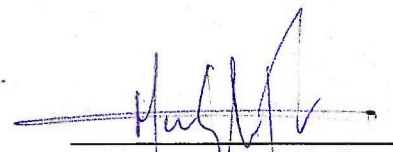
Mrinmoy Pal

Mrinmoy Pal
Munich, 05.12.2024

Confirmation of contribution

I hereby confirm the statement of contribution reproduced above is both truthful and accurate.

Prof. Dr. Maria-Elena Torres-Padilla
Munich, 05.12.2024

Summary

Mammalian gametogenesis and early development involve extensive epigenomic reorganization, offering a unique opportunity to address important yet underexplored questions central to reproductive biology and regenerative medicine. How does the maternal germline reprogram to produce a fertilization-competent egg? How is the epigenome established following fertilization? What molecular pathways orchestrate this process *in vivo*? What are the dependencies between the different layers of the epigenome? How do these dynamic changes functionally influence developmental plasticity and cell fate? My dissertation investigates these questions using mouse oocytes and preimplantation embryos as a model for early mammalian development. First, I adapted the low-input DamID technique to map the genome interactions with the nuclear lamina across different developmental stages, focusing on lamina-associated domains or LADs (**Part I**). Building on our previous work showing that LAD structures are absent in fully grown oocytes, we find that autosomal LADs are already undetectable in growing oocytes. These gene desert regions, usually heterochromatic in other cell types, contain oocyte-specific enhancer elements that regulate folliculogenesis (**Part II**). After fertilization, LADs undergo gradual but dynamic reorganization during the maternal-to-zygotic transition, both after the first mitosis and throughout the progression of the second cell cycle. This repositioning correlates with the expression of genes and transposable elements in 2-cell stage embryos. Inhibition of transcription during zygotic genome activation (ZGA) impairs the correct rearrangement of the LADs, leading to atypical features of lamina-associated chromatin (**Part III**). Next, in a collaborative work, we used single-cell Repli-seq to study the establishment of replication timing (RT) during mammalian embryogenesis. DNA replication occurs according to a less defined pattern in the zygote, and the RT program gradually consolidates with developmental progression. Our findings suggest that LAD formation precedes and potentially predisposes the partitioning of the genome into early and late replicating domains (**Part IV**). We identify RIF1 as a key regulator of replication timing consolidation *in vivo*, with its depletion resulting in a less coordinated RT program in 4-cell stage embryos and beyond. Intriguingly, the changes in RT in RIF1-depleted embryos are uncoupled from changes in genome-lamina

association (**Part V**). Lastly, a significant part of my work involved screening for molecular pathways that regulate *de novo* LAD establishment. This revealed chromatin pathways whose disruption significantly altered nuclear architecture in zygotes and 2-cell stage embryos. Although LADs are not inherited from oocytes, my work suggests that the maternal germline carries epigenetic bookmarking to guide the establishment of nuclear organization in zygotes. Our observations suggest that the absence of a constitutive heterochromatin pathway permits the distinctive LAD fragmentation at the 2-cell stage to coexist with a non-canonical chromatin landscape. We propose that LAD boundaries are reorganized based on positional information from H3K4me3 and H3K9me3 domains which counteract each other. Remarkably, the initial establishment of LADs in zygotes is not essential for preimplantation development, as embryos with disrupted LADs are able to reconstruct their nuclear architecture by the 2-cell stage. However, disruption of LADs in both zygotes and 2-cell stage is associated with a failure to undergo a timely maternal-to-zygotic transition and impaired embryonic development (**Part VI**). In summary, this dissertation provides valuable insights into the molecular understanding of epigenome establishment and highlights hierarchies between embryonic chromatin, 3D nuclear organization and genome function.

Aims of the thesis

- When is the 3D organization of the genome erased during oocyte development?

- Which molecular pathways regulate the establishment and remodeling of nuclear organization in early embryos?

- How do the interdependencies between embryonic chromatin, nuclear architecture and DNA-related processes manifest?

- What are the functional consequences of disrupting epigenome establishment during early embryogenesis?

Introduction

The journey from an egg to implantation during murine development

Oocyte growth and maturation

The mature germline derives from precursors, that are first specified during embryogenesis. In mice, such precursors, referred to as primordial germ cells (PGCs), appear around embryonic day 7.25 (E7.25) (Ginsburg et al. 1990). By E13.5, PGCs in female embryos start entering the prophase of meiosis I, during which germline chromatin undergoes synaptonemal complex assembly, DNA recombination, and subsequent disassembly of the complex. The resulting primary oocytes become arrested at the diplotene stage in prophase I of meiosis for a prolonged period, lasting up to months in mice and decades in humans (Hartshorne et al. 2009; Wang and Pepling 2021). Upon hormonal stimulation, a select few primordial follicles begin the growth phase, increasing in size (*see* **Figure 1**). During this phase, oocytes accumulate essential proteins and RNAs for embryonic development, transitioning from growing oocytes (GOs) to full-grown oocytes (FGOs) (Bachvarova 1985; Li et al. 2010). These oocytes, also known as germinal vesicle (GV) stage oocytes, include two primary types based on nuclear organization: the NSN (non-surrounded nucleolus) type, which is transcriptionally active, and the SN (surrounded nucleolus) type, which is transcriptionally silent (Bouniol-Baly et al. 1999; Miyara et al. 2003).
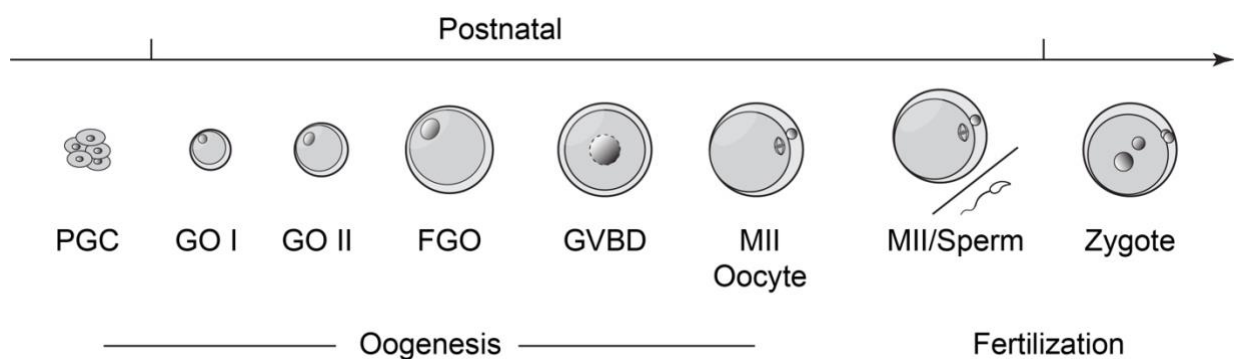
Figure 1. Oocyte development in mice. Mammalian oogenesis includes a growth phase and meiotic maturation. During intraovarian growth, mouse oocytes expand in diameter from ~10 μm to 80 μm, corresponding to a ~500-fold volume increase. Meiotic maturation begins with nuclear envelope breakdown in fully grown oocytes, proceeds with the first meiotic division and extrusion of the first polar body, and arrests in metaphase of the second meiotic division (MII stage). The oocyte is ovulated at the MII stage. Fertilization of the MII egg by sperm forms the one-cell zygote. PGC: primordial germ cells, GO: growing oocyte, FGO: fully grown oocyte, GVBD: germinal vesicle breakdown, MII: metaphase II oocyte.

Fertilization and cleavage

Fertilization marks the beginning of life through the union of an egg and a sperm, two highly specialized cells. The mouse oocyte stores significant cytoplasmic content, including essential proteins and RNA for the early embryo. In contrast, sperm carry minimal cytoplasm, and their genomes are tightly packaged with protamines rather than histones (Rodman et al. 1984). Upon ovulation, the oocyte undergoes nuclear envelope breakdown, completes the first meiotic division, and extrudes the first polar body. The ovulated egg is encased in a zona pellucida and surrounded by cumulus cells. After ovulation, the egg travels through the oviduct towards the uterus. Fertilization occurs in the ampulla of the oviduct when sperm penetrates the zona pellucida and fuses with the egg membrane, introducing the paternal genome into the egg's cytoplasm (Coy et al. 2012). This triggers Ca^{2+} oscillations, leading to the extrusion of the second polar body and the completion of the second meiotic division (Miao and Williams 2012) (*see* **Figure 1**).

The fusion of the gametes forms a zygote containing two parental genomes initially located on opposite sides. These genomes quickly form separate pronuclei, with the maternal chromosomes decondensing and histones replacing protamines in the sperm genome. During the progression of the zygotic cell cycle, the pronuclei migrate towards each other and unite before the first mitosis. Before the first mitosis, the pronuclei undergo nuclear envelope breakdown, and the chromosomes align on a single metaphase plate, resulting in the division of the zygote into a 2-cell stage embryo.

Zygotes and 2-cell embryos in mice are totipotent, meaning each cell can develop into a complete organism, including both embryonic and extraembryonic tissues (Tarkowski 1959; Papaioannou et al. 1989). As the 2-cell embryo continues to divide, the cells reduce in size with each cleavage. The embryo undergoes compaction at the 8-cell stage before cavitating to form the blastocoel in the blastocyst. The late blastocyst implants into the uterine wall after hatching from zona pellucida.

Embryonic transcription and first cell fate decisions

At fertilization, the embryo is transcriptionally silent and relies on proteins and RNA stored in the oocyte. The mouse zygote performs minor zygotic genome activation (ZGA) coinciding with the onset of the first S-phase (Abe et al. 2018). Minor ZGA is characterized by prevalent transcription of intergenic regions. Major ZGA initiates during the S-phase of the 2-cell stage in mice and is marked by increased RNA polymerase II (Pol II) activity (Schultz 2002; Abe et al. 2015). The transcriptional control of major ZGA is more canonical compared to minor ZGA and involves promoter-proximal start sites and splicing. Along with the degradation of maternal RNA and proteins, ZGA enables the switch from maternal to embryonic control of development. This process is called maternal-to-zygotic transition (MZT; *see* **Figure 2**) and a successful MZT is essential for further development of the embryo (Warner and Versteegh 1974; Abe et al. 2018). Key transcription factors such as DUX and OBOX play crucial roles in ZGA regulation (reviewed in Zou et al. 2024). DUX, expressed during minor ZGA, binds promoters of ZGA-associated genes and transposable elements, activating their transcription (De Iaco et al. 2017; Hendrickson et al. 2017). However, loss of DUX causes only minor defects in ZGA and developmental potential (De Iaco et al. 2020; Bosnakovski et al. 2021). OBOX proteins, unique to rodents, redundantly regulate both minor and major ZGA (Ji et al. 2023; Sakamoto et al. 2024). OBOX facilitates chromatin accessibility and RNA polymerase II recruitment, essential for ZGA gene expression and early development (Ji et al. 2023).
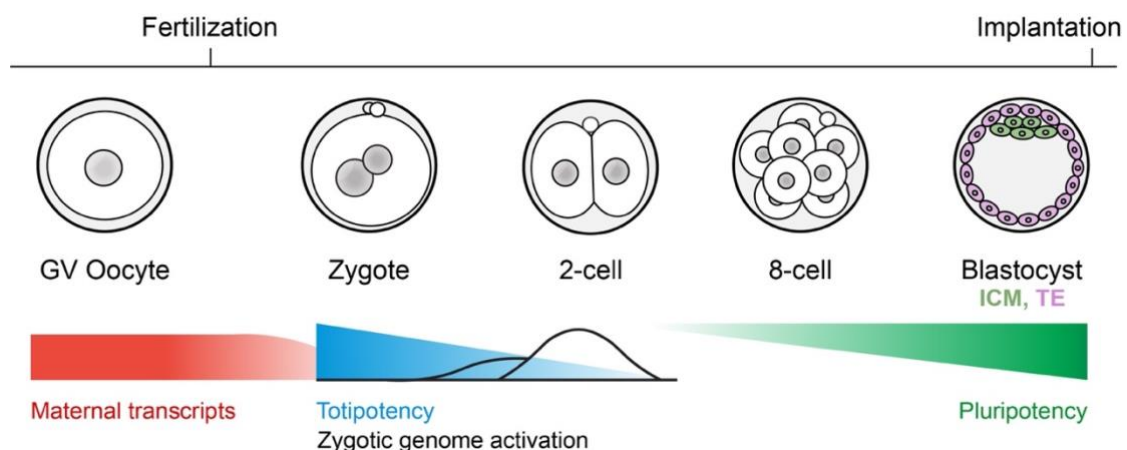
Figure 2. Overview of mouse preimplantation development and onset of embryonic transcription. In mice, zygotic genome activation (ZGA) begins with a minor wave in the zygote, followed by a major wave in the 2-cell embryo. The red bar indicates maternal RNA degradation during the maternal-to-zygotic transition (MZT), while black lines depict embryonic mRNA synthesis during ZGA. At the zygote and 2-cell stages, cells are totipotent, whereas the blastocyst stage features a fully defined pluripotent inner cell mass (ICM) and the differentiated trophectoderm (TE) lineage.

Preimplantation transcription is unique, with several classes of transposable elements transiently transcribed (Peaston et al. 2004), such as murine endogenous retrovirus-L (MERVL) transposons at the early 2-cell stage (Peaston et al. 2004; Svoboda et al. 2004). Their transcription results in chimeric transcripts (Macfarlan et al. 2012) and knockdown of MERVL results in embryonic lethality due to defects in differentiation and genomic stability (Sakashita et al. 2023). Additionally, long interspersed nuclear element-1 (LINE-1) repeats are expressed, regulating gene expression during ZGA, genome-wide chromatin accessibility, and embryo development (Fadloun et al. 2013; Jachowicz et al. 2017; Li et al. 2024). These findings suggest that timely transcription of transposable elements plays a functional role in the early stages.

The blastocyst stage marks the emergence of the first two distinct cell lineages in the embryo: the pluripotent inner cell mass (ICM) and the differentiated trophectoderm (*see* **Figure 2**), identified by distinct transcription factors (e.g., OCT4 for ICM, CDX2 for trophectoderm) (Schöler et al. 1990; Beck et al. 1995; Plachta et al. 2011). These lineages begin to spatially and morphologically segregate at the 16-cell stage morula (Zernicka-Goetz et al. 2009). The extent to which cell fate decisions are pre-patterned or arise from stochastic processes remains debated. Epigenetic differences between individual blastomeres in the embryo at earlier stages, such as histone arginine methylation (Torres-Padilla et al. 2007) and gene expression variances (e.g., Prdm14, Sox2) (Burton et al. 2013; White et al. 2016), observed at the 4-cell stage, suggest early information linked to cell fate.

Epigenetic reprogramming in early embryos

During preimplantation development, embryos undergo extensive epigenetic reprogramming, involving removing and rewriting histone marks and DNA modifications (Burton and Torres-Padilla 2014; Xia and Xie 2020; Rang et al. 2023). This process, which coincides with changes in nuclear organization (Borsos and Torres-Padilla 2016; Pecori and Torres-Padilla 2023; Bondarieva and Tachibana 2024), is critical for early development, although not all aspects are fully understood.

One of the earliest observed epigenetic changes following fertilization is cytosine methylation in DNA (5mC). Initially, both parental genomes exhibit high levels of 5mC. However, 5mC levels in the paternal DNA rapidly decrease, creating an apparent

asymmetry by the late zygote stage, with the maternal genome retaining high levels of 5mC (Mayer et al. 2000). This asymmetry may be due to the DPPA3 protein binding to the H3K9me2 in the maternal pronucleus, protecting it from active demethylation (Nakamura et al. 2007), although this hypothesis has been recently challenged (Li et al. 2018). The reduction of 5mC in the paternal genome involves both active demethylation and passive dilution through replication (Hajkova et al. 2010; Inoue and Zhang 2011). Active demethylation is thought to occur by converting 5mC to hydroxymethylcytosine (5hmC) by the TET3 hydroxylase enzyme (Gu et al. 2011; Wossidlo et al. 2011). Overall, 5mC levels are low during preimplantation, with a more conventional DNA methylation pattern emerging by the blastocyst stage (Smith et al. 2012).

Post-translational modifications of histone tails are also extensively remodelled after fertilization (Burton and Torres-Padilla 2014). Since sperm chromatin is primarily packaged with protamines, most inherited histone modifications are maternal. These marks are erased and rewritten as development progresses. They were initially studied through immunostaining but were recently assessed by genome-wide approaches.

One of the well-characterized histone marks in embryos is H3K4me3. This modification appears in broad, non-canonical domains in the oocyte (Dahl et al. 2016; Liu et al. 2016; Zhang et al. 2016). Although H3K4me3 is typically associated with active genes in differentiated cells, removing broad H3K4me3 domains does not lead to complete transcriptional silencing during oocyte growth (Zhang et al. 2016). These non-canonical maternal domains are inherited by the zygote and remodelled into canonical promoter peaks only upon zygotic genome activation. The function of these broad maternal H3K4me3 domains in the zygote and early 2-cell embryos remains unclear. Notably, the removal of broad H3K4me3 domains seems crucial for ZGA, as knockdown of the H3K4me3 demethylases KDM5A/B leads to developmental arrest before implantation and impairs the activation of a subset of ZGA genes (Dahl et al. 2016). Paternal chromatin, which initially has lower levels of H3K4me3, undergoes *de novo* methylation in a more canonical, promoter-specific pattern (Zhang et al. 2016). Further research is needed to explore the role of H3K4me3 in the early embryos.

Histone acetylation marks are associated with active promoters and enhancers, helping to open chromatin partly by neutralizing the positive charge on histone tails (Bannister and Kouzarides 2011). H3K27ac is observed at promoters before zygotic genome activation (ZGA) in mouse embryos (Dahl et al. 2016; Wang et al. 2022). It is proposed that major ZGA genes are primed by histone acetylation in zygotes and early 2-cell embryos. H3K27ac, initially hypoacetylated in oocytes, appears as non-canonical broad patterns in zygotes and correlates with H3K4me3 and chromatin accessibility. It is established on the paternal genome post-fertilization, forming broad domains. Inhibiting histone acetyltransferase CBP/P300 impairs ZGA and causes 2-cell arrest (Wang et al. 2022). Whether non-canonical H3K27ac-marked regions can act as enhancers in oocytes and early embryos remains to be investigated.

The classical repressive chromatin mark, tri-methylation of H3K9 (H3K9me3), is also extensively remodelled in embryos (Wang et al. 2018; Burton et al. 2020). Post-fertilization, H3K9me3 shows distinct asymmetry, with high levels in the maternal but not the paternal chromatin of zygotes (Santos et al. 2005). The establishment of heterochromatin in early embryos was suggested to occur through the sequential expression and controlled activity of SUV39H enzymes (Burton et al. 2020). SUV39H2 catalyses the *de novo* H3K9me3 deposition in the paternal pronucleus. Forcing early formation of constitutive heterochromatin by ectopically expressing SUV39H1 in early embryos disrupts development and interferes with epigenetic reprogramming. This new H3K9me3 does not repress gene expression but instead marks promoters for future compaction (Burton et al. 2020). Removal of H3K9me2 through depletion of EHMT2 (G9a) also has minimal effects on gene expression in oocytes and 2-cell stage embryos (Au Yeung et al. 2019). However, studies have shown that active deposition of H3K9me3 in embryos helps silence transposable elements by recruiting DNA methylation machinery (Wang et al. 2018).

Another repressive mark, H3K27me3, has been studied genome-wide using ChIP-seq in embryos. Following fertilization, the maternal allele inherits extensive H3K27me3 domains from the oocyte, while these marks are quickly removed from the paternal allele, with new enrichment beginning by the late zygote stage (Liu et al. 2016; Zheng et al. 2016). These newly established paternal H3K27me3 marks form broad, low-level domains primarily in intergenic regions. On the maternal allele, H3K27me3 is lost from

the promoters of typical Polycomb targets after fertilization and is only fully restored post-implantation. In early embryos, H3K9me3 domains overlap significantly with maternally inherited H3K27me3, unlike in later stages where they rarely coincide (Rang et al. 2023). This overlap decreases around the morula stage and canonical H3K27me3 is established at promoters of developmental genes around implantation, contributing to bivalent promoters (Azuara et al. 2006; Bernstein et al. 2006) that bear both activating (H3K4me3) and repressive (H3K27me3) marks. Additionally, allele-specific H3K27me3 peaks have been identified, contributing to a DNA methylation-independent form of imprinting, and resulting in monoallelic expression of several genes (Inoue et al. 2017).

H3K36me2/3 is often correlated with actively transcribed genes. However, in oocytes, the formation of DNA methylation, non-canonical H3K4me3, and H3K27me3 is largely regulated by H3K36 methylation (Xu et al. 2019; Yano et al. 2022). Notably, H3K36me3 has been found to overlap extensively with DNA methylation in oocytes and shows an inverse relationship with H3K27me3 (Xu et al. 2019). The knockout of the methyltransferase Setd2, which depletes H3K36me3, led to the expansion of H3K27me3, suggesting that H3K36me3 plays a role in regulating Polycomb mark distribution in oocytes. However, SETD2 is dispensable for *de novo* DNA methylation in the male germline. Instead, the lysine methyltransferase NSD1 plays a critical role in *de novo* DNA methylation in prospermatogonia, including at imprinted genes. H3K36me2 deposited by NSD1 safeguards a subset of genes against H3K27me3-associated transcriptional silencing (Shirane et al. 2020). In contrast, H3K36me2 in oocytes is predominantly dependent on SETD2 and coincides with H3K36me3. Loss of maternal H3K36me3 disrupts the maternal epigenome, leading to defects in ZGA and embryonic development. After fertilization, maternal H3K36me3 diminishes from the late 2-cell stage and disappears by the 8-cell stage, while zygotic H3K36me3 gradually forms during preimplantation development (Xu et al. 2019).

31

Three-dimensional nuclear architecture in early mouse embryos

Layers of genome and nuclear organization

In differentiated cells, chromatin organization within the nucleus is not random (Cremer and Cremer 2001; Bolzer et al. 2005). The spatial 3D arrangement of genomic regions adds an additional layer of epigenetic regulation (reviewed in Bonev and Cavalli 2016; Willemin et al. 2024). This genome folding renders specific DNA sequences accessible for the transcriptional machinery (Gorkin et al. 2014) and also plays a role in protecting the genome from DNA damage (Hauer and Gasser 2017). Chromosomes form long-range intra-chromosomal interactions, resulting in their folding into highly structured 3D arrangements such as A/B compartments (Lieberman-Aiden et al. 2009) and topologically associating domains (TADs) (Dixon et al. 2012; Sexton et al. 2012; Nora et al. 2012). This organization is closely linked to replication and transcription. A compartments typically contain active chromatin regions that replicate early in the S-phase and are characterized by higher chromatin accessibility. In contrast, B compartments consist of heterochromatic regions that replicate later and are mostly inaccessible (Ryba et al. 2010; Dixon et al. 2012; Pope et al. 2014). TADs, which are smaller organizational units compared to compartments, serve to restrict enhancer-promoter interactions, promoting more precise transcriptional regulation (Nora et al. 2012; Hnisz et al. 2016; Flavahan et al. 2016). Beyond TADs and compartments, the genome is also organized around nuclear landmarks and organelles (Canat et al. 2020; Belmont 2022). One of the key aspects of this spatial arrangement is the division of the genome into regions associated with the nuclear lamina versus those located more centrally in the nucleus. Lamina-associated domains (LADs) (Pickersgill et al. 2006; Guelen et al. 2008; van Steensel and Belmont 2017) are large genomic regions, ranging from 100 kb to 10 Mb, that interact with the nuclear lamina, a meshwork of intermediate filaments composed of lamins (reviewed in Burke and Stewart 2013), which are key structural components of the nuclear envelope.

LADs across cell types share distinctive characteristics, such as high AT content, low gene density, and a tendency to contain functionally repressed chromatin (Meuleman et al. 2013; van Steensel and Belmont 2017; Briand and Collas 2020). LADs can partially overlap with genomic regions associated with the nucleolus (nucleolus-

associated domains or NADs) (Kind et al. 2013; Bizhanova et al. 2020; Bersaglieri et al. 2022), suggesting that the nuclear lamina and the nucleolus constitute interchangeable scaffolds for heterochromatin. Typically, LADs replicate late and correspond to B compartments, while inter-LADs (iLADs) replicate early during S-phase and correlate with A compartments (Guelen et al. 2008; Peric-Hupkes et al. 2010; Pope et al. 2014). iLADs exhibit higher transcriptional activity compared to LADs, with disassociation from the nuclear lamina often seen upon gene activation (Tumbar and Belmont 2001; Therizols et al. 2014). However, tethering a gene to the nuclear periphery doesn't necessarily silence it, indicating nuclear positioning alone does not dictate gene expression (Finlan et al. 2008; Jachowicz et al. 2013). In certain cell types, LAD boundaries are marked by sharp changes in H3K4me2 and H3K27me3 (Guelen et al. 2008; Harr et al. 2015). LADs are enriched with H3K9me2 (Guelen et al. 2008; Wen et al. 2009) and inhibiting the H3K9 methyltransferase EHMT2 (G9a) reduces their nuclear lamina contacts (Bian et al. 2013; Kind et al. 2013). Histone deacetylation plays a role in repression within LADs, facilitated by the interactions of Emerin and LAP2β with HDAC3 (Somech et al. 2005; Demmerle et al. 2013), as well as A-type lamins with the sirtuins SIRT1 and SIRT6 (Ghosh et al. 2015). Finally, LADs are depleted in cytosine methylation and LADs have been linked to loss of DNA methylation in cancer (Berman et al. 2011).

Microscopy-based dynamics of nuclear organization in early embryos

The dynamic reorganization of the 3D genome following fertilization, during embryonic reprogramming, is evident even with simple DAPI staining (*see* **Figure 3**). A notable feature that emerges during oogenesis is the presence of nucleolar precursor bodies (NPBs) or nucleolar-like bodies (NLBs), around which pericentromeric repeats cluster, forming distinct ring-like structures. These structures are associated with efficient reprogramming upon nuclear transfer, although their exact function remains unclear (Martin et al. 2006). This arrangement persists until the 4-cell stage, when centromeric regions re-cluster into chromocenters resembling those found in somatic, differentiated murine nuclei. The 3D localization of centromeric repeats around NPBs at the early 2-cell stage and the formation of chromocenters by the late 2-cell stage are essential for proper embryonic development (Probst et al. 2010; Casanova et al. 2013). Recently, chromocenter formation in the developing embryo has been shown

33

to coincide with a transition from a liquid-like to a more gel-like biophysical state (Guthmann et al. 2023).

Electron microscopy reveals electron-dense heterochromatin at the nuclear periphery in nearly all somatic nuclei (Fawcett 1966; Kind et al. 2013; van Steensel and Belmont 2017). However, in zygotes and early 2-cell stage nuclei, these electron-dense regions are not visible, only becoming apparent at later stages (Ahmed et al. 2010). Additionally, chromatin mobility in the preimplantation embryo has been studied using fluorescence recovery after photobleaching (FRAP) (Bošković et al. 2014; Ooga et al. 2016). These studies demonstrated higher chromatin mobility in 2-cell embryos compared to 8-cell embryos, indicating a gradual transition to more stable chromatin organization as the embryo develops. Overall, the distinct genome structure of zygotes and 2-cell embryos suggests that the unique nuclear architecture of the early embryo may be linked to totipotency.

Insights into genome reorganization in mammalian embryogenesis

Recent advancements in low-input genomics techniques have significantly improved our understanding of chromatin architecture in preimplantation embryos. Chromatin architecture is nowadays widely studied by so called '3C' methods (Chromosome Conformation Capture; Dekker et al. 2002) of which the most commonly used is Hi-C (Lieberman-Aiden et al. 2009; Rao et al. 2014). Hi-C studies have uncovered a gradual establishment of genome folding during early embryogenesis (Du et al. 2017; Ke et al. 2017; Flyamer et al. 2017; Collombet et al. 2020). Oocytes progressively lose compartments and TADs during their maturation (Flyamer et al. 2017), whereas sperm cells exhibit a more canonical structure, characterized by additional long-range interactions (Ke et al. 2017; Jung et al. 2017), likely reflecting increased chromatin compaction. Similarly, compartment strength is higher in the paternal pronucleus compared to the maternal pronucleus in zygotes (Du et al. 2017; Flyamer et al. 2017), with A and B compartments becoming more defined during cleavage stages (*see* **Figure 3**). Embryonic compartments display the expected correlations with transcriptional and chromatin features (Ke et al. 2017), such as H3K4me3 and accessibility in A compartments, and DNA methylation and H3K27me3 in B compartments. These compartments are also distinguished by genomic signatures,

with SINE-B1 elements being more prevalent in A compartments and LINE-1 elements more abundant in B compartments (Lu et al. 2021).

Intriguingly, H3K27me3 plays a significant role in defining genome organization in mouse oocytes, which lack TADs and compartments but instead form cohesin-independent polycomb-associated domains (PADs) (Du et al. 2020). The emergence of TADs, which are only fully established by the blastocyst stage, depends on CTCF. The depletion of both maternal and zygotic CTCF results in developmental failure only around implantation in mice (Moore et al. 2012; Andreu et al. 2022). Interestingly, CTCF depletion does not lead to structure-related transcriptional changes (Andreu et al. 2022), and cell-fate specification remains unaffected, suggesting that TADs are not essential for establishing gene expression programs, at least until implantation.

Characterization and dynamics of LADs in mouse embryos

Using DamID for LaminB1 (van Steensel and Henikoff 2000; Guelen et al. 2008; Kind et al. 2013), which utilizes m6A methylation within GATC motifs by the bacterial Dam enzyme, LADs have been mapped from zygotes to the blastocyst stage in mouse embryos (Borsos et al. 2019; also reviewed in Pecori and Torres-Padilla 2023; Bondarieva and Tachibana 2024). LADs are undetectable in fully grown germinal vesicle (GV) oocytes, indicating that they are not inherited through the maternal germline but are established *de novo* in zygotes immediately after fertilization (*see* **Figure 3**). LADs undergo extensive remodeling throughout preimplantation development, particularly around the time of the maternal-to-zygotic transition. Overall, zygotic LADs exhibit expected genomic features, including high AT content, and correlate with B compartments. However, the maternal and paternal pronuclei display distinct LAD characteristics: maternal LADs are smaller and appear more fragmented, while paternal LADs are larger and resemble those found in mouse embryonic stem cells (mESCs). These allelic differences in LAD structures become equalized only after the 8-cell stage (Borsos et al. 2019).

Additionally, late 2-cell stage LADs are unique, as they tend to be more fragmented, exhibit reduced AT content, and contain higher gene density. While LAD restructuring at the late 2-cell stage correlates with transcriptional activation during ZGA, there is a notable increase in LADs overlapping with A compartments at this stage (Borsos et al.

2019). The significance of this observation is unclear, but it suggests an uncoupling between these two levels of nuclear organization. Furthermore, global demethylation of H3K9me3 does not appear to impact LAD establishment in zygotes for either allele (Borsos et al. 2019). These findings, along with the growing body of research on non-canonical chromatin in early embryos (Dahl et al. 2016; Zhang et al. 2016; Zheng et al. 2016; Wang et al. 2018, 2022; also reviewed in Rang et al. 2023), prompt questions about whether the molecular mechanisms and chromatin dependencies of LAD establishment in embryos differ from those involved in LAD maintenance in cultured cells. Finally, the observation that LAD formation precedes TAD consolidation in early mouse embryos (*see* **Figure 3**) highlights the need for further investigation into the sequential establishment of the epigenome and the temporal and functional hierarchy of its different structural components.
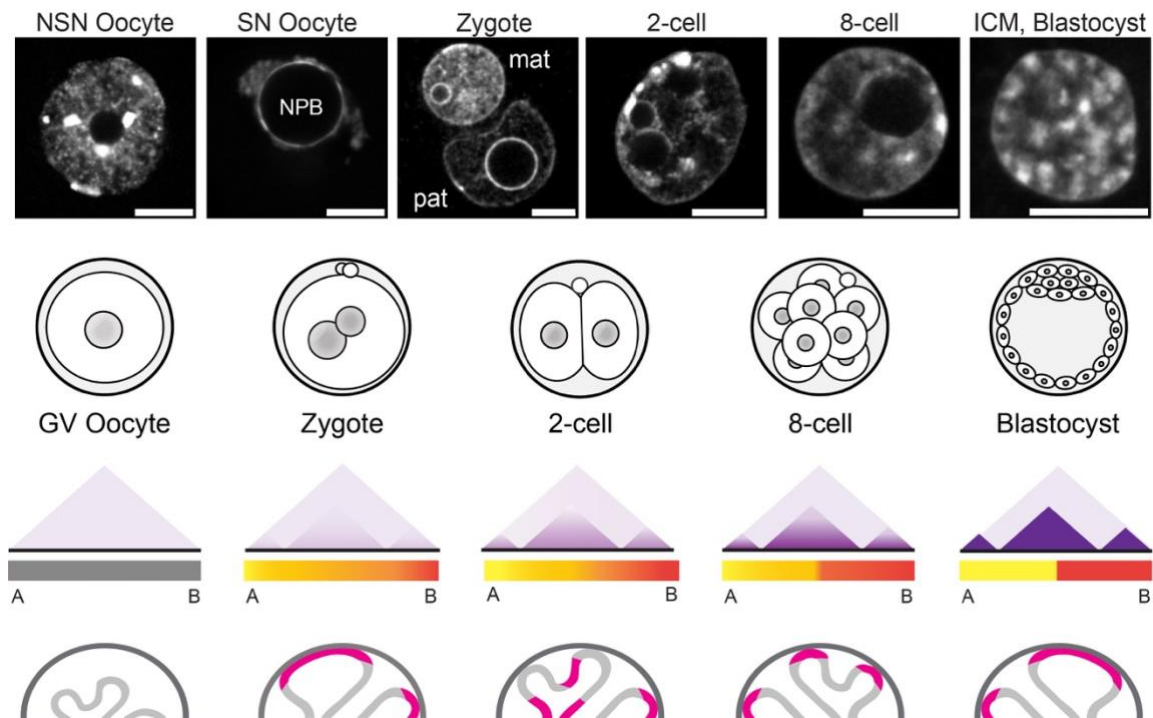
Figure 3: Changes in nuclear organization during mouse preimplantation development. The nuclear organization differences between oocytes and preimplantation stages are readily visualized via DNA (DAPI) staining. During oocyte maturation, centromeric repeats organize into a ring-like structure encircling the nucleolar precursor bodies (NPBs). This association persists until the late 2-cell stage, at which point the repeats begin to recluster, forming chromocenters. Note that in zygotes, the parental genomes remain segregated in separate pronuclei. Scale bars: 10 μm. NSN: non-surrounded nucleolus, SN: surrounded nucleolus, mat: maternal pronucleus, pat: paternal pronucleus), ICM: inner cell mass. In mouse embryos, TADs and compartments consolidate progressively. LADs are established early in zygotes and undergo dynamic reshuffling throughout preimplantation development. TAD: Topologically associating domain, LAD: lamina-associated domain, iLAD: Inter-LAD. TAD strength (insulation score) is shown in purple in diagonal matrices. Compartment strength is indicated by bar intensity: yellow (A compartment), red (B compartment), grey (no compartments). The nuclear periphery is shown in black, genomic regions that are LADs in ICM (or mouse ESCs) are shown in magenta and iLADs in grey.

Establishment of replication timing during epigenome reprogramming

Replication timing (RT) is a major epigenetic feature that refers to the order in which the genome is replicated during the S phase (Aladjem et al. 2002; Ryba et al. 2011). The coordination and order in which specific genomic regions undergo DNA replication leads to the segregation into early and late replicating domains. This is tightly linked to other chromatin features and nuclear organization in differentiated cells or cells in culture. LADs and B compartments are typically late-replicating regions, while inter-LADs, often associated with active A compartments, replicate early during S-phase (Yaffe et al. 2010; Moindrot et al. 2012; Pope et al. 2014). While the mechanisms regulating replication timing are largely unknown, RIF1 has been shown to suppress firing of late replication origins within heterochromatin (Peace et al. 2014). Strikingly, RIF1-depletion in human embryonic stem cells results in a complete erasure of the RT program (Klein et al. 2021). This disruption is associated with changes in the histone modification landscape as well as in intra- and inter-compartment interactions (Klein et al. 2021), suggesting that replication timing has a central role in maintaining the epigenome. However, it is not understood when and how the replication timing program is established during early mammalian embryogenesis. Recent work has demonstrated distinct features of the DNA replication process in early embryos. For example, replication fork speed is slow at the beginning of development in early mouse embryos (Nakatani et al. 2022). Additionally, an increase in DNA replication fork speed appears to be associated with a reduction in cellular plasticity. In human zygotes, this slower replication fork speed is also accompanied by fork stalling and the accumulation of DNA damage (Palmerola et al. 2022). It is tempting to hypothesize that such observations may be linked to the higher aneuploidy rate in human embryos compared to mouse embryos. In summary, while replication timing seems to play a critical role in chromatin and cellular identity, the establishment of RT *in vivo* has yet to be studied.

Open questions about epigenome establishment during early mammalian embryogenesis

Molecular basis of genome folding establishment during early development

The studies relating to nuclear organization in mammalian embryos have provided foundational insights, leading to the first descriptions of chromatin folding during early development. Although significant efforts have been made to identify molecular regulators of TADs, LADs, and other nuclear structures in cell culture models, mechanistic insights into the establishment of genome folding *in vivo* are limited. Recent studies have made interesting discoveries, such as the impact of KDM5B, an H3K4me3 demethylase, in preventing LAD formation in mouse zygotes (Borsos et al. 2019). Additionally, the role of heterochromatin formation in establishing and consolidating compartments and LADs is an open question. In mice, heterochromatin is immature at the beginning of development, and H3K9me3 is non-repressive (Burton et al. 2020). Investigating the role of heterochromatin formation and maturation in nuclear architecture during early development is an exciting prospect. Higher-throughput screenings and functional perturbations will help identify molecular pathways that govern the establishment of the 3D genome at the beginning of mammalian development.

Influence of nuclear organization on chromatin-regulated processes at the beginning of development

Distinguishing correlation and causation in nuclear architecture's role in chromatin-regulated processes remains challenging. The preimplantation embryogenesis provides an outstanding platform to interrogate the interdependencies among different pillars of the 3D genome, transcription, and DNA replication. Inhibition of ZGA in mice did not affect TAD and compartment consolidation, suggesting that transcription is not critical for strengthening TAD insulation (Du et al. 2017; Ke et al. 2017). On the other hand, TAD formation in mouse embryos depends on DNA replication (Ke et al. 2017), and the incorporation of replication-dependent H3.1/3.2 histone variants promotes TAD insulation in 2-cell embryos (Funaya et al. 2024). However, unlike TAD formation, LAD formation in mouse embryos do not depend on DNA replication (Borsos et al. 2019). Additionally, cohesin depletion in somatic cell nuclear transfer

(SCNT) experiments facilitates minor ZGA in mouse embryos (Zhang et al. 2020). Therefore, the relationship of the 3D genome architecture to different DNA-related processes seems more complicated than previously thought. More experimental efforts are required to dissect and disentangle these interdependencies systematically, in particular through perturbation experiments.

Role of nuclear organization in early embryonic development and cell fate

Finally, understanding the connection between nuclear organization and organism development or cellular fate remains a fundamental question. Given nuclear architecture's role in coordinating DNA damage response, DNA replication, and transcription, it is logical to consider its potential regulation of cell fate and development. Understanding the regulation of dynamic and interdependent chromatin processes is critical for providing coherent cellular responses to developmental cues. Combining gain-of-function or loss-of-function approaches for pathways and molecular effectors involved in 3D organization and dissecting the developmental phenotypes will enhance our understanding of the functional role of nuclear organization in development and cellular fate.

References

Abe K, Funaya S, Tsukioka D, Kawamura M, Suzuki Y, Suzuki MG, Schultz RM, Aoki F. 2018. Minor zygotic gene activation is essential for mouse preimplantation development. *Proc Natl Acad Sci* **115**: E6780–E6788.

Abe K, Yamamoto R, Franke V, Cao M, Suzuki Y, Suzuki MG, Vlahovicek K, Svoboda P, Schultz RM, Aoki F. 2015. The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3′ processing. *EMBO J* **34**: 1523–1537.

Ahmed K, Dehghani H, Rugg-Gunn P, Fussner E, Rossant J, Bazett-Jones DP. 2010. Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo. *PLOS ONE* **5**: e10531.

Aladjem MI, Rodewald LW, Lin CM, Bowman S, Cimbora DM, Brody LL, Epner EM, Groudine M, Wahl GM. 2002. Replication Initiation Patterns in the β-Globin Loci of Totipotent and Differentiated Murine Cells: Evidence for Multiple Initiation Regions. *Mol Cell Biol* **22**: 442–452.

Andreu MJ, Alvarez-Franco A, Portela M, Gimenez-Llorente D, Cuadrado A, Badia-Careaga C, Tiana M, Losada A, Manzanares M. 2022. Establishment of 3D chromatin structure after fertilization and the metabolic switch at the morula-to-blastocyst transition require CTCF. *Cell Rep* **41**: 111501.

Au Yeung WK, Brind'Amour J, Hatano Y, Yamagata K, Feil R, Lorincz MC, Tachibana M, Shinkai Y, Sasaki H. 2019. Histone H3K9 Methyltransferase G9a in Oocytes Is Essential for Preimplantation Development but Dispensable for CG Methylation Protection. *Cell Rep* **27**: 282-293.e4.

Azuara V, Perry P, Sauer S, Spivakov M, Jørgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, et al. 2006. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8**: 532–538.

Bachvarova R. 1985. Gene Expression During Oogenesis and Oocyte Development in Mammals. In *Oogenesis* (ed. L.W. Browder), pp. 453–524, Springer US, Boston, MA https://doi.org/10.1007/978-1-4615-6814-8_11 (Accessed August 14, 2024).

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381–395.

Beck F, Erler T, Russell A, James R. 1995. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn Off Publ Am Assoc Anat* **204**: 219–227.

Belmont AS. 2022. Nuclear Compartments: An Incomplete Primer to Nuclear Compartments, Bodies, and Genome Organization Relative to Nuclear Architecture. *Cold Spring Harb Perspect Biol* **14**: a041268.

Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, Noushmehr H, Lange CPE, van Dijk CM, Tollenaar RAEM, et al. 2011. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet* **44**: 40–46.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315–326.

Bersaglieri C, Kresoja-Rakic J, Gupta S, Bär D, Kuzyakiv R, Panatta M, Santoro R. 2022. Genome-wide maps of nucleolus interactions reveal distinct layers of repressive chromatin domains. *Nat Commun* **13**: 1483.

Bian Q, Khanna N, Alvikas J, Belmont AS. 2013. β-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. *J Cell Biol* **203**: 767–783.

Bizhanova A, Yan A, Yu J, Zhu LJ, Kaufman PD. 2020. Distinct features of nucleolus-associated domains in mouse embryonic stem cells. *Chromosoma* **129**: 121–139.

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Müller S, Eils R, Cremer C, Speicher MR, et al. 2005. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* **3**: e157.

Bondarieva A, Tachibana K. 2024. Genome folding and zygotic genome activation in mammalian preimplantation embryos. *Curr Opin Genet Dev* **89**: 102268.

Bonev B, Cavalli G. 2016. Organization and function of the 3D genome. *Nat Rev Genet* **17**: 661–678.

Borsos M, Perricone SM, Schauer T, Pontabry J, de Luca KL, de Vries SS, Ruiz-Morales ER, Torres-Padilla M-E, Kind J. 2019. Genome–lamina interactions are established de novo in the early mouse embryo. *Nature* **569**: 729–733.

Borsos M, Torres-Padilla M-E. 2016. Building up the nucleus: nuclear organization in the establishment of totipotency and pluripotency during mammalian development. *Genes Dev* **30**: 611–621.

Bošković A, Eid A, Pontabry J, Ishiuchi T, Spiegelhalter C, Raghu Ram EVS, Meshorer E, Torres-Padilla M-E. 2014. Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. *Genes Dev* **28**: 1042–1047.

Bosnakovski D, Gearhart MD, Ho Choi S, Kyba M. 2021. Dux facilitates post-implantation development, but is not essential for zygotic genome activation†. *Biol Reprod* **104**: 83–93.

Bouniol-Baly C, Hamraoui L, Guibert J, Beaujean N, Szöllösi MS, Debey P. 1999. Differential Transcriptional Activity Associated with Chromatin Configuration in Fully Grown Mouse Germinal Vesicle Oocytes1. *Biol Reprod* **60**: 580–587.

Briand N, Collas P. 2020. Lamina-associated domains: peripheral matters and internal affairs. *Genome Biol* **21**: 85.

Burke B, Stewart CL. 2013. The nuclear lamins: flexibility in function. *Nat Rev Mol Cell Biol* **14**: 13–24.

Burton A, Brochard V, Galan C, Ruiz-Morales ER, Rovira Q, Rodriguez-Terrones D, Kruse K, Le Gras S, Udayakumar VS, Chin HG, et al. 2020. Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat Cell Biol* **22**: 767–778.

Burton A, Muller J, Tu S, Padilla-Longoria P, Guccione E, Torres-Padilla M-E. 2013. Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep* **5**: 687–701.

Burton A, Torres-Padilla M-E. 2014. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat Rev Mol Cell Biol* **15**: 723–735.

Canat A, Veillet A, Bonnet A, Therizols P. 2020. Genome anchoring to nuclear landmarks drives functional compartmentalization of the nuclear space. *Brief Funct Genomics* **19**: 101–110.

Casanova M, Pasternak M, El Marjou F, Le Baccon P, Probst AV, Almouzni G. 2013. Heterochromatin Reorganization during Early Mouse Development Requires a Single-Stranded Noncoding Transcript. *Cell Rep* **4**: 1156–1167.

Collombet S, Ranisavljevic N, Nagano T, Varnai C, Shisode T, Leung W, Piolot T, Galupa R, Borensztein M, Servant N, et al. 2020. Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**: 142–146.

Coy P, García-Vázquez FA, Visconti PE, Avilés M. 2012. Roles of the oviduct in mammalian fertilization. *Reprod Camb Engl* **144**: 649.

Cremer T, Cremer C. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**: 292–301.

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY, et al. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**: 548–552.

De Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D. 2017. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet* **49**: 941–945.

De Iaco A, Verp S, Offner S, Grun D, Trono D. 2020. DUX is a non-essential synchronizer of zygotic genome activation. *Dev Camb Engl* **147**: dev177725.

Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. *Science* **295**: 1306–1311.

Demmerle J, Koch AJ, Holaska JM. 2013. Emerin and histone deacetylase 3 (HDAC3) cooperatively regulate expression and nuclear positions of MyoD, Myf5, and Pax7 genes during myogenesis. *Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol* **21**: 765.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380.

Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, et al. 2017. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**: 232–235.

Du Z, Zheng H, Kawamura YK, Zhang K, Gassler J, Powell S, Xu Q, Lin Z, Xu K, Zhou Q, et al. 2020. Polycomb Group Proteins Regulate Chromatin Architecture in Mouse Oocytes and Early Embryos. *Mol Cell* **77**: 825-839.e7.

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

Fawcett DW. 1966. On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. *Am J Anat* **119**: 129–145.

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.

Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suvà ML, Bernstein BE. 2016. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**: 110–114.

Flyamer IM, Gassler J, Imakaev M, Brandão HB, Ulianov SV, Abdennur N, Razin SV, Mirny LA, Tachibana-Konwalski K. 2017. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**: 110–114.

Funaya S, Takahashi Y, Suzuki MG, Suzuki Y, Aoki F. 2024. H3.1/3.2 regulate the initial progression of the gene expression program. *Nucleic Acids Res* **52**: 6158–6170.

Ghosh S, Liu B, Wang Y, Hao Q, Zhou Z. 2015. Lamin A Is an Endogenous SIRT6 Activator and Promotes SIRT6-Mediated DNA Repair. *Cell Rep* **13**: 1396–1406.

Ginsburg M, Snow MH, McLaren A. 1990. Primordial germ cells in the mouse embryo during gastrulation. *Dev Camb Engl* **110**: 521–528.

Gorkin DU, Leung D, Ren B. 2014. The 3D Genome in Transcriptional Regulation and Pluripotency. *Cell Stem Cell* **14**: 762–775.

Gu T-P, Guo F, Yang H, Wu H-P, Xu G-F, Liu W, Xie Z-G, Shi L, He X, Jin S, et al. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**: 606–610.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948–951.

Guthmann M, Qian C, Gialdini I, Nakatani T, Ettinger A, Schauer T, Kukhtevich I, Schneider R, Lamb DC, Burton A, et al. 2023. A change in biophysical properties accompanies heterochromatin formation in mouse embryos. *Genes Dev* **37**: 336.

Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. 2010. Genome-Wide Reprogramming in the Mouse Germ Line Entails the Base Excision Repair Pathway. *Science* **329**: 78–82.

Harr JC, Luperchio TR, Wong X, Cohen E, Wheelan SJ, Reddy KL. 2015. Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biol* **208**: 33–52.

Hartshorne GM, Lyrakou S, Hamoda H, Oloto E, Ghafari F. 2009. Oogenesis and cell death in human prenatal ovaries: what are the criteria for oocyte selection? *Mol Hum Reprod* **15**: 805–819.

Hauer MH, Gasser SM. 2017. Chromatin and nucleosome dynamics in DNA damage and repair. *Genes Dev* **31**: 2204–2221.

Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim J-W, Wike CL, Weaver BD, Pflueger C, Emery BR, Wilcox AL, et al. 2017. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat Genet* **49**: 925–934.

Hnisz D, Weintraub AS, Day DS, Valton A-L, Bak RO, Li CH, Goldmann J, Lajoie BR, Fan ZP, Sigova AA, et al. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**: 1454–1458.

Inoue A, Jiang L, Lu F, Suzuki T, Zhang Y. 2017. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**: 419–424.

Inoue A, Zhang Y. 2011. Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos. *Science* **334**: 194–194.

Jachowicz JW, Bing X, Pontabry J, Bošković A, Rando OJ, Torres-Padilla M-E. 2017. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat Genet* **49**: 1502–1510.

Jachowicz JW, Santenard A, Bender A, Muller J, Torres-Padilla M-E. 2013. Heterochromatin establishment at pericentromeres depends on nuclear position. *Genes Dev* **27**: 2427–2432.

Ji S, Chen F, Stein P, Wang J, Zhou Z, Wang L, Zhao Q, Lin Z, Liu B, Xu K, et al. 2023. OBOX regulates mouse zygotic genome activation and early development. *Nature* **620**: 1047–1053.

Jung YH, Sauria MEG, Lyu X, Cheema MS, Ausio J, Taylor J, Corces VG. 2017. Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep* **18**: 1366–1382.

Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, et al. 2017. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell* **170**: 367-381.e20.

Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B. 2013. Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**: 178–192.

Klein KN, Zhao PA, Lyu X, Sasaki T, Bartlett DA, Singh AM, Tasan I, Zhang M, Watts LP, Hiraga S-I, et al. 2021. Replication timing maintains the global epigenetic state in human cells. *Science* **372**: 371–378.

Ladstätter S, Tachibana-Konwalski K. 2016. A Surveillance Mechanism Ensures Repair of DNA Lesions during Zygotic Reprogramming. *Cell* **167**: 1774-1787.e13.

Li L, Zheng P, Dean J. 2010. Maternal control of early mouse development. *Dev Camb Engl* **137**: 859–870.

Li X, Bie L, Wang Y, Hong Y, Zhou Z, Fan Y, Yan X, Tao Y, Huang C, Zhang Y, et al. 2024. LINE-1 transcription activates long-range gene expression. *Nat Genet* **56**: 1494–1502.

Li Y, Zhang Z, Chen J, Liu W, Lai W, Liu B, Li X, Liu L, Xu S, Dong Q, et al. 2018. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* **564**: 136–140.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science* **326**: 289–293.

Liu X, Wang C, Liu W, Li J, Li C, Kou X, Chen J, Zhao Y, Gao H, Wang H, et al. 2016. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**: 558–562.

Lu JY, Chang L, Li T, Wang T, Yin Y, Zhan G, Han X, Zhang K, Tao Y, Percharde M, et al. 2021. Homotypic clustering of L1 and B1/Alu repeats compartmentalizes the 3D genome. *Cell Res* **31**: 613–630.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**: 57–63.

Martin C, Brochard V, Migné C, Zink D, Debey P, Beaujean N. 2006. Architectural reorganization of the nuclei upon transfer into oocytes accompanies genome reprogramming. *Mol Reprod Dev* **73**: 1102–1111.

Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* **403**: 501–502.

Meuleman W, Peric-Hupkes D, Kind J, Beaudry J-B, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B. 2013. Constitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **23**: 270–280.

Miao Y-L, Williams CJ. 2012. Calcium signaling in mammalian egg activation and embryo development: Influence of subcellular localization. *Mol Reprod Dev* **79**: 742.

Misteli T, Soutoglou E. 2009. The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nat Rev Mol Cell Biol* **10**: 243–254.

Miyara F, Migne C, Dumont-Hassan M, Le Meur A, Cohen-Bacrie P, Aubriot F-X, Glissant A, Nathan C, Douard S, Stanovici A, et al. 2003. Chromatin configuration and transcriptional control in human and mouse oocytes. *Mol Reprod Dev* **64**: 458–470.

Moindrot B, Audit B, Klous P, Baker A, Thermes C, de Laat W, Bouvet P, Mongelard F, Arneodo A. 2012. 3D chromatin conformation correlates with replication timing and is conserved in resting cells. *Nucleic Acids Res* **40**: 9470–9481.

Moore JM, Rabaia NA, Smith LE, Fagerlie S, Gurley K, Loukinov D, Disteche CM, Collins SJ, Kemp CJ, Lobanenkov VV, et al. 2012. Loss of Maternal CTCF Is Associated with Peri-Implantation Lethality of Ctcf Null Embryos. *PLOS ONE* **7**: e34915.

Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, Sekimoto T, Ikawa M, Yoneda Y, Okabe M, et al. 2007. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* **9**: 64–71.

Nakatani T, Lin J, Ji F, Ettinger A, Pontabry J, Tokoro M, Altamirano-Pacheco L, Fiorentino J, Mahammadov E, Hatano Y, et al. 2022. DNA replication fork speed underlies cell fate changes and promotes reprogramming. *Nat Genet* **54**: 318–327.

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381–385.

Ooga M, Fulka H, Hashimoto S, Suzuki MG, Aoki F. 2016. Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics* **11**: 85–94.

Palmerola KL, Amrane S, Angeles ADL, Xu S, Wang N, Pinho J de, Zuccaro MV, Taglialatela A, Massey DJ, Turocy J, et al. 2022. Replication stress impairs

chromosome segregation and preimplantation development in human embryos. *Cell* **185**: 2988-3007.e20.

Papaioannou VE, Mkandawire J, Biggers JD. 1989. Development and phenotypic variability of genetically identical half mouse embryos. *Dev Camb Engl* **106**: 817–827.

Peace JM, Ter-Zakarian A, Aparicio OM. 2014. Rif1 Regulates Initiation Timing of Late Replication Origins throughout the S. cerevisiae Genome. *PLOS ONE* **9**: e98501.

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.

Pecori F, Torres-Padilla M-E. 2023. Dynamics of nuclear architecture during early embryonic development and lessons from liveimaging. *Dev Cell* **58**: 435–449.

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SWM, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, et al. 2010. Molecular maps of the reorganization of genome – nuclear lamina interactions during differentiation. *Mol Cell* **38**: 603–613.

Pickersgill H, Kalverda B, de Wit E, Talhout W, Fornerod M, van Steensel B. 2006. Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nat Genet* **38**: 1005–1014.

Plachta N, Bollenbach T, Pease S, Fraser SE, Pantazis P. 2011. Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat Cell Biol* **13**: 117–123.

Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**: 402–405.

Probst AlineV, 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.

Rang FJ, Kind J, Guerreiro I. 2023. The role of heterochromatin in 3D genome organization during preimplantation development. *Cell Rep* **42**: 112248.

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665–1680.

Rodman TC, Pruslin FH, Allfrey VG. 1984. Protamine-DNA association in mammalian spermatozoa. *Exp Cell Res* **150**: 269–281.

Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM. 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res* **20**: 761–770.

Ryba T, Hiratani I, Sasaki T, Battaglia D, Kulik M, Zhang J, Dalton S, Gilbert DM. 2011. Replication timing: a fingerprint for cell identity and pluripotency. *PLoS Comput Biol* **7**: e1002225.

Sakamoto M, Ito A, Wakayama S, Sasaki H, Wakayama T, Ishiuchi T. 2024. Detection of newly synthesized RNA reveals transcriptional reprogramming during ZGA and a role of Obox3 in totipotency acquisition. *Cell Rep* **43**: 114118.

Sakashita A, Kitano T, Ishizu H, Guo Y, Masuda H, Ariura M, Murano K, Siomi H. 2023. Transcription of MERVL retrotransposons is required for preimplantation embryo development. *Nat Genet* **55**: 484–495.

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.

Schöler HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P. 1990. New type of POU domain in germ line-specific protein Oct-4. *Nature* **344**: 435–439.

Schultz RM. 2002. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update* **8**: 323–331.

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell* **148**: 458–472.

Shirane K, Miura F, Ito T, Lorincz MC. 2020. NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts Polycomb-associated silencing. *Nat Genet* **52**: 1088–1098.

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* **484**: 339–344.

Somech R, Shaklai S, Geller O, Amariglio N, Simon AJ, Rechavi G, Gal-Yam EN. 2005. The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. *J Cell Sci* **118**: 4017–4025.

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* **269**: 276–285.

Tarkowski AK. 1959. Experiments on the Development of Isolated Blastomeres of Mouse Eggs. *Nature* **184**: 1286–1287.

Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. 2014. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* **346**: 1238–1242.

Torres-Padilla M-E, Parfitt D-E, Kouzarides T, Zernicka-Goetz M. 2007. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**: 214–218.

Tumbar T, Belmont AS. 2001. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat Cell Biol* **3**: 134–139.

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

van Steensel B, Henikoff S. 2000. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol* **18**: 424–428.

Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, Chen C, Kou X, Zhao Y, Chen J, et al. 2018. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat Cell Biol* **20**: 620–631.

Wang M, Chen Z, Zhang Y. 2022. CBP/p300 and HDAC activities regulate H3K27 acetylation dynamics and zygotic genome activation in mouse preimplantation embryos. *EMBO J* **41**: e112012.

Wang X, Pepling ME. 2021. Regulation of Meiotic Prophase One in Mammalian Oocytes. *Front Cell Dev Biol* **9**: 667306.

Warner CM, Versteegh LR. 1974. In vivo and in vitro effect of α-amanitin on preimplantation mouse embryo RNA polymerase. *Nature* **248**: 678–680.

Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* **41**: 246–250.

White MD, Angiolini JF, Alvarez YD, Kaur G, Zhao ZW, Mocskos E, Bruno L, Bissiere S, Levi V, Plachta N. 2016. Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. *Cell* **165**: 75–87.

Willemin A, Szabó D, Pombo A. 2024. Epigenetic regulatory layers in the 3D nucleus. *Mol Cell* **84**: 415–428.

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J. 2011. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* **2**: 241.

Xia W, Xie W. 2020. Rebooting the Epigenomes during Mammalian Early Embryogenesis. *Stem Cell Rep* **15**: 1158–1175.

Xu Q, Xiang Y, Wang Q, Wang L, Brind'Amour J, Bogutz AB, Zhang Y, Zhang B, Yu G, Xia W, et al. 2019. SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat Genet* **51**: 844–856.

Yaffe E, Farkash-Amar S, Polten A, Yakhini Z, Tanay A, Simon I. 2010. Comparative Analysis of DNA Replication Timing Reveals Conserved Large-Scale Chromosomal Architecture. *PLOS Genet* **6**: e1001011.

Yano S, Ishiuchi T, Abe S, Namekawa SH, Huang G, Ogawa Y, Sasaki H. 2022. Histone H3K36me2 and H3K36me3 form a chromatin platform essential for DNMT3A-dependent DNA methylation in mouse oocytes. *Nat Commun* **13**: 4440.

Zernicka-Goetz M, Morris SA, Bruce AW. 2009. Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. *Nat Rev Genet* **10**: 467–477.

Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q, et al. 2016. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**: 553–557.

Zhang K, Wu D-Y, Zheng H, Wang Y, Sun Q-R, Liu X, Wang L-Y, Xiong W-J, Wang Q, Rhodes JDP, et al. 2020. Analysis of Genome Architecture during SCNT Reveals a Role of Cohesin in Impeding Minor ZGA. *Mol Cell* **79**: 234-250.e9.

Zheng H, Huang B, Zhang B, Xiang Y, Du Z, Xu Q, Li Y, Wang Q, Ma J, Peng X, et al. 2016. Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol Cell* **63**: 1066–1079.

Zou Z, Wang Q, Wu X, Schultz RM, Xie W. 2024. Kick-starting the zygotic genome: licensors, specifiers, and beyond. *EMBO Rep* **25**: 4113–4130.

Results

Part I
Publication (Book chapter)

DamID to map genome-protein interactions in preimplantation mouse embryos

Chapter 18

DamID to Map Genome-Protein Interactions in Preimplantation Mouse Embryos

Mrinmoy Pal, Jop Kind, and Maria-Elena Torres-Padilla

Abstract

Investigating the chromatin landscape of the early mammalian embryo is essential to understand how epigenetic mechanisms may direct reprogramming and cell fate allocation. Genome-wide analyses of the epigenome in preimplantation mouse embryos have recently become available, thanks to the development of low-input protocols. DNA adenine methyltransferase identification (DamID) enables the investigation of genome-wide protein-DNA interactions without the requirement of specific antibodies. Most importantly, DamID can be robustly applied to single cells. Here we describe the protocol for performing DamID in single oocytes and mouse preimplantation embryos, as well as single blastomeres, using a Dam-LaminB1 fusion to generate high-resolution lamina-associated domain (LAD) maps. This low-input method can be adapted for other proteins of interest to faithfully profile their genomic interaction, allowing us to interrogate the chromatin dynamics and nuclear organization during the early mammalian development.

Key words Mouse embryo, Low-input DamID, Single-cell genomics, LADs, Nuclear organization

---

1 Introduction

Unveiling the features and mechanisms behind nuclear organization at the earliest stages of mammalian embryogenesis is essential to understand how the parental genomes are reprogrammed to establish totipotency. In the mouse, the two gametes correspond to very different architectures, in terms of their genome packaging. The sperm is mostly devoid of histones, and the paternal DNA is packaged in a highly compacted configuration through interactions with protamines. In contrast, the oocyte contains histones, which have accumulated a number of histone modifications during oocyte growth. After fertilization, an extensive chromatin remodeling process ensues, which involves changes in histone modifications, de novo deposition of histone variants, transcriptional activation of retrotransposons, but also changes in the 3D genome.

The position of the genome within the 3D nuclear space has emerged as a key epigenetic feature [1–3]. The association with the nuclear lamina, the primary scaffold of the nuclear envelope, is a hallmark of nuclear organization. In higher eukaryotes, chromatin in the proximity of the nuclear lamina tends to be more heterochromatic and gene-poor [4, 5]. These domains are referred to as lamina-associated domains (LADs), ranging in size of 10 kb–10 Mb in mammalian cells (~0.5 Mb median), and display distinctive genomic features. Globally, genes within LADs tend to be lowly expressed, while genes located within the inner nuclear space or inter-LADs (iLADs) display in general much higher transcriptional activity [4–6].

During the last few years, genomic approaches applied to in vivo mouse embryos have enabled us to revisit our molecular understanding of embryonic chromatin [7–13]. However, these approaches have only recently become available, mainly because of the specialist skills required to manipulate the embryo but, most importantly, because of the scarcity of the material available. Low-input methods to investigate the molecular makeup and 3D organization of the chromatin are therefore a valuable tool. Among them, the development of a robust single-cell protocol for DamID was pioneering [14]. This method is based on the ability of the *Escherichia coli* DNA adenine methyltransferase (Dam) to methylate adenines at the N^6 position (^{m6}A) within GATC motifs [15]. Because endogenous ^{m6}A methylation is practically undetectable across most eukaryotes, the methylation catalyzed by ectopic Dam can be identified based on a methylation-sensitive restriction enzyme and the subsequent amplification and sequencing of the methylated genomic DNA. Dam can be tethered to, e.g., a nuclear region of interest by expressing low amounts of a fusion protein between Dam and a component of the nuclear lamina. Indeed, DamID has been used to map interactions of specific proteins, or nuclear compartments, and the genome of several eukaryotes, including *C. elegans* [16], *Drosophila* [17], but also mammalian cells [4, 5]. In fact, a considerable amount of our knowledge on LADs stems from DamID using a LaminB1 fusion protein. As mentioned above, DamID for LaminB1 has been successfully adapted to single cells and can be used for readouts of imaging, using an innovative ^{m6}A-tracer fused to a fluorescent reporter [18], or also of next-generation sequencing [14].

In this chapter, we provide a detailed protocol for performing DamID in preimplantation mouse embryos. We have used this protocol to map LADs in mouse oocytes and embryos [19], but we propose that it can be easily amenable to try with other fusion proteins. Globally, the protocol involves four parts. The first one concerns embryo manipulation, including dissection, microinjection and mRNA production, culture, and collection. The second one includes all the molecular biology steps necessary to produce

high-quality DamID libraries. The third one includes the sequencing protocol and pipelines. Lastly, bioinformatic analyses can be performed to address a number of different questions. Because the sequencing protocols are rather universal based on standard sequencer equipment and bioinformatic methods have been described elsewhere [14], we only provide a brief overview of the third and fourth parts. Most of the chapter is therefore focused on the implementation of the first two parts of the DamID pipeline.

While we have optimized all the above steps for LaminB1 fusions, the protocol may be used for other fusions to interrogate interactions between the genome and other proteins of interest. The main limitation toward this goal can be the natural residence timing of the protein of interest on its target DNA, which may or may not enable efficient methylation. From an experimental viewpoint, using alternative fusion proteins will only require further optimization of the critical steps, which in our view would be (1) optimal concentrations of Dam-fusion (typically DamID experiments are performed under very low concentrations of Dam to avoid nonspecific methylation), (2) determining the optimal time to enable DNA methylation by the Dam-fusion of interest, and (3) determining optimal amplification conditions for library preparation, which may vary depending on the extent of Dam methylation achieved by the fusion of interest.

2 Materials

2.1 Hardware

1. Microinjection system.
2. CO_2 incubator with active humidification.
3. Benchtop centrifuge with tube and plate rotors.
4. Nucleic acid spectrophotometer such as NanoDrop™.
5. Conventional gel electrophoresis equipment.
6. Real-time thermal cycler with 96-well plate format.
7. Thermal cycler with 96-well plate format.
8. Fluorometer such as Qubit.
9. Automated electrophoresis system such as Bioanalyzer or TapeStation.
10. (Access to a facility providing) Illumina sequencer.
11. Optional: UV PCR workstation.
12. Optional: liquid-handling robot (e.g., Nanodrop II).

2.2 Plasmid Constructs and mRNA Preparation

1. pRN3P-^{m6}A-Tracer-EGFP (Addgene plasmid 139403): Insert codes for ^{m6}A-Tracer-EGFP fusion protein. ^{m6}A-Tracer is a C-terminal fragment of DpnI enzyme that specifically recognizes and binds G^{m6}ATC.

2. pRN3P-HA-AID-Dam-LaminB1: Insert codes for *E. coli* DNA adenine methyltransferase (Dam) fused with murine LaminB1 protein. The fusion protein also contains a HA-tag and an auxin-inducible degron (AID) allowing conditional control of protein stability.

3. pRN3P-HA-AID-Dam-only (Addgene plasmid 136065): Insert ORF codes for HA-tagged DNA adenine methyltransferase which contains an AID domain.

4. pRN3P-TIR1–3×Myc (Addgene plasmid 119766): Insert codes for a 3×Myc-tagged plant auxin receptor called transport inhibitor response 1 (TIR1).

5. pRN3P-mbEGFP (Addgene plasmid 139402): Insert codes for a membrane-targeted GAP43-EGFP fusion protein.

6. SfiI enzyme plus 10× CutSmart buffer.

7. Sodium dodecyl sulfate (SDS): 20% SDS prepared in Milli-Q water.

8. Proteinase K: 20 mg/mL stock aliquots stored at −20 °C.

9. TE: 10 mM Tris–HCl pH 7.5 with 1 mM EDTA prepared in nuclease-free water.

10. In vitro transcription kit (e.g., mMESSAGE mMACHINE T3 Transcription Kit).

2.3 Embryo and Oocyte Manipulation and Collection

1. Pregnant mare serum gonadotropin (PMSG) and human chorionic gonadotropin (hCG) 100 IU/mL (store aliquoted at −20 °C).

2. M2.

3. BSA-free M2.

4. Calcium-free M2.

5. M16.

6. Paraffin oil (embryo tested).

7. 3-Isobutyl-1-methylxanthine (IBMX): 200 mM stock prepared in DMSO.

8. 0.5% pronase: diluted in M2 and stored at −20 °C.

9. K-modified simplex optimized medium (KSOM).

10. 500 μM Indole-3-acetic acid (IAA): prepared in KSOM (*see* **Note 1**).

11. Fluorospheres.

12. 35 mm dish.

13. 8-well PCR strips.

14. DamID buffer: 10 mM Tris acetate pH 7.5, 10 mM magnesium acetate, and 50 mM potassium acetate.

2.4 Single-Cell DamID and Next-Generation Sequencing

1. 3× Lysis buffer: 10 mM Tris acetate pH 7.5, 10 mM magnesium acetate, 50 mM potassium acetate, 2% Tween-20, 2% IGEPAL CA-630, and freshly added 2 mg/mL Proteinase K.

2. DpnI enzyme plus 10× CutSmart buffer.

3. T4 DNA ligase (5 U/µL) plus buffer.

4. 50 µM DamID double-stranded adapter: Dissolve Adapter_top and Adapter_bottom to 100 µM in annealing buffer, and then mix equal volumes of both oligonucleotides in a tightly closed tube. Place tube in a container with boiling water and let cool to room temperature to allow slow annealing of adapters.

    Adapter_top   5′   CTAATACGACTCACTATAGGG-CAGCGTGGTCGCGGCCGAGGA 3′
    Adapter_bottom 5′ TCCTCGGCCGCG 3′

5. Annealing buffer: 100 mM potassium acetate and 30 mM HEPES pH 7.5.

6. 25 µM barcoded PCR primers: 5′ NNNNNNBAR-CODGTGGTCGCGGCCGAGGATC 3′.

7. PCR mix (e.g., MyTaq red reaction mix).

8. SPRI beads.

9. Spin column purification kit.

10. End-It DNA End-Repair Kit.

11. Klenow fragment (3′→5′ exo-).

12. Kit for library preparation (onto DNA fragments) (e.g., TruSeq Nano DNA LT library kit).

3 Methods

3.1 Considerations for the Experimental Design

The experimental DamID design involves the expression of a fusion protein of interest and the untethered Dam enzyme as a control. For single-cell DamID (scDamID), the Dam-fusion and the untethered enzyme cannot be simultaneously expressed in the same cell, yet the information obtained from untethered Dam expression can be used to normalize/control for intrinsic Dam methylation activity. Detection of nonspecific contacts will result in interaction profiles that are very similar to profiles obtained with the untethered Dam. As Dam marks primarily open chromatin regions, it also provides reliable insight into single-cell chromatin accessibility [19, 20] (*see* **Note 2**).

DamID involves the in vivo expression of the Dam-fusion protein over a period of time, and therefore protein-DNA interaction represents cumulative interaction profiles of all contacts that occurred within the chosen time frame. This is fundamentally different from methods like chromatin immunoprecipitation

(ChIP) that records only snapshots of current chromatin states. Therefore, it is important to control for the time window in which the lamina-DNA interactions are recorded. Here we used the AID/TIR1 degron system because of the rapid induction upon auxin washouts [21]. An important consideration for determining the time of induction is that, on the one hand, enough time is allowed for sufficient ^{m6}A methylation for single-cell detection, yet on the other hand, the induction should not exceed, e.g., the duration of one interphase (*see* **Note 3**). Another consideration is that upon DNA replication the DamID-mark goes undetected due to the inability of DpnI to digest hemimethylated DNA. Sufficient time should therefore be allowed for the restoration of the fully methylated ^{m6}A state in G2 phase (which is recommended), or cells should be harvested prior to the initiation of DNA replication at G1/S phase. Time windows of Dam-methylation should therefore be carefully chosen depending on the embryonic stage of interest. In our experience, 4–6 h of Dam expression is sufficient to obtain robust methylation profiles.

Optimizing the concentration of mRNA for injections is crucial. The first step in optimizing this condition involves DamID amplifications of embryos injected with mRNA concentration series, followed by gel electrophoresis of 8 μL of DamID PCR products (of a 50 μL reaction), to verify smear intensity and distribution of fragment size. For a typical successful scDamID experiment, a clearly visible smear is expected to appear within 25–30 cycles of PCR (*see* Fig. 3 for an example). Non-injected embryos serve as important PCR amplification controls. Of the experimental conditions that meet this criterion, Illumina sequencing can be performed to obtain information on overall sample quality. For Dam-LaminB1, the genomic profiles are expected to differ from the untethered Dam control and display a genomic organization in large Mb-scale continuous stretches of ^{m6}A enrichment. Such parameters can be assessed by computing autocorrelation (ACF) scores of consecutive genomic regions (e.g., 100 kb bins) and determining the length of runs of continuous stretches of ^{m6}A enrichment on binarized DamID-scores (for details on both methods, *see* ref. 20). Additionally, a very simple metric to assess the quality of a dataset is to determine the enrichment of ^{m6}A reads in target regions (e.g., LADs) over nontarget regions. For the latter, some a priori knowledge about the expected characteristics of the respective interaction profiles is required.

3.2 mRNA Preparation

1. Linearize the plasmids by SfiI enzyme so that T3 promoter site is upstream of the sequences to be transcribed. Prepare reactions with 10 μg of plasmid DNA, 5 μL of 10× CutSmart buffer, and 5 μL of SfiI enzyme, and make up the volume to 50 μL with water.

2. Incubate overnight at 50 °C.

3. Assess complete digestion of plasmid DNA using gel electrophoresis.

4. Eliminate SfiI and possible RNase, by subjecting the sample to Proteinase K treatment. To 50 μL of reaction mixture, add of 50 μL water, 2 μL of 20 mg/mL Proteinase K, and 2.5 μL of 20% SDS.

5. Incubate at 37 °C for 50 min.

6. Heat-inactivate Proteinase K at 72 °C for 10 min.

7. Purify linearized plasmid by phenol-chloroform extraction followed by ethanol precipitation, and resuspend pellet in 20 μL of RNase-free TE.

8. Use 1 μg of linearized plasmid to perform in vitro transcription following manufacturer's instructions of the chosen kit.

9. Purify mRNA by LiCl precipitation to remove unincorporated nucleoside triphosphates and other impurities. Resuspend pellet in 10 μL of RNase-free TE.

10. Assess RNA quality by running it on a freshly prepared 1% agarose gel. Incubate RNA at 70 °C for 10 min to resolve secondary structure, and chill on ice before loading. Quantitate using a NanoDrop or alike.

11. Prepare the mRNA mix as described below and store aliquots at −80 °C. We typically do not reuse injection mixtures, but fresh aliquots can be stored at −80 °C for years. The membrane GFP mRNA is encoded by a GAP43–EGFP cDNA, which contains a dual palmitoylation sequence and serves as a positive control for microinjection, so that only GFP-positive embryos are collected for downstream DamID. An alternative reporter for controlling microinjection can be used.

    mRNA mix for AID-Dam-LaminB1 injections (*see* Fig. 1):

    - 100 ng/μL membrane-EGFP
    - 250 ng/μL TIR1
    - 150 ng/μL m6A-tracer
    - AID-Dam-LaminB1
        - 5 ng/μL for oocyte or zygote collection
        - 10 ng/μL for 2-cell collection
        - 20 ng/μL for 8-cell collection
        - 100 ng/μL for blastocyst collection

    mRNA mix for AID-Dam-Only injections (*see* Fig. 1):

    - 100 ng/μL membrane-EGFP
    - 250 ng/μL TIR1
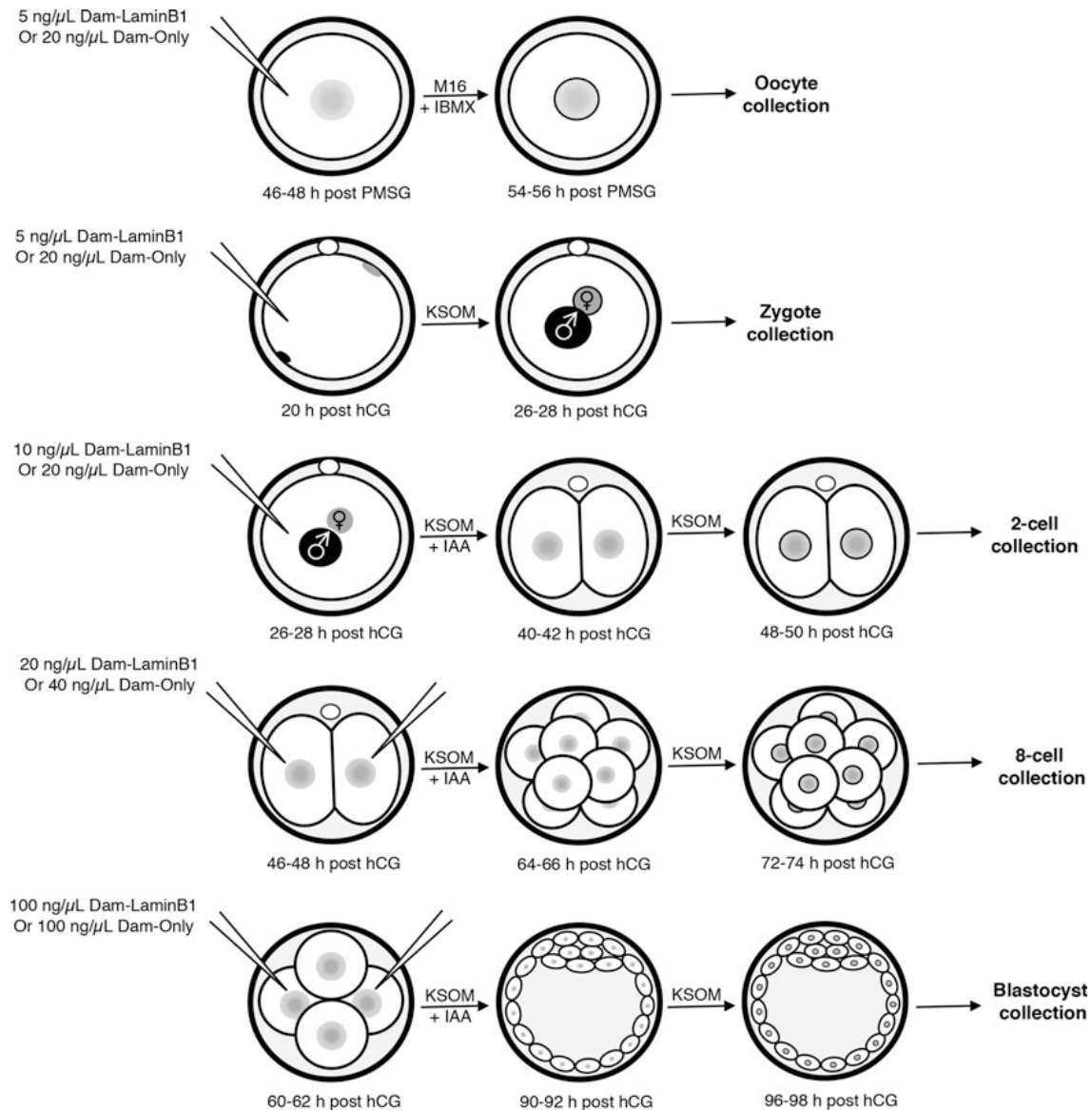
Fig. 1 Schematic of embryo manipulation and collection according to developmental stages. Culture media, recommended concentration of Dam-LaminB1 or Dam-only in the mRNA mixture, the timings of microinjection, auxin washout (through IAA removal), and embryo collection for DamID are indicated

- 150 ng/μL m6A-tracer
- AID-Dam-Only
  - 20 ng/μL for oocyte or zygote collection
  - 20 ng/μL for 2-cell collection
  - 40 ng/μL for 8-cell collection
  - 100 ng/μL for blastocyst collection

3.3 Embryo ***Microinjection*** ***and Culture***	Mate 5–8-weeks-old F1 (CBA × C57BL/6J) females with CAST/ EiJ males for hybrid crosses (*see* **Note 4**) and with F1 males for non-hybrid crosses. Induce superovulation by intraperitoneal injection of 10 IU of PMSG and hCG 46–48 h later. Culture oocytes or embryos in a 37 °C, 5% CO_2 incubator in appropriate media drops covered with paraffin oil prepared on a 35 mm dish. In all embryonic stages from the 2-cell stage, we control DamID temporally, with the addition of auxin, which is done at different times, as described below. *See* Fig. 1 for a schematic summary of the embryo manipulation and collection process.

1. For DamID in oocytes: Isolate GV oocytes 44–48 h after PMSG injection and microinject the mRNA mix. Culture the oocytes in M16 supplemented with 200 μM IBMX. Collect 6–8 h after microinjection.

2. For DamID in zygotes: Obtain early zygotes 20 h post hCG upon natural matings and microinject. Culture zygotes in KSOM drops. Collect at 26–28 h post hCG.

3. For DamID in 2-cell stage embryos: Inject late zygotes collected at 26–28 h post hCG. Culture them in KSOM containing 500 μM IAA. Wash out (*see* **Note 5**) IAA at 40–42 h post hCG and culture them in KSOM drops for 6–8 h before collection.

4. For DamID in 8-cell stage embryos: Inject both blastomeres of late 2-cell embryos harvested 46–48 h post hCG and culture in IAA-containing media. Wash out auxin at 64–66 h post hCG, and collect the 8-cell stage embryos around 72–74 h post hCG.

5. For DamID in blastocysts: Microinject four-cell embryos (at least two blastomeres) collected at 60–62 h post hCG. Wash out IAA at 90–92 h post hCG when blastocysts start to caveat and collect embryos after 6–8 h.

3.4 Embryo ***Collection***	Place 3 mL of DamID buffer into an agar-coated 35 mm dish and keep it at room temperature. Prepare 8-well PCR strips with 2 μL of DamID buffer per tube (*see* **Note 6**).
3.4.1 For Oocyte *Collection*	

1. Remove zona pellucida by incubating the GV oocytes with 0.5% pronase prepared in M2 for 10 min at 37 °C.

2. Transfer the oocytes to the agar-coated dish containing DamID buffer.

3. Take up oocyte(s) just to the tip of a new glass capillary and place into the 8-well PCR strip.

4. Keep full 8-well PCR strips on ice and freeze them at −80 °C until downstream processing.

3.4.2 For Collection of Zygotes, 2-Cell and 8-Cell Stage

1. Remove zona pellucida by treating the embryos with 0.5% pronase prepared in M2 for 10 min at 37 °C.

2. Wash the embryos through calcium-free M2 and incubate for 5 min. Mechanically separate the polar bodies from the embryos with a thin glass capillary by pipetting up and down in calcium-free M2. If single blastomere DamID is performed, mechanical dissociation of individual blastomeres is performed at this step and in the same calcium-free M2 medium, but using an appropriate glass pipette depending on the size of the blastomeres.

3. After removing the polar bodies, transfer the embryos or single blastomeres to the agar-coated dish containing DamID buffer.

4. Take up embryo(s) or single blastomeres with new glass capillary and place into the 8-well PCR strip.

5. Freeze the 8-well PCR strips at −80 °C and store until further processing.

3.4.3 For Blastocyst Collection

1. Remove zona pellucida by treating the embryos with 0.5% pronase containing M2 for 10 min at 37 °C. Pronase treatment in blastocysts may be shorter, due to the natural thinning of the zona pellucida in embryos at later stages.

2. Incubate the embryos in BSA-free M2 containing 1:50 dilution of Fluorospheres in order to label the trophectoderm (TE) (outer layer of cells). Wash out residual Fluorospheres after 2 min. Do not over-incubate; otherwise, labeling of inner cell mass (ICM) may also occur.

3. Keep embryos in calcium-free M2 for 25 min and perform mechanical disaggregation by repeated mouth pipetting with a finely pulled glass pipette (*see* **Note 7**). A heated stage is preferred, in order to maintain the temperature at 37 °C. If this is not available, performing the mechanical separations in small groups (two to three) of embryos is advised.

4. Separate Fluorosphere-positive TE cells from ICM cells under a fluorescent microscope (*see* Fig. 2).

5. Transfer the ICM cells to the agar-coated dish containing DamID buffer. Place them into the 8-well PCR strip using the tip of a fresh glass capillary and freeze them to store.

3.5 Processing Single Embryos or Single Blastomeres for DamID

Following the transfer of single embryos or single blastomeres into PCR strips or 96-well plates, all subsequent additive reactions are performed in the same well without cleaning the sample in between, to keep material loss at a minimum. All subsequent steps are performed at room temperature, unless otherwise specified. If available, using a multistep pipette or liquid-handling robot will highly decrease hands-on time and increase throughput.
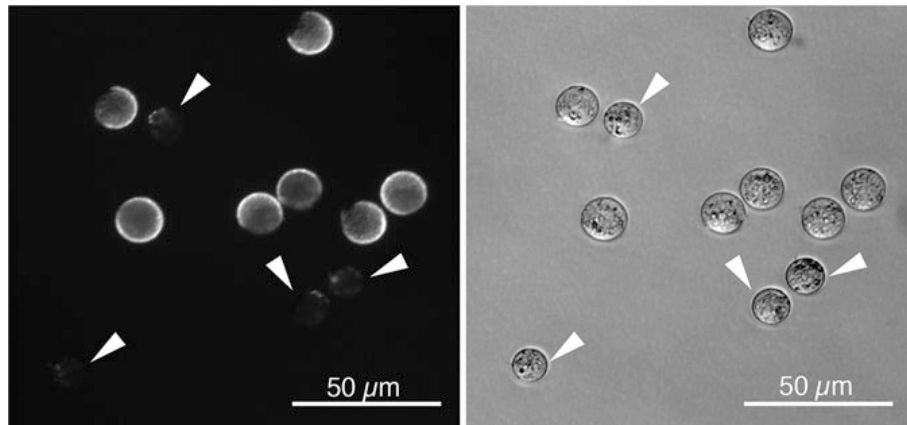
Fig. 2 Representative images showing labeling of inner and outer cells using Fluorospheres. Dissociated cells were imaged under green fluorescence (left panel) or bright-field (right panel). Arrowheads point to inner cells, as can be seen from the lack of fluorescence signal throughout the cell membrane

A decontaminated working environment, such as a UV PCR workstation, is advised.

1. Prepare an appropriate amount of 3× lysis buffer including freshly added Proteinase K, and keep on ice until dispensation. Per sample well, 1 μL of 3× lysis buffer is needed.

2. Dispense 1 μL of 3× lysis buffer per well, and centrifuge at $1000 \times g$ for 1 min to ensure the cells are at the bottom of the well.

3. Incubate plates at 42 °C overnight to lyse cells and digest all proteins.

4. The next day, incubate plates at 80 °C for 20 min to heat inactivate Proteinase K.

5. The plates can now be stored at −20 °C until further downstream processing.

3.6 DamID: Amplification of Dam-Marked Genomic Fragments

Genomic DNA that has been methylated at GATC motifs is specifically digested, leaving blunt ends to which a universal adapter is ligated. Using barcoded primers that hybridize to this adapter, the methylated fragments are specifically enriched for by PCR. To avoid cross contamination, take care not to touch samples between wells or cause spill overs.

1. Add 6.9 μL of 1× CutSmart buffer and 0.1 μL of DpnI enzyme to prepare DpnI digestion mix. Dispense 7 μL of digestion mix to each well.

2. Incubate plates at 37 °C for 8 h to digest methylated DNA.

3. Incubate plates at 80 °C for 20 min to heat inactivate DpnI, and then cool on ice (*see* **Note 8**).

4. Assemble adapter ligation mix and add 10 μL of mix to each well.

   Adapter ligation mix per well: 9.7 μL of 1× T4 ligase buffer, 0.05 μL of 50 μM adapter, and 0.25 μL of T4 ligase (*see* **Note 9**).

5. Incubate plates at 16 °C for 12–16 h.

6. Incubate plates at 65 °C for 10 min to heat inactivate T4 ligase, and then cool on ice.

7. Add 2 μL of 25 μM cell-specific, barcoded PCR primer to each well. Take care because each barcoded primer corresponds to a single sample.

8. On ice, assemble PCR mix and add 28 μL of mix to each well.

   PCR mix per well: 28 μL of PCR buffer including polymerase (e.g. 10 μL of 5× MyTaq Red Reaction buffer, 0.5 μL of MyTaq polymerase, and 17.5 μL of nuclease-free water).

9. Run the assembled reactions in a thermocycler using the program as described:

Step			
1		72 °C for 10 min	
2	94 °C for 1 min	65 °C for 5 min	72 °C for 15 min
3–6	94 °C for 1 min	65 °C for 1 min	72 °C for 10 min
7–35	94 °C for 1 min	65 °C for 1 min	72 °C for 2 min

   When testing new uncharacterized samples, evaluate PCR product by agarose gel electrophoresis after 22 cycles, and adjust cycles as necessary.

10. Run 8 μL of PCR product on 1% agarose gel to check control samples, smear intensity, and distribution of fragment size. Include a 1 kb + DNA ladder. Run more or fewer cycles of PCR if necessary. *See* Fig. 3 for an example of DamID PCR result.

3.7 Preparation of Illumina Sequencing Libraries

The amplified product is multiplexed by pooling together all samples with different barcodes. The pools are subsequently cleaned by gel extraction. Gel extraction is desired because of a frequently observed contaminating product of low molecular weight that impacts on Illumina sequencing efficiency (*see* Fig. 4 for example of pooled samples containing the undesired PCR product). This PCR product is likely caused by the formation and amplification of double-stranded adapter concatemers. After gel extraction, the samples are additionally cleaned with a PCR purification spin column or bead purification step and further processed into libraries for deep sequencing.
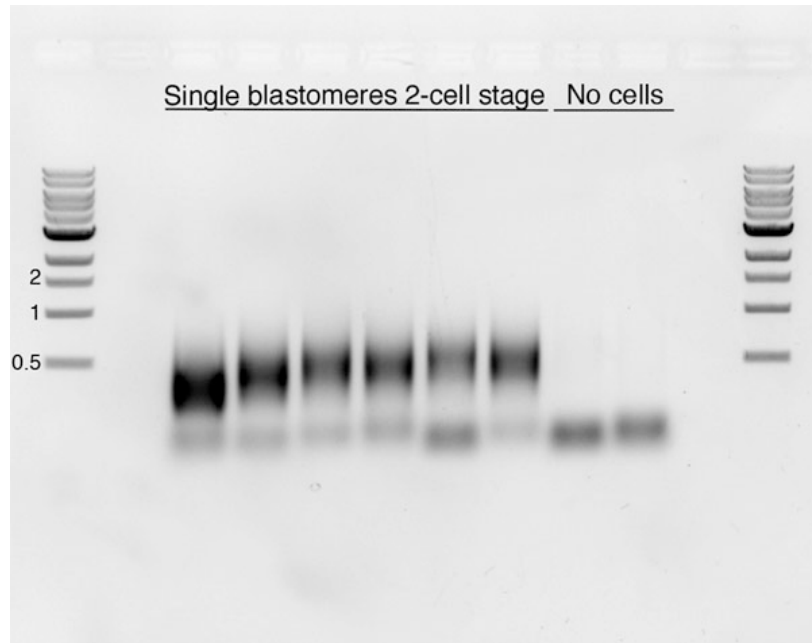
Fig. 3 Representative examples of successful scDamID PCR amplifications of six blastomeres of the 2-cell stage and two negative (empty) controls. The low-molecular-weight material is excess PCR primer
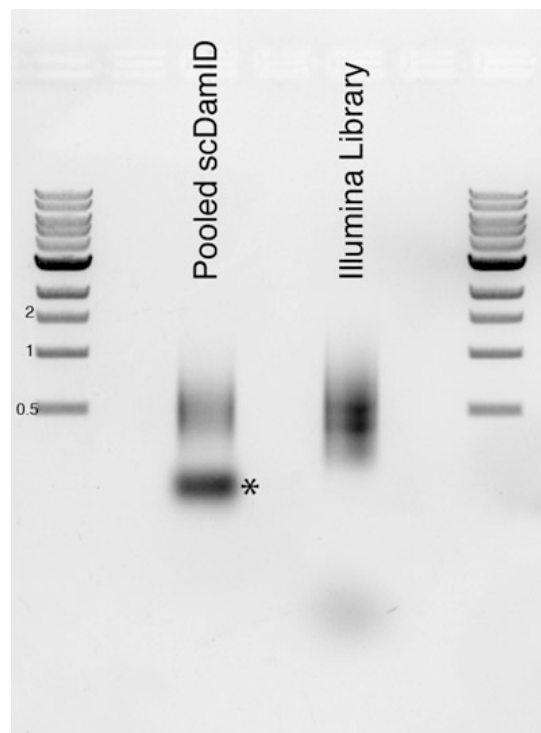
Fig. 4 Example of a successful removal of a common scDamID contaminating product followed by successful Illumina library preparation. On the left is a representative example of pooled scDamID samples prior to gel purification. Gel purification is required to remove the undesired product (indicated with an asterisks). To the right is an Illumina library with the desired molecular weight distribution and without the contaminating product

1. Evaluate PCR product on gel and estimate relative concentration of the different samples.

2. Pool barcoded samples together according to their estimated concentration; the aim is to generate a mixture with equal numbers of molecules across the samples (*see* **Note 10**).

3. Purify the pooled samples by gel extraction followed by PCR purification on spin columns or with SPRI beads (2× bead volume to sample), and elute in 30 μL of nuclease-free water.

4. Measure the concentration of purified PCR products by NanoDrop.

5. Of 300 ng purified PCR product, blunt the 3′ or 5′ overhanging ends in a 50 μL reaction according to DNA End-Repair Kit instructions.

6. Purify the DNA by PCR purification spin columns, and elute in 30 μL of nuclease-free water.

7. Add a 3′ adenine to the DNA ends by incubation at 37 °C for 30 min with Klenow (3′→5′ exo-) (30 μL of DNA, 5 μL of 10× buffer, 0.1 μL of 100 mM dATP, 0.5 μL of enzyme, 14.4 μL of nuclease-free water) followed by heat inactivation at 75 °C for 20 min.

8. Purify the DNA with SPRI beads (1.8× bead volume to sample), and elute in 25 μL of nuclease-free water.

9. Ligate the Illumina indexed Y-shaped double-stranded adapters (provided in the Illumina TruSeq Nano DNA LT library kit) by incubation at room temperature for 2 h (25 μL of DNA, 2.5 μL of double-stranded adapter, 0.5 μL of 5 U/μL T4 DNA ligase, 4 μL of 10× T4 ligase buffer, 8 μL of nuclease-free water) followed by heat inactivation at 65 °C for 20 min.

10. Purify the DNA two times with SPRI beads of 1.8× bead volume to sample, followed by 1.2× bead volume to sample, and elute in 50 μL of nuclease-free water.

11. Perform PCR with 25 μL of the eluted DNA (25 μL of DNA, 10 μL of 5× MyTaq Red Reaction buffer, 0.5 μL of MyTaq polymerase, 1 μL 2.5 μM Illumina oligo mix and 13.5 μL of nuclease-free water) for upto nine PCR amplification cycles (94 °C for 1 min; 94 °C for 30 s, 58 °C for 30 s and 72 °C for 30 s for six to nine cycles, and 72 °C for 2 min).

12. Run 8 μL of the PCR reaction mixture on 1% agarose gel to check smear intensity and distribution of fragment size. Include a 1 kb + DNA ladder. Adjust the number of cycles of PCR if necessary.

13. Purify the DNA by PCR purification spin columns first, followed by a final purification with SPRI beads (1.6× bead volume to sample), and elute in 25 μL of nuclease-free water.

14. Measure the concentration of each library with a Qubit fluorometer, per manufacturer's instructions.

15. Evaluate the fragment distribution of each library with an automated electrophoresis system such as Agilent Bioanalyzer or TapeStation.

16. For Dam-LaminB1 in embryos, sequence single end to a depth of approximately 500 K raw reads per single-cell sample.

17. Typically, for Illumina multiplex sequencing, four to ten libraries are combined in a single sequencing reaction. Each library consists of 20–50 single cells mixed in appropriate equimolar ratios judged from the agarose gel images.

3.8   Raw Data Processing and Visualization

Raw reads are demultiplexed by their library-specific index and by their sample-specific DamID barcode, after which the DamID primer sequence is removed and sequences are aligned to the reference genome. Reads per GATC are counted, summed across sequencing lanes, aggregated in genomic segments, and optionally smoothened for visualization. Dam methylation across the genome is typically calculated using either an "Observed over Expected (OE)" pipeline [14] based on methylation enrichment across genomic bins or an enrichment pipeline based on the log2 ratio of Dam-LaminB1 over Dam-Only [22].

4   Notes

1. Prepare 0.25 M stocks (500×) of IAA in water and store aliquots at −20 °C for up to 2 years. Use one aliquot only once.

2. The preferential methylation of accessible chromatin regions by untethered Dam poses a challenge for Dam-protein fusions for which binding profiles are expected to overlap with open chromatin regions like promoters and enhancers and therefore resemble untethered Dam profiles. A potential solution could be the use of Dam mutants with reduced intrinsic activity and DNA affinity [23].

3. Depending on the construct, the cell type and degron or induction system, timing of the protein stabilization/expression may need to be optimized. We have good experiences with inductions between 4 and 24 h.

4. Hybrid crosses are required to obtain parent-of-origin-specific information. Performing reciprocal crosses (mate CAST/EiJ females with F1 males) is recommended to confirm that the parent-of-origin-specific differences are not a result from a genetic bias derived from different strains.

5. Washing out the auxin involves moving the embryos through an uncovered drop of 500 μL KSOM without IAA followed by

wash through three to four drops of KSOM without IAA in a final dish.

6. During cell collection, it is recommended to include empty wells (0 cell) as negative control (*see* Fig. 3).

7. For mechanical disaggregation, the pipette tip should be flame-polished to remove any sharp edges, and the inner diameter should be almost half of the diameter of the embryo.

8. In conventional DamID, a digestion step with MboI is included to destroy and thereby avoid PCR amplification of fragments with unmethylated GATCs. We do not include this MboI digestion in scDamID, but it is not advised against per se.

9. Lowering the double-stranded adapter concentration from 0.2 μL [14] to 0.05 μL of 50 μM stock concentration helped in reducing this contaminant of embryo samples. It is possible that lowering the double-stranded adapter below 0.05 μL would reduce the contaminating product further without compromising for sample complexity. This has not been tested.

10. We recommend pooling multiple experimental conditions in the same library to avoid batch effects (or, at the very least, enable batch correction). Therefore, if the number of samples exceeds the number of available barcodes, take care to add barcodes to your samples such that multiple conditions can be pooled together. While one library per condition does facilitate future re-sequencing of particular samples, it is best practice not to pool in that manner until after you have established potential differences between experimental conditions.

Acknowledgments

References

1. Bonev B, Cavalli G (2016) Organization and function of the 3D genome. Nat Rev Genet 17:661–678. https://doi.org/10.1038/nrg.2016.112

2. Hug CB, Vaquerizas JM (2018) The birth of the 3D genome during early embryonic development. Trends Genet 34:903–914. https://doi.org/10.1016/j.tig.2018.09.002

3. Zheng H, Xie W (2019) The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Biol 20:535–550. https://doi.org/10.1038/s41580-019-0132-4

4. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453:948–951. https://doi.org/10.1038/nature06947

5. Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SWM, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, Reinders M, Wessels L, van Steensel B (2010) Molecular maps of the reorganization of genome–nuclear lamina interactions during differentiation. Mol Cell 38:603–613. https://doi.org/10.1016/j.molcel.2010.03.016

6. Akhtar W, de Jong J, Pindyurin AV, Pagie L, Meuleman W, de Ridder J, Berns A, Wessels LFA, van Lohuizen M, van Steensel B (2013) Chromatin position effects assayed by thousands of reporters integrated in parallel. Cell 154:914–927. https://doi.org/10.1016/j.cell.2013.07.018

7. Liu X, Wang C, Liu W, Li J, Li C, Kou X, Chen J, Zhao Y, Gao H, Wang H, Zhang Y, Gao Y, Gao S (2016) Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature 537:558–562. https://doi.org/10.1038/nature19362

8. Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q, Liu W, Kou X, Zhao Y, He W, Li C, Chen B, Li Y, Wang Q, Ma J, Yin Q, Kee K, Meng A, Gao S, Xu F, Na J, Xie W (2016) Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature 537:553–557. https://doi.org/10.1038/nature19361

9. Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, Chen H, Du Z, Xie W, Xu X, Huang X, Liu J (2017) 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. Cell 170:367–381.e20. https://doi.org/10.1016/j.cell.2017.06.029

10. Flyamer IM, Gassler J, Imakaev M, Brandão HB, Ulianov SV, Abdennur N, Razin SV, Mirny LA, Tachibana-Konwalski K (2017) Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature 544:110–114. https://doi.org/10.1038/nature21711

11. Gassler J, Brandão HB, Imakaev M, Flyamer IM, Ladstätter S, Bickmore WA, Peters J-M, Mirny LA, Tachibana K (2017) A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. EMBO J 36:3600–3618. https://doi.org/10.15252/embj.201798083

12. Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S, Zhang N, Duan F, Ming J, Zhang X, Niu W, Song W, Jin H, Guo Y, Dai S, Hu L, Fang L, Wang Q, Li Y, Li W, Na J, Xie W, Sun Y (2018) Chromatin analysis in human early development reveals epigenetic transition during ZGA. Nature 557:256–260. https://doi.org/10.1038/s41586-018-0080-8

13. Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, Chen C, Kou X, Zhao Y, Chen J, Wang Y, Le R, Wang H, Duan T, Zhang Y, Gao S (2018) Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nat Cell Biol 20:620–631. https://doi.org/10.1038/s41556-018-0093-4

14. Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, de Graaf CA, Amendola M, Fudenberg G, Imakaev M, Mirny LA, Jalink K, Dekker J, van Oudenaarden A, van Steensel B (2015) Genome-wide maps of nuclear lamina interactions in single human cells. Cell 163:134–147. https://doi.org/10.1016/j.cell.2015.08.040

15. van Steensel B, Henikoff S (2000) Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. Nat Biotechnol 18:424–428. https://doi.org/10.1038/74487

16. Ikegami K, Egelhofer TA, Strome S, Lieb JD (2010) Caenorhabditis elegans chromosome arms are anchored to the nuclear membrane via discontinuous association with LEM-2. Genome Biol 11:R120. https://doi.org/10.1186/gb-2010-11-12-r120

17. Pickersgill H, Kalverda B, de Wit E, Talhout W, Fornerod M, van Steensel B (2006) Characterization of the Drosophila melanogaster genome at the nuclear lamina. Nat Genet 38:1005–1014. https://doi.org/10.1038/ng1852

18. Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B (2013) Single-cell dynamics of genome-nuclear lamina interactions. Cell 153:178–192. https://doi.org/10.1016/j.cell.2013.02.028

19. Borsos M, Perricone SM, Schauer T, Pontabry J, de Luca KL, de Vries SS, Ruiz-Morales ER, Torres-Padilla M-E, Kind J (2019) Genome–lamina interactions are established de novo in the early mouse embryo. Nature 569:729–733. https://doi.org/10.1038/s41586-019-1233-0

20. Rooijers K, Markodimitraki CM, Rang FJ, de Vries SS, Chialastri A, de Luca KL, Mooijman D, Dey SS, Kind J (2019) Simultaneous quantification of protein–DNA contacts and transcriptomes in single cells. Nat Biotechnol 37:766–772. https://doi.org/10.1038/s41587-019-0150-y

21. Nishimura K, Fukagawa T, Takisawa H, Kakimoto T, Kanemaki M (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat Methods 6:917–922. https://doi.org/10.1038/nmeth.1401

22. Marshall OJ, Brand AH (2015) damidseq_pipeline: an automated pipeline for processing DamID sequencing datasets. Bioinformatics 31:3371–3373. https://doi.org/10.1093/bioinformatics/btv386

23. Szczesnik T, Ho JWK, Sherwood R (2019) Dam mutants provide improved sensitivity and spatial resolution for profiling transcription factor binding. Epigenetics Chromatin 12:36. https://doi.org/10.1186/s13072-019-0273-x

Part II
Publication

Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators

nature cell biology

Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators

Bofeng Liu[1,2,8], Yuanlin He[3,4,8], Xiaotong Wu[2,5,8], Zili Lin[1,2,8], Jing Ma[1,2,8], Yuexin Qiu[3], Yunlong Xiang[1,2], Feng Kong[1,2], Fangnong Lai[1,2], Mrinmoy Pal [6], Peizhe Wang [7], Jia Ming[7], Bingjie Zhang[1,2], Qiujun Wang [1,2], Jingyi Wu[1,2], Weikun Xia[1,2], Weimin Shen[2,5], Jie Na [7], Maria-Elena Torres-Padilla [6], Jing Li [3,4] ✉ & Wei Xie [1,2] ✉

Dynamic epigenomic reprogramming occurs during mammalian oocyte maturation and early development. However, the underlying transcription circuitry remains poorly characterized. By mapping *cis*-regulatory elements using H3K27ac, we identified putative enhancers in mouse oocytes and early embryos distinct from those in adult tissues, enabling global transitions of regulatory landscapes around fertilization and implantation. Gene deserts harbour prevalent putative enhancers in fully grown oocytes linked to oocyte-specific genes and repeat activation. Embryo-specific enhancers are primed before zygotic genome activation and are restricted by oocyte-inherited H3K27me3. Putative enhancers in oocytes often manifest H3K4me3, bidirectional transcription, Pol II binding and can drive transcription in STARR-seq and a reporter assay. Finally, motif analysis of these elements identified crucial regulators of oogenesis, TCF3 and TCF12, the deficiency of which impairs activation of key oocyte genes and folliculogenesis. These data reveal distinctive regulatory landscapes and their interacting transcription factors that underpin the development of mammalian oocytes and early embryos.

Enhancers are *cis*-distal regulatory sequences that can activate promoters over great distances[1,2]. They are typically bound by transcription factors (TFs)[3] and are marked by distinct epigenetic signatures[4–6]. Active enhancers frequently bear histone acetylation, such as H3K27ac[7], which helps activate enhancers by attenuating nucleosome stability, increasing chromatin accessibility and promoting enhancer–promoter communication[4,8]. Many active enhancers also exhibit enhancer RNAs (eRNAs), which are often bidirectionally transcribed and could be detected by nascent RNA-seq and cap analysis of gene expression sequencing (CAGE)[9,10].

The oocyte-to-embryo transition (OET) features a transcriptionally quiescent period starting from the end of oocyte growth to zygotic genome activation (ZGA) accompanied by dramatic epigenetic reprogramming[11–13]. Of note, oocytes and early embryos are often subjected to unique or 'non-canonical' transcription and epigenetic regulation. In mice, H3K4me3 exists as widespread non-canonical, broad domains in oocytes and early embryos before ZGA[14–16]. H3K27me3 occurs pervasively in the oocyte genome, occupying most regions without transcription[17]. Such non-canonical H3K27me3 persists after fertilization until the blastocyst stage and plays a critical role in regulating DNA methylation-independent imprinting and X chromosome inactivation[18–20]. Moreover, the higher-order chromatin organization in oocytes and early embryos is also distinct from that in somatic cells. For example, the repressive lamina-associated domains (LADs), which usually

75

occupy gene deserts[21], are absent in fully grown oocytes (FGOs) and are established de novo after fertilization[22]. However, how this epigenetic reprogramming facilitates transcriptional changes during OET remains elusive in mammals. *Cis*-regulatory elements (CREs), such as enhancers, are central in the interplay between chromatin and transcription but are still poorly defined in mammalian oocytes and early embryos. How CREs interact with TFs to elicit transcription in this period remains unclear. Notably, early reports suggested that mouse oocytes and zygotes may even lack enhancer activity based on enhancer reporter assays[23,24]. Here, we presented a complete putative enhancer map from mouse oogenesis to post-implantation development. These data revealed distinctive epigenetic signatures of enhancers in oocytes and early embryos. We further showed that putative enhancers are often bidirectionally transcribed and can drive reporter activities in oocytes. Notably, these *cis*-regulome maps revealed their potential interacting TFs, leading to the identification of key TFs TCF3 and TCF12 that direct oocyte development.

Results

Dynamic CRE activities in mouse oocytes and early embryos

To identify possible CREs in mouse oocytes and early embryos, we performed H3K27ac STAR chromatin immunoprecipitation (IP) sequencing (ChIP-seq)[16] in growing oocytes (GOs) at postnatal day 7 (GO-P7) and day 10 (GO-P10) stages, FGOs at postnatal week 8, metaphase II (MII) oocytes and mouse early embryos at the one-cell, early two-cell, late two-cell and eight-cell stages and inner cell mass (ICM) from blastocysts (Fig. 1a and Methods). Consistent with the immunostaining results (Extended Data Fig. 1a) and the previous studies[25,26], STAR ChIP-seq did not detect H3K27ac enrichment in MII oocytes, which was thus excluded from subsequent analysis (Fig. 1a and Extended Data Fig. 1a,b). H3K27ac ChIP-seq data were well reproduced in replicates (Extended Data Fig. 1b,c). As validations, the promoter H3K27ac levels correlated with stage-specific gene expression (Fig. 1b, left and Extended Data Fig. 2a). Large fractions of H3K27ac peaks (75.4–86.3%) were in distal regions (2.5 kb away from the transcription start sites; TSSs) (Fig. 1a, left), indicating putative enhancers. Distal H3K27ac correlated with chromatin accessibility[19,27] (Extended Data Fig. 2b) and resided near active genes (non-transcribing stages excluded) (Extended Data Fig. 2c). We refer to these distal regions as putative enhancers, hereafter.

Two enhancer transitions around fertilization and implantation

Combined with H3K27ac data in post-implantation embryos that we previously generated[28], we mapped a complete landscape of H3K27ac from oocytes to post-implantation embryos (Fig. 1a). The hierarchical clustering showed two clear transitions of H3K27ac upon OET (as reported[29]) and implantation (Fig. 1c), indicating dramatic epigenetic reprogramming during these periods. This was largely paralleled by the dynamics of the transcriptome during the same period (Fig. 1d and Extended Data Fig. 3). Of note, embryos at the 1-cell and 2-cell stages were clustered with oocytes in transcriptome analyses, presumably due to their inherited maternal RNAs (Fig. 1d). Distal H3K27ac levels seemed to increase from GO-P7 to GO-P10 and elevated further in FGOs at oocyte-specific putative enhancers (Fig. 1b, right), consistent with the immunofluorescence results (Extended Data Fig. 1a). We then identified 63,657, 42,409 and 37,590 distal H3K27ac sites as putative enhancers in mouse oocytes, pre- and post-implantation embryos, respectively. Notably, comparison with those from a panel of tissues (*n* = 94) from ENCODE[30] showed oocytes possessed a large fraction of unique enhancers (*n* = 31,838, 47.7%), in contrast to pre-implantation embryos (*n* = 6,581, 9.7%) and post-implantation embryos (*n* = 7,855, 11.9%) (Fig. 1b, right, discussed below). Similar to global H3K27ac, distal H3K27ac-defined putative enhancers also displayed two waves of transitions during fertilization and implantation (Fig. 1b, right).

Such transitions were also observed at repeats. H3K27ac peaks were enriched for MaLR and ERVK in oocytes (Extended Data Fig. 4), consistent with previous findings[31]. H3K27ac peaks in pre-implantation embryos were enriched for ERV and SINE elements, including B1/B2/B4, agreeing with them being preferentially accessible at these stages[27]. Post-implantation embryos were relatively enriched for mammalian-wide interspersed repeat (MIR) and LINE2 (L2) (Extended Data Fig. 4). Taken together, these data reveal two global regulome transitions centred around fertilization and implantation.

Prevalent H3K27ac in gene deserts in FGOs

Given that oocytes possess many unique enhancers, we sought to characterize them in greater details. During oocyte growth, a transcription switch occurred between GO-P7 and GO-P10, correlated with promoter H3K27ac changes (Extended Data Fig. 5a). For example, *Hexb*, *Sohlh1* and *Sohlh2*, three genes expressed in GO-P7 but not in GO-P10, showed strong promoter H3K27ac only in GO-P7 (Extended Data Fig. 5b). *Oosp1*, *Oosp2* and *Oosp3* genes[32] were highly induced starting from P10, consistent with increased H3K27ac at their promoters (Extended Data Fig. 5b). Distal H3K27ac was also highly dynamic during oocyte growth (Fig. 2a). Compared to GOs and adult somatic tissues, FGOs showed the most stage-specific putative enhancers (44.3%, *n* = 18,200) (Fig. 2b). Further analysis showed that FGOs exhibited increased H3K27ac (Fig. 2c,d), gene expression (Extended Data Fig. 5c, left) and active repeats (Extended Data Fig. 5c, right, and Extended Data Fig. 5d) in gene-poor regions or gene deserts (Methods). For example, the *Oosp* gene cluster is present in gene deserts (Extended Data Fig. 5b, right). In sum, these data revealed prevalent H3K27ac and potential regulatory activities in gene deserts in FGOs linked to oocyte-specific genes and repeats.

Allelic reprogramming of H3K27ac after fertilization

We then examined H3K27ac in early embryos, by separating allelic signals using single-nucleotide polymorphisms (SNPs) present between the two parental strains (Fig. 2e and Methods). Upon meiotic resumption, H3K27ac is quickly erased in MII oocytes (Extended Data Fig. 1a) due to the recruitment of histone deacetylases to metaphase chromatin[25,26,29]. After fertilization, H3K27ac reappears at the pronuclear stage 3 to pronuclear stage 4 (ref. 33). A comparison of sperm and zygote paternal H3K27ac revealed distinct patterns (Fig. 2e,f), suggesting rapid paternal reprogramming upon fertilization, similar to that for H3K4me3 and H3K27me3 (refs. 16,17). Notably, the maternal allele of the PN5 zygote showed H3K27ac enrichment that partially resembled that of FGOs (Fig. 2e,f), suggesting that regulatory elements may be bound by the inherited maternal factors in the one-cell embryos. However, H3K27ac in gene-poor regions showed an evident decrease after fertilization (Fig. 2e,g, 'gene-poor'). H3K27ac peaks lost in the one-cell embryos enriched for similar TF motifs as those that retained H3K27ac compared with FGOs (Extended Data Fig. 6a, discussed later), suggesting that the absence of these peaks is not due to the loss of maternal TFs, but is likely related to the chromatin reorganization. The exact mechanisms underlying such depletion of H3K27ac in gene deserts remained to be determined. Notably, LADs are absent in FGOs and are de novo established after fertilization in mice[22]. Consistently, the strengths of the genome–lamina interaction and H3K27ac were anti-correlated in early embryos and mouse embryonic stem (mES) cells, except on the maternal allele of the one-cell embryos (*R* = 0.20) (Extended Data Fig. 6b), which was shown to feature fragmented LADs[22]. Given the enrichment of H3K27ac in gene deserts appeared as early as GO-P10 (Fig. 2b), we thus asked whether LADs were already absent at GO-P10 stage using DNA adenine methyltransferase identification (DamID)[22]. Indeed, Dam-lamin B1 profiles suggested the absence of LADs at this stage (Extended Data Fig. 6c). We could not map LADs in an even earlier stage such as GO-P7 due to the difficulty of performing robust microinjection given their smaller sizes. In sum, these
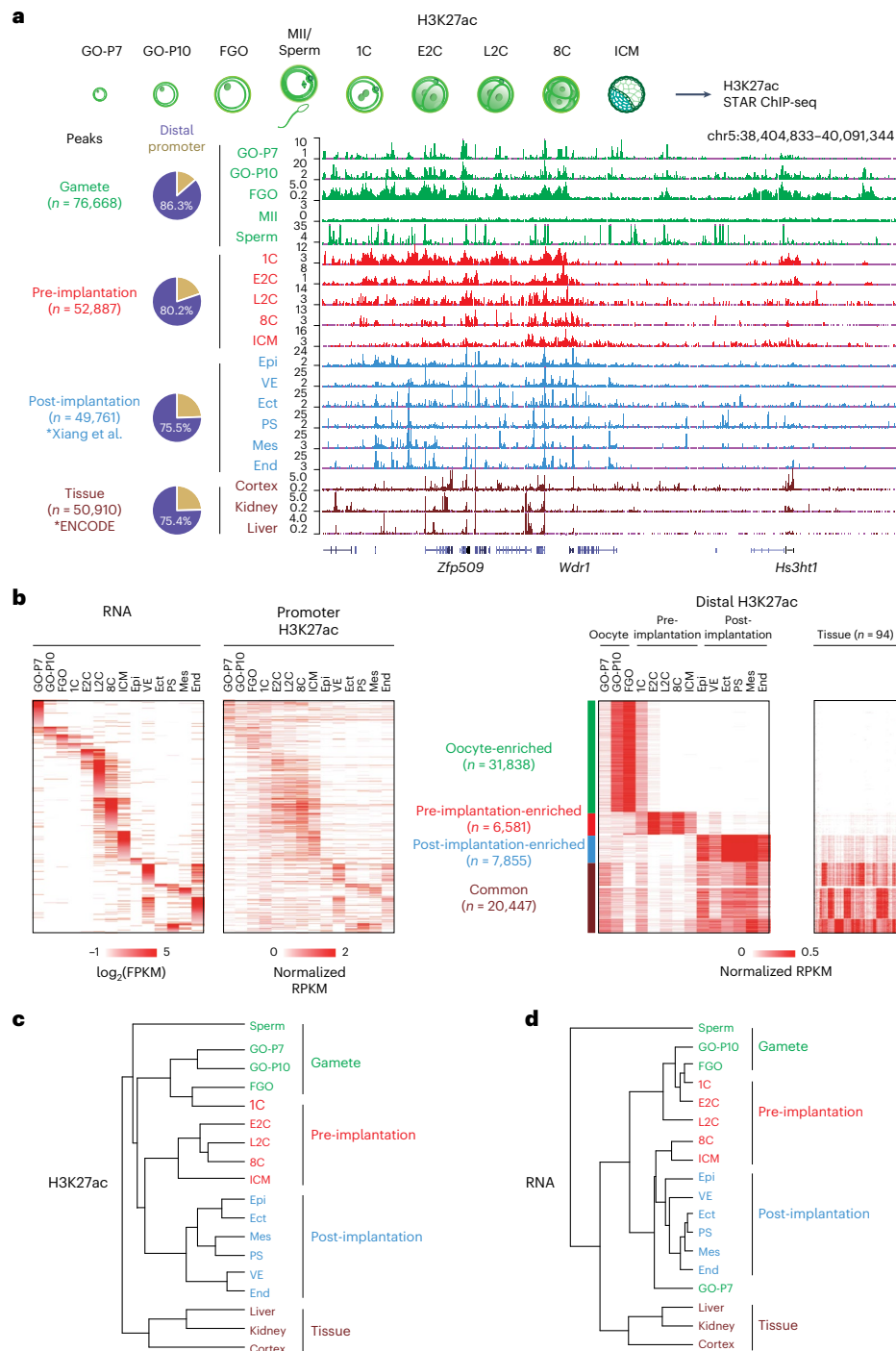
Fig. 1 | H3K27ac landscape in mouse gametes, early embryos and tissues.
a, Schematic showing the overview of H3K27ac STAR ChIP-seq in mouse gametes and early embryos. The UCSC browser view shows H3K27ac signals in gametes (GO-P7, GO-P10, FGO, MII oocyte and sperm), pre-implantation embryos (one-cell PN5, early two-cell, late two-cell, eight-cell and ICM), post-implantation embryos (Epi, VE, Ect, PS, Mes and End) and tissues (cortex, kidney and liver). Pie charts show H3K27ac peak distribution at the promoter and distal regions. Epi, epiblast; VE, visceral endoderm; Ect, ectoderm; PS, primitive streak; Mes, mesoderm; End, endoderm. **b**, Heatmaps showing the stage-specific

gene expression and H3K27ac signals for the corresponding promoters (left). Heatmaps showing the oocyte- (GO-P7, GO-P10 and FGO), pre-implantation- (one-cell, early two-cell, late two-cell, eight-cell and ICM), post-implantation- (Epi, VE, Ect, PS, Mes and End) enriched and common putative active enhancers marked by distal H3K27ac (right). H3K27ac signals of 94 mouse tissues are also mapped. **c**, Hierarchical clustering of global H3K27ac in 2-kb bins. **d**, Hierarchical clustering of gene expression across all stages. Source numerical data are available in source data.
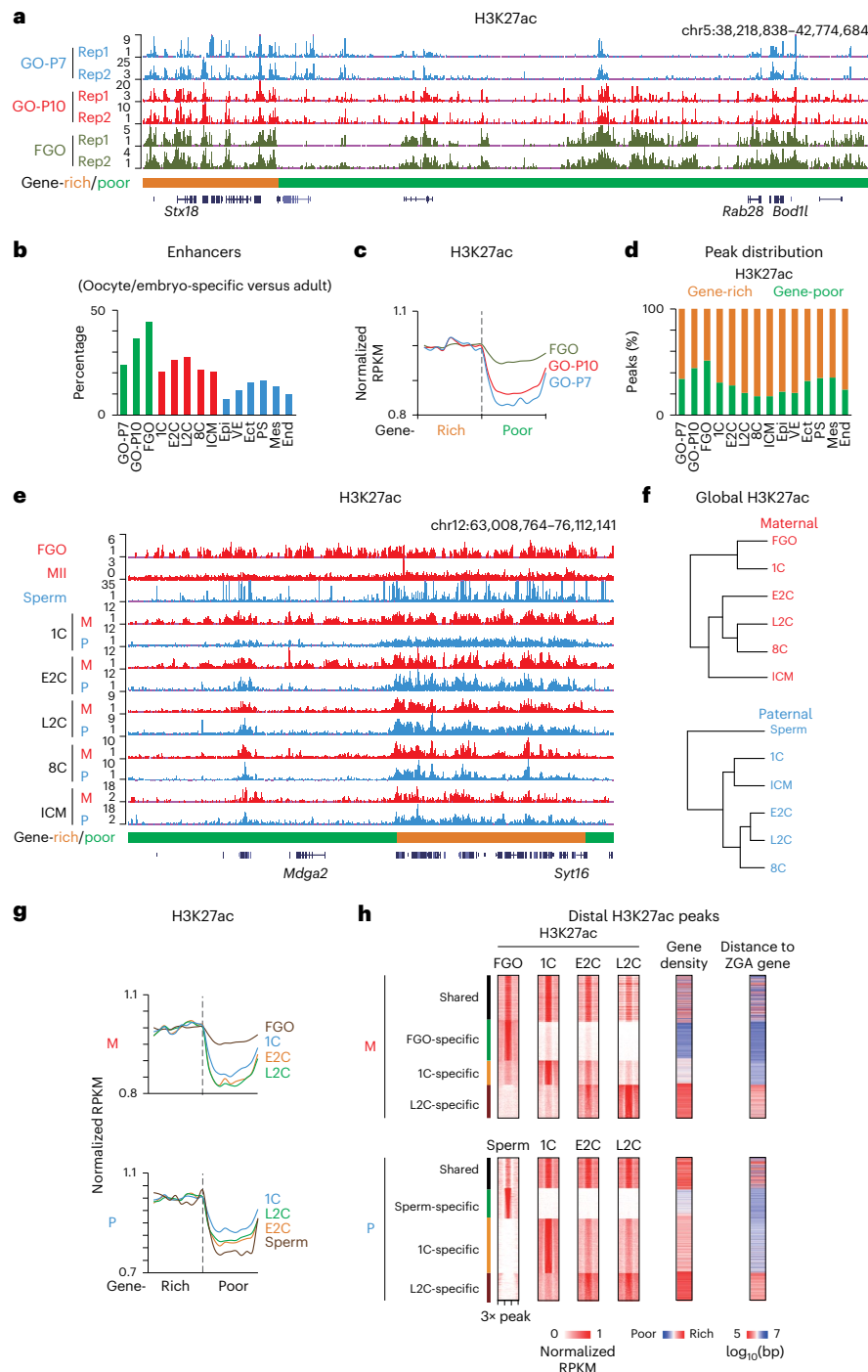
Fig. 2 | Reprogramming of H3K27ac in mouse gametes and early embryos.
a, The UCSC browser view showing H3K27ac signals in GO-P7, GO-P10 and FGOs with two replicates. Gene-rich (orange) and gene-poor (green) regions are also shown. **b**, Bar chart showing the percentages of the unique enhancers at each stage compared with adult tissues. **c**, Line chart showing normalized H3K27ac signals of GO-P7, GO-P10 and FGO at gene-rich regions and the nearby gene-poor regions. **d**, Bar chart showing distributions of H3K27ac peaks in gene-rich (orange) and gene-poor (green) regions at each stage. **e**, The UCSC browser view showing allelic H3K27ac signals in FGO, MII oocytes, sperm, PN5 zygote, early two-cell, late two-cell, eight-cell and ICM. M, maternal (red). P, paternal (blue).

Gene-rich and gene-poor regions are also shown. **f**, Hierarchical clustering of FGOs and early embryos on allelic H3K27ac enrichment. For FGO and sperm H3K27ac data, only regions covered by SNPs were included for analysis. **g**, Line charts showing normalized H3K27ac signals in gene-rich and gene-poor regions in gametes (FGO and sperm) and early embryos (one-cell, early two-cell and late two-cell). M, maternal; P, paternal. **h**, Heatmaps showing distal H3K27ac signals, gene density and distance to nearby ZGA genes at shared, FGO/sperm-specific, one-cell-specific and late two-cell-specific allelic peaks. Source numerical data are available in source data.
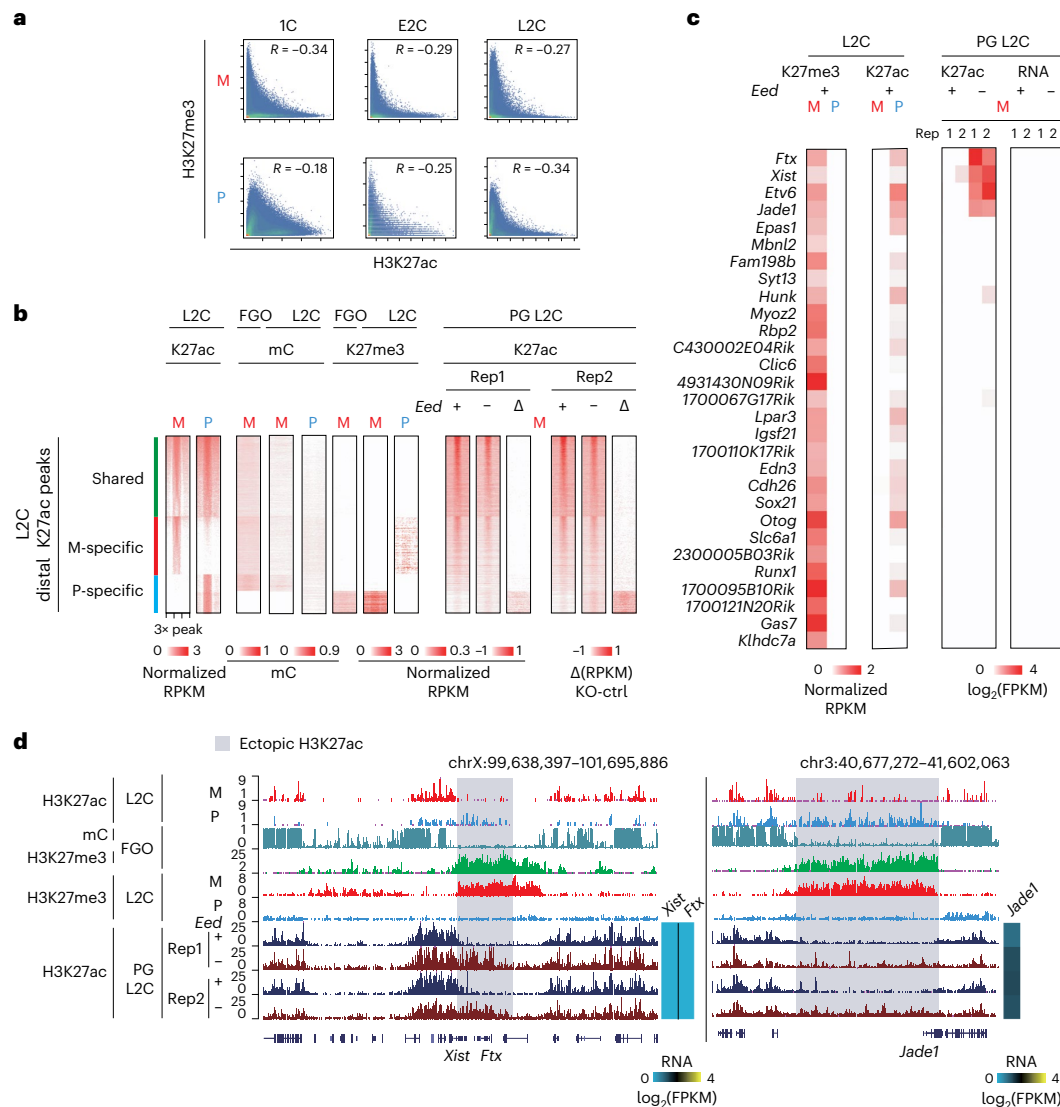
78

Fig. 3 | Reprogramming of H3K27ac and H3K27me3 during ZGA. a, Scatter-plots comparing allelic H3K27ac and H3K27me3 in the one-cell, early two-cell and late two-cell embryos. M, maternal. P, paternal. Pearson correlations of each stage are also shown. **b**, Heatmaps showing H3K27ac, DNA methylation (mC) and H3K27me3 signals at late two-cell H3K27ac maternal-specific (red) and paternal-specific (blue) peaks in WT (left) and parthenogenetic (PG) *Eed* control and KO late two-cell embryos (right). **c**, Heatmaps showing H3K27me3, H3K27ac signals

and RNA at H3K27me3-imprinting genes in WT, parthenogenetic *Eed* control and KO late two-cell embryos. M, maternal. P, paternal. **d**, The UCSC browser view showing H3K27me3, H3K27ac and mC signals of WT late two-cell and FGO, parthenogenetic control and *Eed* knockout late two-cell embryos at *Xist*, *Ftx* and *Jade1*. Heatmaps show RNA expression of related genes. Ectopic H3K27ac is shaded.

results suggest that after fertilization, the paternal H3K27ac undergoes reprogramming and the maternal H3K27ac adopts an FGO-like pattern, while losing enrichment in gene-poor regions.

At the late two-cell stage, H3K27ac distal peaks preferentially resided in gene-rich regions and were proximal to major ZGA genes (Fig. 2d,h). Notably, early two-cell embryos manifested an intermediate H3K27ac state between those of the one-cell and late two-cell embryos, as they already bore H3K27ac in sites destined to be activated in late two-cell embryos (Fig. 2h), suggesting chromatin priming before major ZGA. This finding echoes the 'pre-configuration' of RNA polymerase II (Pol II) to major ZGA genes at the early two-cell stage[34]. Therefore, these data indicate that the activities of H3K27ac-marked regulatory elements also undergo pre-configuration before ZGA.

Maternal H3K27me3 represses putative embryonic enhancers

H3K27ac undergoes erasure in MII oocytes and re-establishment after fertilization (Fig. 2e and Extended Data Fig. 1a). Yet, how H3K27ac is re-established remains elusive. Maternally deposited H3K27me3 represses enhancers during ZGA in *Drosophila*[35]. Oocyte-derived H3K27me3 is also inherited in mouse early embryos and regulates gene imprinting[19]. Notably, H3K27me3 is mutually exclusive with H3K27ac in the one-cell and two-cell embryos (Fig. 3a). Therefore, we asked whether the activities of embryonic enhancers around ZGA are affected by oocyte-inherited H3K27me3 in mouse embryos. We deleted *Eed* in oocytes using *Gdf9-Cre*, erasing H3K27me3 globally[36]. Due to limited SNPs between the parental strains which prevented allele distinction in embryos (Methods), we obtained parthenogenetically activated

(PG) late two-cell embryos from *Eed*[−/−] oocytes and profiled H3K27ac. We first identified regions that showed paternal-specific H3K27ac in wild-type (WT) embryos and further classified them into those marked by H3K27me3 or DNA methylation on the maternal allele. Indeed, we observed ectopic H3K27ac in maternal H3K27me3-marked regions in *Eed*[−/−] PG embryos (Fig. 3b, right, '△'). These regions included, but were not limited to, a subset of the H3K27me3-controlled imprinted regions[19] such as *Xist*, *Etv6* and *Jade1* (Fig. 3c,d). Nevertheless, these imprinted genes remained silenced based on the RNA-seq analyses (Fig. 3c, 'RNA'). We speculate that additional regulators, such as key TFs, are necessary for the ultimate activation of these genes. Thus, oocyte-inherited H3K27me3 represses putative embryonic enhancers in mouse embryos.

H3K4me3 marks enhancers in oocytes and pre-implantation embryos

Previously, we found that somatic enhancers were aberrantly activated and acquired H3K4me3 in *dnmt1*-deficient zebrafish early embryos that lost DNA methylation[37]. As mammalian oocytes and embryos are naturally hypomethylated[38,39], we asked whether their putative enhancers may be also susceptible to acquiring H3K4me3. In line with previous work[29], distal H3K27ac sites were preferentially marked by H3K4me3 in oocytes and pre-implantation embryos, but were less so in post-implantation embryos, which became DNA hypermethylated (Fig. 4a). Consistent with the antagonism between H3K4me3 and DNA methylation in FGOs[40], putative enhancers with high levels of DNA methylation showed low levels of H3K4me3 in WT FGOs but acquired H3K4me3 in *Dnmt3a/b* mutant FGOs (Fig. 4b and Extended Data Fig. 7a). This result echoed a similar finding for enhancers in *Dnmt3a/b* double knockout mES cells (Extended Data Fig. 7b). One exception is GO-P7, where the entire genome is DNA hypomethylated[41] but the enhancers did not show strong H3K4me3 enrichment (Fig. 4a), suggesting that additional mechanisms may exist to prevent H3K4me3 deposition. H3K4me3 is closely associated with RNA Pol II, including that at enhancers[42]. Consistently, enhancers with the H3K4me3-H3K27ac dual mark were more likely to be bound by Pol II and showed shorter distances to nearby active genes compared with H3K27ac-only enhancers (Extended Data Fig. 7c,d). Finally, in oocytes (except GO-P7) and pre-implantation embryos, but not post-implantation embryos, H3K4me3/H3K27ac dual-marked and H3K27ac-only distal regions showed comparable enrichment for distal CREs identified by ENCODE[43] (Extended Data Fig. 7e). Taken together, H3K4me3 also marks a portion of putative active enhancers in oocytes and pre-implantation embryos, a feature that is closely linked to global DNA hypomethylation.

Oocyte enhancers are transcribed and drive reporter expression

We then sought to functionally validate putative enhancers in oocytes. Given eRNAs were shown to be a reliable marker for active enhancers[44,45], we took advantage of a CAGE dataset in mouse GO-P14 oocytes[46] to assess whether putative enhancers were transcribed (Fig. 4c). CAGE allows the mapping of the transcription initiation sites of unidirectional transcribed RNAs at TSSs and bidirectionally transcribed RNAs preferentially at enhancers[10]. We identified 7,157 unidirectionally and 2,786 bidirectionally transcribed sites using CAGEr[47] and CAGEfightR[48] (Fig. 4c and Methods). To strictly exclude promoters, we used an expanded promoter annotation that included defined TSSs of de novo assembled oocyte transcripts using deep-depth RNA-seq data[31] and pooled promoters from a panel of somatic cells (191,499 H3K4me3 sites from 26 tissues from ENCODE[49]). These data confirmed that unidirectionally transcribed sites predominantly (99.5%) enriched for promoters (overlapping with expanded oocyte TSSs, somatic H3K4me3 sites or both). By contrast, only 60.6% of bidirectionally transcribed sites overlapped with annotated promoters. We considered the rest 39.4% as 'CAGE-enhancers' (*n* = 1,097; Methods). The small number of CAGE-enhancers was likely an underestimation of enhancers in oocytes

considering the limited sensitivity of low-input CAGE. Reassuringly, 91.4% of CAGE-enhancers overlapped with H3K27ac (compared with 16.2% of random sites) (Extended Data Fig. 7f). Overall, 90.8% were also occupied by H3K4me3 (compared with 20.5% of random sites) and 83.4% were marked by both marks (compared with 7.3% of random sites). Consistent with the essential roles of histone acetylation in transcription[50], removing histone acetylation in NSN-FGO by Plumbagin, an inhibitor for histone acetyltransferases (HATs)[51], blocked transcription as measured by EU staining (Extended Data Fig. 8a,b). On the other hand, the role of H3K4me3 at enhancers remains elusive. H3K4me3 at enhancers is reported to cause enhancer overactivation[52]. Yet, a mutation in *Mll2*, which encodes an H3K4me3 methyltransferase in oocytes, caused a substantial reduction of non-promoter H3K4me3, which had little correlation with transcription defects[40]. However, a detailed analysis showed that 83.3% of CAGE-enhancers[46] and 52.5% of H3K4me3-marked distal H3K27ac sites still retained H3K4me3 upon the ablation of *Mll2* (Extended Data Fig. 8c,d). Therefore, it remains to be determined whether H3K4me3 is functional at these putative enhancers. Taken together, these data show that putative enhancers in oocytes are often marked by H3K4me3, H3K27ac and bidirectional transcription.

To further validate the putative enhancers in oocytes, we employed self-transcribing active regulatory region sequencing (STARR-seq)[53]. We optimized the STARR-seq protocol for low-input samples with an improved RNA recovery method adapted from Smart-seq2 (ref. 54; Methods and Extended Data Fig. 9a). As the limited oocytes were insufficient to support the assessment of all enhancers, we constructed the STARR-seq plasmid library by manually cloning 70 enhancer candidates with strong bidirectional CAGE signals and distal H3K27ac peaks (84% also carried H3K4me3, *n* = 59) (Fig. 4d, 'CAGE +') (Methods) and 16 negative regions (regions with neither CAGE signals nor H3K27ac in oocytes, including two putative embryo-specific enhancers near *Nanog* and *Fgf3*) (Fig. 4d,e). We also tested whether H3K27ac alone, without CAGE signals, can mark active enhancers by cloning two such putative enhancer sites near *Nobox* and *Bmp15* (Fig. 4d, 'CAGE −'). The STARR-seq plasmid library was injected into the nuclei of FGOs and RNAs were extracted for sequencing. Our results showed that 64% (45 of 70, with 37 out of 45 (82%) carrying H3K4me3) of putative enhancers showed bidirectional activities, compared with none (0 of 16) from the negative controls (*P* = 1 × 10⁻⁶; Fig. 4e and Extended Data Fig. 9b). For example, on chromosome 16, all three candidate sites, but not the negative control, showed strong STARR-seq RNA signals (Extended Data Fig. 9c). While most of these elements enhanced reporter activities in both orientations, some did exhibit stronger activation abilities for one orientation than the other (Fig. 4e), echoing the observation that enhancers are largely but not completely orientation-independent[55,56]. Moreover, the putative enhancers near *Nobox* and *Bmp15* with no CAGE signals only showed weak STARR-seq signals, often for only one orientation (Fig. 4f, 'CAGE −'), raising the possibility that CAGE together with H3K27ac may be a better mark for active enhancers than H3K27ac alone. To further validate these enhancers, we cloned seven candidate enhancers that showed positive STARR-seq signals, four negative controls (including embryo-specific enhancer candidates near *Nanog* and *Fgf3*) and the two 'CAGE −' candidate enhancers near *Nobox* and *Bmp15* into the GFP (mNeonGreen) reporter (Fig. 4g, Extended Data Fig. 9d,e and Methods). A *Zp3* promoter-driven mCherry construct was co-injected as a control to normalize GFP signals. Our data showed that 100% (7 out of 7) of positive candidates, but neither the four negative controls nor the two 'CAGE −' candidates near *Nobox* and *Bmp15*, showed enhanced GFP expression compared with the empty vector (Fig. 4g and Extended Data Fig. 9d,e). Notably, these positive enhancers also preferentially showed Pol II binding in FGOs[34], compared with 'CAGE −' enhancers and negative controls (Fig. 4g and Extended Data Fig. 9d; 'Pol II'). Genome-wide analyses also showed that distal H3K27ac sites associated with Pol II tended to have strong H3K4me3 and CAGE signals and were closer to
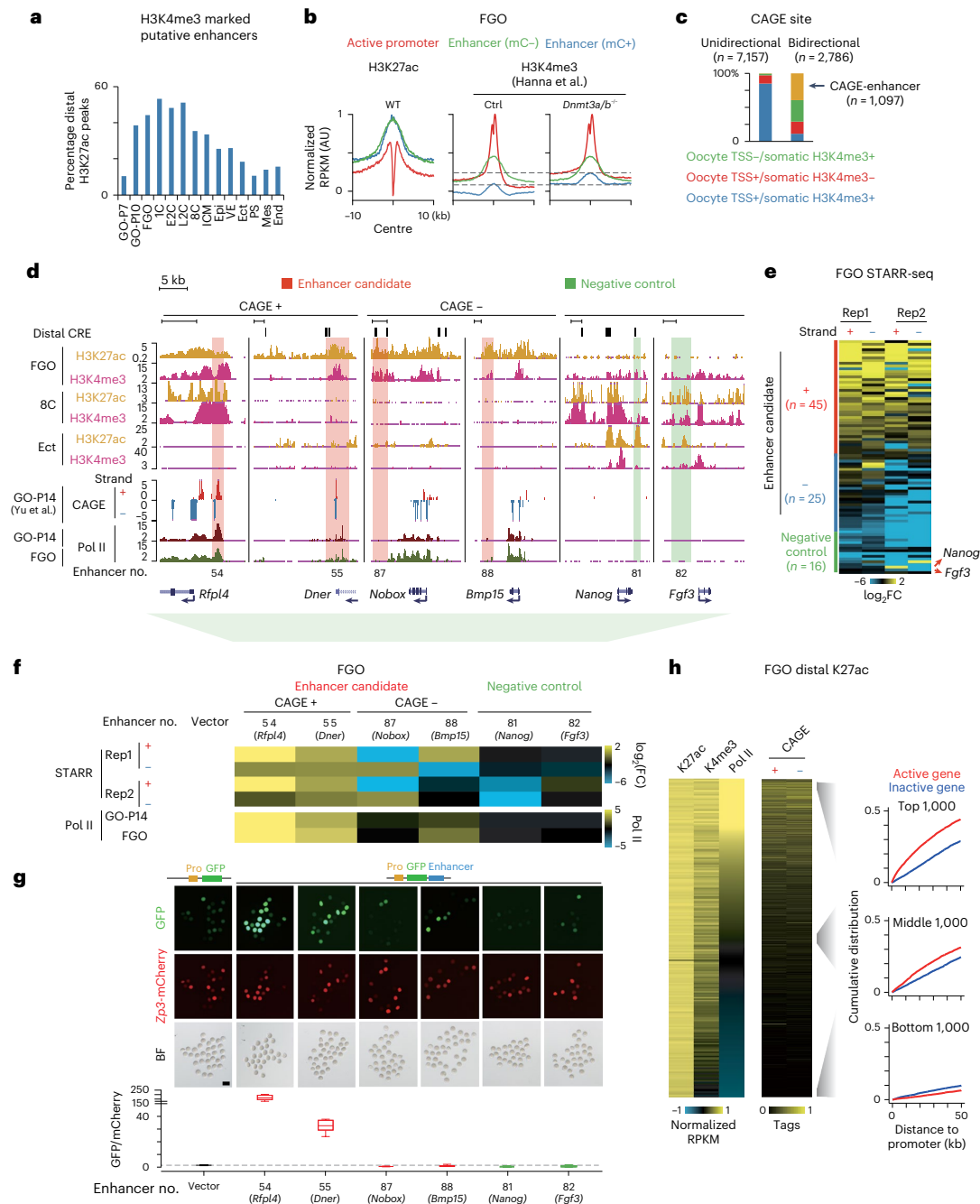
Fig. 4 | Validation of enhancers in FGOs. a, Bar chart showing the percentages of H3K4me3-marked putative enhancers (marked by H3K27ac) at each stage from oocytes to post-implantation embryos. **b**, Line charts showing H3K27ac and H3K4me3 signals at putative enhancers with low (green) or high (blue) DNA methylation and active promoters (red) in WT (left), control (middle) and *Dnmt3a/Dnmt3b* KO (right)[40] FGOs. The dashed lines indicate the peaks of H3K4me3 signals at putative enhancer regions. **c**, Bar chart showing the overlap between uni- or bidirectional CAGE sites and oocyte TSSs or somatic H3K4me3 sites. **d**, The UCSC browser views showing H3K27ac and H3K4me3 enrichment and CAGE signals near oocyte candidate enhancers and negative controls. **e**, Heatmaps showing FGO STARR-seq signals (STARR/input) for enhancer candidates (*n* = 70) and negative controls (*n* = 16, including putative embryo-specific enhancers near *Nanog* and *Fgf3*). **f**, Heatmaps showing STARR-seq

(STARR/input) signals in FGO with two replicates and Pol II signals in GO-P14 and FGO at enhancer candidates (orange shade in **d**) and negative control (green shade in **d**) regions. **g**, Top, fluorescence and bright fields of mouse FGOs in the enhancer reporter assay (Pro, minimal promoter). Scale bar, 100 μm. Boxplot showing the ratio of GFP to mCherry intensity in the enhancer reporter assay (bottom). The numbers of oocytes used in each group: 17, 18, 15, 13, 10, 11 and 6. The median is indicated by the centre line. The bottom, top edges and whiskers represent the 10th and 90th percentiles and 1.5 × interquartile range (IQR), respectively. The dashed line indicates the ratio in the empty vector group. **h**, Heatmaps showing H3K27ac, H3K4me3 and CAGE signals at all distal H3K27ac peaks in FGOs (ranked by Pol II signals) (left). Line charts showing the cumulative distribution of the distances between promoters of active and inactive genes (top 5,000) and nearest distal H3K27ac sites (top, middle and bottom 1,000 peaks) (right).
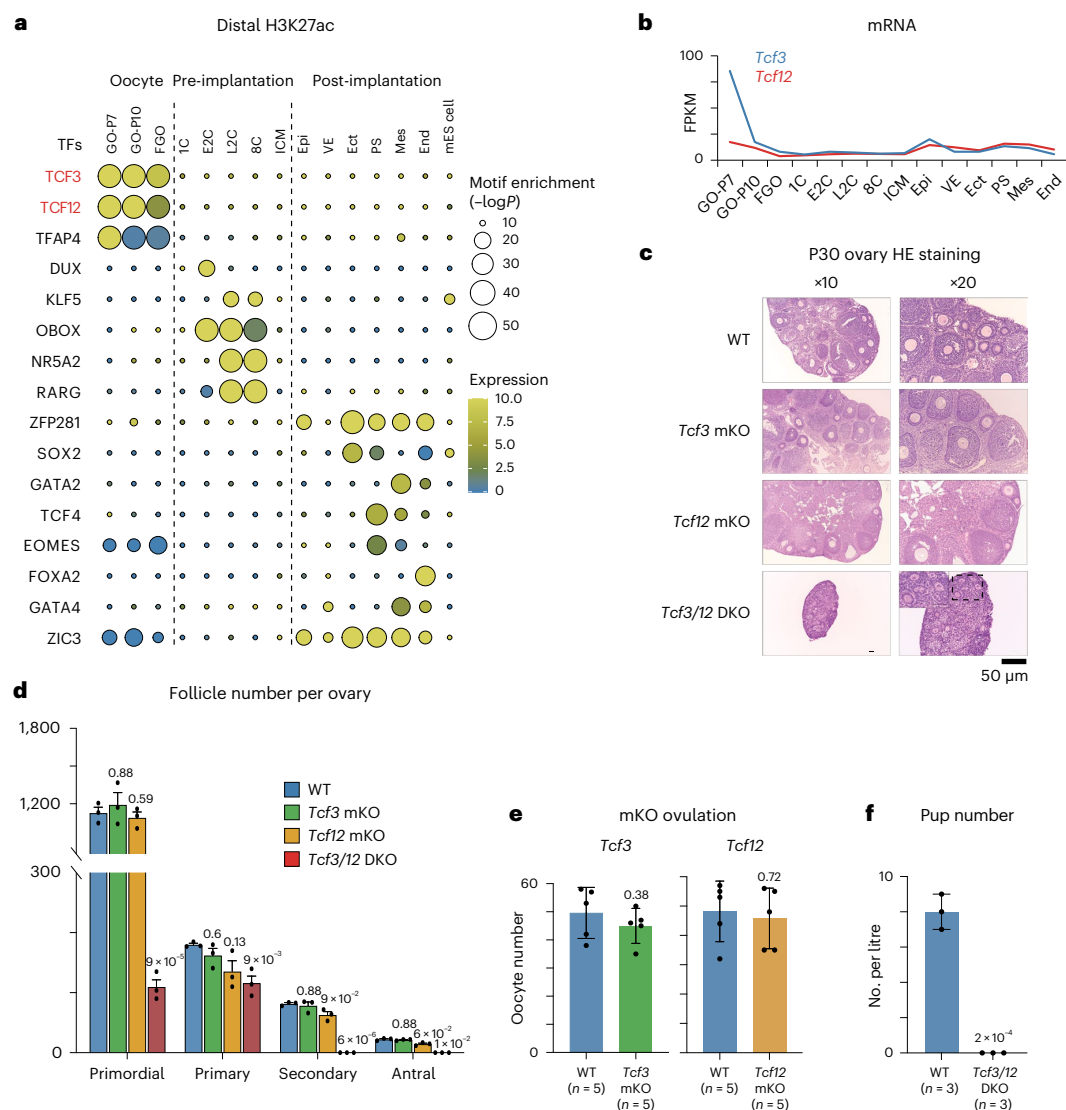
81

Fig. 5 | Identification of candidate TF regulators from putative enhancer maps during oogenesis and pre-implantation embryos. a, Transcription factor motifs identified from distal H3K27ac peaks at each stage in mouse oocytes (GO-P7, GO-P10 and FGO), embryos (one-cell, early two-cell, late two-cell, eight-cell, ICM from blastocyst, Epi, VE, Ect, PS, Mes and End and ES cells. Each circle represents a distinct TF motif, with the colour indicating the RNA expression level and the size indicating the enrichment of the motif ($-\log_{10}$ P value, hypergeometric test with Bonferroni correction, one-sided, from HOMER[57]; Methods). **b**, Line chart showing RNA levels of *Tcf3* and *Tcf12* in oocytes and early embryos. **c**, Haematoxylin and eosin (H&E) staining of ovary sections

from WT, *Tcf3* mKO, *Tcf12* mKO and *Tcf3/12* DKO mice at postnatal day 30. Scale bar, 50 μm. **d**, Bar chart showing the follicle numbers of WT (blue), *Tcf3* mKO (green), *Tcf12* mKO (yellow) and *Tcf3/12* DKO (red) P30 ovaries (n = 3 biological replicates). *P* value (*t*-test, two-sided) is also shown. The error bars represent the s.e.m. **e**, Bar charts showing ovulation rates in WT and *Tcf3* (left) or *Tcf12* (right) mKO mice (n = 5 biological replicates). *P* value (*t*-test, two-sided) is also shown. The error bars represent the s.e.m. **f**, Bar chart showing the number of pups per litter in WT and *Tcf3/12* DKO mice (n = 3 biological replicates, indicated by dots). *P* value (*t*-test, two-sided) is also shown. The error bars represent the s.e.m. Source numerical data and unprocessed blots are available in source data.

active genes, indicating they are more likely to be active enhancers (Fig. 4h). These data indicate that active enhancers exist in FGOs as validated by both STARR-seq and the enhancer reporter assay and Pol II and CAGE association provide additional prediction power for active enhancers than H3K27ac alone.

Enhancer maps identify TCF3/12 as key folliculogenesis regulators

Enhancers are bound by cell-type-specific TFs[7]. To further validate these enhancer maps, we performed a motif analysis using HOMER[57]

in distal H3K27ac peaks in oocytes and early embryos to search for potentially interacting TFs (Fig. 5a). Consistent with the previous studies[27,28], key factors such as DUX[58–60], OBOX[61,62] and NR5A2 (refs. 63–65) were enriched at the pre-implantation stages. SOX2, GATA2, TCF4 and EOMES were enriched at the post-implantation stages. Of note, no strong TF motif enrichment was present at the one-cell stage and ICM. This was possibly due to the dilution of enrichment by different TF motifs as these stages undergo rapid transitions. Indeed, an analysis with finer gene classification identified similar motifs at the one-cell stage as those in oocytes and two-cell embryos (Extended Data Fig. 6a).
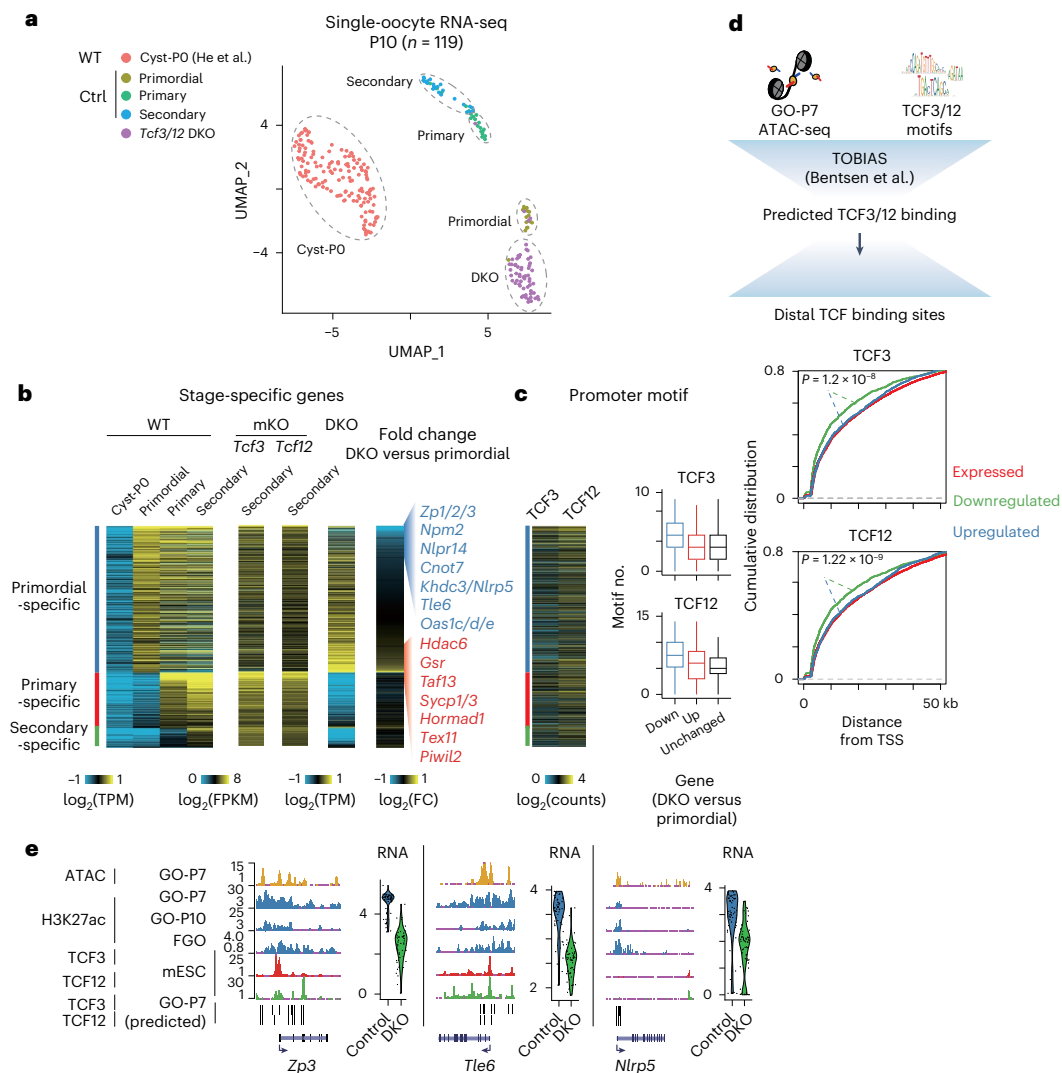
82

Fig. 6 | TCF3 and TCF12 are key folliculogenesis regulators. a, UMAP plot showing the different representations of oocyte clusters from ref. 72, control and *Tcf3/Tcf12* DKO GO-P10 ovary by single-oocyte RNA-seq. **b**, Heatmaps showing the stage-specific gene expression at different stages of WT, *Tcf3* mKO, *Tcf12* mKO and *Tcf3/Tcf12* DKO oocytes. Fold changes of DKO oocytes versus WT oocytes from primordial follicles as well as upregulated and downregulated example genes are also shown. **c**, Heatmap showing TCF3 and TCF12 promoter motif densities (left). Box plots showing TCF3 and TCF12 promoter motif densities for downregulated ($n = 50$), upregulated ($n = 56$) and unchanged genes ($n = 923$)

(right). The median is indicated by the centre line. The bottom, top edges and whiskers represent the 10th and 90th percentiles and 1.5 × IQR, respectively. **d**, Top, schematic showing TCF3/TCF12 binding prediction by TOBIAS[73]. Bottom, line charts showing the cumulative distribution of the distance between TSSs of expressed, downregulated and upregulated genes and nearest distal predicted TCF3 (middle) or TCF12 (bottom) binding sites. *P* values, *t*-test, two-sided. **e**, The UCSC browser views and violin plots showing ATAC-seq, H3K27ac, mES cells TCF3/TCF12 signals and RNA expression (normalized) in control and DKO oocytes at representative downregulated genes.

Notably, motifs of TCF12 (HEB), TCF3 (E2A) and TFAP4 seemed to be specific for oocytes (Fig. 5a). These motifs were shared by putative enhancers in both GOs and FGOs, despite their distinct H3K27ac landscapes, suggesting that the same TFs may actively redistribute in the genome upon oocyte growth. TCF3/TCF12 are basic helix-loop-helix (bHLH) TFs known to play compensatory roles in T cell lineage differentiation and B cell development by forming heterodimers[66–68]. TCF12 also participates in germ layer development in concert with the Polycomb repressive complex 2 (PRC2)[69]. *Tcf3 and Tcf12* were expressed in oocytes and early embryos but their expression culminated in GO-P7 (Fig. 5b). FGO enhancers previously identified by STARR-seq also enriched for the TCF3 and TCF12 motifs (Extended Data Fig. 10a,b). Of note, TCF3 could

interact with FIGLA in vitro, a germ cell-specific TF required for ovarian follicle formation and activate zona pellucida genes (*Zp1/Zp2/Zp3*) in a reporter assay in fibroblast cells[70,71]. However, whether TCF3/TCF12 regulate oocyte development in vivo remains unknown. Therefore, we generated conditional knockout (KO) mice deficient for either *Tcf3* or *Tcf12* in oocytes (driven by *Gdf9-Cre*) (Methods) and confirmed the depletion of *Tcf3/Tcf12* in oocytes (Extended Data Fig. 10c and Methods). However, folliculogenesis and ovulation seemed largely normal for *Tcf3* maternal knockout (mKO) and *Tcf12* mKO oocytes (Fig. 5c–e). Due to the functional compensation and the similar binding motifs of TCF3 and TCF12 (ref. 67), we obtained *Tcf3/Tcf12* double KO (DKO) oocytes (Methods) and confirmed the depletion of
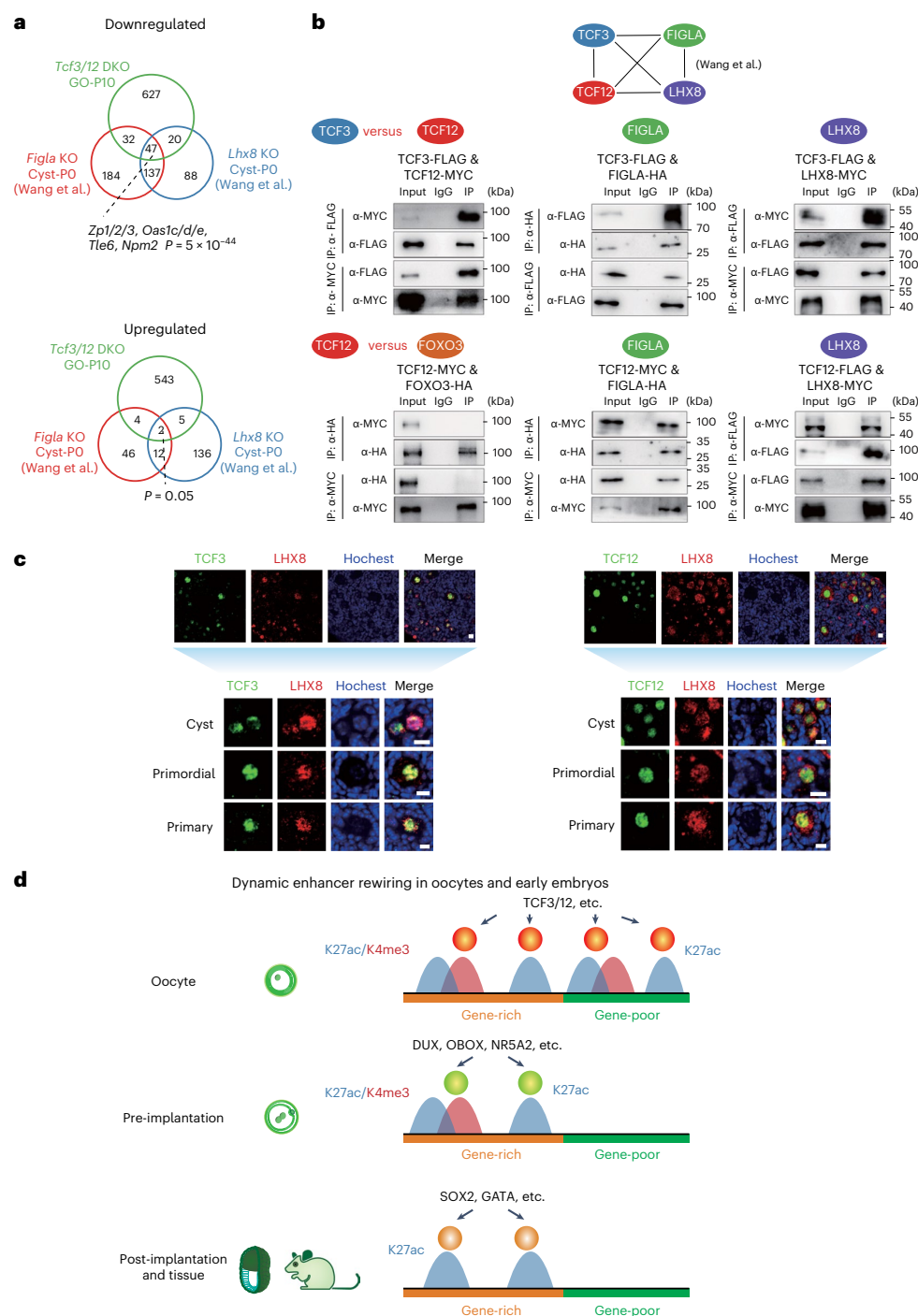
Fig. 7 | Dynamic enhancer rewiring in mouse oocytes and early embryos.
a, Venn diagrams showing the overlap of downregulated (top) and upregulated (bottom) genes among *Tcf3/Tcf12*, *Figla* and *Lhx8* knockout oocytes. *P* values (Fisher's exact test, two-sided) for overlapped genes are also shown. **b**, Schematic of IP results showing pairwise interactions between TCF3, TCF12, FIGLA and LHX8 (top). The interaction between FIGLA and LHX8 is from Wang et al.[78]. IPs for various TF combinations in HEK293T cells: TCF3 with TCF12, FIGLA, LHX8 and TCF12 with FOXO3 (negative control), FIGLA and LHX8 (bottom). **c**, Immunostaining showing TCF3, TCF12 and LHX8 in oocytes from the cyst, primordial and primary follicles of P2.5 ovary (*n* = 3 biological replicates).

Scale bar, 10 μm. **d**, A model illustrating the enhancer dynamics in oocytes and early embryos. In oocytes, putative enhancers reside in both gene-rich and gene-poor regions, potentially interacting with key TFs such as TCF3 and TCF12. A large portion of putative active enhancers bear both H3K4me3 and H3K27ac. After fertilization, putative enhancers are located mainly in gene-rich regions, likely interacting with a new set of TFs, such as DUX, OBOX and NR5A2, in pre-implantation embryos. In post-implantation embryos and adult tissues, putative enhancers enrich for motifs for lineage-specific TFs such as SOX2 and GATA and no longer carry H3K4me3 likely due to genome-wide DNA hypermethylation. Source numerical data and unprocessed blots are available in source data.

84

TCF3/TCF12 (Extended Data Fig. 10d–f). Notably, these mice were infertile (Fig. 5f). A detailed analysis showed ablation of *Tcf3/Tcf12* led to a drastic loss of primordial follicles, with some progressing to a primary follicle-like state but not to the secondary follicles (Fig. 5c,d). To investigate the effect of TCF3/TCF12 deficiency on transcription, we manually performed single-oocyte RNA-seq for *Tcf3/Tcf12* DKO oocytes at the GO-P10 stages (Fig. 6a). Uniform Manifold Approximation and Projection (UMAP) analysis showed that *Tcf3/Tcf12* DKO oocytes were clustered closer to primordial follicles but were distinct from primary or secondary follicles in WT oocytes (Fig. 6a). To further pinpoint the states of these mutant oocytes, we identified stage-specific genes in oocytes from primordial, primary and secondary follicles. We also added RNA-seq data for P0 oocytes in germ cell cysts isolated from newborn ovaries[72] (Fig. 6a,b). Genes specifically activated in oocytes from primary and secondary follicles were expressed normally in *Tcf3* or *Tcf12* mKO oocytes, but completely failed to be expressed in *Tcf3/Tcf12* DKO oocytes (Fig. 6b). Therefore, despite some DKO oocytes morphologically resembling primary follicles (Fig. 5c,d), their transcriptomes were still arrested at the primordial follicle stage (Fig. 6b). The majority of primordial oocyte-specific genes were activated in DKO mutants. However, 20.1% (121 of 601) of them still failed to be expressed, including *Zp1/Zp2/Zp3*, *Oas1c/Oas1d/Oas1e* and *Cnot7* (Fig. 6b, right). These data support compensatory functions between TCF3 and TCF12 and suggest that they, together, play key roles in oocyte development in primordial follicles, although we cannot fully exclude the possibility that they may also have different functions.

We then asked whether these transcription defects were related to TCF3/TCF12 binding. Among differentially expressed genes, the downregulated genes contained more TCF3/TCF12 motifs at their promoters (Fig. 6c). As attempts to profile the binding of TCF3/TCF12 in oocytes failed, we inferred TCF3/TCF12 binding sites based on ATAC-seq in GO-P7 (the earliest stage at which we could collect sufficient oocytes) and TCF3/TCF12 motifs using TOBIAS[73] (Fig. 6d, top and Methods). Encouragingly, the predicted TCF3/12 distal binding was also present near downregulated genes (Fig. 6d, bottom), as exemplified by that near *Zp3*, *Tle6* and *Nlrp5* (Fig. 6e). *Zp3* and *Tle6* were also bound by TCF3/TCF12 in mES cells[69,74] (Fig. 6e). These downregulated genes in *Tcf3/Tcf12* mutant oocytes significantly overlapped with the downregulated genes (*Zp1/Zp2/Zp3*, *Oas1c/Oas1d/Oas1e*, *Tle6* and *Npm2*) in mouse oocytes deficient for *Figla* or *Lhx8*, two TFs required for primordial follicle formation or maintenance[75–78] (Fig. 7a). Furthermore, TCF3 and TCF12 interacted with each other and both also interacted with FIGLA and LHX8 (Fig. 7b). As a negative control, TCF12 did not interact with FOXO3, another oocyte TF regulator[79]. Thus, these data are in line with the notion that TCF3/TCF12 and FIGLA, LHX8 share common targets in oocytes. Immunostaining in the P2.5 ovary for TCF3, TCF12 and LHX8 (we did not find good antibodies for FIGLA) revealed that all three TFs exhibited nuclear localization and both TCF3 and TCF12 showed partial co-localization with LHX8 (Fig. 7c), supporting a model that they may function cooperatively. Taken together, these analyses revealed a drastic transition of the transcription factor network during the OET and identified TCF3/TCF12 as key regulators of mouse oogenesis.

Discussion

Transcription during mammalian gametogenesis and early development undergoes extensive reprogramming. However, the CREs underlying the transcription network in this process remain poorly understood. Here, we mapped putative enhancers across 15 stages during mouse oogenesis and embryogenesis. Our data demonstrate that the enhancer networks undergo drastic transition, bear unique epigenetic signatures and likely interact with stage-specific TFs to wire transcription networks in oocytes and early embryos (Fig. 7d).

Putative enhancers in oocytes and early embryos are globally distinct from those in tissues and cell lines (Fig. 1b). Many putative enhancers appear in gene-poor regions in oocytes, where a number of oocyte-specific genes reside. The activities of these enhancers gradually diminish after fertilization. Notably, this correlated with loss of LADs in mouse FGOs and their subsequent re-establishment after fertilization[22]. It is tempting to speculate that loss of LADs in FGOs may enable an active transcription environment in gene deserts to permit the expression of a subset of oocyte-specific genes that are otherwise repressed. In mouse oocytes and pre-implantation embryos, many putative enhancers are also marked by H3K4me3, a classic promoter mark, likely due to DNA hypomethylation in the genomes (Fig. 4a), as observed in zebrafish *dnmt1*-deficient embryos[37]. These data are consistent with the H3K4me3-H3K4me1 seesaw model previously proposed[80]. In addition, H3K4me3 at enhancers is actively converted to H3K4me1 by the demethylases KDM5B/KDM5C in mES cells to avoid overactivation[52,81]. *Kdm5a/Kdm5b/Kdm5c* are lowly expressed in FGOs[82], which may further contribute to the widespread presence of H3K4me3 on enhancers. Future works are warranted to decipher the function of H3K4me3 at enhancers in oocytes and early embryos.

Earlier studies reported a lack of enhancer activities in oocytes[23,24]. Here, our study presented evidence of enhancer activities in FGOs using STARR-seq[53] and an enhancer reporter assay (Fig. 4d–g and Extended Data Fig. 9a–e). We noticed that at least for one enhancer tested previously[23], the corresponding TF (SP1) is poorly expressed in oocytes (Extended Data Fig. 10g). Moreover, our results indicate that the association of Pol II and CAGE signals may provide additional prediction power for active enhancers in oocytes than H3K27ac alone. Finally, in strong support of the notion that enhancers are actively employed in oocytes and early embryos, we identified key TF regulators of ovarian folliculogenesis, TCF3 and TCF12, inferred from the enhancer maps. Therefore, uncovering these regulatory networks represents an important step towards decoding the genetic circuitry underlying the OET, which resets the life clock to generate a totipotent embryo.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-024-01422-x.

References

1. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170–1187 (2016).
2. Yu, M. & Ren, B. The three-dimensional organization of mammalian genomes. *Annu. Rev. Cell Dev. Biol.* **33**, 265–289 (2017).
3. Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
4. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Mol. Cell* **49**, 825–837 (2013).
5. Ong, C.-T. & Corces, V. G. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* **12**, 283–293 (2011).
6. Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272–286 (2014).
7. Creyghton, M. P. et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107**, 21931–21936 (2010).
8. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349–352 (1997).
9. Guerrini, M. M., Oguchi, A., Suzuki, A. & Murakawa, Y. Cap analysis of gene expression (CAGE) and noncoding regulatory elements. *Semin. Immunopathol.* **44**, 127–136 (2022).

10. Andersson, R. & Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* **21**, 71–87 (2020).

11. Du, Z., Zhang, K. & Xie, W. Epigenetic reprogramming in early animal development. *Cold Spring Harb. Perspect. Biol.* **14**, a039677 (2022).

12. Eckersley-Maslin, M. A., Alda-Catalinas, C. & Reik, W. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19**, 436–450 (2018).

13. Fu, X., Zhang, C. & Zhang, Y. Epigenetic regulation of mouse preimplantation embryo development. *Curr. Opin. Genet Dev.* **64**, 13–20 (2020).

14. Dahl, J. A. et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548–552 (2016).

15. Liu, X. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558–562 (2016).

16. Zhang, B. et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553–557 (2016).

17. Zheng, H. et al. Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol. Cell* **63**, 1066–1079 (2016).

18. Inoue, A., Jiang, L., Lu, F. & Zhang, Y. Genomic imprinting of Xist by maternal H3K27me3. *Genes Dev.* **31**, 1927–1932 (2017).

19. Inoue, A., Jiang, L., Lu, F., Suzuki, T. & Zhang, Y. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**, 419–424 (2017).

20. Harris, C. et al. Conversion of random X-inactivation to imprinted X-inactivation by maternal PRC2. *eLife* **8**, e44258 (2019).

21. van Steensel, B. & Belmont, A. S. Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell* **169**, 780–791 (2017).

22. Borsos, M. et al. Genome-lamina interactions are established de novo in the early mouse embryo. *Nature* **569**, 729–733 (2019).

23. Lawinger, P., Rastelli, L., Zhao, Z. & Majumder, S. Lack of enhancer function in mammals is unique to oocytes and fertilized eggs. *J. Biol. Chem.* **274**, 8002–8011 (1999).

24. Majumder, S., Zhao, Z., Kaneko, K. & DePamphilis, M. L. Developmental acquisition of enhancer function requires a unique coactivator activity. *EMBO J.* **16**, 1721–1731 (1997).

25. Kim, J. M., Liu, H., Tazaki, M., Nagata, M. & Aoki, F. Changes in histone acetylation during mouse oocyte meiosis. *J. Cell Biol.* **162**, 37–46 (2003).

26. Akiyama, T., Nagata, M. & Aoki, F. Inadequate histone deacetylation during oocyte meiosis causes aneuploidy and embryo death in mice. *Proc. Natl Acad. Sci. USA* **103**, 7339–7344 (2006).

27. Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).

28. Xiang, Y. et al. Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. *Nat. Genet.* **52**, 95–105 (2020).

29. Wang, M., Chen, Z. & Zhang, Y. CBP/p300 and HDAC activities regulate H3K27 acetylation dynamics and zygotic genome activation in mouse preimplantation embryos. *EMBO J.* **41**, e112012 (2022).

30. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

31. Veselovska, L. et al. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol.* **16**, 209 (2015).

32. Paillisson, A. et al. Identification, characterization and metagenome analysis of oocyte-specific genes organized in clusters in the mouse genome. *BMC Genomics* **6**, 76 (2005).

33. Santenard, A. et al. Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat. Cell Biol.* **12**, 853–862 (2010).

34. Liu, B. et al. The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature* **587**, 139–144 (2020).

35. Zenk, F. et al. Germ line-inherited H3K27me3 restricts enhancer function during maternal-to-zygotic transition. *Science* **357**, 212–216 (2017).

36. Du, Z. et al. Polycomb group proteins regulate chromatin architecture in mouse oocytes and early embryos. *Mol. Cell* **77**, 825–839 (2020).

37. Wu, X. et al. Methylome inheritance and enhancer dememorization reset an epigenetic gate safeguarding embryonic programs. *Sci. Adv.* **7**, eabl3858 (2021).

38. Wang, L. et al. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991 (2014).

39. Shirane, K. et al. Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. *PLoS Genet.* **9**, e1003439 (2013).

40. Hanna, C. W. et al. MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nat. Struct. Mol. Biol.* **25**, 73–82 (2018).

41. Smallwood, S. A. et al. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**, 811–814 (2011).

42. Pekowska, A. et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J.* **30**, 4198–4210 (2011).

43. The ENCODE Project Consortium et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).

44. Lam, M. T., Li, W., Rosenfeld, M. G. & Glass, C. K. Enhancer RNAs and regulated transcriptional programs. *Trends Biochem. Sci.* **39**, 170–182 (2014).

45. Zhu, Y. et al. Predicting enhancer transcription and activity from chromatin modifications. *Nucleic Acids Res.* **41**, 10032–10043 (2013).

46. Yu, C. et al. TBPL2/TFIIA complex establishes the maternal transcriptome through oocyte-specific promoter usage. *Nat. Commun.* **11**, 6439 (2020).

47. Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P. & Lenhard, B. CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic Acids Res.* **43**, e51 (2015).

48. Thodberg, M., Thieffry, A., Vitting-Seerup, K., Andersson, R. & Sandelin, A. CAGEfightR: analysis of 5′-end data using R/Bioconductor. *BMC Bioinform.* **20**, 487 (2019).

49. Shen, Y. et al. A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).

50. Shahbazian, M. D. & Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* **76**, 75–100 (2007).

51. Dahlin, J. L. et al. Assay interference and off-target liabilities of reported histone acetyltransferase inhibitors. *Nat. Commun.* **8**, 1527 (2017).

52. Shen, H. et al. Suppression of enhancer overactivation by a RACK7-Histone demethylase complex. *Cell* **165**, 331–342 (2016).

53. Neumayr, C., Pagani, M., Stark, A. & Arnold, C. D. STARR-seq and UMI-STARR-seq: assessing enhancer activities for genome-wide-, high-, and low-complexity candidate libraries. *Curr. Protoc. Mol. Biol.* **128**, e105 (2019).

54. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).

55. Lee, D. et al. STARRPeaker: uniform processing and accurate identification of STARR-seq active regions. *Genome Biol.* **21**, 298 (2020).

56. Klein, J. C. et al. A systematic evaluation of the design and context dependencies of massively parallel reporter assays. *Nat. Methods* **17**, 1083–1091 (2020).

86

57. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

58. Whiddon, J. L., Langford, A. T., Wong, C.-J., Zhong, J. W. & Tapscott, S. J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).

59. Hendrickson, P. G. et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).

60. De Iaco, A. et al. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49**, 941–945 (2017).

61. Ji, S. et al. OBOX regulates mouse zygotic genome activation and early development. *Nature* **620**, 1047–1053 (2023).

62. Youjia, G. et al. Obox4 promotes zygotic genome activation upon loss of Dux. Preprint at *bioRxiv* https://doi.org/10.1101/2022.07.04.498763 (2024).

63. Lai, F. et al. NR5A2 connects zygotic genome activation to the first lineage segregation in totipotent embryos. *Cell Res.* **33**, 952–966 (2023).

64. Festuccia, N. et al. *Nr5a2* is essential for morula development. Preprint at *bioRxiv* https://doi.org/10.1101/2023.01.16.524255 (2023).

65. Gassler, J. et al. Zygotic genome activation by the totipotency pioneer factor Nr5a2. *Science* **378**, 1305–1315 (2022).

66. Jones, M. E. & Zhuang, Y. Acquisition of a functional T cell receptor during T lymphocyte development is enforced by HEB and E2A transcription factors. *Immunity* **27**, 860–870 (2007).

67. Barndt, R. J., Dai, M. & Zhuang, Y. Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB. *Mol. Cell. Biol.* **20**, 6677–6685 (2000).

68. Zhuang, Y., Barndt, R. J., Pan, L., Kelley, R. & Dai, M. Functional replacement of the mouse E2A gene with a human HEB cDNA. *Mol. Cell. Biol.* **18**, 3340–3349 (1998).

69. Yoon, S. J., Foley, J. W. & Baker, J. C. HEB associates with PRC2 and SMAD2/3 to regulate developmental fates. *Nat. Commun.* **6**, 6546 (2015).

70. Liang, L. F., Soyal, S. M. & Dean, J. FIG alpha, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes. *Development* **124**, 4939–4947 (1997).

71. Lim, E. J. & Choi, Y. Transcription factors in the maintenance and survival of primordial follicles. *Clin. Exp. Reprod. Med.* **39**, 127–131 (2012).

72. He, Y. et al. Single-cell RNA-Seq reveals a highly coordinated transcriptional program in mouse germ cells during primordial follicle formation. *Aging Cell* **20**, e13424 (2021).

73. Bentsen, M. et al. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun.* **11**, 4267 (2020).

74. Wang, Q. et al. The p53 family coordinates Wnt and nodal inputs in mesendodermal differentiation of embryonic stem cells. *Cell Stem Cell* **20**, 70–86 (2017).

75. Choi, Y., Ballow, D. J., Xin, Y. & Rajkovic, A. Lim homeobox gene, Lhx8, is essential for mouse oocyte differentiation and survival1. *Biol. Reprod.* **79**, 442–449 (2008).

76. D'Ignazio, L. et al. Lhx8 ablation leads to massive autophagy of mouse oocytes associated with DNA damage. *Biol. Reprod.* **98**, 532–542 (2018).

77. Soyal, S. M., Amleh, A. & Dean, J. FIGα, a germ cell-specific transcription factor required for ovarian follicle formation. *Development* **127**, 4645–4654 (2000).

78. Wang, Z., Liu, C. Y., Zhao, Y. & Dean, J. FIGLA, LHX8 and SOHLH1 transcription factor networks regulate mouse oocyte growth and differentiation. *Nucleic Acids Res.* **48**, 3525–3541 (2020).

79. John, G. B., Gallardo, T. D., Shirley, L. J. & Castrillon, D. H. Foxo3 is a PI3K-dependent molecular switch controlling the initiation of oocyte growth. *Dev. Biol.* **321**, 197–204 (2008).

80. Sharifi-Zarchi, A. et al. DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. *BMC Genomics* **18**, 964 (2017).

81. Kidder, B. L., Hu, G. & Zhao, K. KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biol.* **15**, R32 (2014).

82. Shao, G. B. et al. Dynamic patterns of histone H3 lysine 4 methyltransferases and demethylases during mouse preimplantation development. *Vitr. Cell Dev. Biol. Anim.* **50**, 603–613 (2014).

[1]Center for Stem Cell Biology and Regenerative Medicine, MOE Key Laboratory of Bioinformatics, New Cornerstone Science Laboratory, School of Life Sciences, Tsinghua University, Beijing, China. [2]Tsinghua-Peking Center for Life Sciences, Beijing, China. [3]State Key Laboratory of Reproductive Medicine and Offspring Health, Nanjing Medical University, Nanjing, China. [4]Innovation Center of Suzhou Nanjing Medical University, Suzhou, China. [5]Laboratory of Molecular Developmental Biology, State Key Laboratory of Membrane Biology, Tsinghua University, Beijing, China. [6]Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München, Munich, Germany. [7]Center for Stem Cell Biology and Regenerative Medicine, School of Medicine, Tsinghua University, Beijing, China. [8]These authors contributed equally: Bofeng Liu, Yuanlin He, Xiaotong Wu, Zili Lin, Jing Ma. ✉e-mail: ljwth@njmu.edu.cn; xiewei121@tsinghua.edu.cn

87

Methods

Animal maintenance

All animal maintenance and experimental procedures were carried out according to the Institutional Animal Care and Use Committee guidelines of Tsinghua University, Beijing, China or under the authorization of the Upper Bavarian Government (ethical approval protocol no. 21-XW2 and IACUC-1601220). All oocytes were collected from WT C57BL/6N females. PWK/PhJ mice were originally purchased from The Jackson Laboratory and raised in the local core facility. C57BL/6 and ICR mice were purchased from Vital River.

Preparation of mouse oocytes and embryos

In brief, GOs were isolated by mechanical dissection in M2 medium (Sigma, M7167). GOs were isolated from mice at postnatal day 7 or day 10. Pre-implantation embryos were collected from 5–6-week-old C57BL/6N females (Vital River) mated with PWK/PhJ males (The Jackson Laboratory). Adult female mice were superovulated for oocyte and embryo collection. For superovulation, female mice were injected intraperitoneally with 7.5 IU pregnant mare serum gonadotropin (PMSG) and human chorionic gonadotropin (hCG) with 5 IU 46–48 h after PMSG injection. FGOs (>70 μm) were isolated from 8-week-old mice 46–48 h after PMSG injection. MII oocytes were isolated from 6-week-old mice 20 h after hCG injection. Each set of embryos was isolated at a defined period after hCG injection, 27–29 h (PN5 zygotes), 33–35 h (early two-cell), 46–48 h (late two-cell), 62–65 h (eight-cell), 94–96 h (blastocysts) in M2 medium.

At least 200 cells were collected for STAR ChIP-seq experiments. To remove the granulosa cells, the dissociated oocyte and granulosa cell complexes were transferred into M2 medium containing 0.1% hyaluronidase (Sigma, A5177) and digested for 5 min. Oocytes were collected after removing the zona pellucida by acidic Tyrode's solution (Sigma, T1788) treatment. To remove the zona pellucida of embryos, the embryos were incubated with 10 IU ml^{-1} pronase (Sigma, P8811) for several minutes. Samples were washed in PBS quickly and manually picked up into the lysis buffer (0.5% NP-40, 0.5% Tween-20, 0.1% SDS and proteinase inhibitor) for STAR ChIP-seq.

Cell culture of ES cells

The R1 ES cells were cultured on gelatin in DMEM containing 15% FBS, leukaemia inhibiting factor, penicillin/streptomycin, L-glutamine, β-mercaptoethanol and non-essential amino acids.

STAR ChIP-seq library preparation and sequencing

STAR ChIP-seq library preparation was conducted following a protocol described previously[16]. In brief, each sample was subjected to MNase (Sigma, N3755-200UN) digestion at 37 °C. The reaction was terminated by adding stop buffer (110 mM Tris-HCl, pH 8.0 and 55 mM EDTA) and cold 2× RIPA buffer. Each chromatin sample was supplemented with RIPA buffer to make sure the lipid in the tube could flow while rotating it. The IP sample was incubated with antibodies for H3K27ac (Active Motif, 39133, 1:70 diluted) overnight with rotation at 4 °C. The next day, the sample was incubated with protein A Dynabeads (Life Technologies) for 2 h with rotation at 4 °C. Beads were washed with RIPA buffer four times and LiCl buffer once. After washing, tubes were spun briefly and the supernatant was removed. For each IP sample, beads were resuspended with ddH$_2$O and Ex-Taq buffer (TaKaRa). Then, 1 μl proteinase K (Roche, 10910000) was added at 55 °C for 90 min to elute DNA from beads. The supernatant was transferred to a new tube and the proteinase K was inactivated at 72 °C for 40 min. Then, 1 μl rSAP (NEB, M0371) was added to dephosphorylate the 3' end of DNA at 37 °C for 1 h. rSAP was inactivated at 65 °C for 10 min. The resulting sample was subjected to library preparation starting from PolyC tailing as previously described[83]. Mouse sperm ChIP-seq was performed as described previously with modifications[84].

Whole-mount immunofluorescent staining

Oocytes and embryos were fixed in 4% paraformaldehyde at room temperature for 30 min. Subsequently, they were permeabilized for 10 min in PBST (1% Triton X-100). After blocking with 1% BSA for 1 h, samples were incubated with H3K27ac antibody (Active Motif, 39133, 1:200 diluted) at 4 °C overnight. The secondary antibody Alexa Fluor 488-conjugated anti-rabbit (Jackson ImmunoResearch, 611-545-215, 1:200 diluted) was added with 4,6-diamidino-2-phenylindole (DAPI) (Invitrogen, D1306) after washing the primary antibody. Images were acquired on an 880 META laser scanning confocal microscope and manipulated by ZEN software (v.3.9).

DamID library preparation and sequencing

The DamID procedure was performed as previously described[22]. In brief, a messenger RNA mixture containing 100 ng μl^{-1} membrane-eGFP and AID–Dam-lamin B1 or AID–Dam were injected into the cytoplasm of oocytes. Oocytes were isolated and injected with 5 ng μl^{-1} AID–Dam-lamin B1 or 20 ng μl^{-1} AID–Dam and kept in auxin-free M2 medium for 6–8 h to methylate LADs or accessible regions, respectively. Following oocyte collection, library preparation and sequencing were performed as previously described[22].

Inhibitor treatment

To block H3K27ac, FGOs were collected and cultured in M2 medium containing 60 μM plumbagin (Selleck, S4777) for 24 h with milrinone. FGOs cultured in M2 medium with 0.1% DMSO were used as control.

EU staining

Cell-LightTM EU Nascent RNA Detection kit (RiboBIO, C10316) was used to explore the transcription level of Plumbagin or DMSO-treated FGOs. These FGOs were transferred into 100 μl 500 mMEU solution (1:1,000 diluted in medium), soaked at 37 °C for 2 h, washed several times with PBS (GIBCO, 21600-044) and fixed by 4% polyformaldehyde for 30 min at room temperature. FGOs were permeabilized with 1% PBST at room temperature for 10 min and washed with PBS several times before staining in 100 μl Apollo reaction buffer (Apollo 567 nm). Freshly prepared Apollo reaction buffer was maintained in the dark for 30 min at room temperature with gentle shaking, after which DNA-stained embryos with DAPI were mounted and imaged using an 880 META laser scanning confocal microscope.

STARR-seq plasmid library cloning

The candidate enhancer sequences were obtained from PCR of genome DNA (primer sequences in Supplementary Table 3) and cloned between the truncated form of GFP and the polyA site in the hSTARR-seq_ORI vector (Addgene, #99296). Purified PCR products of each candidate enhancer were pooled together and ligated with Illumina adaptors in DNA Library Prep kit (NEB, E7645S). Subsequently, adaptor-ligated PCR products were purified with AMPure XP beads and amplified by PCR with library cloning primers (forward: 5′-TAGAGCATGCACCGGACACTCTTTCCCTA-CACGACGCTCTTCCGATCT, reverse: 5′-GGCCGAATTCGTCGAGT-GACTGGAGTTCAGACGTGTGCTCTTCCGATCT) (1 μl adaptor-ligated PCR products, 2.5 μl 10 μM forward primer, 2.5 μl 10 μM reverse primer, 25 μl KAPA 2× HiFi HotStart Ready Mix (Roche, KK2602) and 19 μl H$_2$O) with the programme of 98 °C for 45 s (98 °C for 15 s, 65 °C for 30 s and 72 °C for 45 s) with five cycles and 72 °C for 2 min, which results in that candidate sequences were flanked by overlap sequences with vector around insertion site. The hSTARR-seq_ORI vector was restriction digested by *Age*I-HF and *Sal*I-HF and purified for Gibson assembly with ClonExpress II One Step Cloning kit (Vazyme, C112). The STARR-seq plasmid library was amplified using Illumina i5 and i7 index primers and its quality and complexity were assessed by deep sequencing.

STARR-seq library preparation and sequencing

The STARR-seq plasmid library was purified and injected into the pronucleus of FGOs cultured with milrinone. After 24 h, the FGOs were lysed in hypotonic lysis buffer (Vazyme, N712) and the polyadenylated mRNAs were captured by the oligonucleotide (dT) primers. After 3 min at 72 °C, the Smart-seq2 reverse transcription reactions were performed. After pre-amplification and AMPure XP beads purification, cDNAs were amplified with junction PCR primers (forward: 5′-TCGTGAGGCACTGGG CAG*G*T*G*T*C, reverse: 5′-CTTATCATGTCTGCTCGA*A*G*C, where * indicates a phosphorothioate bond), which specifically enriches reporter transcripts and excludes STARR-seq plasmids. Then the PCR products were cleaned up with AMPure XP beads and we performed sequencing-ready PCR to add Illumina i5 and i7 indexes for deep sequencing. All libraries were sequenced on an Illumina HiSeq 2500 platform.

Oocyte reporter assay

The hSTARR-seq_ORI vector (Addgene, #99296) was modified for reporter assay constructs: truncated *Gfp* was replaced with a *mNeon-Green* coding sequence. Candidate enhancer sequences were inserted after polyA. The primer sequences used for the amplification of candidate enhancers are listed in Supplementary Table 2. *Zp3* promoter was inserted into the pGL4.23 luciferase reporter vector (Promega, E8411) and the luciferase gene was replaced with an *mCherry* coding sequence, which acts as a control for enhancer reporter. Purified enhancer reporter plasmid for each candidate enhancer and *Zp3-mCherry* plasmid were injected into the pronucleus of FGOs cultured with milrinone. After 24 h, the ratio of mNeonGreen to mCherry fluorescence intensity was recorded as enhancer activity.

Tcf3/*Tcf12* and *Eed* cKO mice

Tcf3^{flox/flox} and *Tcf12*^{flox/flox} transgenic mice were gifted from Y. Zhuang at Duke University. *Gdf9-Cre* mice were gifted from Y. Su at Shandong University. The *Gdf9-Cre* mice were crossed with *Tcf3*^{flox/flox} and *Tcf12*^{flox/flox} mice to establish *Gdf9-Cre Tcf3*^{flox/flox} *Tcf12*^{flox/flox} cKO mice. All mice had a C57BL/6J genetic background. Primers used for genotyping are listed in Supplementary Table 1. *Eed* KO mice were previously described[36]. In brief, *Eed*^{flox/flox} FGOs and *Eed*^{−/−} FGOs were collected from 8-week-old *Eed*^{flox/flox} and *Eed*^{flox/flox} *Gdf9*-Cre mice, respectively.

H&E staining, IHC and IF staining

Ovaries were fixed in 10% formalin overnight, paraffin-embedded and sectioned to a thickness of 5 μm. Sections were deparaffinized and rehydrated and endogenous peroxidase activity was blocked by incubating in 3% hydrogen peroxide in methanol for 15 min.

The sections were stained with H&E. Follicles were classified as primordial follicles (an oocyte surrounded by a partial or complete layer of squamous granulosa cells), primary follicles (a single layer of cuboidal granulosa cells), secondary follicles (more than one layer of cuboidal granulosa cells with nonvisible antrum) and antral follicles (a clearly defined antral space and a cumulus granulosa cell layer).

For immunohistochemistry (IHC), the antigen of ovarian sections was retrieved at 95 °C for 15 min in 10 mM sodium citrate buffer (pH 6.0). Sections were blocked for 60 min at room temperature in PBS contained with 5% BSA and incubated at 4 °C overnight with the primary antibodies: TCF3 (Proteintech, 21242-1-AP, 1:200 dilution) and TCF12 (Proteintech, 14419-1-AP, 1:200 dilution). Subsequently, the sections were washed with TBS 5 min three times and incubated with secondary antibodies for 60 min. The signals were coloured with a diaminobenzidine (DAB) reagent (ZSGB-BIO).

Immunofluorescence (IF) staining was performed with a multi-immunofluorescent kit (Aifang Biological, AFIHC033). Samples were incubated with primary antibodies: TCF3 (Proteintech, 21242-1-AP, 1:2,000 diluted), TCF12 (Proteintech, 14419-1-AP, 1:2,000 diluted) and LHX8 (Abclonal, A2046, 1:2,000 diluted).

Quantitative real-time PCR

The method for preparing cDNA was adapted from the Smart-seq2 protocol[54]. qPCR was performed using the ChamQ SYBR Green qPCR Master Mix (Vazyme Q311) with an ABI StepOnePlus Real-Time PCR system (Applied Biosystems, Life Technologies). Relative mRNA levels were calculated by normalizing them to *β-Actin* mRNA levels. Primer sequences are listed in Supplementary Table 1.

Western blot

Ovaries were lysed in RIPA (Beyotime Institute of Biotechnology, P0013B) with 1% protease inhibitor (MCE, HY-K0012). Proteins were separated by electrophoresis by 10% SDS–PAGE and transferred into polyvinylidene fluoride membranes (Bio-Rad). The membranes were blocked with 5% nonfat dry milk for 60 min and incubated at 4 °C overnight with the following primary antibodies: TCF3 (Proteintech, 21242-1-AP, 1:800 diluted), TCF12 (Proteintech, 14419-1-AP, 1:800 diluted) and β-tubulin (Yifei Xue Biotechnology, YFMA0053, 1:1,000 diluted). The membranes were washed with TBST for 10 min three times and incubated with secondary antibodies for 60 min. The signals were enhanced through enhanced chemiluminescence (Biosharp, BL520A).

RNA-seq library preparation and sequencing for oocytes

The ovaries were digested in 500 μl HBSS supplemented with 0.25% trypsin, 1 mM EDTA and 0.01% DNase I and incubated at 37 °C for 10 min with gentle agitation. After aspirating the supernatant completely, the cells were resuspended in 500 μl HBSS. The dissociated single-cell suspensions were transferred under the microscope (Nikon, SMZ1000) and oocytes were washed three times with 0.1% BSA/PBS, each oocyte was transferred into 0.2-ml PCR tubes containing 2 μl lysis buffer. Transcriptome libraries were prepared following the Smart-seq2 protocol[54]. Sequencing libraries were constructed by using KAPA HyperPlus kit (Kapa Biosystems) according to the manufacturer's instructions. All libraries were sequenced on the Illumina NovaSeq 6,000 platform.

ATAC-seq library preparation and sequencing

The ATAC-seq procedure was performed as previously described[27]. In brief, cells were transferred into 6 μl lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂ and 0.5% NP-40) on ice for 10 min. The ATAC reaction was performed by adding 4 μl ddH₂O, 4 μl 5× TTBL and 5 μl TTE mix V5 (Vazyme, TD502) at 37 °C for 30 min and then stopped by adding 5 μl 5× TS stop buffer at room temperature for 5 min. The DNA product was PCR-amplified with 10 μl index (Vazyme, TD202), 10 μl 5× TAB and 1 μl TAE (Vazyme, TD502) for 16 cycles. The amplified DNA was size-selected using AMPure Beads for 200–800-bp DNA fragments. All libraries were sequenced by an Illumina 2500 or XTen platform, accordingly.

Data analyses

ChIP-seq data processing. The paired-end reads were aligned with the parameters: -t -q -N 1 -L 25 -X 2,000 –no-mixed–no-discordant by Bowtie (v.2.2.2)[85]. All unmapped reads, non-uniquely mapped reads and PCR duplicates were removed. For downstream analysis, we normalized the read counts by computing the numbers of reads per kilobase of bin per million of reads sequenced (RPKM). RPKM values were calculated by merged replicate bam files (SAMtools v.1.3.1)[86]. To minimize the batch and cell-type variation, the RPKM values were further normalized through *z*-score transformation (Python v.2.7.12). To visualize the ChIP-seq signal in the UCSC genome browser, we extended each read by 250 bp and counted the coverage for each base (bedGraphToBigWig v.4). The correlation between ChIP-seq replicates was calculated as follows: ChIP-seq correlation was calculated by deep-Tools[87]. Allele assignment of sequencing reads for mouse embryos was conducted as described previously[27]. The distal H3K27ac peaks (beyond ±2.5 kb from TSSs) were identified as putative enhancers (Supplementary Table 4).

Gene expression data processing. All RNA-seq data were mapped to mm9 genome by Tophat (v.2.4.0)[88]. The gene expression level was calculated by Cufflinks (v.2.2.1)[88] based on the annotation mm9 refFlat database from the UCSC genome browser. The expression matrix (FPKM) for control and Tcf3/Tcf12 DKO single-oocyte RNA-seq was produced in a manner similarly to that of bulk RNA-seq data. Subsequently, the data were imported into Seurat (v.4.2.1)[89] to perform UMAP clustering and compute the average expression.

DNase-seq, ATAC-seq, DNA methylation, DamID and STARR-seq data processing. DNase-seq and ATAC-seq were mapped to mm9 genome by Bowtie with similar parameters as ChIP-seq data. For downstream analysis, we calculated the read counts by computing RPKM on the genome 100-bp bin. DNA methylation data were mapped to mm9 genome by BSMAP (v.2.74)[90] with parameters: -r 0 -p 16 -w 100 -v 0.1. PCR duplicates were removed. For each CG site, the methylation level was calculated as the total methylated counts (combining Watson and Crick strands) divided by the total counts across all reads covering this CG.

DamID was processed as previously described[22]. In brief, DamID was mapped to the mm10 genome and the computation of OE values per bin was carried out as previously described[91]. Reads that precisely flanked an annotated GATC site were associated with GATC fragments and kept for downstream analysis.

STARR-seq was processed as previously described[53]. In brief, STARR-seq was mapped to the mm9 genome by Bowtie with similar parameters as ChIP-seq data. For downstream analysis, we calculated the read counts by computing RPKM on the genome 100-bp bin for input controls and RNA. STARR-seq signals were then calculated as RNA/input. The background level of STARR-seq was estimated using all negative controls ($n = 64$; 16 negative controls × two replicates × both strands). After excluding outliers ($n = 4$), the background cutoff was determined as the values that correspond to mean + 3 × s.d. (fold change = 0.29, 99.7% confidence level). Enhancer candidates with STARR-seq RNA/input higher than the background cutoff in both replicates from at least one strand of enhancers were identified as STARR-seq-positive enhancers. The rest were classified as STARR-seq-negative enhancers.

Analyses of ChIP-seq peaks and peak comparison. H3K27ac peaks were called using HOMER[57] with the parameters -region -size 1,000 -minDist 1,500 -gsize 2.9×10^9 -fdr 0.0005. The peaks with tag count numbers more than 40 were selected as strong peaks for downstream analysis. Peak comparison was conducted using BEDTools (v.2.26.0)[92] intersectBed function.

Identification of stage-specific genes. A Shannon-entropy-based method was used to identify stage-specific genes, as previously described[93]. Genes with entropy score less than 2 were selected as candidates for stage-specific genes. Among these genes, we selected candidates of stage-specific genes for each stage based on the following criteria: the gene is highly expressed at this stage (FPKM > 5) and such high expression cannot be observed in more than two additional stages. These genes were then reported in the final stage-specific gene lists and visualized using Java TreeView (v.1.1.6r4)[94].

The comparison between H3K27ac ChIP-seq peaks and repetitive elements. To identify the enrichment of repetitive elements in distal H3K27ac peaks, the ChIP-seq peaks were compared with the locations of annotated repeats (RepeatMasker) downloaded from the UCSC genome browser. As repeats of different classes vary greatly in numbers, a random set of peaks with identical lengths of ChIP-seq peaks were used for the same analysis as a control. The numbers of observed peaks that overlap with repeats were compared with the number of random peaks that overlap with repeats and a log ratio value ($\log_2$) was generated as the 'observed/expected' enrichment.

Motif analyses for distal H3K27ac ChIP-seq peaks. To find the sequence motif enriched in ChIP-seq peaks, findMotifsGenome.pl from the HOMER program was used[57].

Gene Ontology analysis. The DAVID web-tool (v.6.8)[95] was employed to identify the Gene Ontology terms using databases including molecular functions, biological functions and cellular components[96].

Hierarchical clustering analysis. Hierarchical clustering was performed in R (v.4.2.1) by hclust() function with ChIP-seq RPKM values via Pearson correlation coefficients.

Identification of oocyte enhancers by CAGE. The GO-P14 oocytes GAGE data were obtained from previous work[46] and mapped to mm9 genome using Bowtie (v.2.2.2)[85]. Uniquely mapped reads were kept for downstream analyses using CAGEr Bioconductor package[47]. The unidirectional and bidirectional transcription starting sites were identified using CAGEfightR[48]. The candidate enhancers identified by CAGE in GO-P14 are included in Supplementary Table 5.

Statistical analyses and reproducibility. Statistical analyses were performed in GraphPad Prism v.8.2.0 and R v.4.2.1. No statistical methods were used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized and the investigators were not blinded to allocation during outcome assessment.

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
The generated and analysed data are available in the Gene Expression Omnibus with accession number GSE217970. Source data are provided with this paper.

References

83. Peng, X. et al. TELP, a sensitive and versatile library construction method for next-generation sequencing. *Nucleic Acids Res.* **43**, e35 (2015).
84. Hisano, M. et al. Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nat. Protoc.* **8**, 2449–2470 (2013).
85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
86. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
87. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
88. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
89. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
90. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinform.* **10**, 232 (2009).
91. Kind, J. et al. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell* **163**, 134–147 (2015).
92. Quinlan, A. R. BEDTools: the Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinform.* **47**, 11.12.1–11.12.34 (2014).
93. Schug, J. et al. Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol.* **6**, R33 (2005).
94. Saldanha, A. J. Java Treeview–extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).

95. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

96. Dennis, G. Jr et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, P3 (2003).

97. Lin, S. et al. Comparison of the transcriptional landscapes between human and mouse tissues. *Proc. Natl Acad. Sci. USA* **111**, 17224–17229 (2014).

98. Xiong, Z. et al. Ultrasensitive Ribo-seq reveals translational landscapes during mammalian oocyte-to-embryo transition and pre-implantation development. *Nat. Cell Biol.* **24**, 968–980 (2022).

Author contributions

W. Xie and J.L. conceived and designed the project. Z.L., F.K., P.W., J. Ming, F.L. and J. Ma collected and prepared WT and *Eed* KO embryos and oocytes. Y.H. and Y.Q. collected *Tcf3/12* DKO, *Tcf3* mKO and *Tcf12* mKO oocytes. J. Ma, B.L., Y.X. and B.Z. conducted STAR ChIP-seq in mouse oocytes, embryos and mES cells. B.L. conducted ATAC-seq in mouse oocytes. X.W. modified and conducted STARR-seq in mouse oocytes with the help of W.S. Y.H. conducted single-oocyte RNA-seq. X.W. and Y.Q. conducted immunostaining in WT oocytes and embryos with the help of Z.L. and P.W. X.W. conducted plumbagin treatment, EU staining and reporter assay in FGOs with the help of F.K., F.L. and Z.L. F.K., Z.L. and F.L. conducted microinjection in oocytes. Y.Q. conducted qPCR, western blot, IHC, H&E staining in *Tcf3/12* DKO oocytes and IP in HEK293T cells. Y.Q. analysed fertility and follicle numbers of *Tcf3/12* DKO, *Tcf3* mKO and *Tcf12* mKO mice. Q.W. prepared *Dnmt3a/b*^−/− mES cells. M.P. and M.-E.T.-P. conducted lamin B1 DamID. W. Xia helped with various experiments. B.L. conducted the bioinformatics analysis with the help of J.W., B.Z. and Y.X. B.L., X.W. and W. Xie prepared most figures and wrote the manuscript with help of J.L., Y.Q., M.-E.T.-P., M.P., J.N. and all other authors.

a

H3K27ac

b

c

Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Validation of H3K27ac ChIP-seq data in mouse oocytes and early embryos. a, Top, immunostaining showing H3K27ac signals in GO-P7 (n=9), GO-P10 (n=8), FGO-NSN (non-surrounded nucleolus) (n=10), FGO-SN (surrounded nucleolus) (n=4), MII oocyte (n=8), one-cell PN5 (n=16), early two-cell (n=19), late two-cell (n=9), 8-cell embryos (n=6), and blastocyst (n=5). One representative image from 3 independent experiments is shown. Scale bar, 20 μm. Bottom, boxplot showing the ratio of H3K27ac to DAPI intensity in immunostaining. The median is indicated by the center line. The bottom, top edges, and whiskers represent the 10th and 90th percentiles and 1.5 times the interquartile range (IQR), respectively. **b**, The UCSC browser view showing H3K27ac signals at each stage with two biological replicates. **c**, Heatmap showing the pairwise Pearson correlation between each sample for their H3K27ac levels in oocytes, early embryos, and tissues. Source numerical data and unprocessed blots are available in source data.
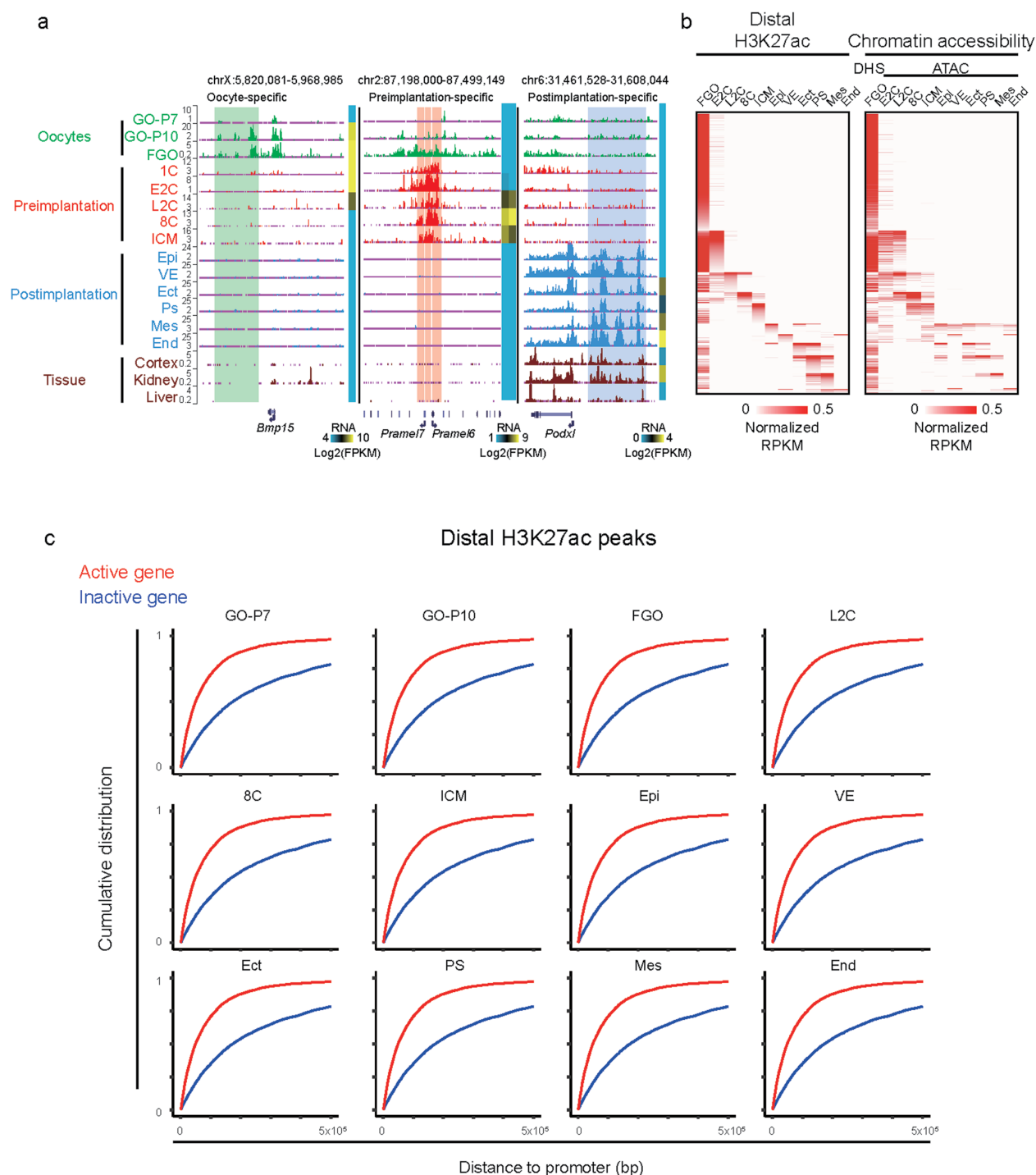
Extended Data Fig. 2 | H3K27ac, gene expression, and chromatin accessibility in mouse oocytes and early embryos. a, The UCSC browser view and heatmaps showing H3K27ac signals and RNA expression at representative genes. Oocyte-, pre-implantation- and post-implantation-specific putative enhancers are shaded green, red, and blue, respectively. **b**, Heatmaps showing the stage-specific distal H3K27ac signals and the mapped chromatin accessibility signals. DHS, DNase I hypersensitive site. **c**, Line charts showing the cumulative distribution of the distances between promoters of active and inactive genes (top 5,000) and nearest distal H3K27ac sites (top 10,000).
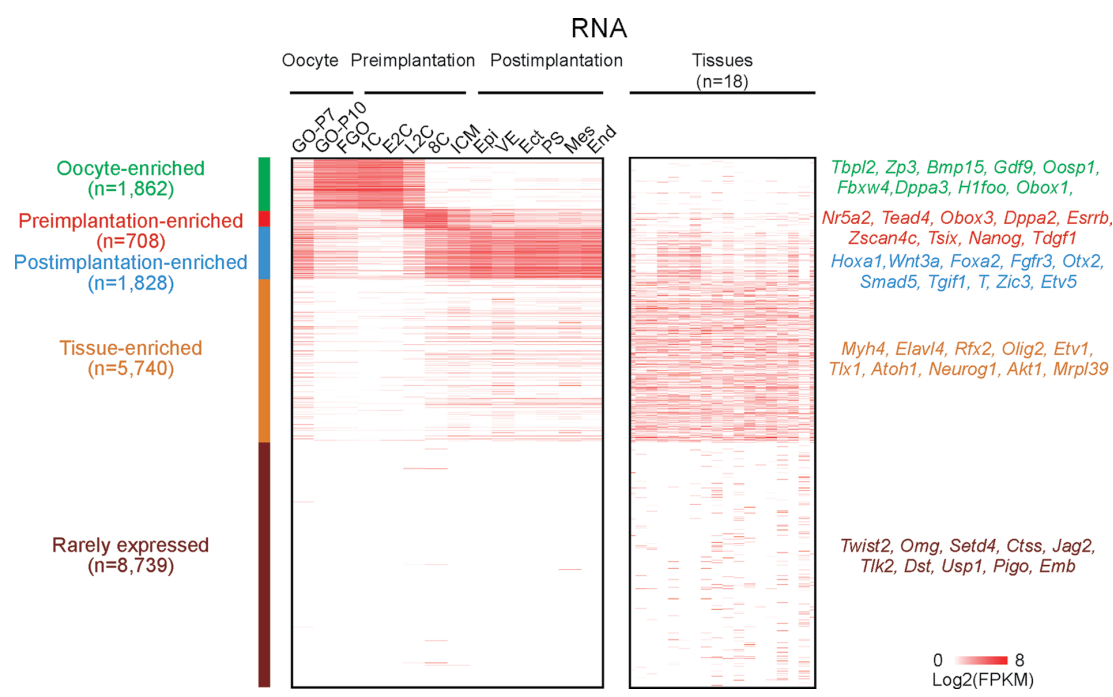
Extended Data Fig. 3 | Gene expression in mouse oocytes, early embryos, and tissues. Heatmaps showing the stage-enriched gene expression in mouse oocytes and early embryos. Mouse adult tissue gene expression[97] is also mapped. Example genes are listed on the right.

Enrichment of repeats in distal H3K27ac peaks
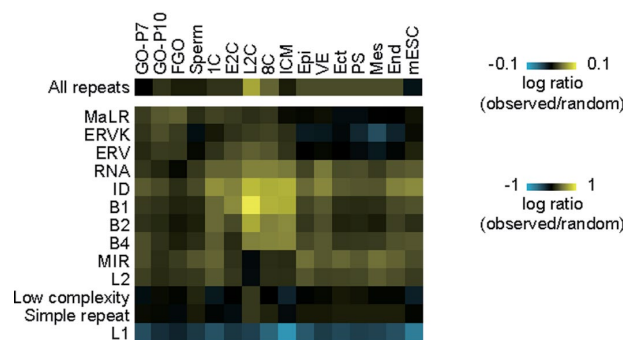
Extended Data Fig. 4 | Repeat enrichment in H3K27ac peaks. Heatmaps showing enrichment of all repeat subfamilies at distal H3K27ac peaks compared to that in random peaks in oocytes, early embryos, and mESCs. The H3K27ac enrichment was calculated as a log2 ratio for the numbers of observed peaks overlapped with repeats divided by those for random peaks (shuffled peaks with lengths matched).
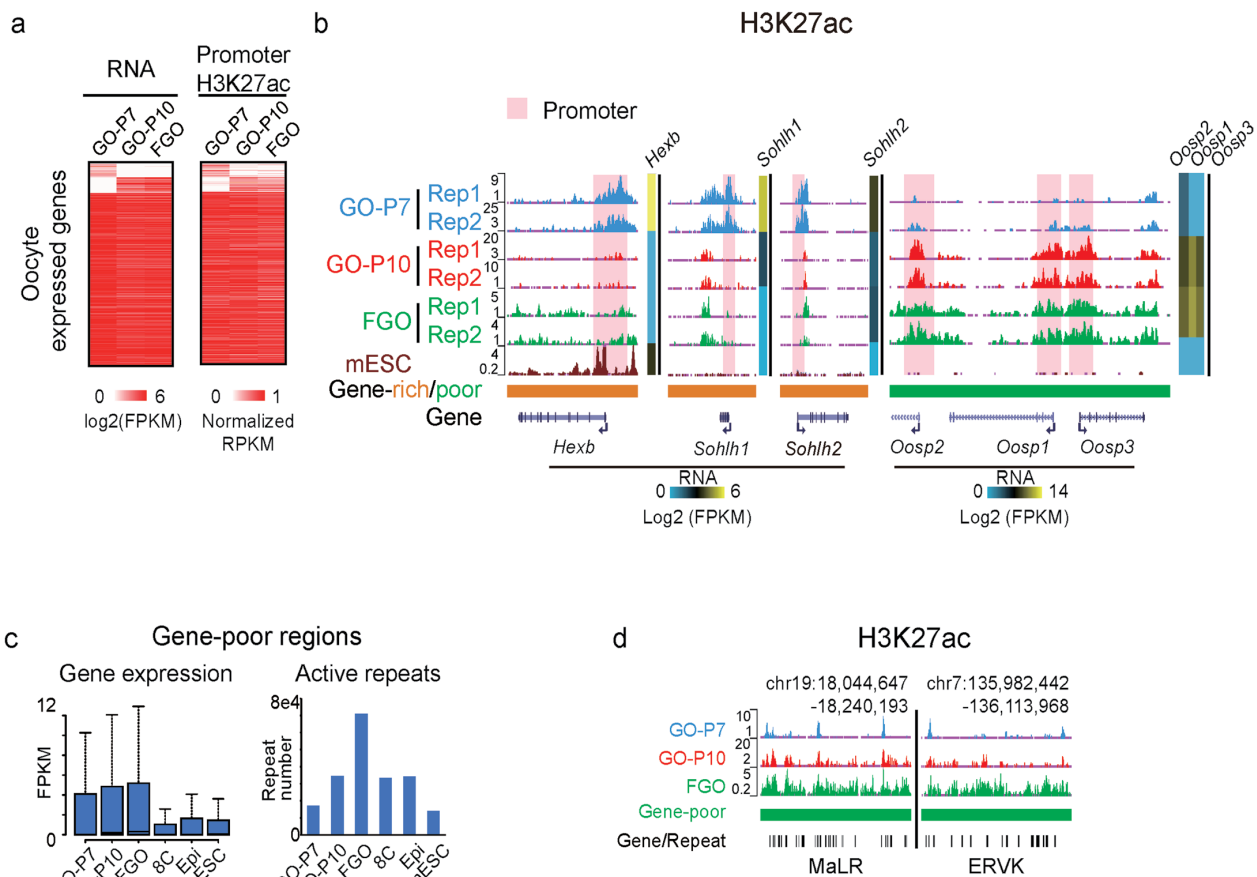
Extended Data Fig. 5 | Dynamics of putative enhancers during mouse oogenesis. a, Heatmaps showing the expression for oocyte stage-specific and common genes with promoter H3K27ac signals mapped. **b**, The UCSC browser views and heatmaps showing H3K27ac signals and RNA expression, respectively, in GO-P7, GO-P10, FGO, and mESCs (ENCODE) with replicates at representative genes. **c**, Bar charts showing the expression levels of genes (left, n=5,505) and the numbers of active repeats (right; GO-P7, n=17,082; GO-P10, n=34,430; FGO, n=70,975; 8C, n=33,384; Epi, n=34,265; mESC, n=14,026) in gene-poor regions at each stage. **d**, The UCSC browser views showing H3K27ac signals at representative repeats in gene-poor regions in GO-P7, GO-P10, and FGOs.
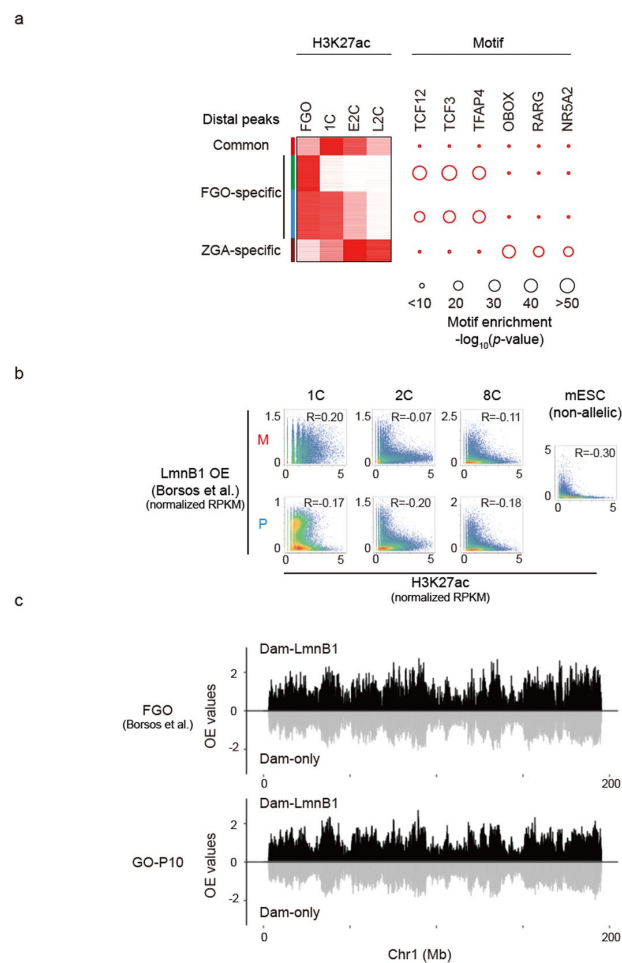
Extended Data Fig. 6 | The transition of putative enhancers during fertilization and ZGA. a, Heatmap and bubble plot comparing distal H3K27ac signals and motif enrichment (-log10 p-value, hypergeometric test with Bonferroni correction, one-sided from HOMER[57], Methods), respectively, at putative enhancers in FGO, one-cell, early two-cell, and late two-cell stage embryos. **b**, Scatter plots comparing allelic H3K27ac and Lamin B1 DamID signals[22] in the one-cell, late two-cell, 8-cell embryos, and mESCs. M, maternal. P, paternal. Pearson correlations of each stage are also shown. **c**, Lamin B1 DamID and DamID-only control signal profiles in FGOs[22] and GO-P10 for chromosome 1 (n=4 biological replicates).
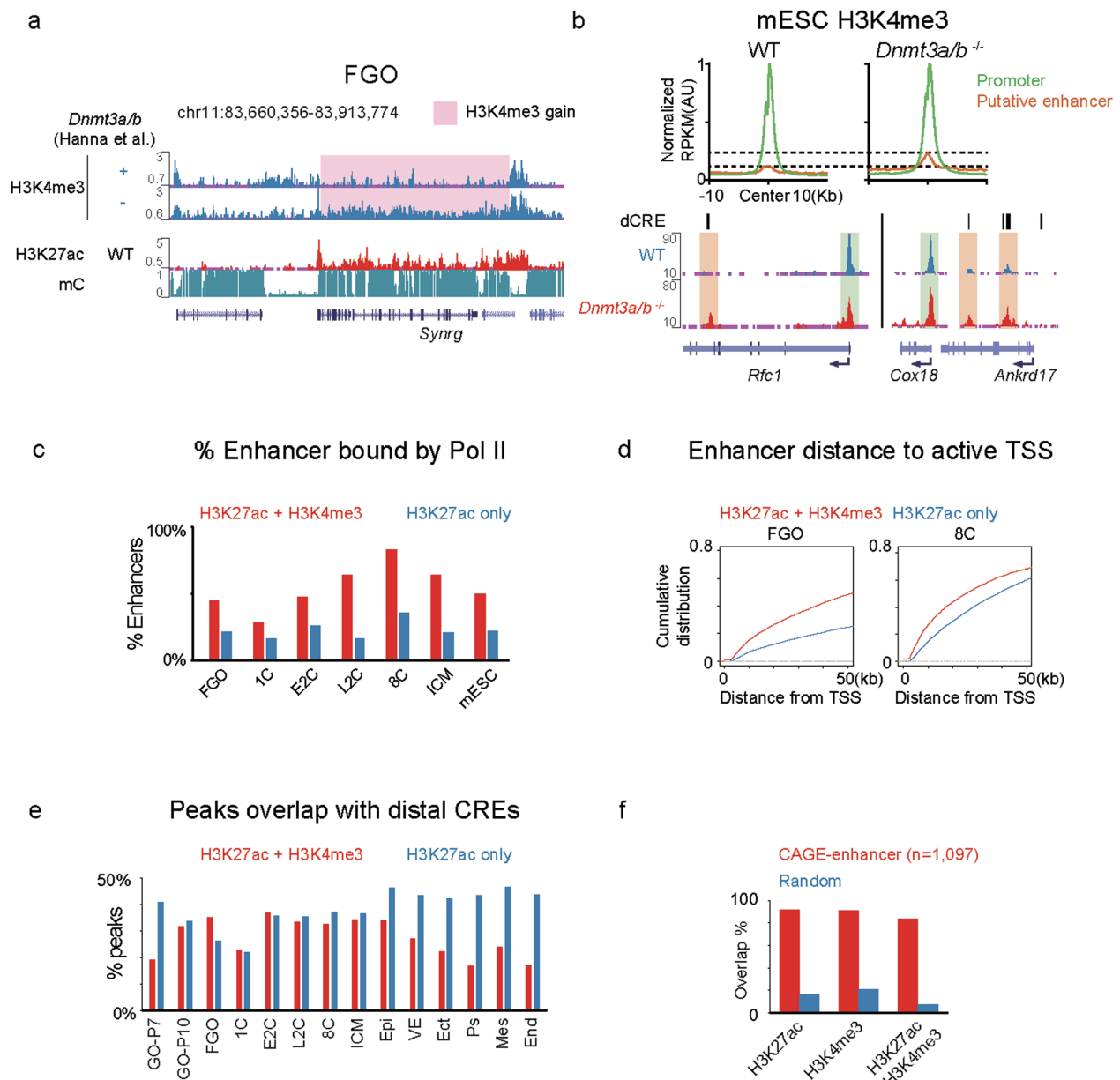
Extended Data Fig. 7 | Analyses of H3K4me3-marked enhancers in oocytes and embryos. a, The UCSC browser view showing H3K4me3 signals in control and *Dnmt3a/b* knockout FGOs[40], and H3K27ac and DNA methylation signals in wild-type FGOs. H3K4me3-gain regions upon *Dnmt3a/b* knockout are shaded. **b**, Top, line charts showing H3K4me3 signals at putative enhancers (orange) and promoters (green) in wild-type (left) and *Dnmt3a/b* KO (right) mESCs. The dashed lines indicate the peaks of H3K4me3 signals at putative enhancer regions. Bottom, the UCSC browser views showing H3K4me3 signals at promoters and putative enhancers (annotated ENCODE dCRE) at representative genes in wild-type and *Dnmt3a/b* KO mESCs. Putative enhancer and promoter regions are shaded orange and green, respectively. **c**, Bar chart showing the percentages of

H3K27ac/H3K4me3 co-marked and H3K27ac only marked enhancers bound by distal Pol II[34] at each stage. **d**, Line charts showing the cumulative distribution of the distance between transcription start sites (TSSs) of active gene and nearest distal putative enhancers marked by either H3K27ac only (blue) or both H3K27ac and H3K4me3 (red) in FGO (left) and the 8-cell embryos (right). **e**, Bar chart showing the percentages of H3K27ac only (blue) and H3K27ac/H3K4me3 (red) peaks that also overlap ENCODE dCREs. **f**, Bar chart showing the percentages of CAGE-defined enhancer sites that also overlap H3K27ac sites, H3K4me3 sites, or both. Random sites with identical lengths and numbers were similarly analysed as controls.
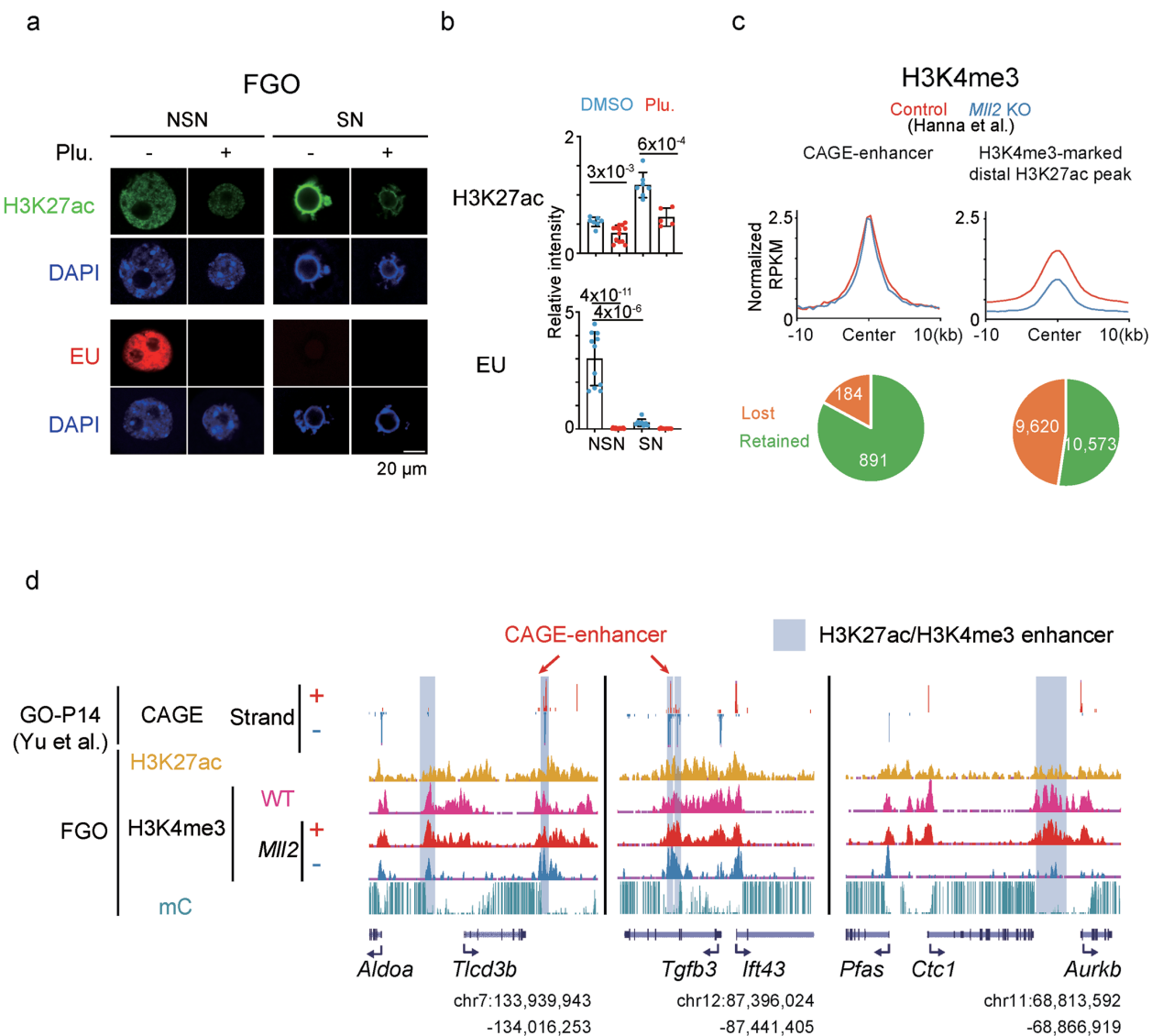
Extended Data Fig. 8 | Analyses of enhancer activities in oocytes.
a, Representative images of H3K27ac immunostaining (green) and EU staining (red) in DMSO or plumbagin-treated FGO-NSN and FGO-SN. DNA was stained with DAPI (blue). Scale bar, 20μm. **b**, Bar charts showing the relative intensities of H3K27ac or EU in DMSO (blue) or plumbagin (red) treated FGO-NSN and FGO-SN. The dots indicate the relative intensities of individual oocytes. The numbers of oocytes used: 7, 13, 7, and 5 for H3K27ac (top), and 11, 17, 8, and 9 for EU staining (bottom). *P*-values (*t*-test, two-sided) are also provided. **c**, Line charts showing

H3K4me3 signals at CAGE-enhancers and H3K4me3-marked distal H3K27ac peaks in *Mll2* control and KO oocytes[40]. Pie charts showing the percentages of CAGE-enhancers and H3K4me3 at distal H3K27ac peaks in *Mll2* KO oocytes compared with wild-type. **d**. The UCSC browser views showing CAGE in GO-P14 (ref. [46]), H3K27ac, and H3K4me3 in FGOs. Putative enhancer regions defined by CAGE (red arrows) or distal H3K27ac (blue shades) are indicated. Source numerical data and unprocessed blots are available in source data.

a FGO STARR-seq

b FGO STARR-seq

c FGO STARR-seq

d FGO

e
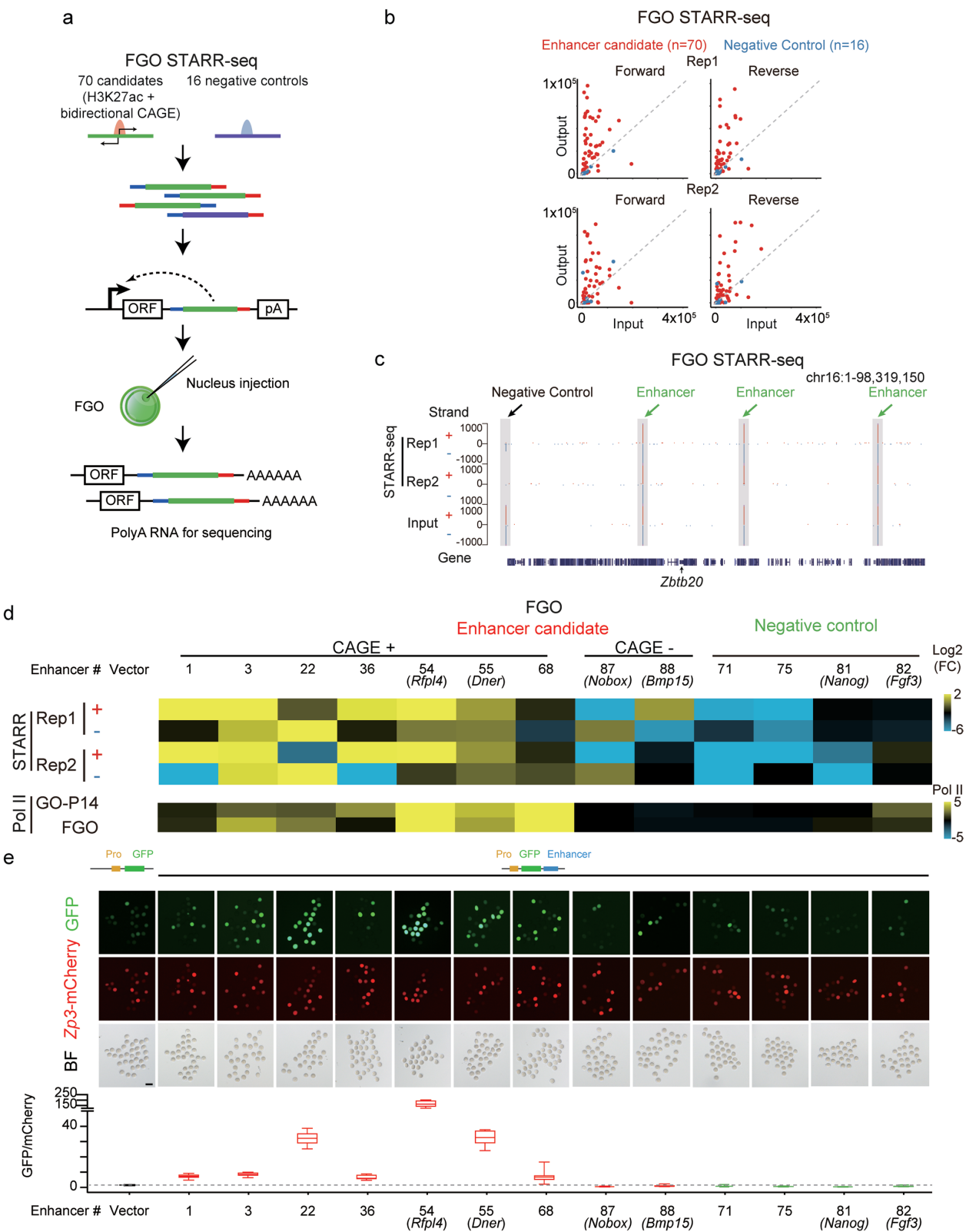
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | STARR-seq and reporter assay in oocytes. a, Schematic of STARR-seq in FGOs. 70 candidates and 16 negative controls were manually cloned into the STARR-seq constructs and then the pooled products were injected into the nuclei of FGOs. The RNA was recovered by a method adapted from Smart-seq2 (ref. 54) (Methods) to suit low-input cells, followed by sequencing. **b**, Scatter plots showing STARR-seq signals (RNA output vs. DNA input) (Methods) in both replicates and strands. Red, enhancer candidates; blue, negative control elements. **c**, UCSC genome browser showing FGO STARR-seq RNA output and DNA input signals on chromosome 16. **d**, Heatmaps showing STARR-seq (STARR/input) signals in FGO with two replicates and Pol II signals in GO-P14 and FGO at enhancer candidates and negative control regions. **e**. Top, fluorescence and bright fields of mouse FGOs in an enhancer reporter assay (Pro, mini promoter). Scale bar, 100 μm. Bottom, boxplot showing the ratio of GFP to mCherry intensity in the enhancer reporter assay. The dashed line indicates the ratio in the empty vector group. The numbers of oocytes used in each group: 17, 14, 17, 18, 10, 18, 15, 17, 13, 10, 12, 19, 11, and 6. The median is indicated by the center line. The bottom, top edges, and whiskers represent the 10th and 90th percentiles and 1.5 times the interquartile range (IQR), respectively. Source numerical data and unprocessed blots are available in source data.

a

FGO STARR-seq

b

TCF3 TCF12

Motif enrichment
-lg (*p*-vlaue)

c

GO mRNA

d

GO mRNA

e

Ovary

f

Ovary

g

Sp1
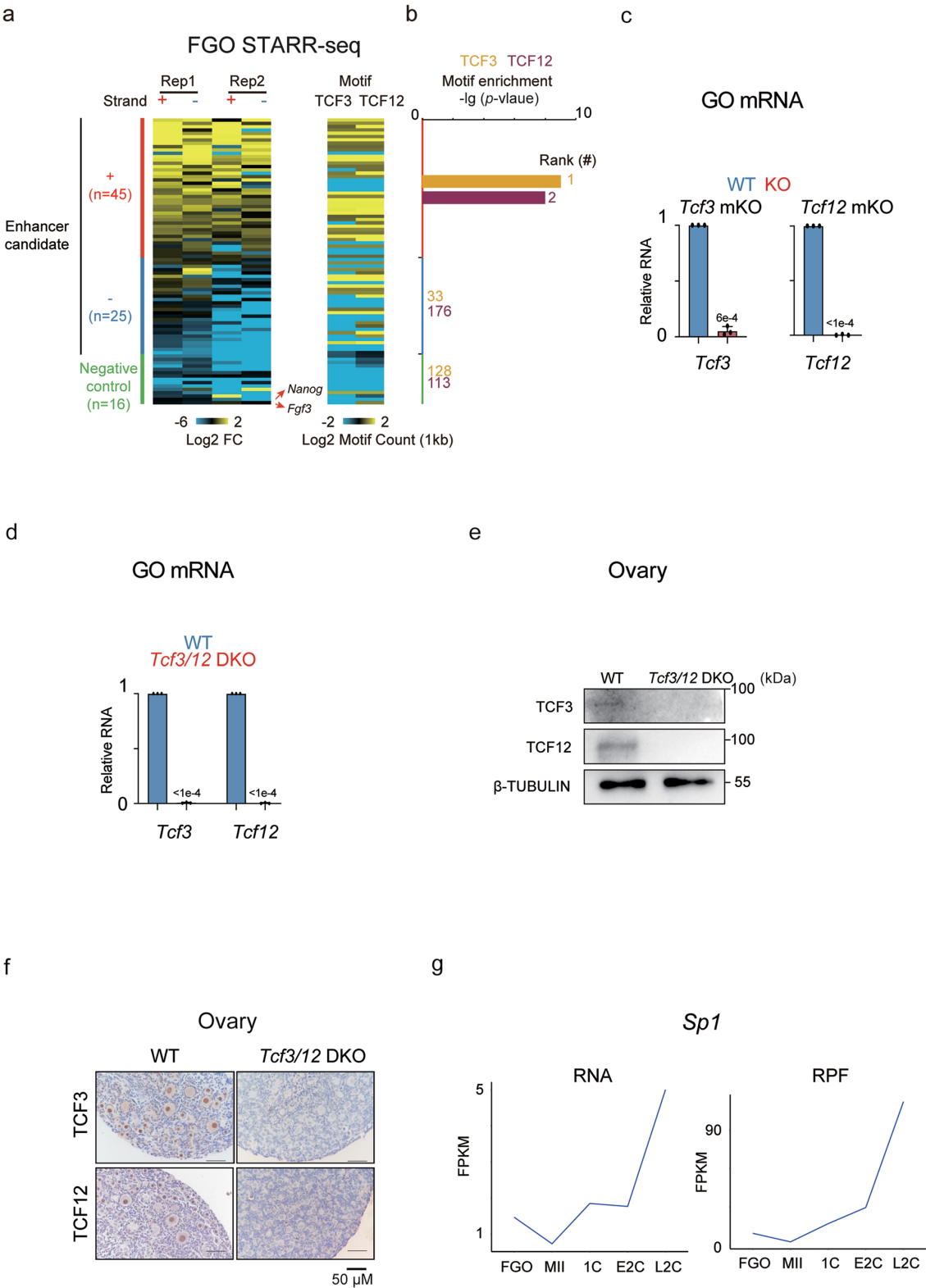
Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | TCF3/12 regulate oocyte transcription and development. a. Heatmaps showing STARR-seq signals (RNA/DNA input) and TCF3/12 motif densities (counts per kb) for enhancer candidates (including those showing positive and negative STARR-seq signals) and negative controls. **b**. Bar charts showing TCF3/12 motif enrichment (-log10 p-value, hypergeometric test with Bonferroni correction, one-sided from HOMER[57], Methods) in enhancers previously evaluated by STARR-seq and negative controls. Ranks of TCF3/12 motifs among all motifs are also shown. **c**, qPCR results showing relative RNA levels of *Tcf3* and *Tcf12* in wild-type and *Tcf3* or *Tcf12* mKO GO-P5 oocytes (n=3

biological replicates). *P*-value (*t*-test, two-sided) is also shown. **d**, qPCR results showing relative RNA levels of *Tcf3* and *Tcf12* in WT and *Tcf3*/12 DKO GO-P5 oocytes (n=3 biological replicates). *P*-value (*t*-test, two-sided) is also shown. **e**, Western blot showing protein levels of TCF3 and TCF12 in WT and *Tcf3*/12 DKO GO-P5 ovaries (n=3 biological replicates). **f**, Immunohistochemistry (IHC) showing TCF3 and TCF12 in WT (n=5) and *Tcf3/12* DKO (n=5) ovaries. Scale bar, 50 μm. **g**, RNA expression and ribosome-protected fragment (RPF, indicating translation level) levels of *Sp1* from FGOs to late two-cell embryos are shown[98]. Source numerical data and unprocessed blots are available in source data.

Part III
Publication

Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity

Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity

Mrinmoy Pal,[1] Luis Altamirano-Pacheco,[1] Tamas Schauer,[1] and Maria-Elena Torres-Padilla[1,2]

[1]Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München, D-81377 München, Germany; [2]Faculty of Biology, Ludwig-Maximilians Universität, D-81377 München, Germany

Fertilization in mammals is accompanied by an intense period of chromatin remodeling and major changes in nuclear organization. How the earliest events in embryogenesis, including zygotic genome activation (ZGA) during maternal-to-zygotic transition, influence such remodeling remains unknown. Here, we have investigated the establishment of nuclear architecture, focusing on the remodeling of lamina-associated domains (LADs) during this transition. We report that LADs reorganize gradually in two-cell embryos and that blocking ZGA leads to major changes in nuclear organization, including altered chromatin and genomic features of LADs and redistribution of H3K4me3 toward the nuclear lamina. Our data indicate that the rearrangement of LADs is an integral component of the maternal-to-zygotic transition and that transcription contributes to shaping nuclear organization at the beginning of mammalian development.

In mammals, development begins upon fertilization of the oocyte by the sperm, two highly differentiated cells, and gives rise to the one-cell embryo, or zygote. After fertilization, the chromatin of the two gametes undergoes an intense period of chromatin remodeling, which is essential to start a new developmental program. As part of this remodeling, histone modifications are reset and re-established genome-wide with different developmental kinetics, a process that continues until at least 3.5 d later at the blastocyst stage (Burton and Torres-Padilla 2014; Xia and Xie 2020). For example, the oocyte is characterized by broad H3K4me3 domains, which are inherited by the zygote and are largely remodeled during the first two embryonic cell divisions (Dahl et al. 2016; Liu et al. 2016; Zhang et al. 2016). In addition, heterochromatin progressively matures and the patterns of H3K27me3 and H3K9me3 emerge gradually during preimplantation development (Liu et al. 2016; Wang et al. 2018; Burton et al. 2020). This time window is also characterized by a robust expression of transposable elements (Peaston et al. 2004; Fadloun et al. 2013) with H3K4me3 peaks known to be enriched for repeats in mouse preimplantation embryos (Zhang et al. 2016).

Early development is initially supported by maternal transcripts, which are synthesized during oocyte growth and inherited by the embryo. The embryo transitions away from the dependence on maternal supplies in a process referred to as maternal-to-zygotic transition (MZT) (Schultz 2002; Li et al. 2013; Vastenhouw et al. 2019). MZT occurs through several processes, including the activation of the embryonic genome and the degradation of maternal transcripts. In mice, embryonic genome activation, referred to as zygotic genome activation (ZGA), occurs in two waves: A minor wave of transcriptional activation occurs at the late zygote stage, and a major, more substantial transcriptional wave occurs at the late two-cell stage (Zeng and Schultz 2005; Abe et al. 2018; Schulz and Harrison 2019).

The 3D nuclear organization in the early embryo is also heavily remodeled (Pecori and Torres-Padilla 2023). In mice, A and B chromatin compartments are initially not well defined and compartment strength increases gradually during preimplantation development. Likewise, topologically associating domains (TADs) also mature progressively during this period, with TAD boundaries becoming progressively insulated as development proceeds (Du et al. 2017; Flyamer et al. 2017; Ke et al. 2017). Lamina-associated domains (LADs) are genomic regions that

make contact with the nuclear lamina (van Steensel and Belmont 2017) and can reshuffle stochastically after mitosis (Kind et al. 2013). Interestingly, LADs are rapidly established after fertilization, and thus LAD formation precedes consolidation of TADs. In fact, 67% of the LADs established in zygotes correspond to "constitutive" LADs (cLADs) (Borsos et al. 2019), which are cell-type-invariable LADs (Meuleman et al. 2013). However, LADs also become rearranged as development progresses, in particular at the late two-cell stage, where LADs have atypical features compared with cLADs and LADs in other cell types (Peric-Hupkes et al. 2010). For example, two-cell stage LADs are smaller and appear fragmented compared with the zygote stage and display distinctive genomic features, including relatively low CpG density and AT content (Borsos et al. 2019). Such unusual spatial arrangement is transient, as LADs in the four-cell stage do not show such features. Almost 42% of the zygotic LADs reposition to the nuclear interior at the two-cell stage, and two-cell-specific LADs emerge (Borsos et al. 2019). However, how such large-scale genome rearrangement is regulated, and whether developmental processes such as ZGA contribute to these changes in nuclear organization is not known.

Here, we analyzed the rearrangement of LADs that occurs during MZT and defined the role of zygotic genome activation in this process. Our work indicates that LAD reorganization in two-cell embryos is gradual and dynamic. By inhibiting ZGA with two different inhibitors, we show that transcriptional activity at ZGA contributes to LAD reorganization. Surprisingly, transcriptional inhibition of RNA polymerase II results in a redistribution of H3K4me3 domains to the nuclear periphery, which is accompanied by the large-scale repositioning of LAD boundaries and the sequestration of major ZGA genes at the nuclear lamina. Our work sheds light onto the molecular determinants of nuclear organization at the beginning of mammalian development.

Results

LAD reorganization in two-cell embryos is gradual and dynamic

To investigate the temporal definition and molecular regulators of the changes in nuclear organization that occur during ZGA, we focused on LADs. We previously mapped LADs in mouse zygotes before mitosis and in late two-cell stage embryos, which are separated by ∼24 h. In order to obtain a better temporal resolution of LADs during development, we first mapped LADs in early two-cell stage embryos using LaminB1-DamID (Borsos et al. 2019), which corresponds to the end of G1 phase of the second embryonic cell cycle, prior to the major ZGA wave (Fig. 1A; Jukam et al. 2017; Schulz and Harrison 2019). Overall, genome-wide DamID values correlated equally to zygotes and to late two-cell stage embryos (Fig. 1B), suggesting that LADs have intermediate features between the two stages. In addition, despite their similar correlation, the range of DamID values in early two-cell stage is more sim-

ilar to zygote than to late two-cell stage embryos (Fig. 1B). To investigate this further, we performed principal component analysis using the DamID methylation values (Fig. 1C), which indicated that early two-cell LaminB1-DamID values indeed separate between late two-cell stage and zygotes along PC2, while PC1 separates later developmental stages, including eight-cell stage and embryonic stem (ES) cells (Fig. 1C).

We next defined LADs based on the LaminB1-DamID values by using a hidden Markov model (HMM) as before (Meuleman et al. 2013). Visual inspection of LADs confirmed the expected fragmented LAD profile characteristic of late two-cell stage embryos, compared with zygotes (Fig. 1D; Borsos et al. 2019). Early two-cell LADs appeared to be of an intermediate nature, showing both small fragmented LADs and larger continuous LADs (Fig. 1D). Indeed, the median LAD length in early two-cell stage embryos was 1.8 Mb (interquartile range 0.9–3.6 Mb), compared with a median length of 2.25 and 0.9 Mb in zygotes and late two-cell stage embryos, respectively (Fig. 1E). In addition, the percentage of the genome located at the nuclear lamina in early two-cell stage embryos (42.12%) was intermediate between zygotes (36.40%) and late two-cell stage embryos (48.34%), potentially suggesting that the genome largely reorganizes by moving toward the nuclear lamina progressively during this stage (Fig. 1F). The average genomic AT content was clearly higher in early two-cell LADs compared with inter-LADs (iLADs) and was similar to the zygote (Fig. 1G). On the other hand, the pattern of gene density between LADs and iLADs in early two-cell embryos was more similar to the late two-cell stage embryo, in contrast to the zygote, in which iLADs displayed a much higher gene coverage compared with LADs (Fig. 1H). Finally, CpG density in early two-cell stage embryos displayed an intermediate enrichment in iLADs compared with zygotes and late two-cell stage embryos (Fig. 1I). Altogether, these data indicate that LADs at the two-cell stage mature gradually during the complete cell cycle with evolving molecular and genomic features. This suggests that dynamic rearrangement of genome–lamina contacts occurs during interphase progression. The latter is in line with recent work that has demonstrated that LADs evolve over the cell cycle in human cells in culture (van Schaik et al. 2020).

Repositioning of LADs following mitosis correlates with gene and TE class expression

To investigate the potential determinants and the dynamics of LAD reorganization upon the first mitosis and during the two-cell stage in more detail, we first examined whether and when LADs dislodge from the nuclear lamina (e.g., when they become iLADs) and vice versa. We found that rearrangement of the genomic regions in iLADs and LADs occurs both between zygotes and early two-cell stage embryos and between early and late two-cell stage embryos (Fig. 2A). For example, 19% of iLADs become LADs between zygotes and early two-cell stage and remain LADs at the late two-cell stage (iL-L-L) (Fig. 2A). However, we also found that 21% of zygotic iLADs

108
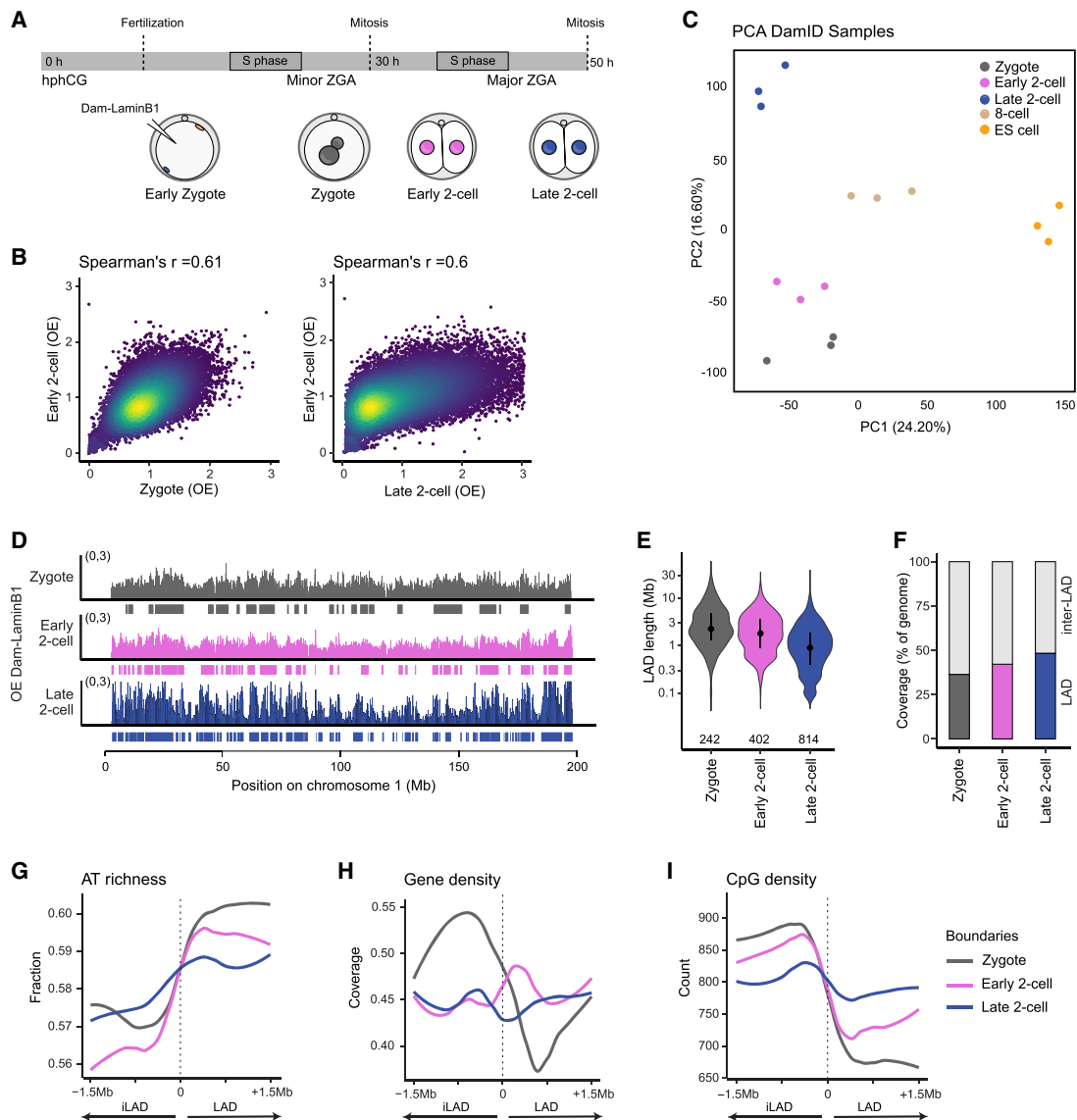
Figure 1. LAD reorganization in two-cell embryos is gradual and dynamic. (*A*) Schematic with time line of early development in mouse embryos and DamID collection time points. (hphCG) Hours post-hCG. (*B*) Genome-wide scatter plots (100-kb bins) of observed over expected (OE) Dam-LaminB1 mean scores from three biological replicates. Spearman's correlation coefficients are indicated. (*C*) Principal component analysis (PCA) of DamID samples. Zygote and late two-cell, eight-cell, and embryonic stem (ES) cell population DamID data analyzed from GSE112551 (Borsos et al. 2019). Each data point represents a biological replicate for the corresponding stages as indicated by the color code. The percentage of variance explained by PC1 and PC2 is indicated in the axis labels. (*D*) Dam-LaminB1 OE value profiles on chromosome 1. Boxes *below* the tracks represent LADs called by HMM. (*E*) Distribution of LAD length. Violin plots show the 25th and 75th percentiles (black lines) and median (circles). *n* indicates the number of LADs called, shown *below* the violin plots. (*F*) The percentage genomic coverage of LADs and iLADs. (*G–I*) Average AT content (*G*), gene coverage (*H*), and CpG density (*I*) (calculated in 100-kb genomic bins) over LAD boundaries of their own developmental stage. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot, and the 1.5-Mb region at the *right* indicates LAD.

remain iLADs in early two-cell stage embryos but become associated with the nuclear lamina at the late two-cell stage (iL-iL-L) (Fig. 2A). Additionally, only a small proportion of the genome (2.2%) behaves as "transient" early two-cell stage iLADs: These are regions that are LADs in zygotes and late two-cell stage embryos but dislodge from the nuclear lamina and are iLADs in early two-cell stage embryos (L-iL-L) (Fig. 2A). However, in line with our previous findings (Borsos et al. 2019), around half of the genome (57%) does not rearrange between zygotes and two-cell stage embryos—35% remain as LADs and 22% remain as iLADs (L-L-L and iL-iL-iL) (Fig. 2A). These data indicate that the rearrangements of the genome are progressive and occur both after the first embryonic

109
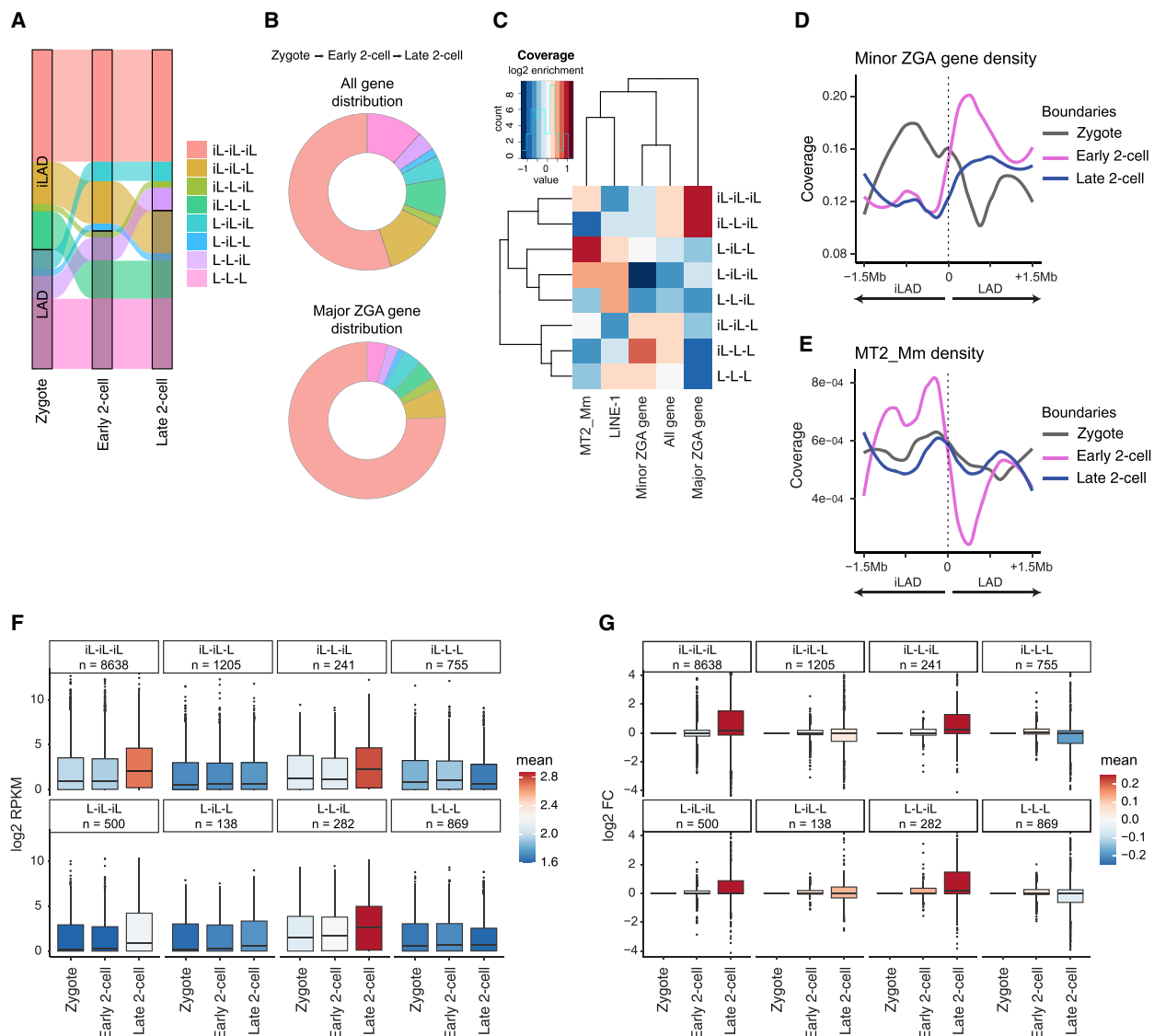
Figure 2. Repositioning of LADs following mitosis correlates with gene and TE class expression. (*A*) Alluvial plot showing LAD reorganization during maternal-to-zygotic transition between zygotes and early and late two-cell stage embryos, respectively. (L) LAD, (iL) inter-LAD. (*B*) Pie charts showing distribution of all genes and major ZGA genes in groups of reorganizing genomic bins. The color code for the groups of reorganizing genomic bins is the same as in *A*. (*C*) Enrichment of genomic features (MT2_Mm, LINE-1, major ZGA, minor ZGA, and all gene coverage) in each group. The average density of all genomic bins was used as expected value. (*D,E*) Metaplot of minor ZGA gene density (*D*) and MT2_Mm enrichment (*E*) on LAD boundaries of the corresponding developmental stage. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot, and the 1.5-Mb region at the *right* indicates LAD. (*F,G*) Gene expression levels (log₂ RPKM) (*F*) and log₂ fold change of gene expression (*G*) contained within each of the LAD/iLAD pattern groups comparing zygotes with early and late two-cell embryos. Note that we excluded transcripts from maternal genes. Single-cell RNA-seq data were analyzed from GSE45719 (Deng et al. 2014). Box plots show median, and the interquartile range and the plots are colored based on mean values. *n* indicates the number of genes analyzed in each group.

mitosis and along with the progression of the second cell cycle at the two-cell stage.

Because major ZGA starts during the two-cell stage (Jukam et al. 2017; Schulz and Harrison 2019), we next asked whether the genome rearrangements that we observed are associated with the genes activated at ZGA (major ZGA genes) (Park et al. 2015). We found that

most major ZGA genes (1111 out of 1462 genes) are located inside constant iLADs (iL-iL-iL; odds ratio 1.4) (Fig. 2B). The remainder of major ZGA genes displayed changes associated with repositioning both from the nuclear interior (iLADs) to the nuclear lamina (LADs) and toward the nuclear interior at the early or late two-cell stage (Fig. 2B). However, ~80% of the major ZGA genes (234 out of

290) that are repositioned in two-cell stage embryos become repositioned to the nuclear interior at either the early or late two-cell stage (Fig. 2B). Thus, while most major ZGA genes are located in iLADs in zygotes and two-cell stage embryos, a proportion of major ZGA genes changes their association with the nuclear lamina at the early and late two-cell stage.

We next performed the opposite analysis and asked whether the genomic regions that become repositioned with respect to the nuclear lamina between the zygote and late two-cell stage embryos are enriched in ZGA genes. Overall, cLADs and iLADs that become LADs in both two-cell stages are depleted of major ZGA genes (L-L-L and iL-L-L) (Fig. 2C) but not minor ZGA genes (Fig. 2C). In addition, the zygotic iLADs that reposition to the nuclear lamina by the end of G1 phase of two-cell stage embryos (iL-L-L) are enriched in minor ZGA genes (odds ratio 1.41) (Fig. 2C). This observation is further supported by the clear switch of minor ZGA gene density from outside the LAD boundary to inside LADs in early two-cell stage embryos (Fig. 2D). Interestingly, we also observed a strong enrichment of MT2_Mm (a MERVL-derived LTR) coverage, but not of LINE-1 elements, in transient early two-cell stage inter-LADs (L-iL-L) (Fig. 2C). Instead, LINE-1 elements are enriched in regions going away from the nuclear lamina from the early to the late two-cell stage (L-iL-iL and L-L-iL) (Fig. 2C), coinciding with their increase in expression levels (Supplemental Fig. S1A; Fadloun et al. 2013; Jachowicz et al. 2017). Notably, MERVL elements shifted their 3D localization completely at this developmental time: Genomic regions just outside the early two-cell stage LAD boundaries became highly enriched in MERVL (MT2_Mm), in contrast to both the zygotic and late two-cell stages, which display neither enrichment nor depletion (Fig. 2E). These data suggest that MERVL elements move toward the nuclear interior (iLADs) at the early two-cell stage, where they are particularly enriched at the LAD–iLAD boundaries. Considering that MERVL elements are highly and transiently expressed at the early two-cell stage (Ishiuchi et al. 2015; Kruse et al. 2019; Liu et al. 2020; Sakashita et al. 2023), these observations establish that the repositioning of MERVL-containing LADs into iLADs at this stage is concordant with their transcriptional activation prior to major ZGA. Thus, we next addressed whether levels of gene expression in general are associated with specific changes of LADs and iLADs that occur during this time window. Because oocytes carry maternal transcripts that accumulate during oocyte growth, we removed maternal genes from our analysis to avoid the confounding effects of oocyte-inherited transcripts present in the zygote. Genes in LADs at the two-cell stage showed the lowest transcript abundance regardless of whether such LADs were cLADs or iLADs prior to the two-cell stage (Fig. 2F). Indeed, changes in gene expression occurred in regions of the genome that repositioned into LADs at the two-cell stage, with a clear reduction in expression from the zygote to the late two-cell stage but notably also when compared with the early two-cell stage (Fig. 2G). Consistent with our observation that most major ZGA

genes remain within iLADs (Fig. 2B), we noted a higher expression of associated genes in the late two-cell stage embryos (iL-iL-iL) (Fig 2F,G). Thus, our data indicate a dynamic repositioning of a subset of LADs and iLADs during the two-cell stage that correlates with the transcriptional activity of genes contained therein. In addition, MERVL elements are a unique feature of early two-cell stage iLADs.

Transcriptional inhibition results in large-scale alterations in LADs at the two-cell stage

TAD borders are known to be remodeled to a certain extent by transcriptional activity in cultured cells and in *Drosophila* embryos (Li et al. 2015; Hug et al. 2017; Rowley et al. 2017). However, the extent to which transcriptional activity remodels nuclear architecture in mouse embryos is not fully characterized. Thus, to address directly whether and how transcription during ZGA affects LADs, we performed DamID for LaminB1 in late two-cell stage embryos after incubation with two different RNA polymerase (Pol II) inhibitors: α-amanitin and DRB. While DRB inhibits transcriptional elongation through inhibition of RNA Pol II serine 2 phosphorylation (Dubois et al. 1994), α-amanitin results in full transcriptional inhibition, including via RNA Pol II degradation (Nguyen et al. 1996; Bensaude 2011; T Nakatani, T Schauer, L Altamirano, et al., in prep.). We incubated embryos with either of the two inhibitors continuously from the early zygote stage until the late two-cell stage under conditions known to prevent ZGA (Abe et al. 2018; Liu et al. 2020; T Nakatani, T Schauer, L Altamirano, et al., in prep.) and mapped LADs at this stage (Fig. 3A). PCA revealed that embryos in which ZGA was inhibited clustered together regardless of the inhibitor used and separately from untreated embryos (Fig. 3B). The samples segregated based on their developmental stage along PC2, and both DRB- and α-amanitin-treated late two-cell stage embryos separated from untreated embryos along PC1 and localized closest to the early two-cell stage untreated samples along PC2 (Fig. 3B). These observations suggest that while inhibition of ZGA results in a genome–nuclear lamina interaction profile that is more similar to embryos in which major ZGA has not yet occurred (early two-cell stage), transcriptional inhibition leads also to additional profound alterations (Supplemental Fig. S1B). This suggests that ZGA contributes to the LAD rearrangement that occurs in late two-cell stage embryos but is not the sole determinant of this process. Globally, the genome-wide interactions with the nuclear lamina were affected to a similar extent upon the treatment of the two inhibitors (Spearman's $r = 0.76$) (Fig. 3C), but we noted slightly a stronger effect upon DRB treatment compared with α-amanitin treatment when correlated to both untreated early and late two-cell stage embryos (Fig. 3D,E). This could potentially be due to the differences in the mode of action of the two inhibitors (Bensaude 2011).

Visual inspection of DamID methylation levels and LADs over chromosome tracks revealed that late two-cell stage embryos treated with both α-amanitin and

DRB contain LADs that are less fragmented than the untreated late two-cell stage embryos (Fig. 3F). These data suggest that transcriptional activity underlies the unusual spatial arrangement of LAD patterns in late two-cell stage embryos, characterized by small fragmented LADs (Borsos et al. 2019). To further investigate this, we asked
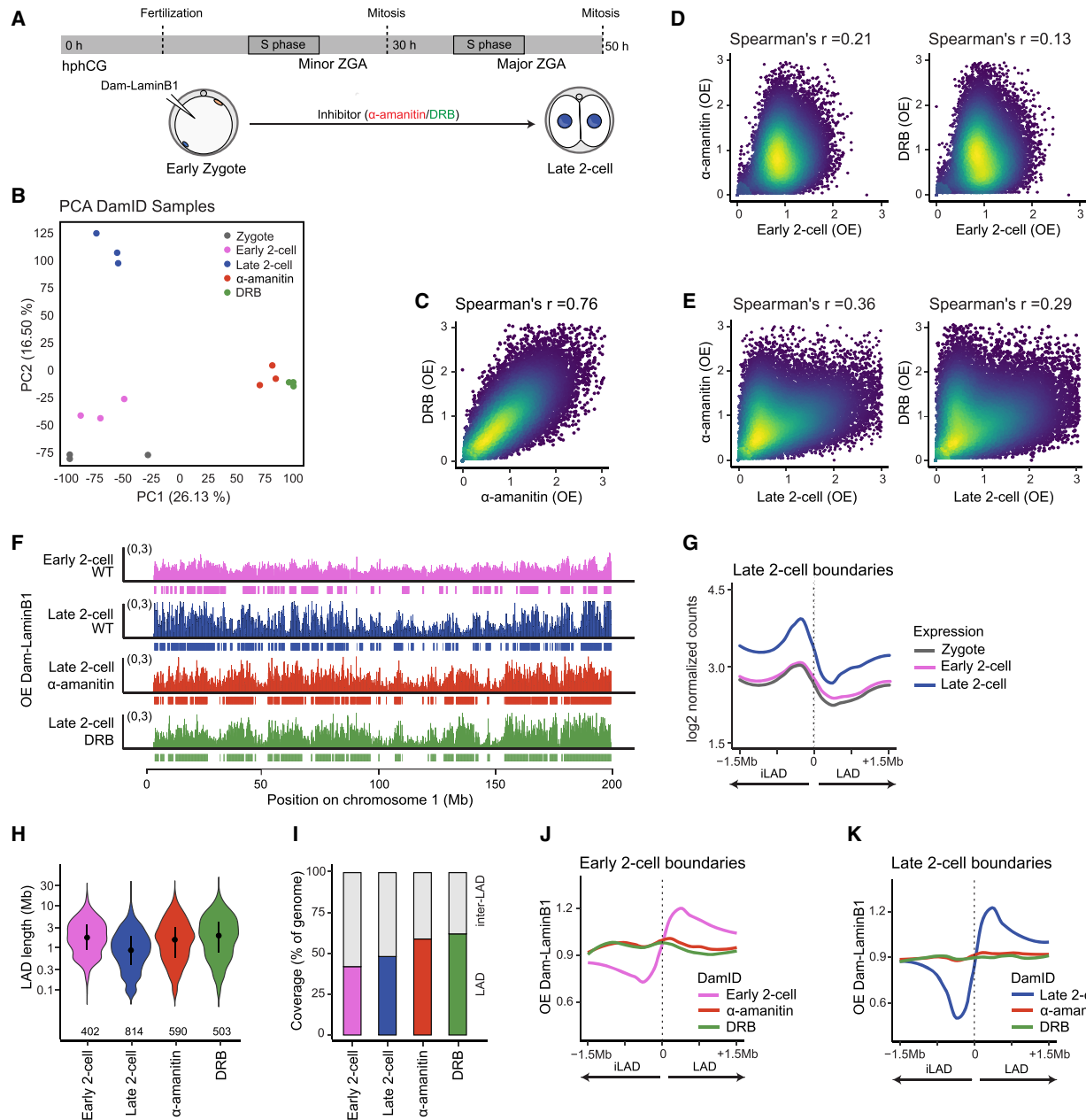
Figure 3. Transcriptional inhibition results in large-scale alterations in two-cell LADs. (*A*) Schematic representation of the experimental design for late two-cell DamID upon inhibition of minor and major waves of ZGA with either α-amanitin or DRB. (hphCG) Hours post-hCG. (*B*) Principal component analysis (PCA) of DamID samples. Each data point represents a biological replicate for the corresponding condition as indicated by the color code. The percentage of variance explained by PC1 and PC2 is indicated in the axis labels. (*C*–*E*) Genome-wide scatter plot of mean OE values. Spearman's correlation coefficients are indicated. (*F*) Dam-LaminB1 OE mean value profiles and called LADs on chromosome 1 in control (WT) early and late two-cell embryos along with α-amanitin- or DRB-treated late two-cell embryos. (*G*) Average log$_2$ normalized counts calculated on 100-kb genomic bins for zygotes and early and late two-cell stage embryos analyzed from GSE45719 (Deng et al. 2014) plotted over late two-cell stage LAD boundaries. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot, and the 1.5-Mb region at the *right* indicates LAD. (*H*) Violin plots showing size distribution of LADs. The number of LADs is indicated *below*. (*I*) The percentage of genome coverage of LADs and inter-LADs. (*J,K*) Average Dam-LaminB1 DamID signal over untreated early (*J*) and late (*K*) two-cell stage LAD boundaries.

112

whether the edges of fragmented LADs are defined by the transcriptional activity, based on RNA-seq, centered over LAD boundaries at the late two-cell stage. Indeed, we found that transcript levels at the late two-cell stage are highest in the proximity of these boundaries compared with the surrounding genomic regions and in comparison with either zygotes or early two-cell stage embryos (Fig. 3G). These analyses suggest that higher transcriptional activity characterizes LAD boundaries at the late two-cell stage. Consistent with a potential role for ZGA in LAD fragmentation, the number of LADs was lower in DRB- and α-amanitin-treated embryos compared with untreated late two-cell stage embryos, while the median LAD length was higher (2 and 1.6 Mb vs. 0.9 Mb, respectively) (Fig. 3H). This resulted in an increased percentage of genome localized into LADs upon DRB and α-amanitin treatment (Fig. 3I). Thus, transcriptional inhibition during ZGA leads to the association of a larger portion of the embryonic genome with the nuclear lamina. Careful examination of LADs on chromosome plots (Fig. 3F) suggested that although the number of LADs after transcriptional inhibition was intermediate between the number of LADs in untreated early and late two-cell stage embryos (Fig. 3H), some changes in the positioning of the LAD boundaries appeared in embryos treated with DRB and α-amanitin. To further investigate this, we plotted the DamID scores of DRB- and α-amanitin-treated two-cell embryos over the early and late two-cell stage boundaries of control embryos. These analyses indicate that treatment with DRB and α-amanitin results in a complete remodeling of wild-type LAD boundaries (Fig. 3J,K).

Atypical features of lamina-associated chromatin emerge upon inhibition of ZGA

To further understand the role of transcription in regulating nuclear organization, we next investigated in more detail the impact of ZGA inhibition on LAD and iLAD rearrangement at the two-cell stage. For this, we first compared LADs and iLADs in α-amanitin-treated embryos with the LADs and iLADs in zygotes and late two-cell stage embryos. While α-amanitin treatment did not affect the repositioning of some genomic regions that typically move toward the nuclear lamina at the two-cell stage (iL-L-L) (Fig. 4A), 36% of zygotic iLADs relocated to the nuclear lamina upon α-amanitin treatment (iL-L-iL) (Fig. 4A). This indicates that α-amanitin treatment precludes the formation of a subset of LADs and iLADs that normally form at the two-cell stage. Interestingly, "de novo LADs" formed in two-cell embryos treated with α-amanitin are enriched in major ZGA genes (iL-L-iL contain 939 out of 1462 major ZGA genes; odds ratio 1.71) (Fig 4B). We obtained similar results with DRB-treated embryos (Supplemental Fig. S1C,D). We found that major ZGA genes are enriched in proximity to zygotic iLAD boundaries (Fig. 4C; Supplemental Fig. S1E), suggesting that regions inside iLADs transform into LADs by transcriptional inhibition. Indeed, we observed that zygotic iLADs, which are enriched for major ZGA genes, become LADs upon α-amanitin or DRB treatment (Fig. 4D; Supplemental Fig. S1E, F). Consequently, major ZGA genes became enriched inside of LAD boundaries in embryos treated with α-amanitin and with DRB, which was not the case in untreated late two-cell stage embryos (Fig. 4C). Thus, we conclude that major ZGA genes relocate to the nuclear lamina upon transcriptional inhibition.

The rearrangement of the genome with respect to the nuclear lamina upon transcriptional inhibition was accompanied by a change in the genomic features of LADs and iLADs. Namely, while LADs are typically characterized by higher AT content compared with iLADs (Meuleman et al. 2013), LADs in α-amanitin- and DRB-treated embryos have lower AT content than iLADs, and this effect is particularly visible close to the LAD boundaries (Fig. 4E). We next investigated the distribution of LINE-1 and SINE B2 elements since they are known to be enriched in LADs and iLADs, respectively, in differentiated cells (Meuleman et al. 2013; Lenain et al. 2017). We found that, in contrast to the controls, LADs become depleted in LINE-1 elements and enriched in SINE B2 elements upon transcriptional inhibition (Fig. 4F; Supplemental Fig. S1G). This was surprising, as it suggested that LINE-1 elements become repositioned toward the nuclear interior in spite of their transcriptional silencing (Supplemental Fig. S1H). We confirmed these observations by performing DNA-FISH for LINE-1, which showed a global visual redistribution of LINE-1 toward the nuclear interior in two-cell stage embryos upon transcriptional inhibition (Supplemental Fig. S1I), validating the DamID data. Thus, the relocalization of LINE-1 elements into iLADs occurs in the absence of transcription.

We also analyzed the levels of H3K4me3, which we previously showed was involved in LAD establishment in zygotes (Borsos et al. 2019). We found that levels of H3K4me3 anticorrelate with LaminB1-DamID methylation levels in control two-cell stage embryos (Fig. 4G), in agreement with our previous observations (Borsos et al. 2019). Because of the known association of H3K4me3 with transcriptional activation (Santos-Rosa et al. 2002), we next asked whether H3K4me3 levels are affected upon inhibition of ZGA and whether those potential alterations relate to the LAD rearrangements that we observed upon transcriptional inhibition. For this, we first reanalyzed publicly available data sets of H3K4me3 from late two-cell stage embryos incubated with α-amanitin (Zhang et al. 2016). Remarkably, visual inspection of H3K4me3 levels across chromosome tracks suggested that α-amanitin treatment led to a redistribution of H3K4me3-marked regions, which highly corresponded to LaminB1-DamID methylation levels (Fig. 4H). Indeed, genome-wide analysis of H3K4me3 enrichment across all LADs and iLADs indicates that while H3K4me3 levels are higher in iLADs in control two-cell stage embryos (Fig. 4I), treatment with α-amanitin and DRB led to a complete reversion of this pattern, with H3K4me3 accumulation at LADs and depletion in iLADs (Fig. 4H,J). These observations suggest that inhibition of ZGA leads to a distribution of the genomic regions that contain H3K4me3 toward the nuclear lamina. We further confirmed this by performing immunostaining for H3K4me3 in late two-cell stage embryos

113
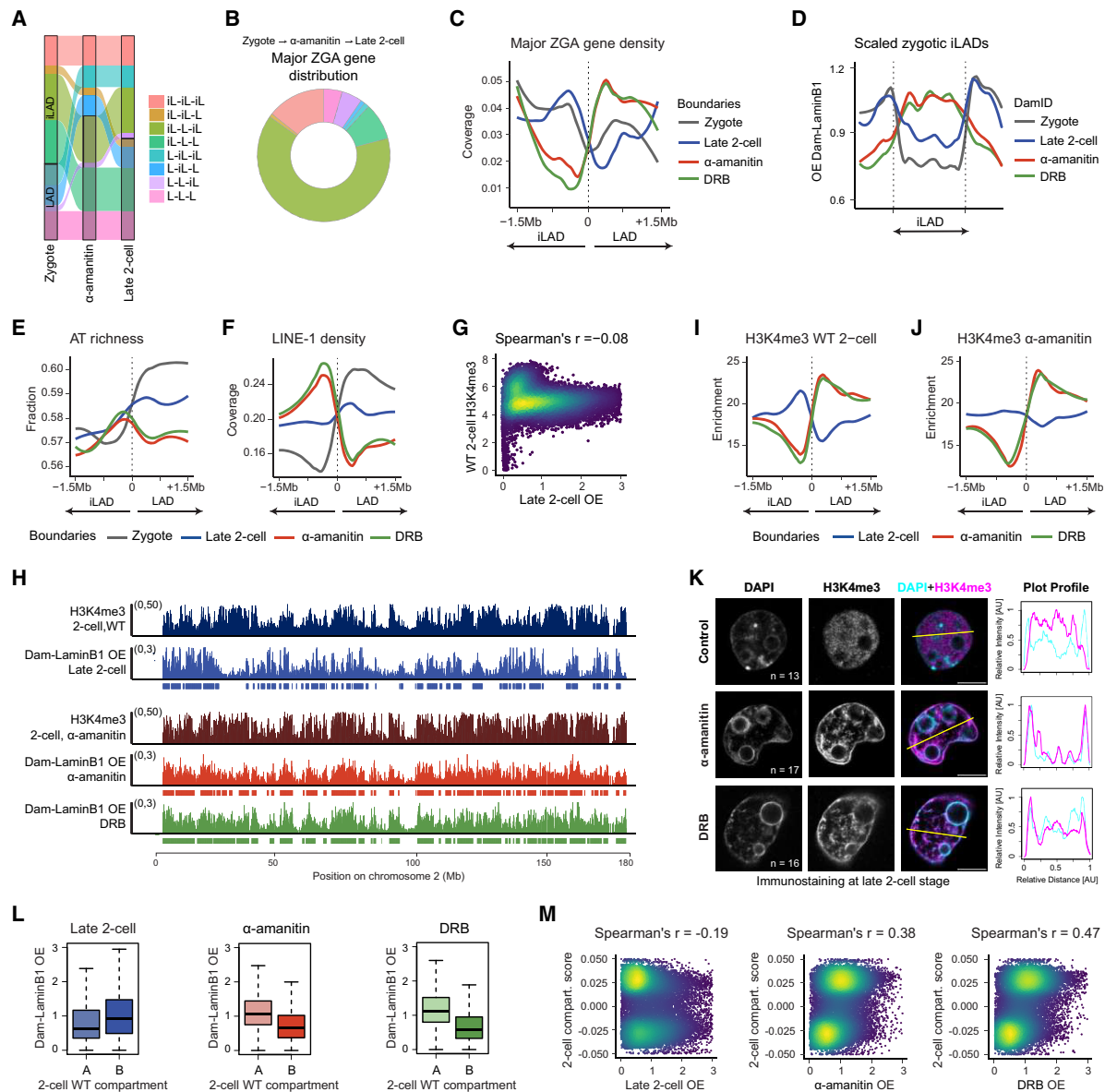
Figure 4. Atypical features of lamina-associated chromatin emerge upon inhibition of ZGA. (*A,B*) Alluvial plot and major ZGA gene distribution in reorganizing genomic regions upon transcriptional inhibition with α-amanitin. (*C*) Metaplot of major ZGA gene density over LAD boundaries. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot, and the 1.5-Mb region at the *right* indicates LAD. (*D*) Average LaminB1-DamID signal over scaled zygotic inter-LADs. (*E,F*) Average AT content (*E*) and LINE-1 element density (*F*) at LAD boundaries. (*G*) Genome-wide scatter plot of mean DamID OE values and $\log_2$ transformed H3K4me3 enrichment in late two-cell stage embryos. H3K4me3 ChIP-seq data from control (WT) embryos was analyzed from GSE71434 (Zhang et al. 2016). Spearman's correlation coefficient is indicated. (*H*) DamID OE value signal and H3K4me3 enrichment visualized on chromosome 2. H3K4me3 ChIP-seq data sets from control (WT) and α-amanitin-treated late two-cell embryos were analyzed from GSE71434 (Zhang et al. 2016). (*I,J*) Metaplots showing average H3K4me3 enrichment on LAD boundaries of untreated late two-cell stage or DRB- and α-amanitin-treated two-cell stage embryos. Please note that in *I*, H3K4me3 data were derived from untreated (WT) late two-cell stage embryos, and in *J*, H3K4me3 data were derived from α-amanitin-treated embryos, but in both panels, the LAD boundary coordinates are from the same samples (late two-cell control [WT] or α-amanitin- or DRB-treated embryos, as indicated by the colored lines). (*K*) Representative single confocal sections from immunostaining of H3K4me3 in control and α-amanitin- and DRB-treated late two-cell embryos (48 h post-hCG). DAPI stains for DNA. *n* = total number of embryos analyzed across three independent experiments. The intensity profiles for the lines shown on the merged images are plotted at the *right*. Scale bars, 10 μm. (*L*) Box plots of Dam-LaminB1 OE values in control (WT) A and B compartment regions. Compartment coordinates were taken from GSE82185 (Du et al. 2017). Box plots show median and the interquartile range (IQR), and whiskers depict the smallest and largest values within 1.5× IQR. (*M*) Genome-wide scatter plot of DamID OE values and compartment score in two-cell stage embryos derived from Hi-C data. The positive compartment (compart.) scores correspond to the A compartment. Spearman's correlation coefficients are indicated.

following DRB and α-amanitin treatment (Fig. 4K). These experiments revealed a drastic change in the localization of H3K4me3: While H3K4me3 is distributed throughout the nucleoplasm in control embryos, it becomes strongly enriched in the nuclear periphery, forming a clear rim around the nucleus in embryos treated with DRB and α-amanitin (Fig. 4K; Supplemental Fig. S2A). We also observed an apparent enrichment of H3K4me3 around the nucleolus precursors (NLBs) after transcriptional inhibition (Fig. 4K; Supplemental Fig. S2A), reflecting the known overlap between LADs and nucleolus-associated domains (NADs) (Bizhanova et al. 2020; Bersaglieri et al. 2022). Analysis of additional histone modifications by immunostaining indicated that other marks typically linked with active transcription, such as H3K9ac, displayed behavior similar to that of H3K4me3 and became visually enriched at the nuclear periphery (Supplemental Fig. S2B). However, a classical repressive histone modification, H3K9me3, did not show this behavior (Supplemental Fig. S2C). Overall, we conclude that inhibition of RNA polymerase II activity in early embryos leads to the rearrangement of H3K4me3-enriched regions and, perhaps more globally, of regions marked by active histone modifications toward the nuclear periphery and the remodeling of LAD boundaries. The mechanism behind this remodeling and whether this may reflect chromatin condensation changes upon global transcriptional inhibition remain to be established. Thus, inhibition of ZGA leads to a change in the genomic and chromatin features of LADs.

Previous work has shown that remodeling of H3K4me3 broad domains to a more canonical pattern after fertilization requires transcriptional activation at ZGA (Zhang et al. 2016). In addition, H3K4me3 broad domains have been postulated to be inhibitory for transcription (Dahl et al. 2016). Thus, we hypothesize that the impaired remodeling of H3K4me3 upon transcriptional inhibition of ZGA results in the sequestration of these domains to the nuclear lamina, in keeping with their transcriptional silent state. While TAD borders are remodeled by transcriptional activity in cultured cells and in *Drosophila* embryos (Li et al. 2015; Rowley et al. 2017), in mouse embryos transcription does not appear to be necessary to consolidate TAD borders and compartments in preimplantation embryos (Du et al. 2017; Ke et al. 2017). However, we found that inhibiting ZGA leads to a drastic remodeling of genome–lamina interactions and of LAD boundaries. To further investigate the relationship between LADs and compartments upon transcriptional inhibition, we analyzed DamID values in A and B compartments in control embryos at the two-cell stage. DamID-LaminB1 values are higher in the B compartments at the two-cell stage in control embryos, as expected (Fig. 4L). However, this pattern is reversed upon α-amanitin and DRB treatment, primarily due to an increase of DamID values within the A compartments but also due to a reduction in lamina interactions of the B compartment regions (Fig. 4L). Indeed, the A compartment regions move toward intermediate Dam-LaminB1 values genome-wide, resulting in a global positive correlation between compartment score and OE values in α-amanitin- and DRB-treated embryos (Fig. 4M). Thus, the A compartment regions gain lamina interactions upon transcriptional inhibition, and overall, compartments display an altered pattern of genome–nuclear lamina interactions upon transcriptional inhibition.

Discussion

Altogether, our data provide novel temporal resolution to the rearrangement of LADs during the maternal-to-zygotic transition and demonstrate that inhibition of transcription during ZGA leads to major changes in LAD organization. Whether a complete transcriptional inhibition in cells in culture also affects LADs has not been investigated and will be interesting to address in the future. It is intriguing that before the major wave of ZGA, in zygotes, LADs show genomic features that are more similar to constitutive LADs across cultured cell types, including LINE-1 enrichment at the nuclear periphery (Fig. 4F; Supplemental Fig. S2D). However, transcriptional inhibition during ZGA leads to unusual features of LADs. LINE-1-enriched genomic regions relocalize toward the nuclear interior, and H3K4me3-marked chromatin starts associating with the nuclear lamina in late two-cell embryos, which is strikingly different from cultured cells or premajor ZGA wild-type zygotes (Supplemental Fig. S2D). This suggests that during maternal-to-zygotic transition, the naturally evolving molecular characteristics of embryonic nuclear organization—in this case of LADs—are dependent on ZGA. From a broader perspective, this implies that the nuclear rearrangement of LADs is an integral component of MZT. The remodeling of nuclear organization after fertilization is considered to be a major event of epigenetic reprogramming occurring at these stages and is not restricted to mice but occurs in other mammals and vertebrates (Pecori and Torres-Padilla 2023). Our results indicate that transcription contributes to the remodeling of one of the pillars of nuclear organization; that is, LAD rearrangement. Interestingly, unlike TADs, LADs are globally unaffected upon inhibition of replication in both zygotes and two-cell stage embryos (Borsos et al. 2019). In contrast, transcriptional inhibition does not affect TAD consolidation (Du et al. 2017; Ke et al. 2017), and thus the contribution of ZGA toward the different pillars of nuclear organization may differ, as well as their dependency toward the different DNA and chromatin-related processes. Future work will determine whether and how other chromatin processes affect nuclear organization.

Our work sheds light onto the molecular mechanisms that occur during fundamental developmental process and how they shape the epigenomic landscape in early mammalian embryogenesis.

maintained and bred in accordance with institutional guidelines. To obtain embryos, 5- to 8-wk-old F1 (C57BL/6J × CBA/H) female mice were mated with DBA/2J males. To induce ovulation, females were injected with 10 IU of pregnant mare serum gonadotropin (PMSG; Ceva) and then 46–48 h later with human chorionic gonadotrophin (hCG; MSD Animal Health). Collected embryos were cultured in KSOM drops under paraffin oil (Sigma) at 37°C with 5% CO_2 as previously described. For DamID in early two-cell embryos, zygotes (18 h post-hCG) were isolated and injected with 250 ng/µL *Tir1*, 50 ng/µL membrane-eGFP, and 10 ng/µL AID-Dam-LaminB1 and cultured in medium containing 500 µM auxin. Auxin was removed just after mitosis for 4–6 h, and early two-cell embryos were collected at 34–36 h post-hCG. For DamID in transcription-inhibited late two-cell embryos, zygotes (18 h post-hCG) were injected with 250 ng/µL *Tir1*, 50 ng/µL membrane-eGFP, and 10 ng/µL AID-Dam-LaminB1-coding mRNA and cultured in KSOM containing 500 µM auxin and either 0.1 mg/mL α-amanitin (BioChemica) or 100 µM DRB (Santa Cruz Biotechnology). To allow methylation of LADs in the late two-cell stage, auxin was washed out from 42 to 48 h post-hCG, and embryos were cultured in KSOM containing either 0.1 mg/mL α-amanitin or 100 µM DRB. For immunofluorescence, zygotes (18 h post-hCG) were isolated and cultured in KSOM containing 0.1 mg/mL α-amanitin or 100 µM DRB until 48 h post-hCG.

DamID sample processing and library preparation

The zona pellucida was removed by treatment with 0.5% pronase in M2 at 37°C. Polar bodies were separated from the embryos by gentle pipetting after trypsin treatment and discarded. For each replicate, a pool of 10–20 blastomeres (five to 10 two-cell embryos) was collected in 2 µL of DamID buffer (10 mM TRIS acetate at pH 7.5, 10 mM magnesium acetate, 50 mM potassium acetate) and stored at −80°C until downstream processing. All experiments were performed in three independent biological replicates. Sample processing and library preparation were done as described previously (Borsos et al. 2019; Pal et al. 2021).

DamID sequencing and analysis

Samples were sequenced using Illumina HiSeq4000 or HiSeq2500 platforms in 150-bp PE mode, but only read1 was used for downstream analysis. For preprocessing of reads, the first six random bases were discarded using trimmomatic (version 0.39). Subsequently, the reads were demultiplexed according to DamID indexes using a Fastx barcode splitter, and the additional 15 bp of adaptors was trimmed using trimmomatic. The preprocessed reads starting with GATC were then mapped to mm10 using bowtie2 (version 2.3.5) with default parameters. Reads aligning to the genome with a quality score <30 were discarded using samtools (version 1.3). Duplicates were removed using picard (version 2.21.1) to finally obtain unique GATC reads. The computation of OE (observed/expected) values per 100-kb bin was carried out as described previously (Kind et al. 2015). LaminB1-DamID data from zygotes and late two-cell stage embryos were obtained from our previous study (GSE112551; Borsos et al. 2019). For data visualization and LAD calling, the OE mean signal of all three replicates was used. To distinguish LADs from inter-LADs, a two-state hidden Markov model (HMM) (Filion et al. 2010) was applied to nonzero OE mean values.

Immunofluorescence

Embryos were fixed in 4% PFA for 20 min at room temperature and permeabilized in PBS containing 0.5% Triton-X for 20 min.

Embryos were kept in blocking buffer (3% BSA in PBS) for 4–5 h and then incubated overnight in primary antibody (H3K4me3, 1:250 [Diagenode C15410003]; H3K9ac, 1:250 [Abcam ab4441]; and H3K9me3, 1:100 [Active Motif 39286]) diluted in blocking buffer. After overnight incubation, samples were washed three times in PBS and stained with secondary antibodies conjugated with Alexa fluor 555 or Alexa fluor 647 in blocking buffer for 2–3 h. After three washes in PBS, embryos were mounted in 3D using VectaShield (Vector Laboratories) containing DAPI. Confocal imaging was performed using a 63× oil objective in a Leica SP8 confocal microscope.

Whole-mount DNA-FISH

LINE-1 DNA-FISH was performed as previously described (Jachowicz et al. 2017). LINE-1 probes (L1spa) were labeled with homemade TAMRA-dATP with a nick translation kit (Roche) and purified with a QIAquick PCR purification kit (Qiagen). Confocal imaging was performed using a 63× oil objective in a Leica SP8 confocal microscope.

Image analysis

Image analysis was performed with ImageJ software's plot profile function. Example lines for obtaining intensity profiles were drawn manually, avoiding NLBs. The results were then processed and plotted using R (version 4.1.2). Intensity values were smoothed using the rollmean function with $k = 21$ from the zoo package (version 1.8–12). Smoothed values were minimum/maximum-scaled such that the final values ranged between 0 and 1.

RNA-seq analysis

The RNA-seq data set for wild-type zygotes and early and late two-cell stage embryos was obtained from GEO with accession number GSE45719 (Deng et al. 2014), processed expression values (RPKM) were downloaded, and the mean RPKM for each developmental stage was calculated. DBTMEE maternal RNA genes (Park et al. 2015) were excluded from RPKM and log_2 fold change analysis of RNA-seq data. For metaplot analysis of RNA-seq data, reads were aligned to the GRCm38 reference genome using STAR (version 2.7.6a), and mapped reads were counted in 100-kb genomic bins using the GenomicAlignments (version 1.30.0) and GenomicRanges (version 1.46.1) R packages. Read counts were normalized to the total number of reads and multiplied by 1 million. Metaplots at LAD boundaries were generated on the log_2 normalized counts using custom R scripts. RNA-seq data for α-amanitin and control two-cell stage embryos were obtained from GEO with accession number GSE72784 (Dahl et al. 2016).

Analysis of transposable elements and ZGA genes

TE annotation for the mm10 genome was obtained from the Hammell laboratory repository (https://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/mm10_rmsk_TE.gtf.gz). A list of minor and major ZGA genes was considered according to DBTMEE (Park et al. 2015) classification. TE and ZGA gene density was calculated using the bedtools (version 2.31.0) coverage function in 100-kb genomic bins (same resolution as DamID). Metaplots on LAD boundaries were generated using deepTools. For enrichment analysis of TEs and ZGA genes in reorganizing genomic bins, the average density of all genomic bins was used as the expected value. TE expression analysis from RNA-seq data was performed using TEtranscripts (version 2.2.3, https://github.com/mhammell-laboratory/TEtranscripts). Briefly, reads were aligned

to the GRCm38 reference genome using STAR (version 2.7.6a) with parameters --outFilterMultimapNmax 100 and --winAnchorMultimapNmax 100. Reads were counted at genes and TEs using TEcount with parameters --mode multi and --stranded no. Read counts were normalized by a normalization factor that was the total sum of the reads per sample divided by the mean total sum of all samples. After $\log_2$ transformation, the median of all LINE-1 family elements was taken for each sample and visualized as a dot plot.

Hi-C data analysis

Hi-C compartment coordinates and scores were obtained from GEO with accession number GSE82185 (Du et al. 2017) and analyzed as described in Borsos et al. (2019).

Analysis of H3K4me3 data sets

H3K4me3 ChIP-seq data sets were downloaded from GEO with accession number GSE71434 (Zhang et al. 2016). After trimming, reads were aligned to the GRCm38 reference genome using bowtie2 (version 2.3.5). Reads were filtered by mapping quality score using samtools (version 1.3) with parameter -q 12. Read pairs were read into R using the readGAlignmentPairs function from the GenomicAlignment package (version 1.30.0) and were filtered for unique fragments. Fragments were counted in 50-kb consecutive genomic bins, normalized by the sum of the fragments counts, and multiplied by 1 million. Metaplots on LAD boundaries were generated using deepTools.

Data availability

DamID data sets generated in this study have been deposited in GEO under accession number GSE241483.

References

Abe K, Funaya S, Tsukioka D, Kawamura M, Suzuki Y, Suzuki MG, Schultz RM, Aoki F. 2018. Minor zygotic gene activation is essential for mouse preimplantation development. *Proc Natl Acad Sci* **115:** E6780–E6788. doi:10.1073/pnas.1805239115

Bensaude O. 2011. Inhibiting eukaryotic transcription. which compound to choose? How to evaluate its activity? *Transcription* **2:** 103–108. doi:10.4161/trns.2.3.16172

Bersaglieri C, Kresoja-Rakic J, Gupta S, Bär D, Kuzyakiv R, Panatta M, Santoro R. 2022. Genome-wide maps of nucleolus interactions reveal distinct layers of repressive chromatin domains. *Nat Commun* **13:** 1483. doi:10.1038/s41467-022-29146-2

Bizhanova A, Yan A, Yu J, Zhu LJ, Kaufman PD. 2020. Distinct features of nucleolus-associated domains in mouse embryonic stem cells. *Chromosoma* **129:** 121–139. doi:10.1007/s00412-020-00734-9

Borsos M, Perricone SM, Schauer T, Pontabry J, de Luca KL, de Vries SS, Ruiz-Morales ER, Torres-Padilla M-E, Kind J. 2019. Genome–lamina interactions are established de novo in the early mouse embryo. *Nature* **569:** 729–733. doi:10.1038/s41586-019-1233-0

Burton A, Torres-Padilla M-E. 2014. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat Rev Mol Cell Biol* **15:** 723–735. doi:10.1038/nrm3885

Burton A, Brochard V, Galan C, Ruiz-Morales ER, Rovira Q, Rodriguez-Terrones D, Kruse K, Le Gras S, Udayakumar VS, Chin HG, et al. 2020. Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat Cell Biol* **22:** 767–778. doi:10.1038/s41556-020-0536-6

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY, et al. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537:** 548–552. doi:10.1038/nature19360

Deng Q, Ramsköld D, Reinius B, Sandberg R. 2014. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343:** 193–196. doi:10.1126/science.1245316

Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, et al. 2017. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547:** 232–235. doi:10.1038/nature23263

Dubois M-F, Bellier S, Seo S-J, Bensaude O. 1994. Phosphorylation of the RNA polymerase II largest subunit during heat shock and inhibition of transcription in HeLa cells. *J Cell Physiol* **158:** 417–426. doi:10.1002/jcp.1041580305

Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P, Torres-Padilla M-E. 2013. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol* **20:** 332–338. doi:10.1038/nsmb.2495

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143:** 212–224. doi:10.1016/j.cell.2010.09.009

Flyamer IM, Gassler J, Imakaev M, Brandão HB, Ulianov SV, Abdennur N, Razin SV, Mirny LA, Tachibana-Konwalski K. 2017. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544:** 110–114. doi:10.1038/nature21711

Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. 2017. Chromatin architecture emerges during zygotic genome activation independent of transcription. *Cell* **169:** 216–228.e19. doi:10.1016/j.cell.2017.03.024

Ishiuchi T, Enriquez-Gasca R, Mizutani E, Bošković A, Ziegler-Birling C, Rodriguez-Terrones D, Wakayama T, Vaquerizas JM, Torres-Padilla M-E. 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin

assembly. *Nat Struct Mol Biol* **22:** 662–671. doi:10.1038/nsmb
.3066

Jachowicz JW, Bing X, Pontabry J, Bošković A, Rando OJ, Torres-Padilla M-E. 2017. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat Genet* **49:** 1502–1510. doi:10.1038/ng.3945

Jukam D, Shariati SAM, Skotheim JM. 2017. Zygotic genome activation in vertebrates. *Dev Cell* **42:** 316–332. doi:10.1016/j.devcel.2017.07.026

Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, et al. 2017. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170:** 367–381.e20. doi:10.1016/j.cell.2017.06.029

Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B. 2013. Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153:** 178–192. doi:10.1016/j.cell.2013.02.028

Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, de Graaf CA, Amendola M, et al. 2015. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell* **163:** 134–147. doi:10.1016/j.cell.2015.08.040

Kruse K, Díaz N, Enriquez-Gasca R, Gaume X, Torres-Padilla M-E, Vaquerizas JM. 2019. Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. bioRxiv doi:10.1101/523712

Lenain C, de Graaf CA, Pagie L, Visser NL, de Haas M, de Vries SS, Peric-Hupkes D, van Steensel B, Peeper DS. 2017. Massive reshaping of genome-nuclear lamina interactions during oncogene-induced senescence. *Genome Res* **27:** 1634–1644. doi:10.1101/gr.225763.117

Li L, Lu X, Dean J. 2013. The maternal to zygotic transition in mammals. *Mol Aspects Med* **34:** 919–938. doi:10.1016/j.mam.2013.01.003

Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong C-T, Cubeñas-Potts C, Hu M, Lei EP, Bosco G, et al. 2015. Widespread rearrangement of 3D chromatin organization underlies Polycomb-mediated stress-induced silencing. *Mol Cell* **58:** 216–231. doi:10.1016/j.molcel.2015.02.023

Liu X, Wang C, Liu W, Li J, Li C, Kou X, Chen J, Zhao Y, Gao H, Wang H, et al. 2016. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537:** 558–562. doi:10.1038/nature19362

Liu B, Xu Q, Wang Q, Feng S, Lai F, Wang P, Zheng F, Xiang Y, Wu J, Nie J, et al. 2020. The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature* **587:** 139–144. doi:10.1038/s41586-020-2847-y

Meuleman W, Peric-Hupkes D, Kind J, Beaudry J-B, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B. 2013. Constitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **23:** 270–280. doi:10.1101/gr.141028.112

Nguyen VT, Giannoni F, Dubois M-F, Seo S-J, Vigneron M, Kédinger C, Bensaude O. 1996. In vivo degradation of RNA polymerase II largest subunit triggered by α-amanitin. *Nucleic Acids Res* **24:** 2924–2929. doi:10.1093/nar/24.15.2924

Pal M, Kind J, Torres-Padilla M-E. 2021. DamID to map genome–protein interactions in preimplantation mouse embryos *Methods Mol Biol* **2214:** 265–282. doi:10.1007/978-1-0716-0958-3_18

Park S-J, Shirahige K, Ohsugi M, Nakai K. 2015. DBTMEE: a database of transcriptome in mouse early embryos. *Nucleic Acids Res* **43:** D771–D776. doi:10.1093/nar/gku1001

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. doi:10.1016/j.devcel.2004.09.004

Pecori F, Torres-Padilla M-E. 2023. Dynamics of nuclear architecture during early embryonic development and lessons from liveimaging. *Dev Cell* **58:** 435–449. doi:10.1016/j.devcel.2023.02.018

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SWM, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, et al. 2010. Molecular maps of the reorganization of genome–nuclear lamina interactions during differentiation. *Mol Cell* **38:** 603–613. doi:10.1016/j.molcel.2010.03.016

Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, Wang P, Ruan Y, Corces VG. 2017. Evolutionarily conserved principles predict 3D chromatin organization. *Mol Cell* **67:** 837–852.e7. doi:10.1016/j.molcel.2017.07.022

Sakashita A, Kitano T, Ishizu H, Guo Y, Masuda H, Ariura M, Murano K, Siomi H. 2023. Transcription of MERVL retrotransposons is required for preimplantation embryo development. *Nat Genet* **55:** 484–495. doi:10.1038/s41588-023-01324-y

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NCT, Schreiber SL, Mellor J, Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407–411. doi:10.1038/nature01080

Schultz RM. 2002. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update* **8:** 323–331. doi:10.1093/humupd/8.4.323

Schulz KN, Harrison MM. 2019. Mechanisms regulating zygotic genome activation. *Nat Rev Genet* **20:** 221–234. doi:10.1038/s41576-018-0087-x

van Schaik T, Vos M, Peric-Hupkes D, Celie PHN, van Steensel B. 2020. Cell cycle dynamics of lamina-associated DNA. *EMBO Rep* **21:** e50636. doi:10.15252/embr.202050636

van Steensel B, Belmont AS. 2017. Lamina-associated domains: links with chromosome architecture, heterochromatin and gene repression. *Cell* **169:** 780–791. doi:10.1016/j.cell.2017.04.022

Vastenhouw NL, Cao WX, Lipshitz HD. 2019. The maternal-to-zygotic transition revisited. *Development* **146:** dev161471. doi:10.1242/dev.161471

Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, Chen C, Kou X, Zhao Y, Chen J, et al. 2018. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat Cell Biol* **20:** 620–631. doi:10.1038/s41556-018-0093-4

Xia W, Xie W. 2020. Rebooting the epigenomes during mammalian early embryogenesis. *Stem Cell Rep* **15:** 1158–1175. doi:10.1016/j.stemcr.2020.09.005

Zeng F, Schultz RM. 2005. RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. *Dev Biol* **283:** 40–57. doi:10.1016/j.ydbio.2005.03.038

Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q, et al. 2016. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537:** 553–557. doi:10.1038/nature19361

118

Supplemental Figures and Legends

Pal et al., Reorganization of Lamina Associated Domains in early mouse embryos is regulated by RNA Polymerase II activity

(Contains Supplemental Figures S1 and S2)

A LINE-1 expression

B

C

D Zygote → DRB → Late 2-cell
Major ZGA gene distribution

E Major ZGA gene density

F

G SINE B2 density

H LINE-1 expression

I

LINE-1 DNA FISH at late 2-cell stage
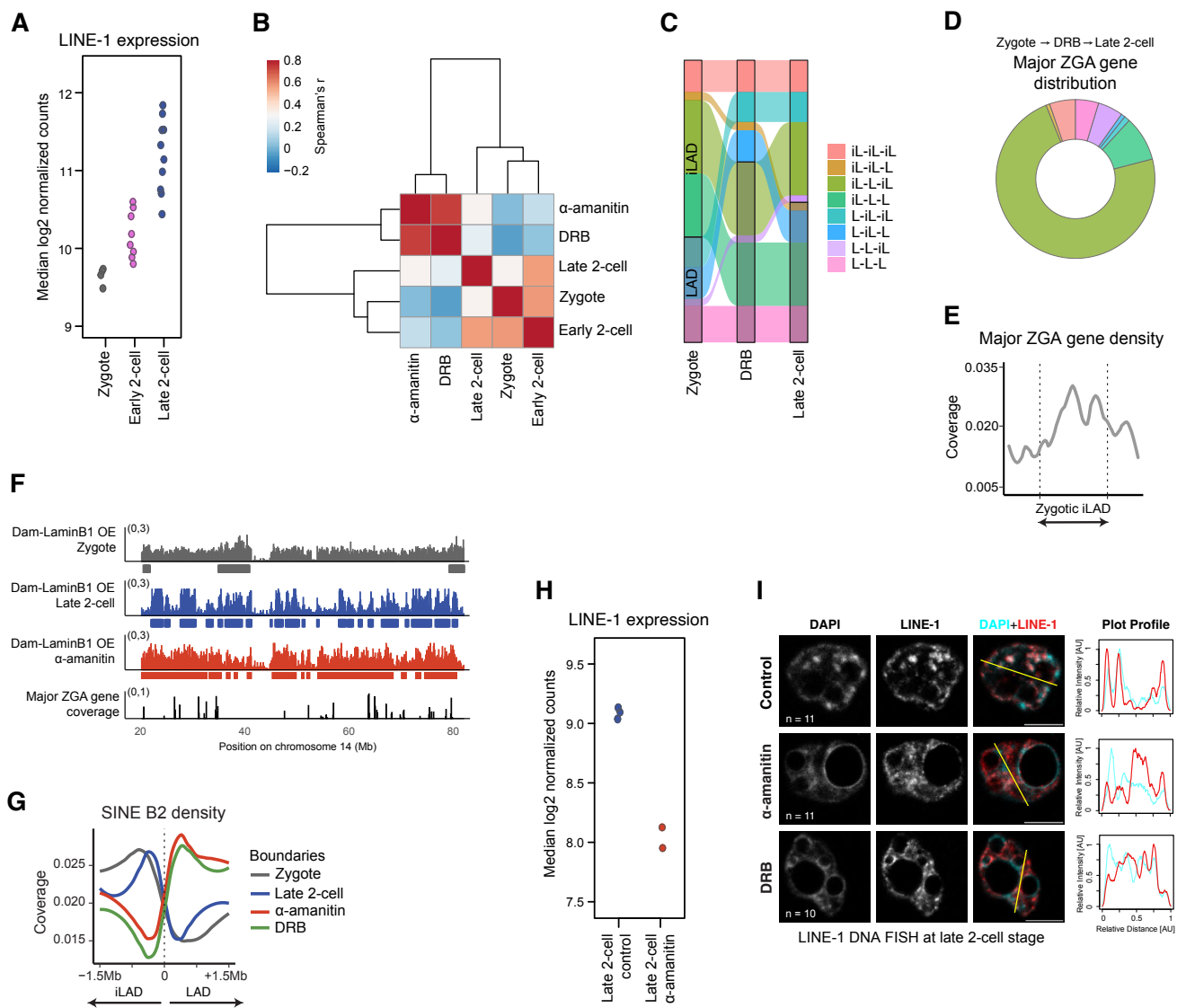
Figure S1. Analysis of DRB and α-amanitin treated samples and changes in LINE-1 expression and localisation.

(*A*) LINE-1 element expression in early mouse embryos. RNA-seq data analyzed from GSE45719 (Deng et al. 2014).

(*B*) Correlation and hierarchical clustering of genome-wide mean Dam-LaminB1 OE values using Spearman's R.

(*C,D*) Alluvial plot and major ZGA gene distribution in reorganizing genomic regions when transcriptional elongation is inhibited with DRB treatment.

(*E*) Average major ZGA gene density on scaled zygotic inter-LADs.

(*F*) DamID OE value signal and major ZGA gene coverage visualized on part of chromosome 14.

(*G*) Average SINE B2 element density at LAD boundaries.

(*H*) LINE-1 element expression in control and α-amanitin treated 2-cell mouse embryos. RNA-seq data analyzed from GSE72784 (Dahl et al. 2016).

(*I*) Representative single confocal sections from LINE-1 DNA FISH in control, α-amanitin and DRB treated late 2-cell embryos (48 hphCG). DAPI stains for DNA. n = total number of embryos analysed across two independent experiments. The intensity profiles for the lines shown on the merged images are plotted at the right. Scale bars, 10 μm.

A — DAPI | H3K4me3 | DAPI+H3K4me3 | Plot Profile
Control, α-amanitin, DRB
Immunostaining at late 2-cell stage

B — DAPI | H3K9ac | DAPI+H3K9ac | Plot Profile
Control (n = 9), α-amanitin (n = 10), DRB (n = 10)
Immunostaining at late 2-cell stage

C — DAPI | H3K9me3 | DAPI+H3K9me3 | Plot Profile
Control (n = 12), α-amanitin (n = 13), DRB (n = 11)
Immunostaining at late 2-cell stage

D — Immunostaining: DAPI | H3K4me3 | DAPI+H3K4me3 (n = 15)
DNA FISH: DAPI | LINE-1 | DAPI+LINE-1 (n = 17)
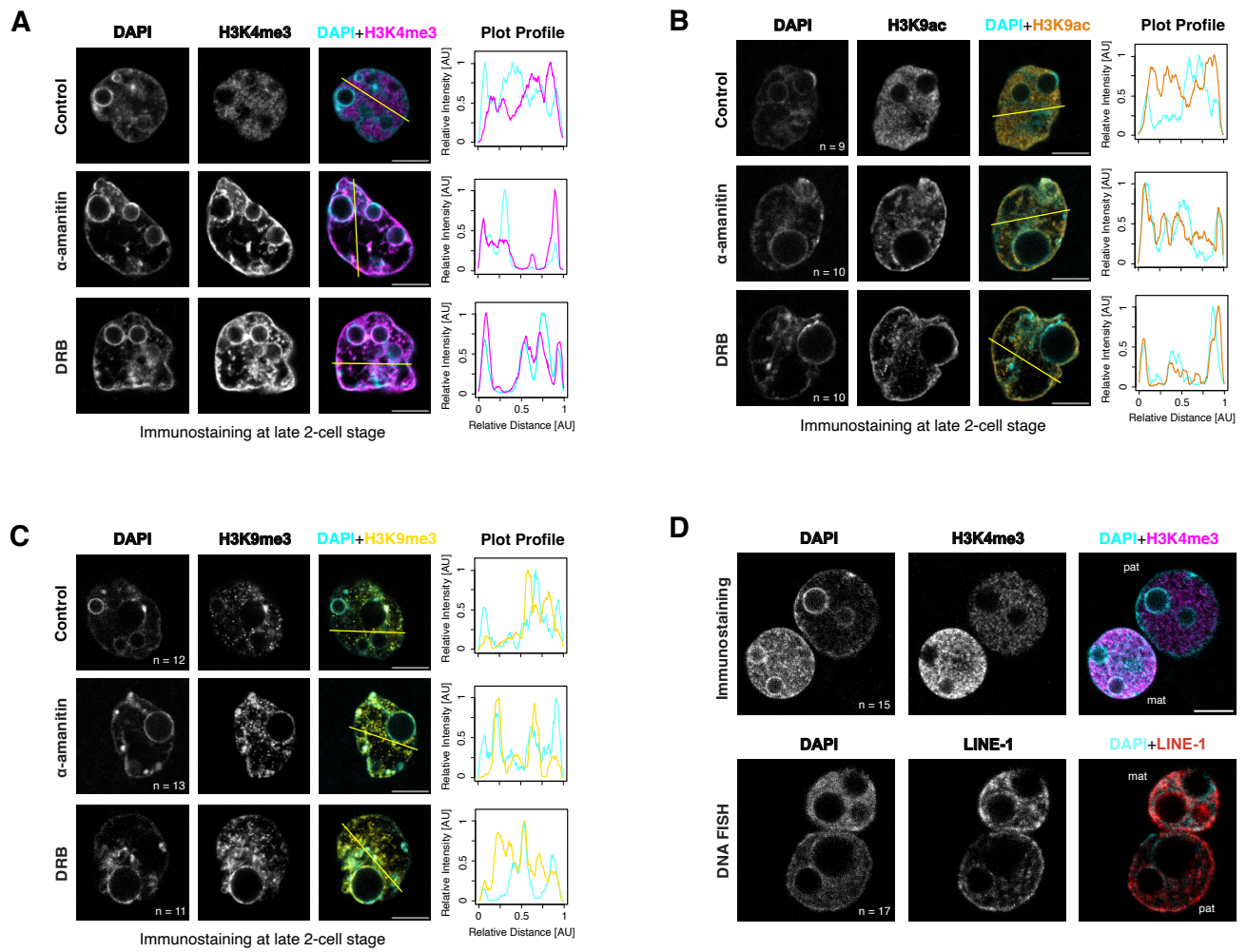
Figure S2. Analysis of histone modifications upon DRB and α-amanitin treatment in 2-cell stage embryos.

(*A*) Additional representative single confocal sections from immunostaining of H3K4me3 in control, α-amanitin and DRB treated late 2-cell embryos (48 hphCG). DAPI stains for DNA. The intensity profiles for the lines shown on the merged images are plotted at the right. Scale bars, 10 μm.

(*B*) Representative single confocal sections from immunostaining of H3K9ac in control, α-amanitin and DRB treated late 2-cell embryos (48 hphCG). n = total number of embryos analysed across two independent experiments. The intensity profiles for the lines shown on the merged images are plotted at the right. Scale bars, 10 μm.

(*C*) Representative single confocal sections from immunostaining of H3K9me3 in control, α-amanitin and DRB treated late 2-cell embryos (48 hphCG). n = total number of embryos analysed across two independent experiments. The intensity profiles for the lines shown on the merged images are plotted at the right. Scale bars, 10 μm.

(*D*) Representative single confocal sections from H3K4me3 immunostaining and LINE-1 DNA FISH in control late zygotes (28 hphCG). n = total number of embryos analyzed across two independent experiments. mat: maternal pronucleus. pat: paternal pronucleus. Scale bar, 10 μm.

Part IV

Publication

Emergence of replication timing during early mammalian development

Article

Emergence of replication timing during early mammalian development

Tsunetoshi Nakatani[1], Tamas Schauer[1,5], Luis Altamirano-Pacheco[1,5], Kyle N. Klein[2], Andreas Ettinger[1], Mrinmoy Pal[1], David M. Gilbert[3] & Maria-Elena Torres-Padilla[1,4 ✉]

DNA replication enables genetic inheritance across the kingdoms of life. Replication occurs with a defined temporal order known as the replication timing (RT) programme, leading to organization of the genome into early- or late-replicating regions. RT is cell-type specific, is tightly linked to the three-dimensional nuclear organization of the genome[1,2] and is considered an epigenetic fingerprint[3]. In spite of its importance in maintaining the epigenome[4], the developmental regulation of RT in mammals in vivo has not been explored. Here, using single-cell Repli-seq[5], we generated genome-wide RT maps of mouse embryos from the zygote to the blastocyst stage. Our data show that RT is initially not well defined but becomes defined progressively from the 4-cell stage, coinciding with strengthening of the A and B compartments. We show that transcription contributes to the precision of the RT programme and that the difference in RT between the A and B compartments depends on RNA polymerase II at zygotic genome activation. Our data indicate that the establishment of nuclear organization precedes the acquisition of defined RT features and primes the partitioning of the genome into early- and late-replicating domains. Our work sheds light on the establishment of the epigenome at the beginning of mammalian development and reveals the organizing principles of genome organization.

Replication timing (RT) is a fundamental epigenetic feature[6], yet how and when RT is established during mammalian development is unknown. During S phase the genome must replicate once and only once. Replication occurs through a coordinated programme whereby origins of replication fire in a temporally defined order, giving rise to replication patterns characteristic of each cell type[7,8]. Early- and late-replication domains correlate with accessible, actively transcribed euchromatin and silent heterochromatin, respectively[9]. RT is interconnected with other epigenetic features, although their temporal and functional dependency has not been fully established. For example, RT is tightly associated with three-dimensional genome organization, with lamina-associated domains (LADs) and B-type compartments typically corresponding to late-replication domains. Whereas mammalian cells do not possess strongly defined genetic sequences specifying replication origins, replication commences within initiation zones, which are regions of about 40 kb that comprise one or more sites of stochastic origin firing[10,11]. Generally, initiation zones of high efficiency tend to replicate early whereas low-efficiency initiation zones replicate late during S phase. Thus, RT is primarily driven by the probability of initiation within initiation zones. How initiation zones are specified at the beginning of development, and whether cells of the early embryo share a similar structure and features of the RT programme with differentiated cells, remain to be established.

Mammalian development begins with fertilization and is followed by an intense period of chromatin remodelling[12]. Major epigenome features are defined for the first time during this developmental time window: LADs are established de novo in mouse zygotes and the A and B compartments, although detectable in zygotes, gradually become more defined as development progresses towards the blastocyst[13]. Topological-associating domains (TADs) are barely detectable before the 8-cell stage and emerge only at late cleavage stages[14–16]. In mice, zygotic genome activation (ZGA) occurs during this time with minor ZGA occurring in zygotes and the major wave of ZGA in late-2-cell-stage embryos[17]. However, when RT programmes first emerge is unknown. In *Drosophila*, microscopy studies indicate that the onset of late replication emerges after ZGA[18] but our understanding of this process—and how and when RT is first established in mammals—is unknown.

RT emerges gradually during preimplantation development

To understand when and how RT emerges during development, we used single-cell Repli-seq[5,19] in preimplantation mouse embryos (Fig. 1a,b). We collected 529 individual cells of which 53, 54, 50, 49, 34, 44 and 55 passed quality control for zygotes, 2-cell, 4-cell, 8-cell, 16-cell, morula and blastocyst-stage inner cell mass (ICM), respectively (Extended Data Fig. 1a,b, Supplementary Table 1 and Methods). Plotting individual cells based on their replication score, which reflects the percentage of their replicated genome (Fig. 1c), showed a clear replication domain structure consistent with progression of replication, with typical
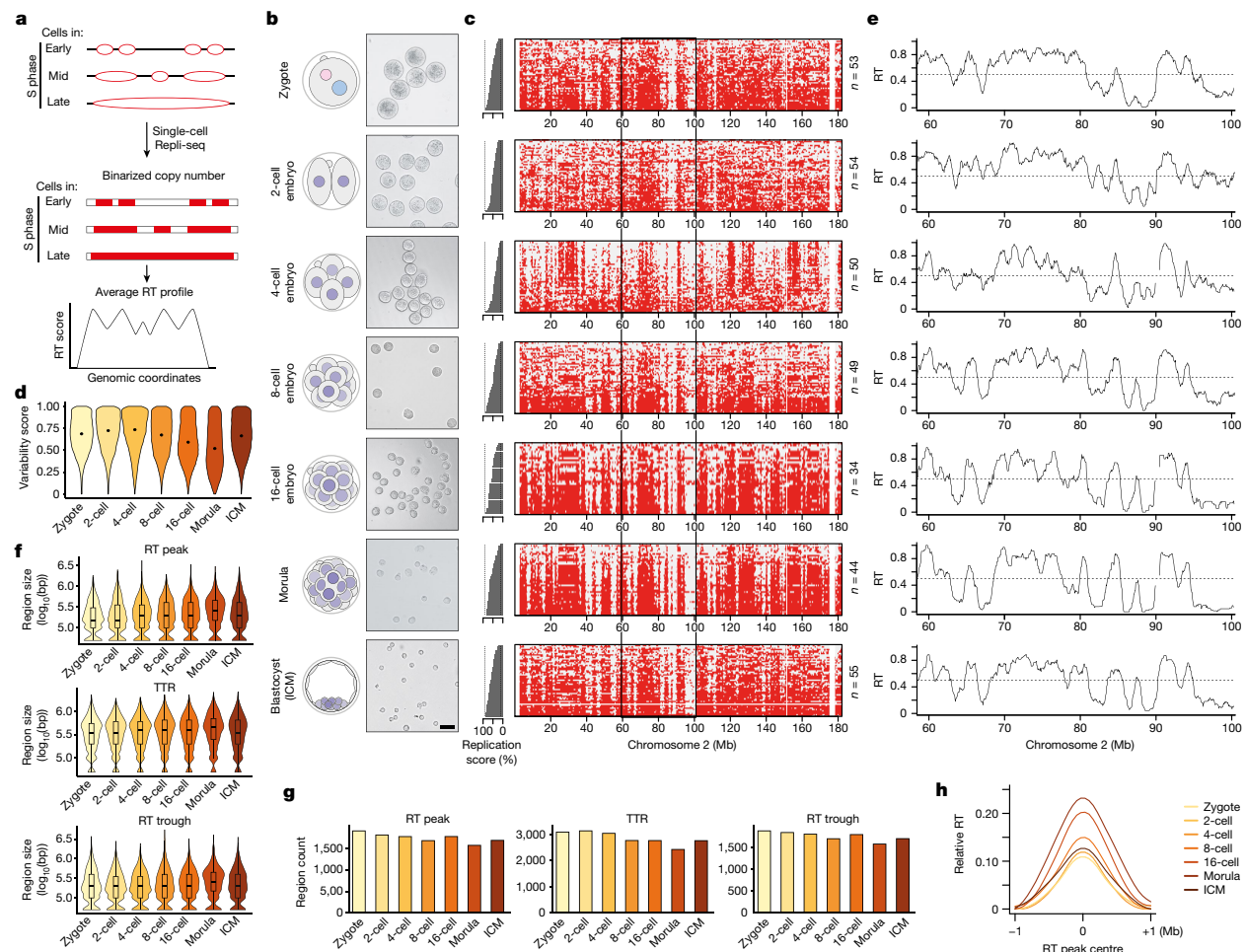
Fig. 1 | RT emerges gradually during mouse preimplantation development.
a, Overview of single-cell Repli-seq used to generate RT profiles from single cells in mouse preimplantation embryos based on copy number variation.
b, Schematic of sampling of embryos and corresponding images of dissociated blastomeres at each stage. The numbers of independent blastomere collections for each stage with similar results are as follows: zygote (3), 2-cell (4), 4-cell (3), 8-cell (3), 16-cell (3), morula (2), ICM (4). Scale bar, 50 μm. **c**, Heatmaps of single cells indicating replication status based on binarized copy number during preimplantation embryogenesis (red, replicated; grey, not replicated). Cells are ranked by their percentage of replicated genome (replication score), which indicates progress in S phase and is plotted as a bar plot on the left. **d**, Variability score during embryonic development; the score is 1 when 50% of cells replicated the genomic bin and 0 when all cells are either replicated (100%) or

non-replicated (0%). Each violin plot shows the distribution of scores for all genomic bins. **e**, RT profiles of preimplantation embryos over a representative region on chromosome 2, denoted by black rectangle in **c**. Black line indicates RT profiles, calculated as the average of overlapping intervals defined by genome-wide replication score. **f,g**, Size (**f**) and number (**g**) of replication features RT peaks (also known as initiation zones), and RT troughs (also known as termination zones) during preimplantation development. Box plots show median and interquartile range (IQR), and whiskers depict the lowest and highest values within 1.5× IQR. bp, base pair. **h**, Relative RT values centred at RT peaks during embryonic development compared with their neighbouring regions. Note that curves for the 2- and 4-cell stages overlap considerably and, to some extent, with that of zygotes.

early–late transitions across most stages (Fig. 1c and Extended Data Fig. 1c). Zygotes and 2-cell embryos were an exception and showed a less defined replication pattern across cells and throughout the genome, suggesting a more variable and less coordinated programme (Fig. 1c). This was due to neither absence of DNA synthesis nor embryonic heterogeneity in the progression of DNA synthesis, because we verified microscopically that zygotes showed an expected and consistent spatial pattern of DNA synthesis through S phase (Extended Data Fig. 1d,e). To provide a quantitative metric of the RT programme we computed a variability score, which measures the variance of the replication programme across cells. RT variability score was highest in zygotes and 2-cell and 4-cell embryos but decreased progressively from the 4-cell stage (Fig. 1d). RT of the ICM appeared more variable

compared with morula, which may reflect the ICM undergoing cell fate decisions towards epiblast and primitive endoderm[20], and thus greater heterogeneity in cell identity is likely to be present therein. Overall, the RT programme at the earliest stages of development is less well defined.

Embryonic RT profiles showed both early- and late-replication domains, visible as valleys and plateaus (Fig. 1e). Visual inspection showed a progressive delineation of replication domains as development proceeds (Fig. 1e and Extended Data Fig. 2a). This is independent of S-phase length because length is relatively constant until the blastocyst stage[21]. To address whether and how RT changes during development, we compared 'early' (RT ≥ 0.5) and 'late' (RT ≤ 0.5) RT values from the zygote to the blastocyst ICM. In general, RT values

increased towards earlier or later (Extended Data Fig. 2b; increase), indicating definition of the early and late RT programme during development. A portion of the genome showed constant early or late RT throughout (33.1% of the genome replicates early and 16.0% replicates late in all seven stages; Extended Data Fig. 2b; constant). However, some regions shift from early to late RT values and vice versa (Extended Data Fig. 2b; shuffle). For example, 20.9% of the genome switches from early to late RT from 2-cell to morula and 11.1% does so between 8-cell and 16-cell. Likewise, 3.1% changes from late to early RT between 8-cell and morula. This analysis also showed that, whereas some genomic regions do shift RT between early and late values, the most common trend is a progressive definition of RT values towards more early and more late (Extended Data Fig. 2b). Indeed, whereas most of the genome in zygotes and 2-cell embryos (73 and 77%, respectively) shows intermediate RT values ($0.4 \leq RT \leq 0.8$), the genome partitions into RT values spanning the complete S phase as development progresses, resulting in stratification into more extreme early and late RT values after the 2-cell stage (Extended Data Fig. 3a). This behaviour resembles A and B compartments[14], which undergo progressive increase in compartment strength during cleavage stages[14,16], suggesting that preimplantation serves as period of gradual establishment of three-dimensional nuclear architecture and RT. We conclude that, although approximately half of the genome preserves its RT, the remaining half undergoes changes in RT as development proceeds and becomes more defined over time.

Next we characterized embryonic RT features by extracting initiation zones, but also zones in which opposing replication forks convene (termination zones) and timing transition regions (TTRs), which are regions located between initiation zones and termination zones[5,7]. Because of the resolution of scRepli-seq. and to distinguish these features from those in methods such as OK-seq and EdU-seq[22,23], we refer to initiation zones as 'RT peaks' and to termination zones as 'RT troughs'. We defined RT peaks as consecutive bins of local maxima and RT troughs as consecutive bins of local minima of RT values (Extended Data Fig. 3b)[10]. Globally, RT peaks increase in size ($P = 0.01$) with more, smaller RT peaks at early cleavage stages compared with later stages (Fig. 1f,g). Similarly, albeit to a lesser extent, TTRs increase in size ($P = 0.01$; Fig. 1f). The size of RT troughs remains overall stable ($P = 0.19$; Fig. 1f) and, similar to embryonic stem (ES) cells; RT troughs have higher AT content than RT peaks and TTRs (Extended Data Fig. 3c). RT peaks can reshuffle into TTRs and TTRs into RT peaks during each cell division (Extended Data Fig. 3d). Similarly, RT troughs converted into TTRs and TTRs into RT troughs but changes from RT peaks into RT troughs and vice versa are extremely rare (Extended Data Fig. 3d). Approximately half of RT peaks and RT troughs changed into TTRs at the subsequent developmental stage, suggesting remodelling of replication features between each stage following cell division. Because TTRs are regions in which potential changes in RT occur[24,25], such remodelling may provide the basis for the gradual developmental progression of the RT programme. In addition, the concomitant decrease in the number of RT peaks and their increase in size suggests a progressive consolidation of the RT programme[7] whereby more adjacent regions with similar RT merge. Indeed, RT peaks become progressively larger and acquire more distinct, earlier relative RT values compared with their genomic surrounding from the 4-cell stage (Fig. 1h). Our data support a gradual consolidation of RT features during preimplantation development and suggest that the shaping of RT occurs at the level of RT peaks and TTRs.

RT in zygote and 2-cell-stage embryos is distinct from later stages

Genome-wide correlation analysis of RT across all stages established that zygotes and 2-cell embryos cluster apart from all other stages (Fig. 2a), suggesting that, despite a similar variability score, the 4-cell-stage RT programme differs from zygotes and 2-cell embryos in other features. To determine the basis of the differences in RT behaviour in zygotes and 2-cell embryos we investigated three alternative explanations. First, to determine whether the unusual RT patterns resulted from asynchrony due to different fertilization times, we performed Repli-seq in zygotes produced by in vitro fertilization (IVF), allowing timely control of fertilization. IVF zygotes showed RT profiles similar to those of zygotes arising from natural fertilization (Extended Data Fig. 4a,b). Second, we considered whether unusual RT patterns result from disparate RT of maternal and paternal genomes, which are thought to replicate asynchronously[26], are physically separated as two pronuclei during the first cell cycle and remain topologically segregated in 2-cell-stage nuclei[27]. To address this we performed Repli-seq in parthenogenetic zygotes containing only one copy of the maternal genome. The replication profiles in parthenotes and normal zygotes were similar (Fig. 2b,c). Genome-wide correlations of RT values confirmed that RT values in parthenogenetic and naturally fertilized zygotes were comparable, and also with IVF zygotes (Fig. 2d and Extended Data Fig. 4c,d). This analysis confirmed that RT separates into two major groups containing zygotes and 2-cell embryos versus all other stages (Extended Data Fig. 4d). We further generated Repli-seq from physically isolated pronuclei (Extended Data Fig. 4e), which showed overall similar RT profiles in maternal and paternal pronuclei (Fig. 2e,f). Both pronuclei exhibited genome-wide correlations similar to natural zygotes (Spearman's $R = 0.65$ and 0.67 for maternal and paternal, respectively; Fig. 2g) and to IVF zygotes (Extended Data Fig. 4c). Maternal RT values correlated slightly better with parthenotes than paternal RT values (Spearman's $R = 0.62$ and 0.49, respectively; Fig. 2h) suggesting that, while highly similar, differences exist between the RT profiles of parental genomes. Finally we investigated whether allele-specific differences can bias RT patterns by performing single-nucleotide polymorphism (SNP)-based analysis of RT in zygotes from hybrid ($F_1 \times DBA$) crosses. Specifically we asked whether the subtle RT differences between parental genomes are consistent across individual embryos. We find that overall there is no consistent allelic-specific bias in zygotes (Extended Data Fig. 4f,g). This indicates that, although maternal and paternal genomes differ slightly in their RT profiles, these differences do not bias zygotic RT. In agreement, RT peaks, TTRs and RT troughs from both genomes have similar RT behaviour (Fig. 2i and Extended Data Fig. 4h,i). In addition, analysis of imprinted genes indicated no replication asynchrony, in line with findings from ES cells[28] (Extended Data Fig. 5). We conclude that RT profiles in zygotes are not due to parental asynchrony but rather reflect inherent properties of RT in both genomes at early developmental stages. Therefore, early embryos show a RT programme that is initially less well defined and becomes progressively more defined from the 4-cell stage.

Segregation between early and late RT increases as development proceeds

Next, we investigated whether the robustness of RT (cell-to-cell heterogeneity) changes during development. We asked whether and how RT heterogeneity fluctuates throughout S phase. We generated a sigmoid model[29] and computed the relationship between RT values and T_{width} (Extended Data Fig. 6a), which quantifies the time difference at which 25–75% of cells replicated a given genomic bin[10,30], for each stage. The T_{width} value thus reflects the variation in RT across cells within the same stage. T_{width} values decreased during development, indicating an overall more coordinated RT programme (Fig. 3a). However, T_{width} increased again for ICM, reflecting the heterogenous nature of the ICM preceding its segregation into epiblast and primitive endoderm lineages (Fig. 3a). Regions replicating early and late were relatively homogenous (Extended Data Fig. 6b). Overlapping of RT features onto T_{width} values indicated that RT peaks and RT troughs are less heterogeneous
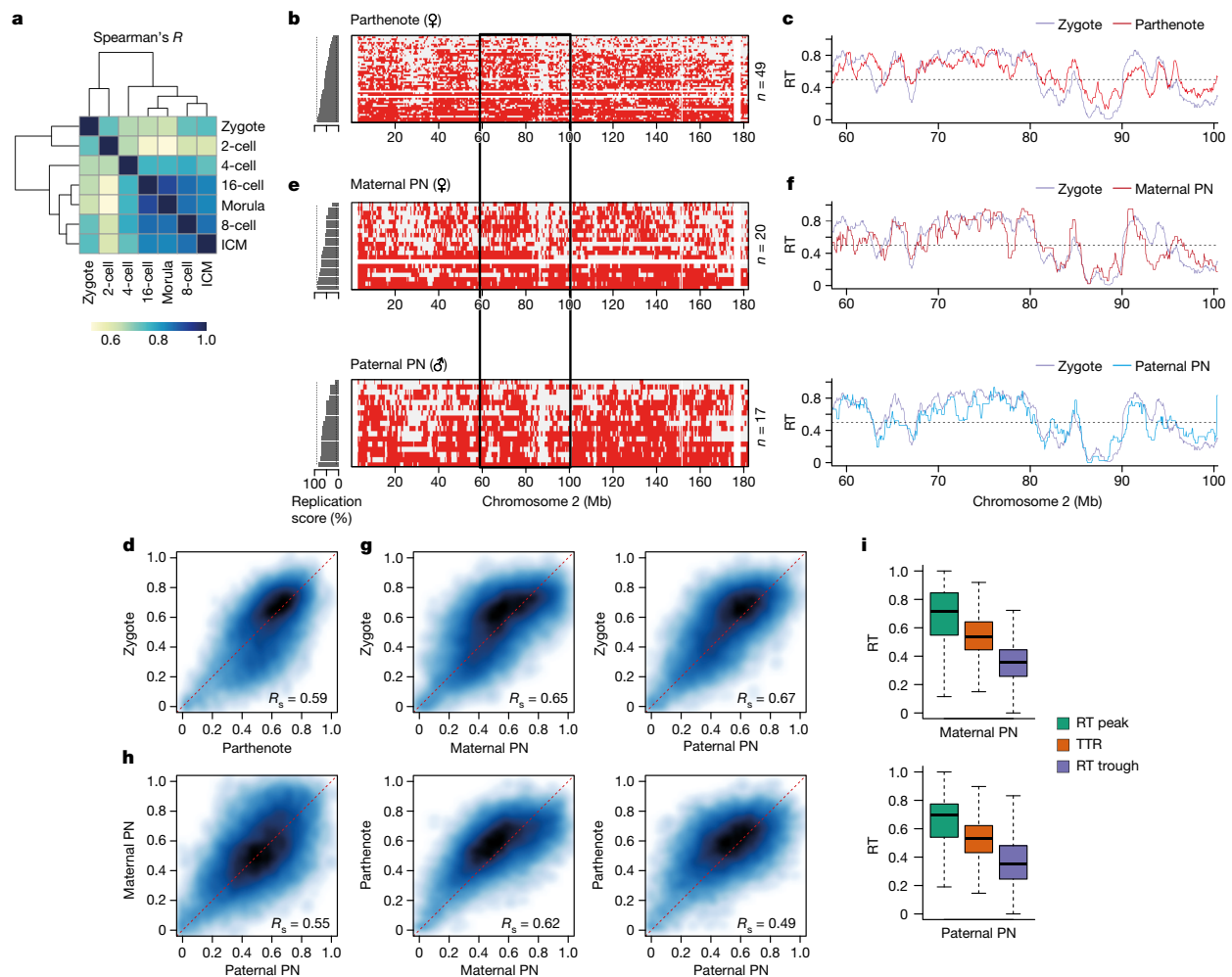
Fig. 2 | RT in zygotes is less well defined compared with that in later embryonic stages and does not exhibit global allelic differences.
a, Correlation of genome-wide RT values between the indicated stages of preimplantation embryos using Spearman's R. **b,c**, Characterization of RT in parthenogenetic zygotes. Heatmaps show replication states in parthenogenetic embryos (**b**) and corresponding average RT profiles (**c**). **d**, Smoothed scatterplot of RT values in normal versus parthenogenetic zygotes. Spearman's correlation (R_s) is indicated. **e,f**, Characterization of RT in physically isolated maternal and paternal pronuclei (PN). Heatmaps show replication states in each pronucleus compared with TTRs (Fig. 3b). In addition, RT peaks and RT troughs are

(**e**) and corresponding average RT profiles (**f**) of the chromosome 2 region, indicated by the black rectangle. **g**, Smoothed scatterplot of RT values in zygotes compared with isolated maternal and paternal pronuclei. Spearman's correlation is indicated. **h**, Smoothed scatterplot of RT values comparing maternal or paternal pronuclei and parthenogenetic zygotes, as indicated. Spearman's correlation is indicated. **i**, RT values of RT peaks, TTRs and RT troughs in isolated maternal and paternal pronuclei. Box plots show median and IQR, whiskers depict the lowest and highest values within 1.5× IQR.

remarkably uniform across cells of the same stage. We also calculated M, which is the replication score at which 50% of cells have replicated a given genomic bin. Thus, the distribution of M-values indicates how well partitioned into early and late are RT values across the genome. M values for mouse ES cells depicted a clear bimodal distribution, reflecting well-defined early and late RT patterns (Fig. 3c). This was not the case for early embryonic stages (Fig. 3c). Instead, a bimodal distribution became apparent after the 2-cell stage, reflecting the emergence of a RT programme that separates the genome towards early (earlier) and late (later) RT values (Fig. 3c and Extended Data Fig. 6c). We conclude that RT heterogeneity fluctuates during S phase within each developmental stage in the same manner as it does in all previously studied systems, and that segregation between early and late RT values increases as development proceeds.

Consolidation of RT is characterized by specific changes in histone modifications

The relationship between RT and transcription remains unclear, with often contradictory reports on RT instructing transcription or vice versa[9,31]. Because the embryo starts transcription de novo following a period of transcriptional silence in the germline, the embryo provides an outstanding opportunity to disentangle the role of transcriptional activation in the establishment of RT. Our above results indicate that the RT programme becomes progressively more defined, particularly after the 2-cell stage (Fig. 1d,h), which corresponds to the time of ZGA[17]. Thus we first asked whether chromatin features of active transcription relate to the progressive definition of RT. H3K36me3 became enriched at RT peaks from the 8-cell stage (Fig. 4a) (no available data for H3K36me3 at the 4-cell stage), indicating that H3K36me3 marks emerging RT
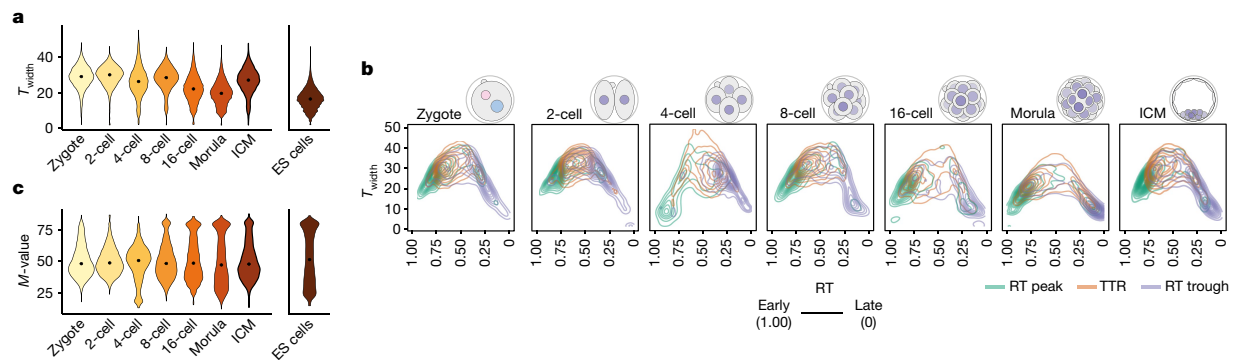
Fig. 3 | RT heterogeneity decreases with developmental progression, and segregation between early and late RT values increases. a, Violin plots showing relative RT heterogeneity (T_{width}), which is the replication score difference between 25 and 75% of cells replicating the 50 kb bin, in embryonic development and in mouse embryonic stem (ES) cells. Dots indicate median.

b, Contour plot showing T_{width} along progression of RT in mouse embryos. RT peaks, TTRs and RT troughs are indicated. **c**, Violin plots showing RT mid-point value (M-value), which is the replication score at which 50% of cells replicated the 50 kb bin during embryonic development and in mouse ES cells. Dots indicate median.

peaks (Extended Data Fig. 7a). Whereas H3K36me3 is associated with gene bodies and is thus typically excluded from replication origins in other cells[23], H3K36me3 does not necessarily reflect transcription elongation kinetics during development[32] and thus our findings may reflect specific embryonic chromatin features. H3K4me3 levels were relatively stable across RT peaks, TTRs and RT troughs, with slightly higher levels at RT peaks and a depletion in RT troughs in zygotes and 2-cell embryos compared with later stages (Fig. 4b and Extended Data Fig. 7a). Because oocytes have distinctive broad H3K4me3 domains, which are remodelled by demethylases KDM5A/5B upon ZGA[33,34], we asked whether H3K4me3 inheritance is linked to RT in embryos. For this we expressed KDM5B[14], known to remove H3K4me3 broad domains[34], in mouse zygotes and performed scRepli-seq at the 2-cell stage (Extended Data Fig. 7b). RT profiles following KDM5B expression showed a similar global pattern in control of 2-cell embryos (Extended Data Fig. 7c,d). In addition, KDM5B expression did not affect RT of major ZGA genes, nor of genes expressed in oocytes (Extended Data Fig. 7e,f), indicating that removal of H3K4me3 following fertilization does not majorly impact RT at regions containing major ZGA genes.

Next we examined whether RT relates to gene expression levels. Genome-wide correlation of RT values and steady-state transcript abundance were low in zygotes and 2-cell embryos (Spearman's correlation, R_s; Fig. 4c). In fact, RT in zygotes and 2-cell embryos correlated similarly with the transcriptome of non-fertilized oocytes and zygotes (Extended Data Fig. 7g). This suggests that either the presence of maternally inherited transcripts from oocytes, which dominates the early transcriptome, overrides a possible relationship with RT or that transcriptional activity does not correlate strongly with RT at these stages. We favour the latter interpretation because 2-cell embryos, which undergo massive transcriptional activation and degradation of maternal transcripts, show a similar correlation between their RT and transcriptome to zygotes (Fig. 4c). Both transcript abundance and RT values change significantly during developmental progression and thus the increasing correlation between RT and transcription during development stems from changes in both transcript abundance and RT (Extended Data Fig. 7h,i). From the 4-cell stage, the correlation between RT and transcript levels increases and the typical relationship between transcription and early replication emerges, with genes expressed at high levels replicating early (Fig. 4c and Extended Data Fig. 7i). Indeed, the correlation between transcript abundance and RT values is significantly greater from the 4-cell stage onwards (Extended Data Fig. 7j). This correlation is similar to ES cells, albeit at a lower extent (Extended Data Fig. 7k). These data show that the known correlation between RT

and gene expression emerges gradually from the 4-cell stage, with genes showing the highest expression replicating early during S phase.

RNA polymerase II at ZGA contributes to fine-tuning of the RT programme

We next addressed directly whether transcription regulates the establishment of RT. We incubated zygotes with α-amanitin under conditions that prevent minor and major ZGA but do not affect RNA polymerase (Pol) I transcription, and performed scRepli-seq at the 2-cell stage (Extended Data Fig. 8a,b). Evaluation of RT at later stages is not feasible because inhibition of ZGA prevents development beyond the 2-cell stage[17]. RT values in α-amanitin-treated embryos showed a moderate correlation with control embryos (Fig. 4d), suggesting that prevention of ZGA with α-amanitin may affect RT at the 2-cell stage. Indeed, we observed changes in RT towards earlier and later following α-amanitin treatment (Extended Data Fig. 8c). Further examination showed localized RT changes in α-amanitin-treated embryos (Fig. 4e), with a statistically significant delay in RT of genomic bins overlapping with major ZGA genes but not of regions containing genes expressed in oocytes (maternal genes) or control regions (Fig. 4f and Extended Data Fig. 8d). To better understand how transcription at ZGA affects RT, we sought to distinguish the effects of general transcription inhibition versus transcription elongation. We took advantage of another RNA Pol II inhibitor, 5,6-dichlorobenzimidazone-1-β-D-ribofuranoside (DRB), which inhibits transcriptional elongation by inhibition of RNA Pol II Ser2 phosphorylation, whereas α-amanitin results in full transcriptional inhibition[35], including via RNA Pol II degradation (Extended Data Fig. 8e,f). DRB treatment during the same period as α-amanitin led to milder changes in RT compared with α-amanitin (Fig. 4d,e). Interestingly, DRB did not significantly change RT of genomic bins containing ZGA genes (Extended Data Fig. 8g,h), suggesting that transcriptional elongation of ZGA genes does not affect their RT. However, DRB and α-amanitin led to similar changes in RT of regions without genes expressed at the 2-cell stage (Fig. 4e and Extended Data Fig. 8i). Thus, we next explored whether other chromatin features relate to the RT phenotype following ZGA inhibition. Prevention of ZGA with α-amanitin alters accessibility in 2-cell embryos[36,37]. Analysis of assay for transposase-accessible chromatin using sequencing (ATAC-seq) datasets showed a significant, positive correlation with RT in 2-cell embryos, indicating that regions replicating early are, in general, more accessible than those that replicate late (Extended Data Fig. 9a,b). This correlation was lost following α-amanitin treatment (Extended Data Fig. 9a,b). Globally, the changes
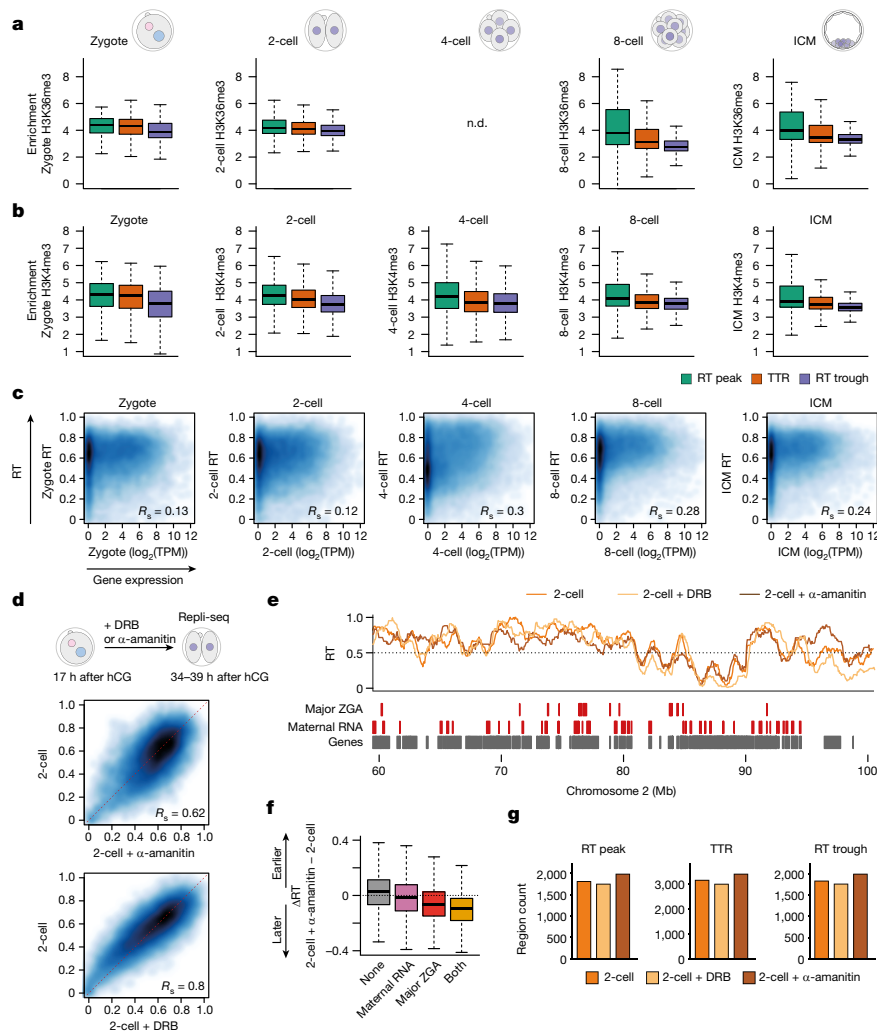
Fig. 4 | Consolidation of RT is characterized by specific changes in histone modifications at RT troughs and RT peaks and is influenced by ZGA. **a**, H3K36me3 coverage at the indicated replication features at different embryonic stages. **b**, H3K4me3 coverage at replication zones at each embryonic stage. **c**, Smoothed scatterplots showing correlation between RT values and transcript abundance ($\log_2(TPM)$) at the indicated embryonic stages. Spearman's correlation is indicated. Note that Spearman's R measures not only linear, but also monotonic, relationships and is robust to outliers. **d**, Smoothed scatterplots showing correlation of RT values between control 2-cell embryos and those treated with α-amanitin (top) or DRB (bottom). DRB or α-amanitin was applied continuously from the zygote stage (17 h after human chorionic gonadotropin

(hCG)) until collection at the 2-cell stage, to block both minor and major ZGA, as indicated in the schematic. Spearman's correlation is indicated. **e**, RT profiles of 2-cell embryos overlayed with those from α-amanitin- and DRB-treated 2-cell embryos. Genomic positions of indicated gene classes according to DBTMEE[54] are shown as rectangles. **f**, Box plots showing the difference in RT values (ΔRT) between α-amanitin-treated and untreated 2-cell embryos at genomic bins overlapping only major ZGA genes, only maternal RNA genes or both gene classes compared with non-overlapping bins. **g**, Number of replication features in control 2-cell embryos and in embryos treated with α-amanitin or DRB. Box plots show median and IQR, whiskers depict the lowest and highest values within 1.5× IQR. n.d., not determined (data not available).

in RT elicited by α-amanitin anticorrelated with sites of genome-wide accessibility in 2-cell control embryos (Extended Data Fig. 9c). Indeed, we find that regions that gain ATAC-seq signal following α-amanitin treatment become replicated later; likewise, regions that lose accessibility become replicated earlier (Extended Data Fig. 9d).

To further understand how transcription during ZGA influences RT, we examined RT features in 2-cell embryos treated with α-amanitin or DRB. Prevention of transcription at ZGA using α-amanitin, but not DRB, led to more TTRs, RT peaks and RT troughs with a concomitant decrease in the size of RT troughs (Fig. 4g and Extended Data Fig. 9e). The increase in their number and the smaller RT troughs suggests a more fragmented, less consolidated RT programme after α-amanitin

treatment. These data also suggest that replication may initiate and terminate at different locations in the absence of embryonic transcription. In support of this, RT troughs in α-amanitin-treated embryos do not show AT content enrichment, in contrast to controls (Extended Data Fig. 9f). In addition, de novo RT peaks in α-amanitin-treated embryos contain fewer genes normally expressed at the 2-cell stage compared with those insensitive to α-amanitin (Extended Data Fig. 9g). Thus, perturbation of RNA Pol II globally at ZGA contributes to fine-tuning of initiation and termination sites at the 2-cell stage.

Finally, we characterized silent chromatin features of the embryonic replication programme. RT troughs contain higher levels of H3K9me3 compared with RT peaks and, to a lesser extent, with TTRs,
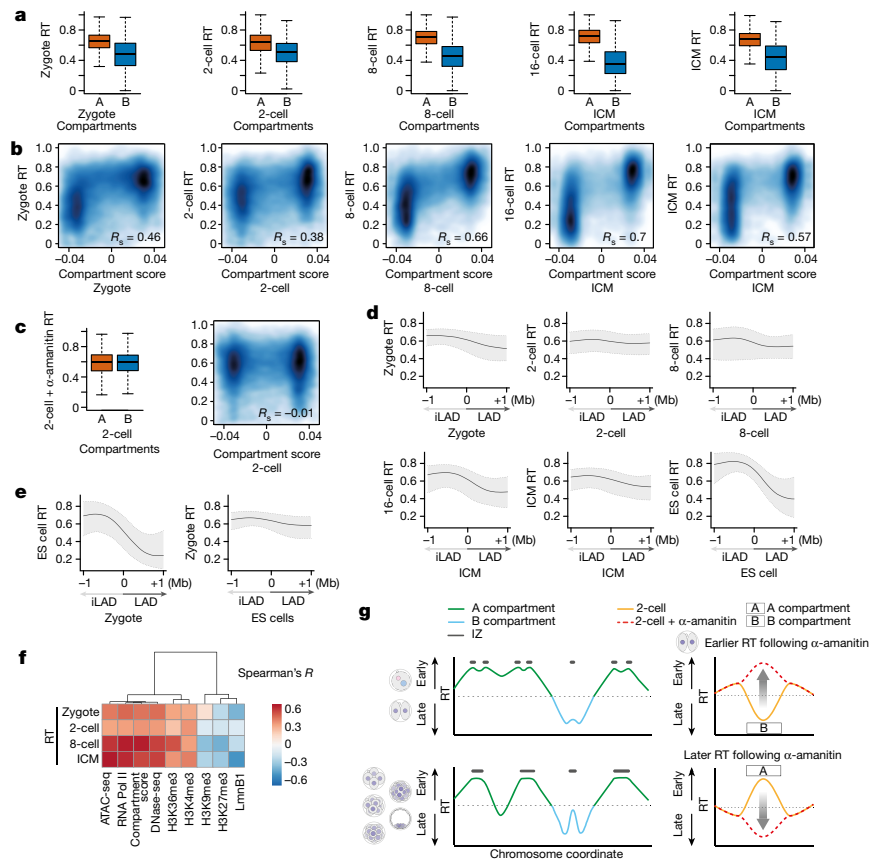
Fig. 5 | The distinctive RT between A and B compartments is dependent on ZGA, and three-dimensional genome organization precedes partitioning of early- and late-replication dynamics. a, Box plots showing RT values in A and B compartments at the indicated stages. Note that, because HiC (high-throughput chromosome conformation capture) data for the 16-cell stage were unavailable, we used the closest developmental stage (ICM) for this comparison. **b**, Smoothed scatterplots showing correlation between RT values and compartment score at the indicated stages. Spearman's correlation is indicated. **c**, Box plots showing RT values in A and B compartments (left) and correlation between RT values and compartment score (right) in α-amanitin-treated, 2-cell-stage embryos. **d**, Composite plots depicting RT values computed against LADs and iLADs at their corresponding developmental stage. Zero indicates the position of LAD–iLAD boundaries. Because DamID data for the 16-cell stage were not available, we used the closest developmental stage (ICM) for this comparison. **e**, Composite plots depicting RT values of mouse ES cells plotted against zygotic LADs (left) and RT values of zygotes against LADs in ES cells (right). Zero indicates the position of LAD–iLAD boundaries. **d**,**e**, Shading and lines indicate IQR and median, respectively. **f**, Correlation (Spearman's *R*) heatmap between RT and distinctive chromatin features. When data for the same stage as RT are not available, those of the closest stage are used for analysis. **g**, Model summarizing our findings indicating progressive resolution of RT following the 2-cell stage. Left, RT peaks merge over time, resulting in changes in both number and size. Right, the effect of ZGA inhibition on RT and its relationship to A and B compartments. **a**,**c**, Box plots show median and IQR, whiskers depict the lowest and highest values within 1.5× IQR.

but these differences emerge only from the 2-cell stage and H3K9me3 levels across RT peaks, TTRs and RT troughs are equivalent in zygotes (Extended Data Fig. 10a). H3K27me3 levels are lowest at RT peaks at all developmental stages and, similarly to H3K9me3, RT peaks and RT troughs acquire gradually different histone modifications during development, with RT peaks showing a depletion of H3K27me3 compared with TTRs and RT troughs by the morula stage (Extended Data Fig. 10b,c). These findings may relate to the progressive heterochromatin maturation of early embryos[38,39]. Overall, maturation of the RT programme is accompanied by a progressive, relative increase in H3K9me3 at RT troughs and a gradual decrease at RT peaks.

Organization into LADs and inter-LADs precedes partitioning of early and late replication

Finally we investigated the dependency between three-dimensional genome architecture and the establishment of RT. In differentiated and stem cells, early and late replication correlate with the A and B compartments, respectively[3,40], and TADs tend to correspond to replication domains[2]. However, because TADs are not clearly detected in early cleavage stages[14,16] we focused on compartments and asked whether the A and B compartments already differ in their RT at the earliest developmental stages. A compartments consistently showed an earlier RT profile compared with B compartments (Fig. 5a and Extended Data Fig. 10d). The distinction between early and late RT values in both compartments was less pronounced in zygotes and became clearer as development proceeds (Fig. 5a). In line with only minor differences in the RT of parental genomes (Fig. 2), RT values were only slightly different in maternal and paternal A and B compartments (Extended Data Fig. 10e). RT differed more between paternal A and B compartments than in maternal compartments, potentially because of the weaker structure of the latter[14–16] (Extended Data Fig. 10e,f). The difference in RT values between A and B compartments increased during development due to both better segregation of RT values and increase in

Article

compartment score (Fig. 5b). Inhibition of ZGA with α-amanitin completely eliminated RT differences between A and B compartments but the compartment score remained similar (Fig. 5c)[14]. Globally, A compartments replicated later and B compartments replicated earlier in α-amanitin-treated embryos compared with controls (Extended Data Fig. 10g). Because B compartments are less accessible than A compartments (Extended Data Fig. 10h), these observations can be explained by our results indicating that α-amanitin leads to a shift towards earlier replication of less accessible regions. We conclude that partitioning of early and late RT during early development coincides with the maturation of A and B compartments. In addition, whereas ZGA does not contribute to compartment strength[14], transcriptional inhibition equalizes differences in RT between compartments.

The genetic constitution of mammalian A and B compartments is largely demarcated by repetitive elements[41,42], which are expressed in the mouse embryo[43,44]. Namely, LINE1 are highly transcribed at the 2-cell stage[43,45] and are enriched in LADs and B compartments[41,42,46]. In fact, LINE1 and SINE segregate mostly exclusively into B and A compartments, respectively[41]. Thus we investigated the replication features of major transposable element families. Overall, LINE1 were enriched in RT troughs and depleted in RT peaks (Extended Data Fig. 10i). This enrichment was stronger for evolutionarily young LINE1, L1Md_A and L1Md_T, contrasting with older LINE2, which showed depletion from RT troughs (Extended Data Fig. 10i). SINE B2 are enriched in RT peaks and depleted in RT troughs, and this tendency became clearer from the 4-cell stage (Extended Data Fig. 10i). MERV-L (MT2_Mm), highly transcribed in 2-cell embryos[44,47], was more homogeneously distributed across RT peaks, TTRs and RT troughs. However, MERV-L enrichment in RT features, albeit low, changed throughout development (Extended Data Fig. 10i). Thus the RT of domains containing MERV-L, unlike LINEs, is dynamic (Extended Data Fig. 10i). Indeed, a change in RT of MERV-L occurs during reprogramming of 2-cell-like cells (2CLCs)[48].

Finally we examined the relationship between LADs and RT. LADs are established in zygotes immediately following fertilization and are reorganized during preimplantation development, but a large proportion of LADs remains constant and is similar to ES cell LADs[13]. In general, LADs, unlike inter-LADs (iLADs), replicate late[2,49]. However, and in sharp contrast to ES cells, RT in zygotes is not clearly distinguishable between LADs and iLADs (Fig. 5d). Zygotic LADs differ between parental genomes[13] and, accordingly, paternal LADs and iLADs exhibit a slight segregation of RT values and maternal ones to a lesser extent (Extended Data Fig. 10j). RT in zygotes did not exhibit a strong bias towards either paternal or maternal LADs/iLADs (Extended Data Fig. 10k). The separation of RT values in LADs and iLADs increases as development proceeds, reaching a clear distinction in ES cells (Fig. 5d). These observations raise the possibility that nuclear organization into LADs and iLADs temporally precedes establishment of the RT programme. To address this, we asked whether RT in ES cells corresponds to LADs/iLADs in zygotes. Remarkably, RT values in embryonic stem cells plotted against the LAD boundaries of zygotes indicated a clear demarcation of RT in embryonic stem cells according to zygotic LAD boundaries (Fig. 5e), indicating that LAD organization in zygotes predisposes RT at later stages of development. In contrast, plotting the RT values of zygotes over ES cell LAD boundaries did not show such a correlation (Fig. 5e). We conclude that organization of LADs and iLADs at the beginning of development precedes the partitioning of early- and late-replication dynamics.

Discussion

Our data indicate that the establishment of RT occurs progressively following fertilization, hand-in-hand with the gradual acquisition of distinctive chromatin features and similarly to other epigenomic features (Fig. 5f). The less well-defined, more heterogeneous RT programme in zygotes and 2-cell embryos may reflect a higher plasticity in the chromatin structure in general and could also be related to

changes in histone deposition occurring at these stages[50]. RNA Pol II in zygotes and 2-cell-stage embryos contributes to the definition of RT. The comparatively milder effects on RT elicited by DRB compared with α-amanitin suggest that RNA Pol II itself influences the RT programme in 2-cell-stage embryos to a greater extent than transcriptional elongation. Although further investigation is warranted to determine whether additional, non-transcription-related effects contribute to these observations—for example via structural proteins[51]—our findings align with work showing that ZGA transcription may be less affected by DRB than by α-amanitin[34,52].

The correlation between transcriptional activity and RT emerges after the 2-cell stage, coinciding with progressive lengthening of the G1 phase[53], known to be important in the definition of RT[6]. Although we observed large-scale changes in RT, for example, with around 20% of the genome switching from early to late RT during preimplantation development, fine-scale changes through the gradual acquisition of histone modifications are also likely to contribute to tuning of RT as cell types emerge. Remarkably, our data indicate that transcription and RNA Pol II function contribute to the definition of the epigenetic features of compartments, in this case their RT (Fig. 5g), but not to their segregation[14]. Our observations that the genome structuring into LADs and iLADs precedes the partitioning of RT at later developmental stages establishes an exciting temporal dependency between these two pillars of the epigenome.

Our work lays the foundations for understanding how genome replication is regulated during development and sheds light on how the epigenome is remodelled at the beginning of mammalian development.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-06872-1.

1. Emerson, D. J. et al. Cohesin-mediated loop anchors confine the locations of human replication origins. *Nature* **606**, 812–819 (2022).
2. Pope, B. D. et al. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–405 (2014).
3. Ryba, T. et al. Replication timing: a fingerprint for cell identity and pluripotency. *PLoS Comput. Biol.* **7**, e1002225 (2011).
4. Klein, K. N. et al. Replication timing maintains the global epigenetic state in human cells. *Science* **372**, 371–378 (2021).
5. Dileep, V. & Gilbert, D. M. Single-cell replication profiling to measure stochastic variation in mammalian replication timing. *Nat. Commun.* **9**, 427 (2018).
6. Gilbert, D. M. & Gasser, S. M. in *DNA Replication and Human Disease* (ed. DePamphilis, M. L.) (Cold Spring Harbor Laboratory Press, 2006).
7. Hiratani, I. et al. Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol.* **6**, e245 (2008).
8. Fragkos, M., Ganier, O., Coulombe, P. & Mechali, M. DNA replication origin activation in space and time. *Nat. Rev. Mol. Cell Biol.* **16**, 360–374 (2015).
9. Farkash-Amar, S. et al. Global organization of replication time zones of the mouse genome. *Genome Res.* **18**, 1562–1570 (2008).
10. Zhao, P. A., Sasaki, T. & Gilbert, D. M. High-resolution Repli-Seq defines the temporal choreography of initiation, elongation and termination of replication in mammalian cells. *Genome Biol.* **21**, 76 (2020).
11. Petryk, N. et al. Replication landscape of the human genome. *Nat. Commun.* **7**, 10208 (2016).
12. Burton, A. & Torres-Padilla, M. E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–734 (2014).
13. Borsos, M. et al. Genome-lamina interactions are established de novo in the early mouse embryo. *Nature* **569**, 729–733 (2019).
14. Ke, Y. et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381 (2017).
15. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
16. Du, Z. et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017).
17. Schultz, R. M. Regulation of zygotic gene activation in the mouse. *Bioessays* **15**, 531–538 (1993).
18. Seller, C. A. & O'Farrell, P. H. Rif1 prolongs the embryonic S phase at the *Drosophila* mid-blastula transition. *PLoS Biol.* **16**, e2005687 (2018).
19. Bartlett, D. A., Dileep, V., Baslan, T. & Gilbert, D. M. Mapping replication timing in single mammalian cells. *Curr. Protoc.* **2**, e334 (2022).

20. Chazaud, C., Yamanaka, Y., Pawson, T. & Rossant, J. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev. Cell* **10**, 615–624 (2006).
21. Streffer, C., van Beuningen, D., Molls, M., Zamboglou, N. & Schulz, S. Kinetics of cell proliferation in the pre-implanted mouse embryo in vivo and in vitro. *Cell Tissue Kinet.* **13**, 135–143 (1980).
22. Tubbs, A. et al. Dual roles of Poly(dA:dT) tracts in replication initiation and fork collapse. *Cell* **174**, 1127–1142 (2018).
23. Petryk, N. et al. MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science* **361**, 1389–1392 (2018).
24. Dileep, V., Rivera-Mulia, J. C., Sima, J. & Gilbert, D. M. Large-scale chromatin structure-function relationships during the cell cycle and development: insights from replication timing. *Cold Spring Harb. Symp. Quant. Biol.* **80**, 53–63 (2015).
25. Miura, H. et al. Single-cell DNA replication profiling identifies spatiotemporal developmental dynamics of chromosome organization. *Nat. Genet.* **51**, 1356–1368 (2019).
26. Bouniol-Baly, C., Nguyen, E., Besombes, D. & Debey, P. Dynamic organization of DNA replication in one-cell mouse embryos: relationship to transcriptional activation. *Exp. Cell. Res.* **236**, 201–211 (1997).
27. Barton, S. C. et al. Genome-wide methylation patterns in normal and uniparental early mouse embryos. *Hum. Mol. Genet.* **10**, 2983–2987 (2001).
28. Rivera-Mulia, J. C. et al. Allele-specific control of replication timing and genome organization during development. *Genome Res.* **28**, 800–811 (2018).
29. Du, Q. et al. DNA methylation is required to maintain both DNA replication timing precision and 3D genome organization integrity. *Cell Rep.* **36**, 109722 (2021).
30. Yang, S. C., Rhind, N. & Bechhoefer, J. Modeling genome-wide replication kinetics reveals a mechanism for regulation of replication timing. *Mol. Syst. Biol.* **6**, 404 (2010).
31. Kupper, K. et al. Radial chromatin positioning is shaped by local gene density, not by gene expression. *Chromosoma* **116**, 285–306 (2007).
32. Abe, K., Schauer, T. & Torres-Padilla, M. E. Distinct patterns of RNA polymerase II and transcriptional elongation characterize mammalian genome activation. *Cell Rep.* **41**, 111865 (2022).
33. Dahl, J. A. et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548–552 (2016).
34. Zhang, B. et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553–557 (2016).
35. Bensaude, O. Inhibiting eukaryotic transcription: which compound to choose? How to evaluate its activity? *Transcription* **2**, 103–108 (2011).
36. Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).
37. Wu, J. et al. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**, 256–260 (2018).
38. Burton, A. et al. Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat. Cell Biol.* **22**, 767–778 (2020).
39. Liu, X. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558–562 (2016).
40. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
41. Lu, J. Y. et al. Homotypic clustering of L1 and B1/Alu repeats compartmentalizes the 3D genome. *Cell Res.* **31**, 613–630 (2021).
42. Wijchers, P. J. et al. Characterization and dynamics of pericentromere-associated domains in mice. *Genome Res.* **25**, 958–969 (2015).
43. Fadloun, A. et al. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat. Struct. Mol. Biol.* **20**, 332–338 (2013).
44. Peaston, A. E. et al. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**, 597–606 (2004).
45. Jachowicz, J. W. et al. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* **49**, 1502–1510 (2017).
46. Meuleman, W. et al. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res.* **23**, 270–280 (2013).
47. Evsikov, A. V. et al. Systems biology of the 2-cell mouse embryo. *Cytogenet. Genome Res.* **105**, 240–250 (2004).
48. Nakatani, T. et al. DNA replication fork speed underlies cell fate changes and promotes reprogramming. *Nat. Genet.* **54**, 318–327 (2022).
49. Peric-Hupkes, D. et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
50. Ishiuchi, T. et al. Reprogramming of the histone H3.3 landscape in the early mouse embryo. *Nat. Struct. Mol. Biol.* **28**, 38–49 (2021).
51. Zhang, S. et al. RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis. *Sci. Adv.* **7**, eabg8205 (2021).
52. Liu, B. et al. The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature* **587**, 139–144 (2020).
53. Smith, R. K. & Johnson, M. H. Analysis of the third and fourth cell cycles of mouse early development. *J. Reprod. Fertil.* **76**, 393–399 (1986).
54. Park, S. J., Shirahige, K., Ohsugi, M. & Nakai, K. DBTMEE: a database of transcriptome in mouse early embryos. *Nucleic Acids Res.* **43**, D771–D776 (2015).

135

Article

Methods

Embryo collection and culture

All experiments were performed under the authorization of the authorities from Upper Bavaria (Tierversuchsantrag von Regierung von Oberbayern). The temperature, humidity and light cycle of mouse cages were maintained at 20–24 °C, 45–65% and 12/12 h dark/light, respectively. F_1 female mice (C57BL/6J × CBA) under 10 weeks of age were superovulated by intraperitoneal injection of 10 U of pregnant mare serum gonadotropin, followed by 10 U of hCG 48 h later, and were then mated with DBA/2J male mice. Zygotes were collected from the oviduct and cumulus cells removed following brief incubation in M2 medium containing hyaluronidase (Sigma-Aldrich). Zygotes were placed in drops of KSOM (potassium simplex optimized medium) and cultured at 37 °C with 5% CO_2 as previously described. For induction of parthenogenetic embryos, MII-stage oocytes were collected, as described above, from superovulated females without mating. Following removal of cumulus cells, oocytes were treated with 10 mM Sr^{2+} for 2 h in Ca^{2+}-free CZB medium and then incubated in KSOM. For generation of IVF-derived zygotes, MII oocytes from F_1 female mice (C57BL/6J × CBA) were inseminated with activated spermatozoa obtained from the caudal epididymides of adult DBA/2 J male mice.

Detection of 5-ethynyl-2′-deoxyuridine incorporation

Cells were incubated with 50 μM 5-ethynyl-2′-deoxyuridine (EdU) for 1 h for each time window, as indicated, and processed for quantification of signal intensity. Incorporated EdU was visualized by Click-iT chemistry (Thermo Fisher Scientific) followed by permeabilization as described in the manufacturer's instructions. Images were acquired on a SP8 confocal laser-scanning microscope (Leica). EdU was coupled to Alexa 594 and images acquired with a Plan-Apochromat ×63/1.4 numerical aperture 1.4 oil-immersion objective (Leica) at 561 nm excitation.

Analysis of EdU incorporation

To quantify EdU incorporation we manually cropped confocal stacks containing several embryos so that each image contained only one single embryo. Only embryos that looked fertilized and with normal pronuclei following visual inspection were included in this analysis. From embryo images we then automatically obtained the maximum intensity value in the EdU channel of the whole stack by ImageJ (v.1.53k) with a custom-made ImageJ macro. We plotted and analysed the resulting EdU intensity values for each time bin with R.

Inhibition of ZGA

For inhibition of both minor and major ZGA, embryos were treated with either 0.1 mg ml^{-1} α-amanitin or 100 μM DRB from the zygote stage at 17 h after hCG injection until their collection for single-cell Repli-seq at the 2-cell stage. Validation of the α-amanitin effect on transcriptional silencing was done using a Click-iT RNA Alexa Fluor 594 Imaging Kit (Thermo Fisher Scientific) at the 2-cell stage (at 40 h after hCG injection).

Gene expression analyses following treatment with α-amanitin and DRB

Twelve embryos were treated with either 0.1 mg ml^{-1} α-amanitin or 100 μM DRB from 17 to 40 h after hCG to inhibit both minor and major ZGA, then flash-frozen in liquid nitrogen in 5 μl of 2× reaction buffer (CellsDirect One-Step qRT–PCR kit, no. 11753100, Thermo Fisher). Next, 0.5 μl of a 1:200 dilution of ERCC spike-in mix (Thermo Fisher) was added to each group and TaqMan Gene Expression assays were performed according to previous work[38]. Complementary DNA was diluted tenfold before analysis with Universal PCR Master Mix and TaqMan Gene Expression assays (Applied Biosystems). All raw C_t values were normalized by those acquired from the ERCC spike-in specific primer set, and relative expression levels of each gene were determined by the ddCt method. We assigned C_t values below the detection range as expression level 0. Primers and probes for ribosomal DNA (*Hsa1*) were produced by TIB MolBiol (custom design)[45]. Primers and probes for Zscan4 cluster and ERCC spike-in were purchased from Applied Biosystems.

Immunostaining following either treatment by α-amanitin and DRB or expression of KDM5B

Embryos were treated with either 0.1 mg ml^{-1} α-amanitin[55,56] or 100 μM DRB from 17 to 40 h after hCG and fixed with 4% paraformaldehyde (PFA) for 20 min at room temperature. For KDM5B expression, 2 μg μl^{-1} KDM5B of in vitro synthesized messenger RNA was microinjected into zygotes at 18 h after hCG and fixed with 4% PFA for 20 min at room temperature at 48 h after hCG, similar to previous experiments[13,33]. Embryos were then permeabilized with 0.5% Triton X-100 containing PBS for 20 min. For immunostaining following Triton pre-extraction, embryos were first permeabilized with pre-extraction buffer (50 mM NaCl, 3 mM $MgCl_2$, 300 mM sucrose, 25 mM HEPES, pH adjusted to 7.4) with 0.5% Triton X-100 for 10 min on ice and washed three times in pre-extraction buffer before fixing in 4% PFA at room temperature for 20 min. Following blocking for 1 h at room temperature in blocking solution (5% normal goat serum in PBS), embryos were incubated with either anti-RNA polymerase II (no. sc-899, 1:100), anti-RNA polymerase II CTD repeat YSPTSPS (phospho S2, no. ab5095, 1:1,000) or anti-H3K4me3 (Diagenode, no. C15410003, 1:250) antibody in blocking solution overnight at 4 °C. Embryos were incubated for 1.5 h at room temperature in blocking solution containing goat anti-rabbit IgG highly cross-adsorbed secondary antibody, Alexa Fluor 488 (Thermo Fisher Scientific, no. A11034, 1:1,000). After washing, embryos were mounted in Vectashield (Vector Laboratories). Confocal microscopy was performed using a ×40 oil objective on an SP8 confocal microscope (Leica) and images acquired with LAS X software.

Repli-seq

Single-cell Repli-seq was performed as previously described[19] based on ref. 5. In brief, early-stage zygotes were collected and cultured until they reached the S phase at each developmental stage, based on their time following hCG injection. Embryos were collected at different time points at each developmental stage to achieve sampling over the entire S phase. Collection times are indicated in Supplementary Table 1. For parthenogenetic embryos and IVF-derived zygotes, the timing of S phase was calculated based on the time elapsed since activation and insemination, respectively. For KDM5B experiments, 2 μg μl^{-1} KDM5B of in vitro synthesized mRNA was microinjected into zygotes at 18 h after hCG as previously described[13]. For each developmental stage, embryos were obtained from several litters and embryos from different litters were collected across different dates to ensure robust data collection. The number of mice used for collection of samples for each developmental stage is indicated in parentheses, as follows: zygote (20), 2-cell (30), 4-cell (27), 8-cell (20), 16-cell (15), morula (16), ICM (19), parthenotes (14), IVF zygotes (14), 2-cell + α-amanitin (14), 2-cell + DRB (24) and 2-cell + KDM5B (24). Zona pellucida was removed by exposure to acid Tyrode, and each blastomere was dissociated by gentle pipetting following trypsin treatment. For Repli-seq with physically isolated pronuclei we distinguished maternal and paternal pronuclei based on their size and relative position to the second polar body, and isolated them using micromanipulation. The remaining zygote containing a single pronucleus was also collected following removal of the polar body so that both pronuclei from the same zygote were further processed for Repli-seq. ICM cells were collected following trypsin digestion as previously described[57], with repeated oral pipetting in 0.5% trypsin and 1 mM EDTA; collection times are indicated in Supplementary Table 1. To distinguish ICM from trophectoderm cells, blastocysts were labelled with Fluoresbrite YG Microspheres (0.2 μm, Polysciences) before incubation with trypsin, and individual cells were

sorted according to either positive (trophectoderm) or negative (ICM) fluorescence under a fluorescence microscope following disaggregation. Individual blastomeres or pronuclei were placed in eight-strip PCR tubes containing lysis buffer, and extracted DNA was fragmented by heat incubation. Fragmented DNA was tagged by the universal primer 5′-TGTGTTGGGTGTGTTTGGKKKKKKKKKKKKNN-3′ and amplified with whole-gene amplification primer sets, which have individual barcodes. This whole-genome amplification procedure was successfully used for single-cell Repli-seq in cell culture[4,5]. Amplified DNA was purified using the QIAquick 96 PCR Purification Kit (QIAGEN), and concentration determined by NanoDrop (Thermo Scientific). Equal amounts of DNA from each sample (up to 96 samples) were pooled and 1 μg of each was ligated with Illumina adaptors using the NEBNext Ultra II DNA Library Prep Kit (NEB). Illumina sequences (NEBNext Multiplex Oligos for Illumina, NEB) were added to adaptor-ligated samples by PCR. Clean-up and size selection of the PCR product was done using SPRIselect (Beckman Coulter), and the quality of the library was confirmed using a 2100 Bioanalyzer with the High Sensitivity DNA Kit (Agilent).

Single-cell Repli-seq read alignment and quality control filtering

An overview of sample collection, mapping statistics and quality control is included in Supplementary Table 1. The quality control parameters we used were (1) the number of reads, which we set as 750,000 aligned reads as minimum; and (2) a coefficient of variation, which we established as a measure of equal/balanced coverage between chromosomes, thus filtering out potential cells with aneuploidy. At early stages, the reason for failure was equally the low number of reads or a high coefficient of variation (typically due to either lack of reads on a complete chromosome or in fragments of the genome; for example, zygotes 13 and 8 were excluded due to low number of reads and zygote 56 to a high coefficient of variation). At later stages, chromosome imbalances were the most common reason for failure (59 cells with high coefficient of variation versus three with low reads in the blastocyst stage), which reflects the known aneuploidy of cells at this embryonic stage. Sequencing reads were aligned to the mm10 genome using bowtie2 (v.2.3.5)[58] with the '--local' option. Duplicates were marked using SAMtools (v.1.9) 'markdup' as described by SAMtools[59] documentation (the commands 'fixmate' and 'sort samtools' were used for this purpose accordingly). Using SAMtools view, reads were filtered by retaining only properly paired reads, removing duplicates and selecting those whose mapping quality was higher than or equal to 20. BED files of the read coordinates were generated with the BEDtools[60] (v.2.29.0) command 'bamtobed'. Using BEDtools intersect, read counts were obtained for contiguous 50 kb genomic bins. For each cell the average of the bin counts was calculated for chromosomes 1–19; these 19 values were then next used to calculate the coefficient of variation as standard deviation divided by the mean. Cells with a coefficient of variation greater than 0.1 were removed from analyses due to chromosome imbalance. To maximize the number of samples used, the coefficient of variation was recalculated, excluding chromosomes one at a time. Cells were considered for further analysis if they passed the threshold when only one specific chromosome was removed. This chromosome was subsequently masked in downstream analyses; this filter removes abnormal genotypes and cells with aneuploidy.

Assignment of replication status

Using the read counts obtained for contiguous 50 kb genomic bins, we used the single-cell Repli-seq bioinformatic pipeline previously described[5], which we followed with some modifications for each embryonic stage as summarized below. Window counts were first normalized to reads per million, and then each bin by its respective average of all samples within the same stage, aiming to correct for mappability biases intrinsic to genomic regions. Outlier regions were then masked, specifically the windows of the lower fifth percentile and upper first

percentile values. To correct for low mappability, windows were segmented with the R package copy number (v.1.28.0, R v.4.0.0)[61] to retain segments with the highest 95% of values. We did not perform the G1/G2 normalization described previously[5], but we verified that this did not impact the results of these analyses. In brief, we used the validated mouse ES cell scRepli-seq datasets in ref. 5 and ran the analysis pipeline as described in their methods section with and without G1 control cells. Subsequently we compared the generated matrix of ones and zeros (that is, bins replicated and not replicated, respectively) by determining the percentage of windows that remained the same (for example, their 1 or 0 replication state did not change) after running the pipeline versus without G1 control. These analyses showed a high concordance between the two pipelines, with over 91% identity of genomic bins with zeros and ones on average across cells (Extended Data Fig. 1b). Importantly, those cells classified as outliers based on our analysis correspond to those that were removed in the original publication[5] based on their 'Removing outlier cells', and were not considered for further analyses. Data were centred by the mean, scaled by the IQR for each cell and smoothed using a median filter with a running width of 15 windows, followed by segmentation with the R package copynumber. Finally, using the function normalmixEM in the R package mixtools (v.1.2.0)[62], segmented values were used to fit a mixture model with two components to identify replicated and non-replicated window populations. To do this, two normal distribution functions were used to select a cutting threshold that better separated distributions; this value is located where the two individual normal distribution functions intersect. If no intersection was found between the means of the two normal distribution functions, the mid-point of the means was used as a threshold.

Computing replication scores, RT values and variability scores

Genome-wide replication score was defined as the percentage of replicated genomic bins for each cell. Throughout the manuscript we have used a 50 kb bin size, but we obtained similar results when using 25 and 100 kb bin size. Cells with a replication score greater than 90% and less than 10% were excluded from downstream analyses. We used the replication score to rank cells by S-phase progression for visualization of their replication status on heatmaps (Fig. 1c). Next we calculated raw RT values as the fraction of cells that replicated the given genomic bin for each stage, respectively. A RT value indicates earlier RT, because a higher proportion of cells replicated the bin. To correct for potential sampling bias of cells, we calculated the fraction of replicated cells in overlapping intervals of the genome-wide replication score with interval size of 35% and increment of 4.33% (for example, 0–35%, 4.33–39.33% and so on) for each genomic bin. The average of these 16 intervals served as the interval RT value that was used for both visualization of RT profiles (Fig. 1e) and downstream analyses. Raw and interval-averaged RT values looked similar overall (Extended Data Fig. 1c; RT raw versus interval), except for some stages in which the number of cells within replication score intervals showed a different distribution. Variability score was calculated using the following formula: $score = 1 - (abs(p - 0.5)/0.5)$, where p is the fraction of replicated cells (ones) for the given bin; note that p is corrected for sampling (as described above). The variability score is therefore a measure of variation in the RT programme across cells, because it represents the number of cells that either replicated or did not replicate a given bin. A value of 1 means that one-half of the cells replicated a given bin and corresponds to the highest variance; likewise, a value of 0 means that either all cells replicated or did not replicate a given bin, which corresponds to the lowest variance and/or no variance.

Identification of initiation zones (referred to as RT peaks), TTRs and termination zones (referred to as RT troughs)

To distinguish the features of RT, initiation zones, TTRs and termination zones were defined based on RT values. Genomic bins were grouped

137

Article

into 15 clusters by their RT values using the Mclust function from the R package mclust (v.5.4.10, R v.4.1.2). Clusters were ranked by their average RT values following analysis similar to that described previously[10], except that we used RT values for clustering as opposed to the 16 Repli-seq fractions. Initiation zones and termination zones were defined as consecutive bins with local maxima or minima of their cluster ranks, respectively, in sliding windows of 21 genomic bins using the rollappy function from the R package zoo (v.1.8-10). Regions between initiation zones and termination zones were defined as TTRs (Extended Data Fig. 3b). The number of initiation zones, which we refer to as RT peaks, recorded previously[10] (approximtely 2,200 in neuronal progenitor cells) is similar to that reported here. To determine the significance of the changes in the number or region size of initiation zones, TTRs and termination zones throughout development, a linear model was fitted using the lm function in R (v.4.1.2). The rank of the developmental stages (that is, 1–7) served as the independent variable. The dependent variable was either the number of regions or the upper quartile of region sizes (75th percentile) for each region type. The P value of the coefficient corresponding to the slope indicates the significance of the linear trend. For composite plots, RT values were centred at the middle point of RT peak coordinates in 2 Mb windows and the median of RT values was calculated per position (Fig. 1h). To visualize relative RT compared with the neighbouring region, the minimum value of the 2 Mb window was subtracted for each stage.

Analysis of RT heterogeneity

Heterogeneity analysis was performed using the sigmoidal model formula as described previously[5,63]. A sigmoidal curve was fitted for each genomic bin by the nls function from the R package stats (v.4.1.2), such that nls(y - 100/(1 + exp($-g \times (x - M)$))), start = list($g = 0.1, M = m0$)) (Extended Data Fig. 6a). The average genome-wide replication score of each of the 16 overlapping intervals (see above) served as the independent variable (x), with the percentage of cells that replicated the bin within the same replication score interval as dependent variable (y). Model parameters were M = mid-point, g = slope (gain) and $m0$ = initial value for M (100 minus the mean of y values). By this method, the replication status of the given genomic bin was related to the overall S-phase progression of cells (measured in intervals of replication score). To anchor the start and end points of the curve, 16 data points of 0 and 100 values were added to the x and y variable, respectively. Two parameters were calculated from the curve fitting, M-value and T_{width}. The M-value (RT mid-point, sometimes also referred to as T_{rep} in the literature[10]) is the replication score (roughly S-phase time) at which 50% of the cells replicated the given bin. A higher M-value indicates later RT. T_{width} is a measure of RT heterogeneity and is defined as the replication score difference (approximate S-phase time difference) of between 25 and 75% of the cells that replicated the given genomic bin. A higher T_{width} value indicates higher heterogeneity, because the transition from non-replicated to replicated status is greater.

Allele-specific analyses

To address any bias that could have been caused by SNPs during alignment, reads were realigned to a SNP-masked genome sequence containing an 'N' anywhere in which a SNP between any of the paternal (DBA) or maternal genomes (C57BL/6 × CBA) is located. The bam files were subsequently divided into paternal and maternal reads. Importantly, not all potential SNPs between strains were used. Splitting considered only SNPs that were different for the three genomes or those whose nucleotide was the same for both maternal genomes but different compared with the paternal one. Both reference preparation and splitting were performed with SNPsplit[64] (v.0.5.0). Reads were filtered using the same tools and thresholds as described above for non-allelic analyses— that is, taking into account read duplication, properly paired criteria and a mapping quality filter. Finally, as previously described, BEDtools intersect was used to count the number of reads for each contiguous

50 kb window. All subsequent analyses were performed on genomic bins, with at least five reads assigned either to the maternal or paternal genome of the same sample.

To determine allelic bias, the $\log_2$ ratio of maternal:paternal read counts was calculated for each bin. The majority of physically separated maternal or paternal pronuclei showed a high positive (over +2) or negative (below −2) $\log_2$ ratio, respectively. Pronuclei with a $\log_2$ ratio of the opposite sign were exchanged for downstream analyses. We identified several parthenogenic examples among IVF zygotes ($\log_2$ ratio above 1), which were excluded from further analyses. Finally we calculated Spearman's correlation coefficients on $\log_2$ maternal:paternal ratios pairwise across single zygotes and visualized these as a correlation heatmap (Extended Data Fig. 4f). A high correlation value between two zygotes indicates that, if a genomic bin has a high allelic bias in one of the zygotes it also has a high bias in the other.

Analysis of imprinted genes

Lists of maternally and paternally imprinted genes were downloaded from the Geneimprint database (https://www.geneimprint.com/site/genes-by-species.Mus+musculus). RT values were extracted for genomic bins overlapping imprinted genes. If multiple bins overlapped the same gene, RT values were averaged. For expression level and allelic bias analysis, supplementary data were downloaded from Gene Expression Omnibus (GEO) (GSE38495 and GSE45719)[65]. A gene was considered expressed when its average fragments per kilobase exon per million mapped reads value in the given stage was greater than zero. Allelic bias was calculated as the $\log_2$-transfomed ratio between read counts assigned to Cast or C57BL/6 genomes. A gene was considered maternally biased if the average $\log^2$ allelic ratio was greater than zero, and paternally biased if less than zero. RT values at imprinted genes were visualized on heatmaps and ordered by their expression and allelic bias status. In total we analysed 49 maternally and 37 paternally imprinted genes, corresponding to 98 and 100 genomic bins, respectively.

Analysis of transposable elements

Transposable element annotation for the mm10 genome was obtained from Hammell's laboratory repository (https://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/mm10_rmsk_TE.gtf.gz).

Enrichment of transposable elements in RT peaks, TTRs or RT troughs was estimated by calculating the $\log_2$ ratio of the number of transposable elements of the given type overlapping with RT peaks, TTRs or RT troughs relative to the overlap of randomly shifted transposable elements with RT peaks, TTRs or RT troughs, respectively. The final enrichment value was the average of 1,000 iterations.

Statistical and genome-wide enrichment analysis

For statistical analyses of single-cell RT data we established a bootstrapping approach and calculated 95% confidence intervals to judge statistical significance[66]. We chose this method to avoid the inflation of P values when n is large due to a large number of genomic bins (n = approximately 49,000) and thus we applied bootstrapping to samples, in this case single cells (n = approximately 30–70), rather than to genomic bins. Namely, we iteratively resampled individual cells with replacement 1,000 times for each stage or condition. For each iteration we recalculated RT values and any subsequent statistic—for example, Spearman's correlation coefficient or ΔRT between conditions, as described above. We constructed confidence intervals from the bootstrap distribution using the percentile method. The 95% confidence interval is the interval between the 2.5th and 97.5th percentiles of the distribution; when 95% confidence intervals do not include zero or two intervals do not overlap, they are significantly different from zero or different from each other, respectively. For enrichment analysis of overlapping regions or gene classes, genomic bins were grouped by significantly differential RT values to increasing (earlier),

decreasing (later) or non-significant (no change) bins. Enrichments were visualized on heatmaps by calculating the ratio of the observed number of overlapping bins relative to the expected value, which is the product of the row and column sums divided by the total number of bins in the corresponding contingency table.

Analysis of public chromatin datasets

Published datasets were downloaded from GEO with accession numbers GSE66581, GSE101571 (ATAC-seq[36]), GSE71434 (H3K4me3 chromatin immunoprecipitation sequencing (ChIP)[34]), GSE112834 (H3K36me3 ChIP[67]), GSE98149 (H3K9me3 ChIP[68]), GSE73952 (H3K27me3ChIP[39]) GSE76687 (H3K27me3 ChIP[69]) and GSE135457 (Pol2 Stacc-seq[52]) and GSE76642 (DNase I hypersensitive sites sequencing[70]). Paired-end reads were trimmed by cutadapt (v.3.4) with parameters -a CTGTCTCTTA TA -A CTGTCTCTTATA -a AGATCGGAAGAGC -A AGATCGGAAGAGC --minimum-length=20. Following trimming, reads were aligned to the mouse reference (GRCm38) using bowtie2 (v.2.3.5) with parameters --end-to-end --very-sensitive --no-unal --no-mixed --no-discordant -I10 -X 500. Reads were filtered by mapping quality score using SAMtools (v.1.3) with the parameter -q 12. Read pairs were read into R (v.3.6.3) using the readGAlignmentPairs function from the GenomicAlignment package (v.1.22.0) and were filtered for unique fragments. Fragments aligned to the mitochondrial genome or small scaffolds were not considered in analyses. Fragments were counted in 50 kb consecutive genomic bins (same bins as for RT profiles), normalized by the sum of fragment counts and multiplied by 1 million. Finally, normalized counts were $\log_2$ transformed following the addition of a pseudocount of 1. Note that, for the analysis of H3K27me3 in Extended Data Fig. 10b,c the dataset used was that of Liu et al. (GSE73952)[39] whereas in Fig. 5f the dataset used was that of Zheng et al.[69] (GSE76687). For the correlation analysis shown in Fig. 5f we used the following stages when the actual stage was not available: early 2-cell ATAC-seq for zygote, morula DNase I hypersensitive sites sequencing for ICM and ES cell LmnB1 DamID for ICM. Differential genomic bins between conditions (for example, ATAC-seq following α-amanitin treatment) were called by DESeq2 (v.1.34.0) with an adjusted P value cutoff of 0.05. For ATAC-seq analysis in α-amanitin-treated embryos, 2-cell-stage embryos administered α-amanitin treatment by Wu et al.[37] (GSE101571) were compared with untreated 2-cell-stage embryos derived from Wu et al.[36] (GSE66581).

Analysis of public HiC and LAD datasets

HiC compartment coordinates and scores (GSE82185)[16], as well as LAD coordinates (GSE112551)[13], were analysed as previously described[13].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The scRepli-seq data for the present study are available from the GEO database, accession GSE218365. Previously published RNA sequencing datasets reanalysed in the present study are available under accessions GSE38495, GSE45719 and GSE98063. Chromatin datasets reanalysed in the present study are available under accessions. GSE66581, GSE101571, GSE71434, GSE112834, GSE98149, GSE73952, GSE76687, GSE135457 and GSE76642. All other data supporting the findings of the present study are available from the corresponding author on reasonable request.

HiC and LAD datasets reanalysed in the present study are available under accessions GSE82185 and GSE112551.

Code availability

Next-generation sequencing data were analysed with publicly available programmes and packages, as detailed in Methods. Additional details on specific code used to generate scRepli-seq workflows are available on request.

55. Aoki, F., Worrad, D. M. & Schultz, R. M. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev. Biol.* **181**, 296–307 (1997).
56. Goddard, M. J. & Pratt, H. P. Control of events during early cleavage of the mouse embryo: an analysis of the '2-cell block'. *J. Embryol. Exp. Morphol.* **73**, 111–133 (1983).
57. Burton, A. et al. Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep.*, **5**, 687–701 (2013).
58. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
59. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
60. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
61. Nilsen, G. et al. Copynumber: efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics* **13**, 591 (2012).
62. Benaglia, T. et al. Mixtools: an R package for analyzing finite mixture models. *J. Stat. Softw.* jstatsoft.org/article/view/v032i06 (2009).
63. Takahashi, S. et al. Genome-wide stability of the DNA replication program in single mammalian cells. *Nat. Genet.* **51**, 529–540 (2019).
64. Krueger, F. & Andrews, S. R. SNPsplit: allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res.* **5**, 1479 (2016).
65. Deng, Q., Ramskold, D., Reinius, B. & Sandberg, R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193–196 (2014).
66. Efron, B. The Jackknife, the Bootstrap and other resampling plans. In *CBMS-NSF Regional Conference Series in Applied Mathematics* (1982); https://doi.org/10.1137/1.9781611970319.
67. Xu, Q. et al. SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat. Genet.* **51**, 844–856 (2019).
68. Wang, C. et al. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat. Cell Biol.* **20**, 620–631 (2018).
69. Zheng, H. et al. Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol. Cell* **63**, 1066–1079 (2016).
70. Lu, F. et al. Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* **165**, 1375–1388 (2016).

Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Quality control of scRepli-seq samples. a. Scatter plots comparing the coefficient of variation calculated on the average read counts per chromosomes and the number of reads for each cell at the indicated embryonic stages. Horizontal and vertical lines indicate cutoffs for filtering cells. **b**. High concordance of replication state between with or without normalization by cells in G1. **c**. Comparison between two computational methods to calculate RT profiles. Shown are representative RT profiles derived from either raw or interval averaged replication timing values in the morula stage. **d**., **e**. Analysis of DNA replication in zygotes by EdU incorporation (d). Representative images of incorporated EdU and the corresponding quantifications are shown in e. Female and male pronuclei are indicated; the white dotted line depicts the nuclear periphery; note the EdU incorporation at the characteristic ring-shaped heterochromatic regions surrounding the nucleoli precursors between the 24 h and 26 h time. Approximate early, mid, and late S-phase times are indicated based on earlier work. Box plots show median of maximum intensity value and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR. n, and N, number of analysed embryos and number of independent biological replicates, respectively. Scale bar, 10 μm.

a
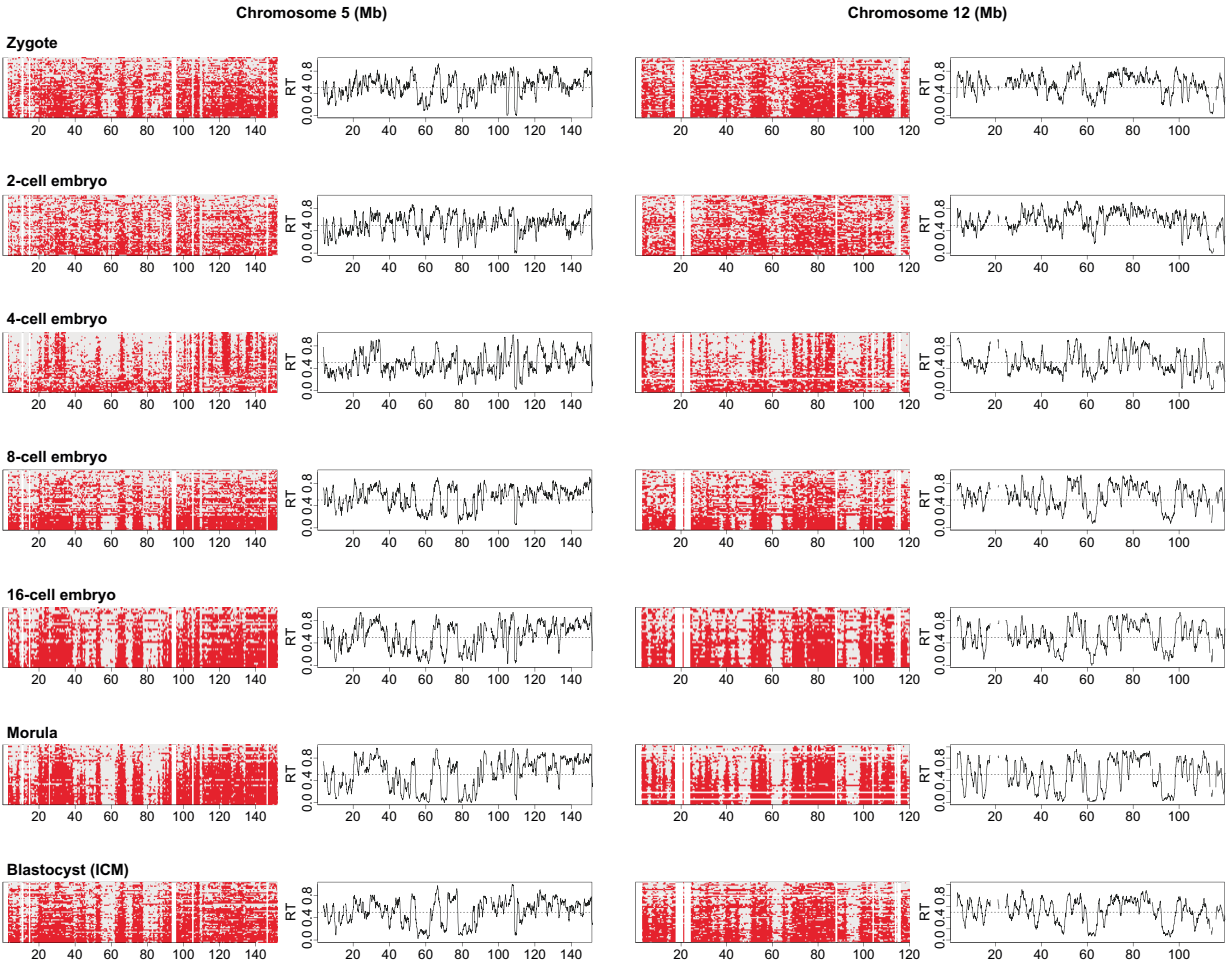
b
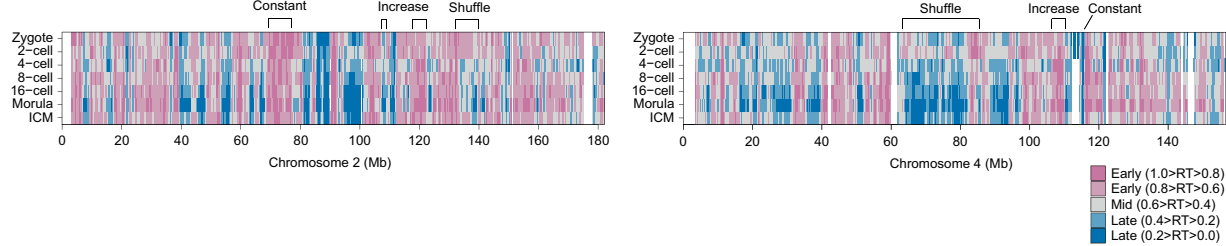
Extended Data Fig. 2 | Heatmaps of replication status and RT profiles of preimplantation embryos over a representative region on Chromosome 5 and 12. a. Cells are ranked by their replication score. The black line indicates RT profiles calculated as the average of overlapping intervals defined by the genome-wide replication score. **b**. Comparison of RT values in bins of 50 kb across embryonic stages. Representative changes in the RT (increase, shuffle, and constant) are indicated. White regions are regions of no coverage in the corresponding sample.
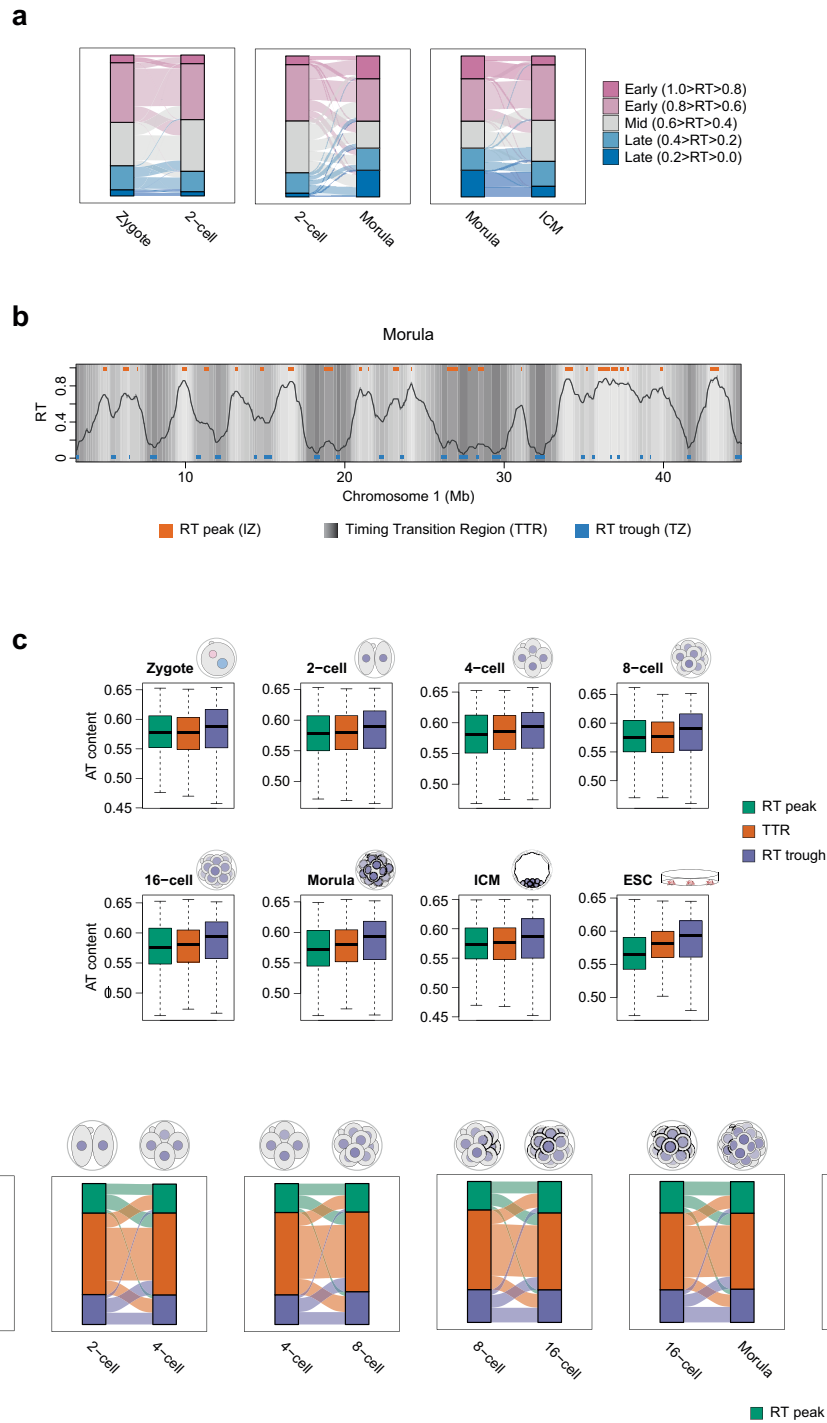
Extended Data Fig. 3 | Dynamics of RT peaks, TTRs, and RT troughs during preimplantation development. a. Alluvial plot showing the changes of RT phases. RT values were categorised in 5 groups from the earliest (1.0 ≥ RT > 0.8) to latest RT (0.2 > RT ≥ 0.0) across the genome. **b**. Representative replication timing profile in the morula stage depicting RT peaks, TTRs, and RT troughs. Grey shading represents 15 clusters of the RT values that were used to call local maxima (RT peaks or IZs) or local minima (RT troughs or TZs). Regions in transition between RT peaks and RT troughs were called as TTRs. The line indicates RT values. **c**. Fraction of A + T nucleotides in RT peaks, TTRs, and RT troughs during preimplantation development. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR. **d**. Alluvial plot showing the relative changes of RT peaks, TTRs, and RT troughs at each cell division during preimplantation development. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR.
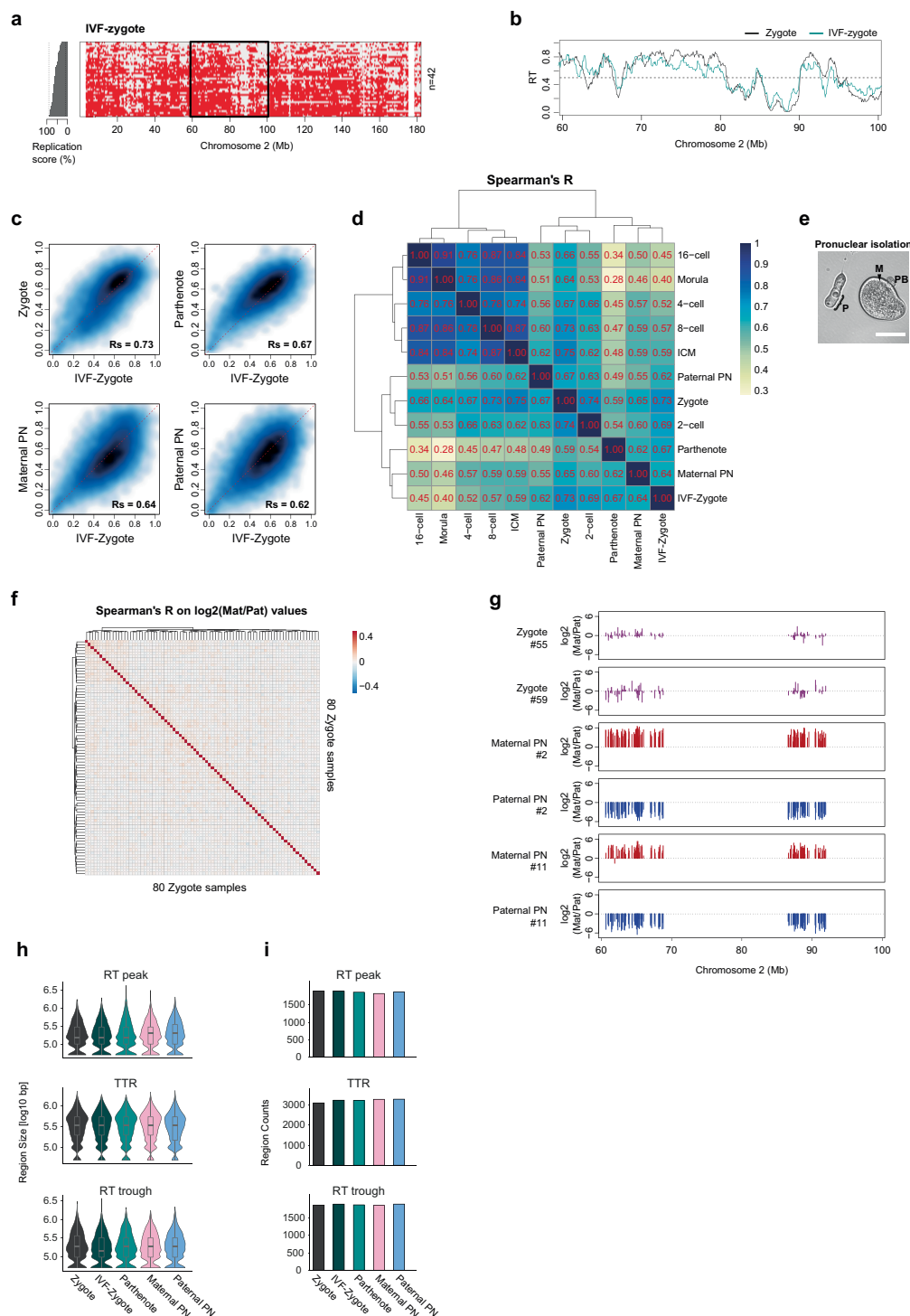
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Unique pattern of RT in zygote is not due to differences in replication between maternal and paternal alleles. a. Representative heatmap depicting binarized replication status of all single cells in zygotes produced by IVF. Cells are ranked by their percentage of replicated genome (replication score), which is plotted as a bar plot on the left. **b**. Average RT profile of IVF-derived zygotes at the chromosome 2 region indicated by a black rectangle in a. The lines indicate RT profiles calculated as the average of overlapping intervals defined by the genome-wide replication score. **c**. Smoothed scatterplot comparing the RT values in zygotes, parthenogenetic zygotes, and isolated pronuclei (PN) compared to that of IVF-derived zygote. Rs indicate Spearman's R. **d**. Correlation of genome-wide RT values between normal zygotes, zygotes produced by IVF, parthenogenetic zygotes and isolated maternal and paternal pronucleus (PN) embryos and later developmental stages using Spearman's R. **e**. Representative brightfield image of isolated paternal pronucleus and remaining maternal pronucleus in the ooplasm. M, P, and PB indicate maternal pronucleus, paternal pronucleus, and polar body, respectively. Pronuclear isolation was repeated twice independently with similar results. Scale bar, 50 μm. **f**. Correlation heatmap of log2 maternal to paternal ratios between individual zygotes after discrimination of parental origins of sequencing reads using SNPs. Allele-specific bias was calculated by computing correlation coefficients of the maternal to paternal ratios across all genomic bins in which SNPs enabled identification of parent-of-origin allele. **g**. Representative genomic tracks of the log2 maternal to paternal ratio in zygotes (magenta) and in physically isolated maternal (red) or paternal pronucleus (blue) samples after assigning parental origin based on SNPs. Regions in which there are no reads (e.g. ~65–85 Mb) correspond to regions with no SNPs. **h**., **i**. Analysis of the size (h) and number (i) of the replication features in normal zygotes compared to zygotes produced by IVF, parthenogenetic zygotes and isolated maternal and paternal pronucleus (PN). Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR.

a
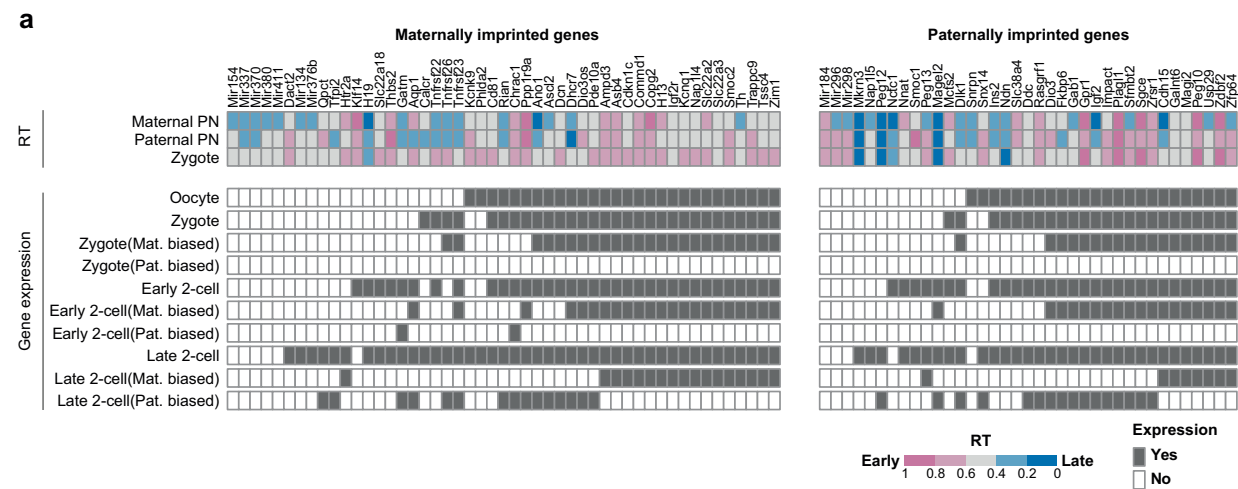
Extended Data Fig. 5 | Analysis of imprinted genes indicated no replication asynchrony. a. Analysis of RT at maternally (left) or paternally (right) imprinted genes in zygotes and in mechanically isolated paternal and maternal pronuclei (PN). RT values are shown on the top and expression data as 'yes' (detected) or 'no' (undetectable) is shown on the bottom. Expression data derives from GSE45719 and indicates expression in the indicated stage or maternal and paternal expression bias detected (grey) or undetected/absent (white).

a

Chr8 (84600001–84650001)

Tw = 20.2
M = 48.9

b

Zygote   2−cell   4-cell   8-cell   16-cell   Morula   ICM

T width

Replication Timing
Early   Late
(1.00)   (0.00)

c M-value multimodality

Dip statistic
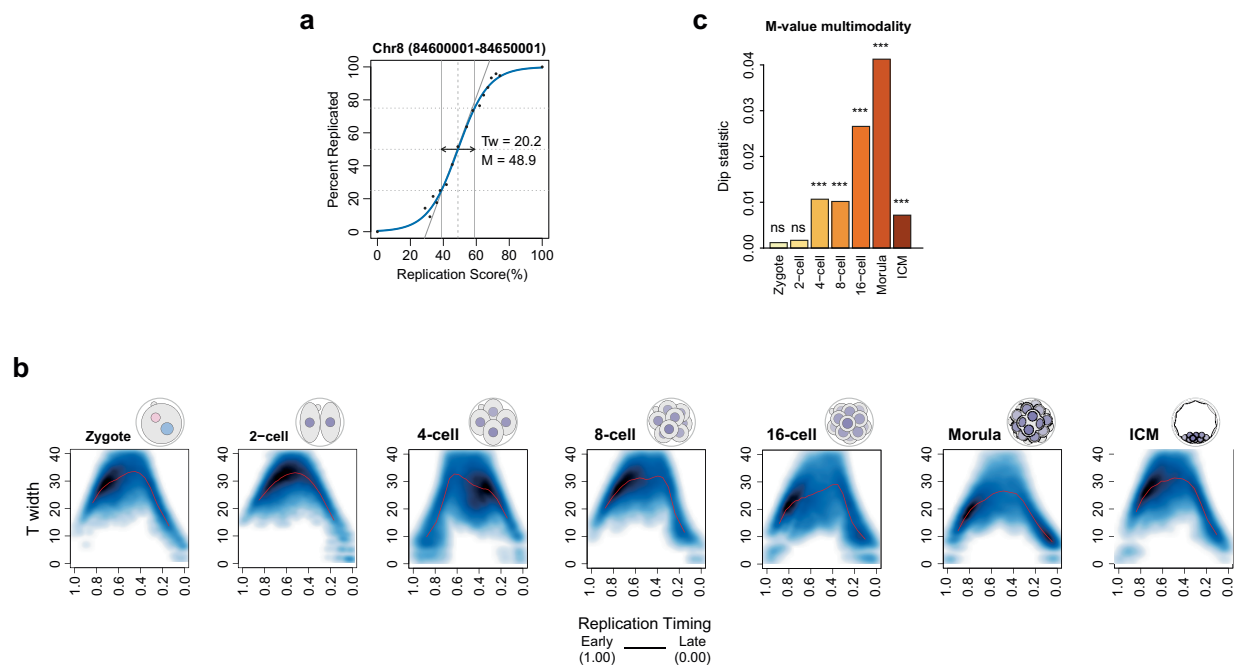
Extended Data Fig. 6 | Early and late replicating regions have less heterogeneity than that of mid replicating region. a. Representative genomic bin on chromosome 8 depicting the parameters of sigmoid curve fitting. M-value (M) represents the replication score at which 50% of the cells replicated the bin. T width (Tw) was calculated based on replication score difference between 25% and 75% of cells replicated the bin. **b**. Smoothed scatter plots of T width along progression of replication timing for the indicated developmental stages. **c**. Measure for multimodality of M-values during embryonic development by Hartigans' dip test. A greater dip value suggests a greater deviation from unimodal distribution (at least bimodal). Asterisks indicate significance levels (***, p-value < 0.001; ns, non-significant, as follows: for zygotes 9.8e − 01; for 2-cell 5.7e − 01; for 4-cell <2.2e − 16; for 8-cell <2.2e − 16; for morula <2.2e − 16; for ICM < 2.2e − 16).
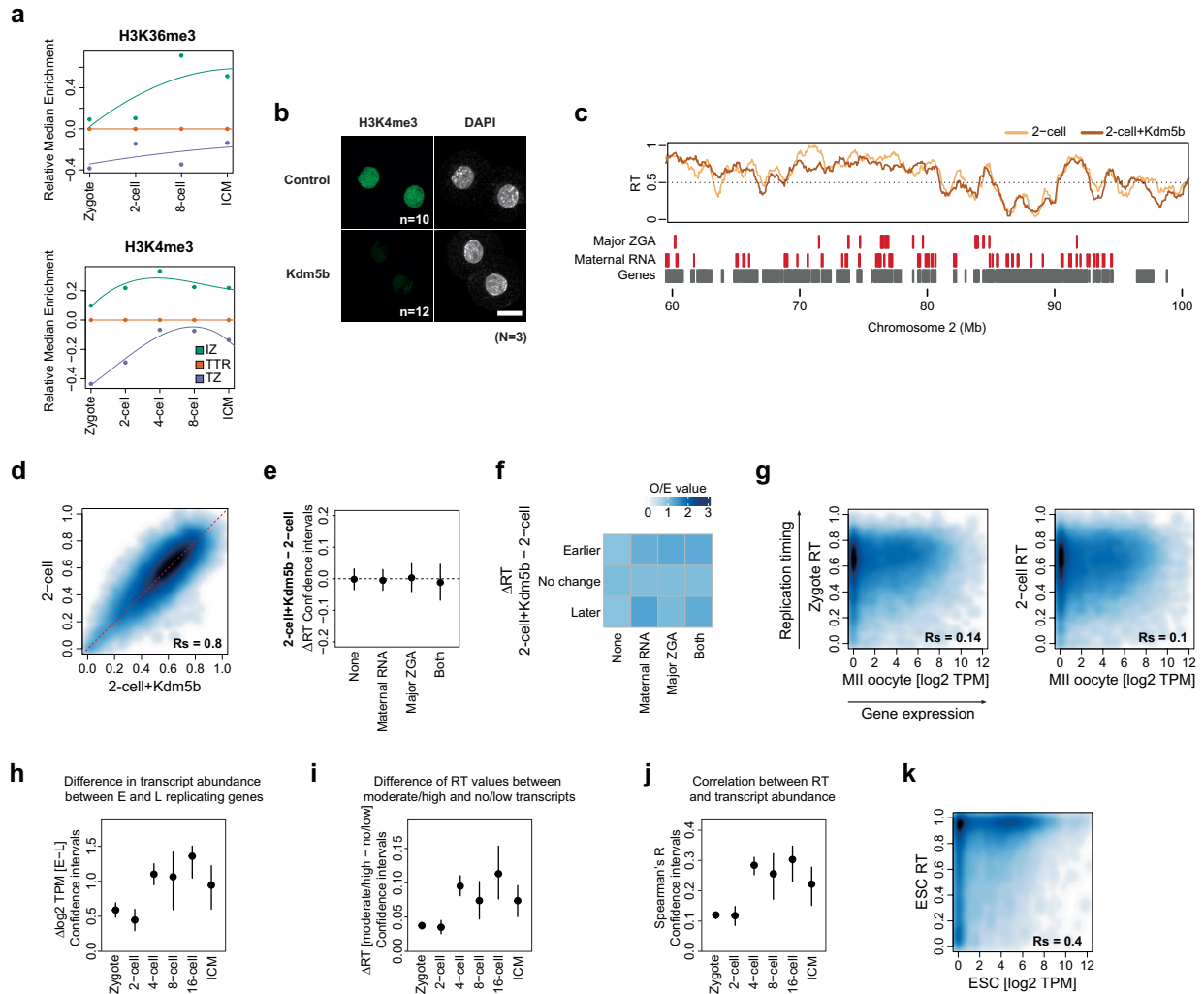
Extended Data Fig. 7 | Correlation between RT values in zygotes and maternal transcripts and analysis of histone modifications in RT peaks, TTRs and RT troughs. a. Kinetics of the relative changes in the enrichment of H3K36me3 and H3K4me3 at RT peaks and RT troughs normalised to TTRs from the zygote to the morula stage. **b.** Immunostaining of histone H3 lysine 4 trimethylation (H3K4me3) after overexpression of Kdm5b. Representative maximal projection images are shown. Total number of embryos (n) analysed in each condition from three independent experiments (N) are shown. Scale bar, 25 µm. **c.** RT profiles of 2-cell stage embryos overlayed with those from Kdm5b-overexpressed 2-cell embryos. Genomic positions of indicated gene classes according to DBTMEE[54] are shown as rectangles. **d.** Smoothed scatterplot of RT values in normal 2-cell embryos versus Kdm5b-overexpressed 2-cell embryo. Spearman's correlation (Rs) is indicated. **e.** Confidence intervals for the changes of RT (ΔRT) between Kdm5b-overexpressed and untreated 2-cell embryos of genomic bins containing maternally expressed genes or Major ZGA genes. 'Both' refers to bins containing ZGA genes and maternally expressed genes, whereas 'None' does not overlap with any of the two categories. **f.** Enrichment of genomic regions displaying a significant change in RT upon Kdm5b expression in bins containing maternally

expressed genes or Major ZGA genes. 'Both' refers to bins containing ZGA genes and maternally expressed genes, whereas 'None' does not overlap with any of the two categories. Observed over expected number of bins is shown (O/E). **g.** Smoothed scatterplots showing correlations between transcript levels (log2 TPM) of Metaphase II (MII) stage oocytes with the RT values of zygote and 2-cell stage embryos. Rs indicates Spearman's R. **h.** Confidence intervals for the difference of transcript levels (Δlog2 TPM) between early (E) vs. late (L) replicating genes. Genomic bins with an RT value greater than 0.5 were considered as Early and with RT value lower than 0.5 as Late. **i.** Confidence intervals for the difference of replication timing (ΔRT) between genes with moderate/high vs. no/low transcript levels. Genes with a transcript level (log2 TPM) greater than 1 were considered moderate/high and with a value lower than 1 as no/low expressed. **j.** Confidence intervals for the Spearman's correlation between RT and transcript abundance. **k.** Smoothed scatterplot showing correlation between transcript levels (log2 TPM) and RT values in mouse ES cells. Rs indicates Spearman's R. In e, h-j the dot represents the mean of 1000 bootstrapped values. Error bars indicate the 95% bootstrap confidence interval.
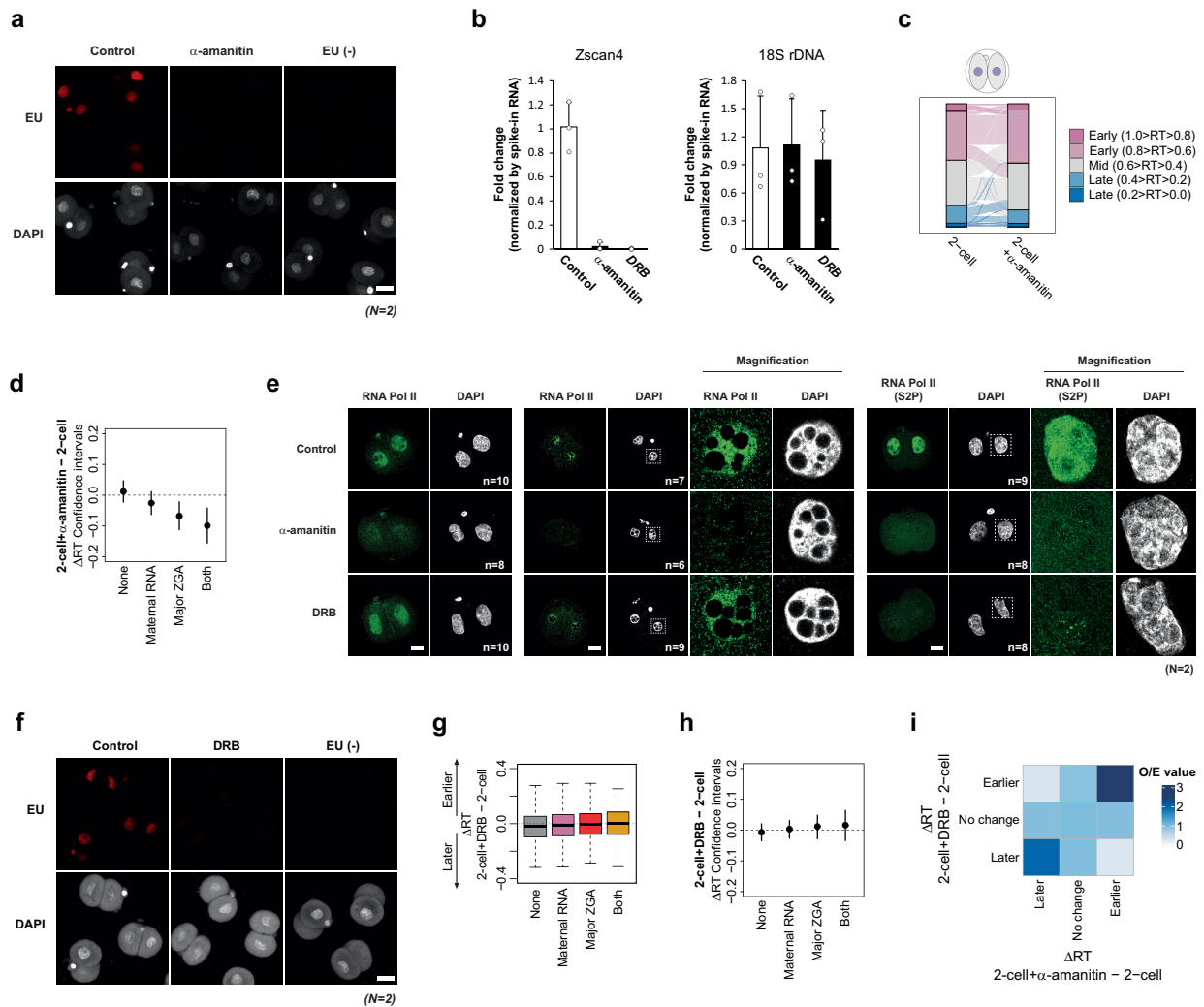
Extended Data Fig. 8 | Effect of RNA Pol II inhibition by α-amanitin and DRB on the embryonic RT programme. a. Visualisation of global transcription during minor and major ZGA by EU click chemistry and efficient inhibition of ZGA using α-amanitin. Representative embryos of a total of 24 (control), 19 (α-amanitin treated) or 19 non-EU treated embryos (EU-) are shown. Scale bar, 50 μm. **b**. Taqman RT-qPCR analysis for Zscan4 cluster and rDNA after α-amanitin and DRB treatment. Barplots show mean ± s.d and dots indicate the values of independent biological replicates. **c**. Alluvial plot indicating the RT values categorised in 5 groups from the earliest (1.0 ≥ RT > 0.8) to latest RT (0.2 > RT ≥ 0.0) across the genome in control 2-cell embryos and their changes upon α-amanitin treatment. **d**. Confidence intervals for the changes in RT (ΔRT) upon α-amanitin treatment of genomic bins containing maternally expressed genes or major ZGA genes. 'Both' refers to bins containing ZGA genes and maternally expressed genes, whereas 'None' does not overlap with any of the two categories. **e**. Immunostaining of RNA Pol II using an antibody recognizing all forms of RNA Pol II or an antibody against its CTD Serine 2 phosphorylated form (S2P) after α-amanitin or DRB treatment with (right) and without (left) Triton pre-extraction. Representative single confocal sections are shown. Total number of embryos (n) analysed in each conditions from two

independent experiments (N) are shown. Scale bars, 25 μm. We note that α-amanitin leads to degradation of RNA PolII in our experimental conditions. **f**. Visualisation of global transcription during minor and major ZGA by EU click chemistry and efficient inhibition of ZGA upon DRB treatment. Representative embryos from two independent experimetns (N) are shown. Scale bar, 50 μm. **g**. Difference of RT values (ΔRT) between DRB-treated and untreated 2-cell embryos at genomic bins overlapping only major ZGA genes, only maternal RNA genes, or both genes compared to non-overlapping bins (None). Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR. **h**. Confidence intervals for the changes in RT (ΔRT) upon DRB treatment of genomic bins containing maternally expressed genes or Major ZGA genes. 'Both' refers to bins containing ZGA genes and maternally expressed genes, 'None' does not overlap with any of the two categories. **i**. Enrichment of genomic regions displaying significant changes in RT upon α-amanitin treatment with bins that display changes in RT upon DRB treatment in 2-cell stage embryos. Observed over expected number of bins is shown (O/E). In d and h, the dot represents the mean of 1000 bootstrapped values. Error bars indicate the 95% bootstrap confidence interval.
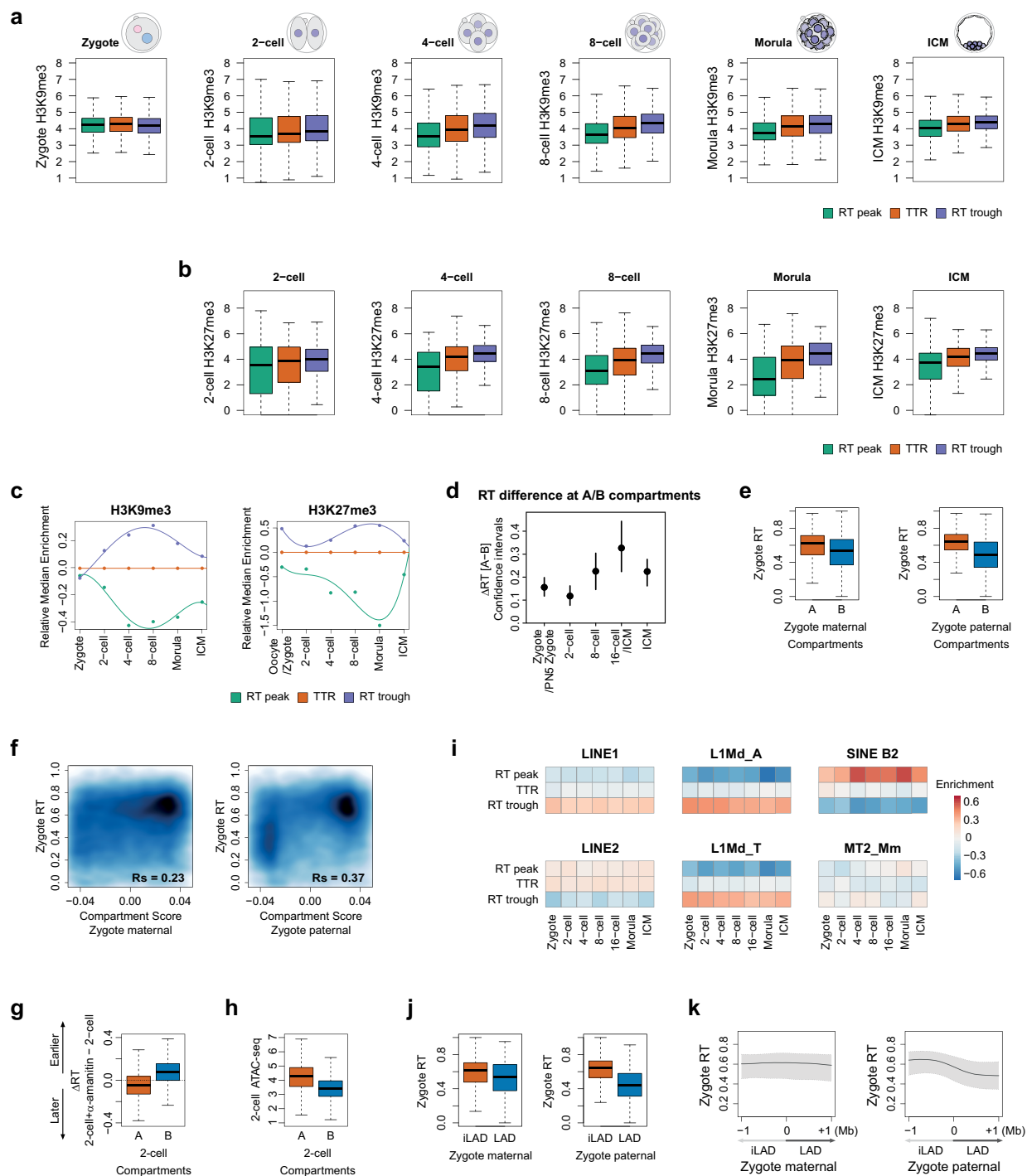
149
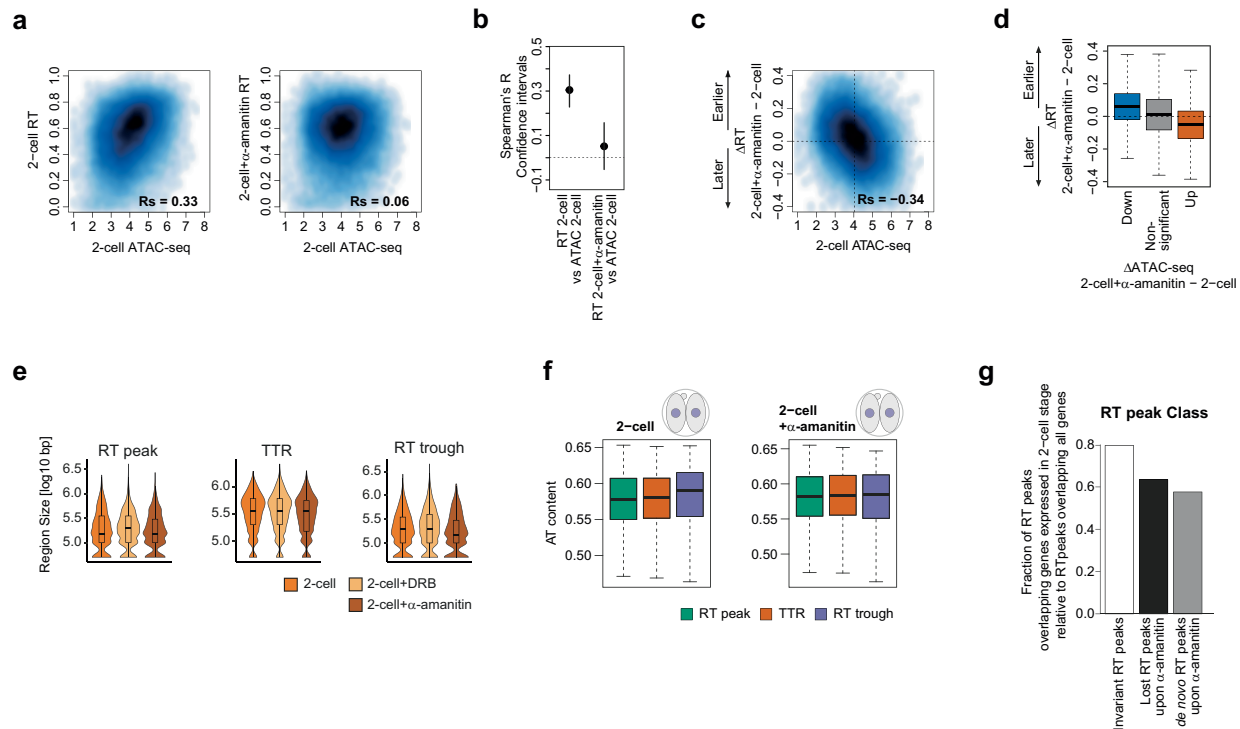
Extended Data Fig. 9 | Relationship between ATAC-seq and RT changes upon transcriptional inhibition. a. Smoothed scatterplot showing correlation between ATAC-seq signal and RT values in 2-cell stage embryos (left) and in α-amanitin treated 2-cell stage embryos (right). Rs indicates Spearman's R. **b**. Pairwise Spearman´s correlation coefficients (R) between RT and ATAC-seq signal in untreated and in α-amanitin treated 2-cell stage embryos. Error bars indicate the 95% bootstrap confidence interval. Dot represents the mean of 1000 bootstrapped values. **c**. Smoothed scatterplot depicting the difference of RT values (ΔRT) between α-amanitin treated and untreated 2-cell embryos against ATAC-seq signal in control 2-cell stage embryos. **d**. Difference of RT values (ΔRT) at genomic bins that significantly lose accessibility (down), gain

accessibility (up) or remain unchanged (non-significant) upon α-amanitin treatment in 2-cell stage embryos. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR. **e**. Size of RT peaks, TTRs and RT troughs in control versus α-amanitin or DRB treated 2-cell embryos. **f**. A + T content in RT peaks, TTRs, and RT troughs in 2-cell and α-amanitin treated 2-cell embryos. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR. **g**. Fraction of RT peaks containing genes expressed at the 2-cell stage relative to all genes in RT peaks specific to 2-cell stage embryos upon α-amanitin treatment (de novo), in RT peaks specific to control 2-cell stage embryos (lost) and RT peaks present in both 2-cell control and α -amanitin treated embryos.

Zygote

Zygote H3K9me3

2-cell

2-cell H3K9me3

4-cell

4-cell H3K9me3

8-cell

8-cell H3K9me3

Morula

Morula H3K9me3

ICM

ICM H3K9me3

■ RT peak ■ TTR ■ RT trough

b

2-cell

2-cell H3K27me3

4-cell

4-cell H3K27me3

8-cell

8-cell H3K27me3

Morula

Morula H3K27me3

ICM

ICM H3K27me3

■ RT peak ■ TTR ■ RT trough

c

H3K9me3

Relative Median Enrichment

Zygote 2-cell 4-cell 8-cell Morula ICM

H3K27me3

Relative Median Enrichment

Oocyte /Zygote 2-cell 4-cell 8-cell Morula ICM

■ RT peak ■ TTR ■ RT trough

d

RT difference at A/B compartments

ΔRT [A−B] Confidence intervals

Zygote /PN5 Zygote 2-cell 8-cell 16-cell /ICM ICM

e

Zygote RT

A B
Zygote maternal
Compartments

Zygote RT

A B
Zygote paternal
Compartments

f

Zygote RT

Rs = 0.23

Compartment Score
Zygote maternal

Zygote RT

Rs = 0.37

Compartment Score
Zygote paternal

i

LINE1 **L1Md_A** **SINE B2**

RT peak
TTR
RT trough

Enrichment

LINE2 **L1Md_T** **MT2_Mm**

RT peak
TTR
RT trough

Zygote 2-cell 4-cell 8-cell 16-cell Morula ICM

g

Earlier

ΔRT 2-cell+α-amanitin − 2-cell

Later

A B
2-cell
Compartments

h

2-cell ATAC-seq

A B
2-cell
Compartments

j

Zygote RT

iLAD LAD
Zygote maternal

Zygote RT

iLAD LAD
Zygote paternal

k

Zygote RT

−1 0 +1 (Mb)
iLAD LAD
Zygote maternal

Zygote RT

−1 0 +1 (Mb)
iLAD LAD
Zygote paternal

Extended Data Fig. 10 | See next page for caption.

Article

152

Part V
Unpublished manuscript

RIF1 regulates the consolidation of replication timing in early mouse embryos independently of changes in nuclear organization towards the nuclear lamina

RIF1 regulates the consolidation of replication timing in early mouse embryos independently of changes in nuclear organization towards the nuclear lamina

Tsunetoshi Nakatani[1], Tamas Schauer[1], Mrinmoy Pal[1], Andreas Ettinger[1], Luis Altamirano-Pacheco[1], Julia Zorn[2], David M. Gilbert[3] & Maria-Elena Torres-Padilla[1, 4] *

[1] *Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München D-81377 München, Germany*

[2] *Core Facility Laboratory Animal Services, Helmholtz Zentrum München D-81377 München, Germany*

[3]*Laboratory of Chromosome Replication and Epigenome Regulation, San Diego Biomedical Research Institute, San Diego, CA 92121*

[4] *Faculty of Biology, Ludwig-Maximilians Universität, München, Germany.*

* Correspondence: torres-padilla@helmholtz-muenchen.de

Summary

Our cells must duplicate their genome before they divide and ensure that genetic and epigenetic information is faithfully passed to their daughters. The genome is replicated with a defined temporal order referred to as Replication Timing (RT). RT is cell-type specific and is tightly linked to the 3D-organization of the genome. During development in mammals, RT is initially not well-defined and becomes progressively consolidated from the 4-cell stage. However, the molecular regulators are unknown. Here, by performing loss-of-function analysis coupled with genome-wide investigation of RT in mouse embryos, we identify RIF1 as a regulator of the progressive consolidation of RT *in vivo*. Embryos depleted for RIF1 show DNA replication features characteristic of an early more totipotent state. RIF1 regulates the progressive stratification into early and late RT values during development and its depletion leads to global RT changes and a more heterogenous RT program. Remarkably, developmental changes in RT are disentangled from changes in transcription and nuclear organization, specifically association with the nuclear lamina. Our work provides molecular understanding into the regulators of replication timing, transcription and genome organization at the beginning of mammalian development.

Keywords: replication timing, RIF1, early mouse embryos, single-cell Repli-seq

Introduction

Genome duplication prior to each cell division is fundamental for the transmission of genetic information. The process of DNA replication is tightly regulated to ensure that replication occurs only once per cell cycle and to guarantee the faithful propagation of the genome. DNA replication occurs in a temporally-coordinated manner, whereby specific regions of the genome are replicated at a given time during S-phase. This defined temporal order at which genome replication occurs is referred to as replication timing (RT).[1,2] RT is also closely linked to chromatin restoration during S-phase, presumably because the chromatin modifier protein complexes associated with the replisome during early S-phase and late S-phase differ.[3,4] For example, histone H3K4me3 methyltransferases are more abundant in replicated chromatin from euchromatic, early replicating regions.[3] Accordingly, genomic regions replicating earlier are typically euchromatic whereas heterochromatin tends to replicate late in S-phase across most cell types.[5,6]

RT is thought to be established in G1 of the cell cycle[7] and is executed by the initiation of replication at specific regions, referred to as initiation zones, in which origins of replication are activated in a coordinated fashion.[8,9] RT is considered an epigenetic fingerprint and is cell type specific.[2] RT has also been associated with the restoration of chromatin states[3,4] during replication and thus understanding the molecular regulation of RT is fundamental for our understanding of the faithful transmission and re-establishment of chromatin states. During embryonic development in mammals, RT is initially not well defined in zygotes and 2-cell stage embryos, but becomes progressively defined as development proceeds, from the 4-cell stage onwards.[10] In mouse embryonic stem cells, replication domains also consolidate upon differentiation, primarily by fusing together into larger domains.[5] In mouse embryos, the emergence of the RT program involves a progressive decrease in the heterogeneity of RT and the fusion into larger initiation zones and is accompanied by the segregation into well partitioned early and late RT values throughout the S-phase.[10,11] Thus, the establishment of RT is an integral part of the extensive chromatin remodelling period at fertilization in mammals.

The regulatory mechanisms of RT are largely unknown but are likely to occur by regulating the local probability of initiation. Amongst the factors regulating origin firing, RIF1 has been shown to suppress firing of late replication origins,[12] and recent studies have also indicated a role for RIF1 in ensuring early replication of highly transcribed genes.[13] RIF1 is dispensable for embryonic stem cell renewal in both mouse and human embryonic stem cells.[14,15] However, RIF depletion in human embryonic stem cells leads to a complete erasure of the RT program and downstream effects on histone modifications and 3D genome organization.[15] Most

importantly, this work demonstrated that RT acts upstream of the epigenetic make-up of human cells.[15] In mouse ES cells, RIF1 loss also leads to altered RT, activation of DNA replication checkpoint response and decreased cell viability without arresting proliferation.[16] While female null embryos die irrespective of the genetic background, male RIF1 null embryos survive, albeit at reduced frequencies, in some genetic backgrounds.[17-20]

In zebrafish, RIF1 is not essential for embryonic development, but zygotic depletion of RIF1 impairs female sex determination,[21] suggesting species and cell-specific function of RIF1. RT has been mapped during early development in zebrafish, where a defined temporal RT order is already detected in pre mid-blastula-transition stage embryos, that is, prior to zygotic genome activation.[22] This contrasts to the mouse, in which RT is not yet fully defined at the time of zygotic genome activation.[10] Recent whole-genome RT analyses conducted before and after zygotic transcription indicated that RIF1 'sharpens' RT profile genome-wide during zebrafish development.[21] In mice, RIF1 is deposited as a maternal protein[23] and is present as several isoforms of different length, which have been recently documented using different antibodies.[14] However, whether RIF1 plays a role *in vivo* at the earliest developmental stages after fertilization in mammals has not been addressed.

The separation into early and late replication domains is also associated with the 3D organization of the genome, with late replication domains corresponding to B-compartments and to Lamina-Associated Domains (or LADs)[2,24] and early replication domains corresponding to A-compartments and inter-LADs (iLADs).[25,26] The control of RT and 3D genome organization may obey to independent and convergent mechanisms, and is currently an area of intense research.[27] In mouse embryos, the relationship between RT and LADs emerges at distinct developmental times. LADs are established immediately after fertilization potentially priming early and late replicating domains.[28] In contrast, A- and B- compartments, although detectable in zygotes, undergo developmental maturation by a progressive increase in compartment strength[29,30] and partitioning of early and late RT during early development coincides with the maturation of A and B compartments.[10] While there is a clear structural correlation between these three pillars of nuclear organization, recent work indicates that they can be molecularly disentangled during embryonic development. For example, while RT is only mildly affected upon transcriptional inhibition at zygotic genome activation,[10] LADs are severely remodelled in the absence of transcription and embryonic LADs at the 2-cell stage are fully dependent on transcriptional activity.[31] Thus, the molecular dependencies between genome organization and RT remain unclear and whether they are regulated by common molecular pathways await investigation.

2

In spite of the importance of RT for epigenome maintenance, the molecular regulators of RT during development *in vivo* are not known. Here, we report that RIF1 regulates RT in mouse embryos. Depletion of RIF1 leads to global changes in RT that are characteristic of a more immature, less defined RT program. These changes are accompanied by an increase in the heterogeneity of the RT program and by a reduction in replication fork speed, which is also characteristic of earlier developmental stages,[32] indicating that RIF1 orchestrates the process of DNA replication at different levels. By addressing RIF1 function at three distinct developmental times we demonstrate that RIF1 regulates RT independently of changes in gene expression and lamina association. Our work identifies a key regulator of the developmental consolidation of RT during the establishment of the epigenome at the beginning of development and provides evidence for a non-interdependence of the layers of genome organization.

Results

RIF1 depletion in mouse embryos results in a less coordinated replication timing program

We recently reported that the RT program is progressively consolidated during pre-implantation development,[10] aligning with the gradual increase in compartment strength.[29,30] To gain a molecular understanding of how RT is established, we aimed to identify molecular regulators of embryonic RT. We asked whether proteins known to regulate RT in other model systems are involved in this process. A strong candidate to be involved in this process is RIF1 because RIF1 can regulate RT in human and mouse ES cells and has been shown to regulate replication timing of heterochromatin after ZGA in *Drosophila*.[15,16,33-35] In addition, RIF1 has been shown to regulate the maturation of RT in zebrafish.[21]

RIF1 is present as a maternally inherited cytoplasmic protein in mouse zygotes and is expressed throughout pre-implantation development.[14,23] To address whether RIF1 orchestrates RT establishment, we performed loss-of-function experiments for RIF1 in embryos and investigated potential changes in RT at three different times during development. We aimed to deplete RIF1 from the 4-cell stage, the time at which RT starts to consolidate in mouse embryos,[10] which also coincides with the detection of RIF1 isoforms in the nucleus.[14] We performed siRNA for RIF1 in zygotes and generated single cell Repli-seq (scRepli-seq) data at the 4-cell, 8-cell and morula stages (Fig. 1A). Overall, we sequenced a total of 416 single cells (Table S1). We confirmed that RIF1 protein was depleted from the 4-cell stage onwards and until the blastocyst stage by performing immunostaining using an antibody recognising the nuclear RIF1 isoforms (Fig. 1B-D and Fig. S1A-D).[14] . This antibody is

3

expected to recognize full-length and some other RIF1 isoforms that are not full-length as well[14]. Of note, depletion of RIF1 did not affect developmental progression to the blastocyst stage (Fig. S1E) and we did not observe overt morphological abnormalities in these embryos. However, we noted an increase in chromosome imbalance as development proceeds to later developmental stages as determined by the coefficient of variation for the average sequencing read coverage per chromosome (Fig. S1F), suggesting a potential effect on genome stability upon RIF1 depletion. To address this and considering that RIF1 plays a role in the DNA damage response pathway, we quantified the levels of phosphorylated H2A.X (γH2A.X) at the morula stage, a proxy for checkpoint activation and DNA-damage response downstream of ATR activity.[36] γH2A.X levels did not increase – and in fact decreased - upon RIF1 depletion (Fig. S1G), suggesting that there is no detectable DNA damage or activation of the DNA damage response in embryos in the absence of RIF1. Importantly, RIF1 depletion did not alter the total number of cells per embryo at the blastocyst stage compared to controls (Fig. S1H), indicating no effect in cellular proliferation. Instead, we noted an increase in the number of cells in mitosis per embryo, suggesting that while RIF1 depletion does not majorly affect cellular proliferation per se, its absence may lead to a prolonged mitosis. This is in line with previous observations in RIF1-depleted human cells, which show an accumulation of cells with a G2/M DNA content [15,37] Thus, we conclude that RIF1 is largely dispensable for development until the blastocyst stage.

We next generated RT profiles in 4-cell, 8-cell and morula stage embryos depleted of RIF1, compared to siRNA controls at the corresponding developmental stage. Sorting cells by their extent of genome replication (replication score) revealed expected progression through the S-phase in RIF1-depleted embryos, with the typical early and late replication patterns (Fig. 1E-G). A visual inspection of RT profiles suggests that RT was globally maintained in embryos in spite of RIF1 depletion (Fig. 1E-G). However, this analysis also indicated a less defined, fuzzier replication pattern across cells and throughout the genome in RIF1-depleted embryos, in particular at the 8-cell and morula stages (Fig. 1E-G). This suggests a less coordinated RT program upon RIF1 depletion after the 4-cell stage. To address this quantitatively, we computed a matrix with the Manhattan distance across cells between all pairs of genomic bins based on the binarized, replicated/unreplicated data. A lower Manhattan distance indicates more similar bins overall and higher distances indicate more dissimilar bins and therefore less coordinated RT program. In line with previous work, we observe a higher coordination as development proceeds from the 4-cell stage to the morula, in controls (Fig. 1H-J). This analysis also indicates that the coordination of the RT program is decreased upon RIF1 depletion at all 3 stages analyzed, as the distances between bins are overall higher upon RIF1

4

depletion (Fig. 1H-J). Interestingly, comparing developmental stages suggests that, while the coordination of the RT program also increases in the absence of RIF1, it does so to a lesser extent than in the corresponding controls (Fig. 1H-J). We also computed the variability score, which measures the variance of the replication program across cells for each genomic bin.[10] In line with previous findings, this analysis confirmed that the variability of the RT program decreases progressively from the 4-cell stage and further confirmed that depletion of RIF1 increases the variability score at the 8-cell and morula stages (Fig. 1K). Statistical analysis using bootstrap (1000 iterations) indicated that the average variability score is similar between controls and RIF1-depleted embryos at the 4-cell stage but is significantly increased upon RIF1 depletion at the 8-cell and morula stages (Fig. S1I). These results establish that RIF1 is required for the progressive acquisition of the coordinated RT program that normally occurs during development from the 4-cell stage. We conclude that RIF1 depletion results in a more variable, less defined RT program in pre-implantation embryos, suggesting that RIF1 mediates the consolidation of the embryonic RT program.

RIF1 regulates the consolidation of RT in mouse embryos

The above data indicates that RIF1 regulates the coordination of RT in mouse embryos, which prompted us to address whether RIF1 mediates the developmental consolidation of the RT program. To address this directly and with further depth, we next examined replication features of RIF1-depleted embryos. In particular, since the developmental consolidation of RT occurs primarily at the level of RT peaks (also known as initiation zones) and timing transition regions (TTRs),[10] we extracted these features from our scRepli-seq datasets as previously described.[10,8] We also analyzed RT troughs (also known as termination zones), which are the regions in which replication forks converge.[38] Control embryos showed the expected consolidation trend as development proceeds, with a reduction of the number of RT peaks, RT troughs and TTRs overall (Fig. 2A).[10] The number of RT peaks, TTRs and RT troughs was not affected upon RIF1 depletion at the 4-cell stage (Fig. 2A). However, the number of all these three features was higher in RIF1-depleted 8-cell and morula stage embryos compared to controls (Fig. 2A). Indeed, bootstrapping (1000 iterations) and calculation of confidence intervals[10] indicated that the number and size of RT features were significantly different between controls and RIF1-depleted embryos at the 8-cell and morula stages, but not at the 4-cell stage (Fig. S2A-B). These data indicate that the consolidation of the RT features past the 4-cell stage is prevented upon RIF1 depletion. This was accompanied by a consequent reduction in the size of RT peaks, TTRs and RT troughs compared to the controls (Fig. 2A), pointing towards a more fragmented RT program in the absence of RIF1, in line with our interpretation of a less consolidated program upon RIF1 depletion.

5

We next asked whether RIF1 regulates the order in which the genome replicates during S-phase. For this, we investigated the distribution of the genome into early and late replication. Globally, the RT patterns were maintained across all genomic bins in RIF1-depleted embryos compared to controls at all stages analyzed (Fig. S2C). However, while the genome-wide correlations of RT values did not differ much between controls and RIF1-depleted embryos, the skewed distribution of the RT values along the diagonal suggested a deviation from the RT values, particularly at the 8-cell and morula stages (Fig. S2C). Indeed, genomic regions replicating early shift to later replication, and late replicating regions replicate earlier upon RIF1 depletion (Fig. S2D). To further investigate this, we stratified the genome into RT values from the earliest (RT > 0.8) to the latest (RT < 0.2) with increments of 0.2 RT values. We first reanalyzed previous data from non-manipulated embryos, which indicate a progressive partitioning into more extreme RT values across the complete S-phase (Fig. S2E).[10] Our siRNA controls reproduced previous findings showing that while most of the genome at the 4-cell stage (44%) shows intermediate RT values (0.6 > RT > 0.4), the genome partitions into values spreading into the complete S-phase progressively thereafter (24% and 21% of the genome in the 8-cell and morula stage displays RT values greater than 0.4 and less than 0.6, respectively) (Fig. 2B). In contrast, the same analysis in RIF1-depleted embryos indicated that the distribution of the genome across the S-phase in 8-cell and morula stage embryos resembled that of the 4-cell stage instead of its corresponding 8-cell stage control (Fig. 2B). A more detailed analysis of RT values of all genomic regions indicated that while overall early replicating regions remain so in all conditions, they shift towards earlier replication timing in control 8-cell stage embryos but not in 8-cell embryos depleted for RIF1 (Fig. 2C). This suggests that RIF1 depletion interrupts the naturally occurring developmental shift to earlier replication of those regions. We observed a similar pattern in morula stage embryos (Fig. 2C). Likewise, "mid" S-phase replicating regions shift to later replication in control 8-cell embryos but not in embryos upon RIF1 depletion (Fig. 2C). Examining RT profiles across the genome revealed that indeed RT values tend to move towards mid values, with some regions that replicate early in controls shifting to later and regions that replicate late shifting to earlier in the absence of RIF1 (Fig. 2D). In fact, overlaying the genome-wide distribution of RT values indicated that while control 8-cell and morula embryos RT values separate towards earlier and later RT values, the RIF1-depleted embryos do not (Fig. 2E). These differences were statistically significant (Fig. S2F). Instead, their distribution resembles that of 4-cell stage embryos (Fig. 2E). We also calculated the M-value, which is a measure of the replication score at which half of the cells replicate a particular genomic bin.[10] The distribution of the M-values reflects the partitioning of the RT values across the genome. For example, M-values for mouse embryonic stem cells and for differentiated cells depict a bimodal distribution, indicating a well-spread partitioning of the genome into early versus late RT.[8,10] In contrast, mouse embryos

6

prior to the 4-cell stage show a unimodal distribution.[10] In agreement, siControl and siRif1 embryos showed a similar, largely unimodal distribution at the 4-cell stage (Fig. S2G). Interestingly, in the absence of RIF1, 8-cell and morula stage embryos the spread of M-values (IQR) was similar to control 4-cell stage embryos but differed from control 8-cell and morula embryos (Fig. S2G-H). This indicates that RIF1 regulates the progressive segregation of the genome into early and late replicating domains during development *in vivo*.

Lastly, we computed the T_{width}, a parameter that reflects the heterogeneity in RT across cells.[38] We find that depletion of RIF1 significantly increases the T_{width} value in comparison to controls at the 8-cell and morula stages, but not at the 4-cell stage (Fig. S2I-J). Thus, RIF1 depletion alters the heterogeneity of RT, suggesting that RIF1 contributes to the robustness of the emerging embryonic replication program by limiting cell to cell variability. Notably, in all the above analyses, the replication features of RIF1-depleted embryos resembled those of control embryos at earlier stages. Overall, the above data suggest that RIF1 depletion results in an immature RT program.

We then asked whether the developmentally earlier, more immature RT program upon RIF1 depletion also involves the molecular properties of the DNA replication process itself, for example, the replication fork dynamics. To address this, we performed DNA fibre analyses to measure replication fork speed in control and RIF1-depleted embryos at the morula stage. In normal development, replication fork speed is initially slow and increases as development proceeds.[32,39] Remarkably, we find that depletion of RIF1 leads to a slower replication fork speed (Fig. 2F), with morula embryos depleted of RIF1 replicating with the same fork speed as control 8-cell stage embryos.[32] This is accompanied by a reduction in the inter-origin distance (Fig. 2G). These observations suggest that RIF1-depleted embryos fire more origins, which characterises the earliest stages of development,[32] and supports our interpretations of a more immature DNA replication program upon RIF1 depletion. Thus, we conclude that depletion of RIF1 slows replication fork, resulting in features of the replication fork that are characteristic of an earlier developmental stage. These data also suggest that the developmental acquisition of an orderly RT program and the increase in replication fork speed during pre-implantation development may be functionally related.

RIF1 establishes developmental patterns of RT

Considering the temporal, developmental specific phenotypes upon RIF1 depletion, we next investigated whether RIF1 regulation of RT is stage-specific. Specifically, we asked whether changes in RT are inherited to the next developmental stage or whether the RT changes elicited upon RIF1 depletion are specific to each developmental stage. To reveal potential

7

developmental patterns, we first performed a principal component analysis (PCA) of the RT values from all control (siControl) and RIF1-depleted (siRif1) embryos together with non-manipulated, wild-type embryos from the zygote to the morula stages.[10] As expected, the zygote and 2-cell stage clustered away from all other later stages (Fig. 3A). In addition, this analysis revealed that, while control embryos clustered with their respective non-manipulated stage, RIF1-depleted embryos consistently clustered with earlier developmental stages (Fig. 3A). Namely, 8-cell RIF1-depleted embryos clustered together with 4-cell stage embryos and morula RIF1-depleted embryos clustered closer to the 8-cell controls, than to the morula controls (Fig. 3A). Overall, this suggests that RIF1 may function to set developmental specific RT changes that, upon depletion, are not properly established. Interestingly, all 4-cell stage embryos clustered largely together, whether controls or upon depletion of RIF1 (Fig. 3A). This suggests that the function of RIF1 in establishing stage-specific developmental RT programs occurs concomitantly with the consolidation of the RT program, which is known to take place from the 4-cell stage.[10]

To further examine a potential role of RIF1 in establishing developmental RT programs, we next analyzed RT changes between subsequent stages. We first compared the differences in RT (ΔRT) that occur between the morula and the 8-cell stage with those occurring between the 8- and the 4-cell stage in control embryos. Genome-wide analysis of the differences in RT between stages revealed no correlation between changes in RT from the 4- to the 8-cell stage and changes in RT between 8-cell and morula stages (Fig. S3A). In other words, regions that become replicated earlier at one stage do not become replicated earlier at the subsequent developmental stage (Fig. S3A). Likewise for genomic regions replicating later between stages (Fig. S3A). This suggests that different genomic regions 'mature' their RT towards earlier or later at subsequent developmental stages. This indicates that each stage undergoes a maturation program and that changes in RT are not necessarily related to changes in RT in the immediate earlier developmental stage. We then asked whether RIF1 regulates the same genomic regions at different developmental stages. For this, we compared the differences in RT values across all genomic bins between RIF1-depleted embryos and controls between stages. This analysis indicated that the changes in RT elicited upon RIF1 depletion at the 8-cell stage are not correlated to those at the 4-cell stage (Fig. S3B). In other words, RIF1 regulates the RT program of these two stages in a stage-specific manner. Interestingly, however, the same analysis between the 8-cell and the morula stage revealed a greater positive correlation (Fig. S3B), indicating that while most genomic regions are similarly regulated by RIF1 at these stages, some other regions are not. These data suggest that RIF1 regulates both, shared as well as stage-specific parts of the RT program in morula and 8-cell stage embryos.

8

In line with the observations above, we note that regions that change to later replication in the absence of RIF1, for example at the 8-cell stage, are not necessarily changing to later replication at the 4-cell or the morula stages (Fig. 3B). These observations raised the possibility that RIF1 acts in the *establishment* rather than in the *maintenance* of RT once it has been set-up. To address this possibility, we asked if RIF1 regulates RT of those genomic regions in particular, which change RT for the first time between subsequent developmental stages. For this, we compared the differences in RT values elicited by depletion of RIF1 in the 8-cell stage with the differences in RT emerging between the 8-cell and the 4-cell stage. We find that changes in RT that occur as development proceeds anticorrelate with RT changes elicited upon RIF1 depletion (Fig. 3C). For example, regions shifting towards earlier replication from the 4-cell to the 8-cell stage in control embryos shift towards later replication in 8-cell upon RIF1 depletion (Fig. 3C and Fig. S3C). We obtained similar results at the morula stage (Fig. 3C and Fig. S3D). These data indicate that depletion of RIF1 affects genomic regions that undergo developmental RT changes and suggest that RIF1 regulates RT changes that emerge normally between each developmental stage.

To provide a molecular understanding for these findings and considering that embryonic transcription fine tunes the RT program,[10] we next investigated whether changes in RT elicited upon RIF1 depletion are associated with transcriptional changes in previous developmental stages. We performed RNA-seq in control and RIF1-depleted 4- and 8-cell stage embryos. RIF1 depletion resulted in the mis-regulation of 175 genes at the 4-cell stage but of only one gene at the 8-cell stage (Fig. 3D, Tables S2 and S3). However, we found no correlation between the differential expressed genes and changes in RT within the same developmental stage: regions that shifted either towards earlier or later replication in the absence of RIF1 did not display changes in transcript abundance (Fig. S3E). This indicates that the changes in RT elicited upon RIF1 loss can occur independently of changes in transcriptional activity and therefore changes in RT are disconnected from changes in expression patterns globally. In addition, in control embryos, the differences in RT that emerge between the 4- and the 8-cell stage did not show any correlation with the differences in gene expression that naturally occur between these two stages (Fig. 3E and Fig. S3F). This indicates that the developmental changes in RT between these two stages are not related to changes in their gene expression profiles.

We next addressed whether RIF1 depletion affects more specifically genes, which are highly transcribed. To address this, we stratified genomic bins in quintiles based on their expression levels in wild-type embryos and calculated the differences in RT elicited by RIF1 depletion in

9

each of these quintiles. We find that, at the 4-cell stage, RIF1 depletion does not affect RT of any quintile (Fig. 3F). However, at both the 8-cell and morula stages, genomic bins that are more highly transcribed become replicated later in the absence of RIF1 (Fig. 3G-H). Thus, these data suggest that while changes in gene expression elicited by RIF1 depletion do not necessarily lead to a change in RT, some loci are more sensitive to RIF1 depletion as development proceeds, at the 8-cell and morula stages, and this can correlate with their expression levels.

Lastly, we examined whether the changes in RT at the 8-cell stage following RIF1 depletion are related to changes in transcription at the 4-cell stage but found no correlation between them (Fig. 3I). The same analysis comparing RIF1-mediated changes of RT at the morula stage indicated no relationship between changes in transcription at the 8-cell stage (Fig. 3I). Overall, these data indicate that the developmental changes in RT are disentangled from changes in transcription and that the changes elicited upon RIF1 depletion are also unrelated to changes elicited in gene expression. Considering that RT often follows changes in gene expression, we find these observations particularly relevant since they establish that changes in RT are not necessarily dependent on changes in transcription *in vivo*.

Thus, we conclude that RIF1 establishes *de novo*, stage-specific developmental RT programs, which are unrelated to changes in gene expression.

Lamina association and RT are uncoupled upon RIF1 depletion in early embryos
We next explored the relationship between the establishment of RT by RIF1 and nuclear organization. In particular, since it has been proposed that early embryonic LADs can prime early and late RT at later developmental stages,[10] we focused on LADs. We asked whether changes in RT could be explained by potential changes in lamina association elicited upon RIF1 depletion at earlier stages. For this, we first investigated whether RIF1 depletion affects nuclear organization by mapping LADs in control and RIF1-depleted embryos using DNA adenine methyltransferase identification (DamID) for LaminB1 in 4- and 8-cell stage embryos.[28,40] We generated LAD profiles using a previously established pipeline based on enrichment of methylation values by Dam-LaminB1 as a proxy for genome interactions with the nuclear lamina.[28,31,41] Interestingly, RIF1 depletion led to changes in the interactions between the genome and the nuclear lamina in both 4- and 8-cell stage embryos (Fig. 4A). We observed both regions that increased and regions that decreased interactions with the nuclear lamina (Fig. 4A). The effects of RIF1 depletion were larger at the 8-cell compared to the 4-cell stage (Fig. 4A). Of note, such changes in lamina association occur in spite of virtually no changes in gene expression in RIF1-depleted embryos at the 8-cell stage and only less

10

than two-hundred de-regulated genes at the 4-cell stage (Fig. 3D). While the total number of LADs remained similar between control and RIF1-depleted embryos at the 8-cell stage (746 and 738 LADs, respectively), the number of LADs increased upon loss of RIF1 at the 4-cell stage (601 to 644 LADs) (Fig. 4B). LADs at the 4-cell stage displayed a median size of 0.9 Mb (mean 1.30 Mb) in controls, compared to 1.0 Mb (mean 1.61 Mb) in RIF1-depleted embryos (Fig. 4B), which overall resulted in a larger proportion of the genome associated with the nuclear lamina at the 4-cell stage upon RIF1 loss (Fig. 4C). Indeed, pair-wise comparison of LADs and iLADs between controls and RIF-1 depleted embryos indicated that RIF1 depletion leads to alterations in LADs at both the 4- and the 8-cell stage (Fig. 4D).

We next asked whether the effects on RT elicited upon RIF1 loss are related to the changes in LADs and iLADs. Comparing global, genome-wide differences in RT between control and RIF1-depleted embryos against differences in lamina association at the 4-cell stage indicated no correlation between changes in RT and nuclear positioning (Fig. 4E). We obtained similar results at the 8-cell stage (Fig. 4E). Indeed, plotting the changes in RT of regions with significantly increased or decreased LaminB1 DamID methylation values indicates no changes in RT according to whether a region is repositioned towards or away from the nuclear lamina (Fig. S4A). Interestingly, we find that alterations in RT can occur regardless of the direction in which repositioning with respect to the nuclear lamina occur. That is, regions, which shift towards early replication upon RIF1 depletion can both increase and decrease the strength of their association with the nuclear lamina (Fig. 4E and Fig. S4B). The same occurs in regions shifting towards late replication in RIF1-depleted embryos (Fig. 4E and Fig. S4B). This phenotype is more marked at the 8-cell than at the 4-cell stage, presumably because the changes in RT are larger at the 8-cell stage. We also observed that genomic regions with lower LaminB1 DamID methylation levels in both control and RIF1-depleted embryos, shift to later replication upon RIF1 knockdown (Fig. S4B, see region ~165 to ~170Mb). This would suggest a role of RIF1 in RT regulation independently of lamina interactions. Indeed, overall, control 8-cell stage iLADs tend to replicate later upon RIF1 depletion (Fig. S4C) and the RT difference between LADs and iLADs becomes more equal in RIF1-depleted embryos (Fig. S4D).Thus, altogether these data indicate that there is no strict relationship between RT changes caused by loss of RIF1 and changes in lamina association within each developmental stage.

Next, we addressed whether alterations in LADs and iLADs upon RIF1 depletion could account for changes in RT at the subsequent developmental stage. Genome-wide analysis of the changes in RT emerging upon RIF1-depletion at the 8-cell stage revealed a positive, yet low correlation with changes in lamina association at the 4-cell stage (Fig. S4E). However,

this correlation decreased to practically zero when comparing RT changes (ΔRT) between control and RIF1-depleted embryos at the morula stage, with changes in LaminB1 DamID methylation levels upon loss of RIF1 at the 8-cell stage (Fig. S4E). Thus, the association to the nuclear lamina is not a determining factor for the outcome of RT changes upon RIF1 depletion. We conclude that RIF1 regulates RT independently of radial nuclear positioning, and that RT and the positioning at the nuclear lamina are molecularly disentangled in early embryos. Instead, A- and B-compartments seem to display a more consistent relationship with the changes in RT. Notably, we find that most regions shifting towards earlier RT upon RIF1 loss have a strong B compartment score and those shifting towards late have a strong A compartment score (Fig. 4F). Indeed, analysis of the chromatin features of genomic regions that change RT upon RIF1 depletion indicates that a strong compartment definition in the regions that shift towards earlier or later replication at all developmental stages analyzed (Fig. 4G). This is also reflected in the strong demarcation by a higher chromatin accessibility and higher levels of transcripts overall, in the regions replicating later upon RIF1 depletion, while the opposite is true for regions replicating earlier, which are characterised by a less accessible and less abundant transcriptome (Fig. 4G). In line with these features, the genomic regions that change RT upon RIF1 depletion towards later are also characterized by higher levels of H3K4me3 but lower levels of H3K27me3 and H3K9me3 compared to those regions, which are not affected by RIF1 loss and this both, at the 8-cell and morula stages (Fig. 4G). In contrast, genomic regions that shift RT towards earlier upon RIF1 depletion are overall enriched in H3K27me3 but depleted of H3K4me3 compared to insensitive-RIF1 regions (Fig. 4G). We propose that the organization of the genome into A- and B-compartments may have a larger influence on RT regulation -or vice versa- than the lamina association. This implies that compartments and RT may act as the core factors for chromatin organization during early development.

Finally, we investigated whether RIF1 depletion leads to global changes in histone modifications. For this, we performed immunostaining for H3K4me3, H3K9me3 and H3K27me3 at the morula stage, which is the stage where we observe the largest effects on RT following RIF1 loss. We did not observe detectable changes in any of these three histone modifications in RIF1-depleted embryos compared to controls (Fig. S4F), suggesting that RIF1 depletion does not lead to global disruption in the levels of histone modifications analyzed.

Discussion

RT is closely linked to the establishment and restoration of chromatin states.[3,42] At the same time, the establishment of new epigenetic programs relies on changes in the epigenetic make up of individual cells. The developmental consolidation of such epigenetic landscapes is

12

therefore of fundamental importance both, to enable changes in cell fate during development but also for the plasticity of cells in the early embryo required for generating several fates. Earlier work indicated that RT is not well defined initially and that consolidation and emergence of a more deterministic RT program emerges progressively from the 4-cell stage. This less-well defined property is in line with a higher chromatin plasticity, whereby specific chromatin and transcriptional programs are not yet locked in early on. Molecularly, this resembles the progressive maturation of A- and B-compartments, reflected by the gradual increase of inter-compartment genome interactions during development.[30] How the consolidation of the RT program is regulated is therefore essential to understand the dependencies between other layers of genome organization. Our work indicates that RIF1 acts as such regulator and that absence of RIF1 prevents the emergence of a consolidated RT program. Our data also provide insights into the relationship between RT and nuclear organization as well as transcription.

RIF1 plays multiple roles during replication.[17,33,43] Interestingly, however, the phenotypes elicited upon RIF1 depletion in different cell types are not fully consistent, suggesting a degree of cell-type specific regulation and/or the presence of different regulators in different cells. It is also possible that phenotypic differences resulting from loss of RIF1 may be due to different cell types and timing at which analyses were performed. For example, mouse fibroblasts depleted of RIF1 show reduced EdU incorporation suggesting defects in S-phase progression, but no detectable phenotype in the proportion of cells in G2/M.[17] They also accumulate DNA damage during S-phase.[17] Moreover, Cornacchia *et al.* showed that RIF1 deletion in pMEFs increases p21 levels, suggesting a delayed entry into S-phase. Whether S-phase progression is regulated by RIF1 in human cells is less clear, as some work has shown that S-phase progression remains overall unaffected upon depletion of RIF1 in HeLa cells[35] while other work indicates that *siRIF1* HeLa cells progress more quickly through S-phase.[43] In addition, by focusing on a 42 Mb segment of human chromosome 5, early replicating sequences replicate later in HeLa cells depleted of RIF1, whereas those replicating late advance their RT.[35] In MEFs, RIF1 deletion induces a genome-wide deregulation of RT.[33] Remarkably, human embryonic stem cells display an almost complete erasure of the RT program, primarily by increasing the RT heterogeneity between individual cells.[15] In agreement with this, our analysis indicate that RIF1 depletion in mouse embryos leads to a more heterogeneous RT program. Thus, while in mouse fibroblasts, where the G1/S checkpoint is active, RIF1 causes a delay in entry into S-phase,[33] in HeLa cells, that are p53 negative, this does not happen,[35] nor it does in immortalised fibroblasts, where p53 has been inactivated.[16] In the latter, it is rather the DNA replication checkpoint that is activated. All these raises the interesting possibility that some of the phenotypes elicited upon RIF1 depletion depend upon the checkpoint machinery of each cell type.

13

In *Drosophila,* developmental progression is accompanied with an increase in the length of S-phase, which is prevented upon RIF1 depletion and occurs at the mid-blastula transition.[34] Elegant cell biology approaches inferred delayed replication of heterochromatin upon RIF1 depletion based on imaging data documenting sequential appearance of satellite sequences.[34] Similarly, the mid-blastula transition in zebrafish is characterized by an initial S-phase lengthening and the appearance of a G2-phase.[22] Furthermore, the zebrafish RT program does not mature as in the wild type, from the shield to the 24hpf stage[21] upon RIF1 depletion. Our work extends these observations to mammals by reporting a role of RIF1 in regulating RT in mouse embryos and provides an in-depth molecular characterization of RT at the genome-wide level and on the impact of RIF1 loss on other features of nuclear organization and transcription.

In contrast to human embryonic stem cells, in which cell proliferation is not majorly impaired albeit an accumulation of cells with a G2/M DNA content upon RIF1 depletion,[15,16] mouse embryonic stem cells in which RIF1 is knocked-down show reduced proliferation, in part due to decreased cell viability[16] and, accordingly, are unable to form teratomas.[44] Interestingly, however, mouse embryonic stem cells depleted of RIF1 display a telomere phenotype but without detectable signs of DNA damage assessed for example by monitoring levels of γH2A.X and 53BP1. However, RIF1 knock-out mice do not show telomere shortening.[17] The effect on telomeres in RIF1-depleted mouse stem cells is indirect and results from the upregulated expression of ZSCAN4, which is negatively regulated by RIF1.[44] Indeed, we and others have subsequently reported that RIF1 depletion in mouse embryonic stem cells leads to the upregulation of *Zscan4* and of a '2C' transcriptional program.[45,46] This contrasts to our results in mouse embryos, in which we did not detect changes in *Zscan4* upon RIF1 depletion. Importantly also, our results indicate that changes in the RT program occur in embryos upon RIF1 depletion and this occurs in the absence of changes in transcription. Our data indicate that RIF1 regulates RT maturation, e.g. at the times during development when cell fates are acquired and that these RT changes are unrelated to the transcriptional changes in gene expression typical of those new cell identities.

RIF1 has been suggested to regulate transcription, in addition to RT, through a potential role in chromatin architecture and loop formation.[35] Evidence in human ES cells indicate that upon the first cell cycle after RIF1 depletion, only few genes (~50) genes affected, but different genes are affected in different cells.[15] However, upon further cell cycle passages, ~2000 genes become affected, which tend to be more consistent between cells. This suggests that transcriptional changes result primarily from several cycles of disrupted RT and that continued

14

proliferation in the absence of RIF1 induces progressive gene deregulation. Our data further suggest that the genome also reorganises in the absence of RIF1, with altered LAD/iLAD boundaries and changes in the nuclear lamina interactions of several LADs and iLADs. A change in chromatin architecture has also been observed in mouse ES cells, where RIF1 mediates inter-replication domain contacts.[16] Recent work in B cells in culture also suggest that RIF1 promotes early replication but has minor effect in gene expression and genome organization.[13] These findings are particularly important, as they add to the observations that these two pillars of nuclear organization, that is the radial positioning towards the nuclear lamina and RT, can be disentangled.

An interesting conclusion from our data is that preventing the consolidation of the RT (through RIF1 depletion) also leads to slower fork speed, which is a feature observed at earlier developmental stages.[32] These observations suggest that the consolidation of RT may be related to the speed at which DNA replication occurs and also, potentially, to the number of origins that are fired. However, at this point it is impossible to ascertain whether they are directly, causally related or whether these are two separate features that co-occur phenotypically in the totipotent cells of the early embryo.

Finally, the expression pattern of RIF1 in the early embryo may alone explain the normal developmental consolidation of the RT program. siRNA screenings in mouse embryonic stem cells identified RIF1 as a factor which, upon depletion, leads to a very efficient reprogramming into 2-cell-like cells (2CLCs).[44,46,47] Complementation studies showed that expression of RIF1 lacking the N-terminus HEAT repeat induces 2CLCs through a dominant negative effect, presumably by competing endogenous, full-length RIF1 function.[47] Mouse oocytes express multiple short isoforms that are primarily derived from the N-terminus of RIF1 and only a small fraction of full-length RIF.[14] This is similar at the 2-cell stage, but the proportion of full-length RIF1 increases drastically at the morula stage. While we cannot rule out the existence of additional isoforms, not detectable with currently available antibodies, our data suggest that the emergence of a more consolidated RT program correlates with the detection of RIF1 in the nucleus, which occurs from the 4-cell stage.[14] Thus, the embryo may have evolved an effective mechanism to regulate the consolidation of the RT program by regulating the isoforms and localization of RIF1 during development. We propose that a less consolidated RT program, in line with the high chromatin and cellular plasticity of the early totipotent embryonic cells, is promoted by the differential isoform localization of RIF1.

It is also interesting to note that the effects of RIF1 depletion at the 4-cell stage are milder compared to cell culture systems, including human embryonic stem cells, in which RIF1

15

170

depletion results in a clear erasure of RT.[15] This may be linked to the biology of the early mammalian embryo, considering that zygotes and 2-cell stage embryos possess an RT program that is very distinctive compared to cells in culture.[10,48,49] In contrast, as development proceeds, when embryonic cells display a more consolidated RT programme, effects upon RIF1 loss are larger. Whether other factors are in play in embryos to further control RT and/or whether the distinctive chromatin configuration of early embryos render the RT more robust to RIF1 perturbation remains to be investigated. These phenotypic differences between embryos and cells in culture also highlight the importance of investigating regulatory mechanisms during embryogenesis *in vivo*, in physiologically relevant contexts.

Overall, our work has identified a molecular regulator of replication timing and nuclear organization of the genome with the nuclear lamina during early mammalian development, two fundamental layers of the epigenome.

Limitations of the study

Our work in early embryos uses single cell Repli-seq both, because blastomeres are asynchronous and also because low input approaches are required. However, it is important to keep in mind that the scRepli-seq data is of relatively low resolution (50kb) and thus individual replication origins cannot be defined. This information would greatly propel our understanding of the regulation of origin firing and RT during development. While all the features of the RT program that we report are consistent with a less coordinated and altered RT in embryos, we cannot formally rule out that the increase in T_{witdh} values, and thus heterogeneity, may be due to sampling of some cells with lower depletion levels of RIF1. Because the use of siRNA may not target all potential isoforms of *Rif1,* the role that we document for RIF1 may be underestimated. Also, our immunostaining analysis suggests that RIF1 depletion does not lead to a global disruption in the levels of histone modifications that we analyzed but cannot rule out that changes may occur at specific genomic loci. Likewise, the cell-type specific phenotypic outcome upon RIF1 depletion remains to be studied. For this, mass-spectrometry to identify RIF1-interacting partners is a plausible approach in cells in culture but is unlikely to yield robust data in early embryos in which the amount of material is limiting. Finally, our observations that RIF1 depletion does not affect developmental progression until the blastocyst stage leave a long-standing open question, as it remains unclear whether and how essential the RT program is for development.

Acknowledgments

Author contributions

T.N. designed, performed and analyzed most of the experiments; T.S. performed all bioinformatics analyses; M.P. performed DamID experiments and contributed to their bioinformatic analysis; A.E. performed image analyses; L.A.P. contributed to code generation; J.Z. provided animal protocol expertise towards approval by the Regierung von Oberbayern; D.M.G. provided study support; M.E.T.-P. conceived, designed and supervised the study. All authors contributed to manuscript preparation and read, commented on and approved the manuscript.

Declaration of interests

M.-E.T.-P. is member of the ethics advisory panel of MERCK. The other authors declare no conflict of interest.

17
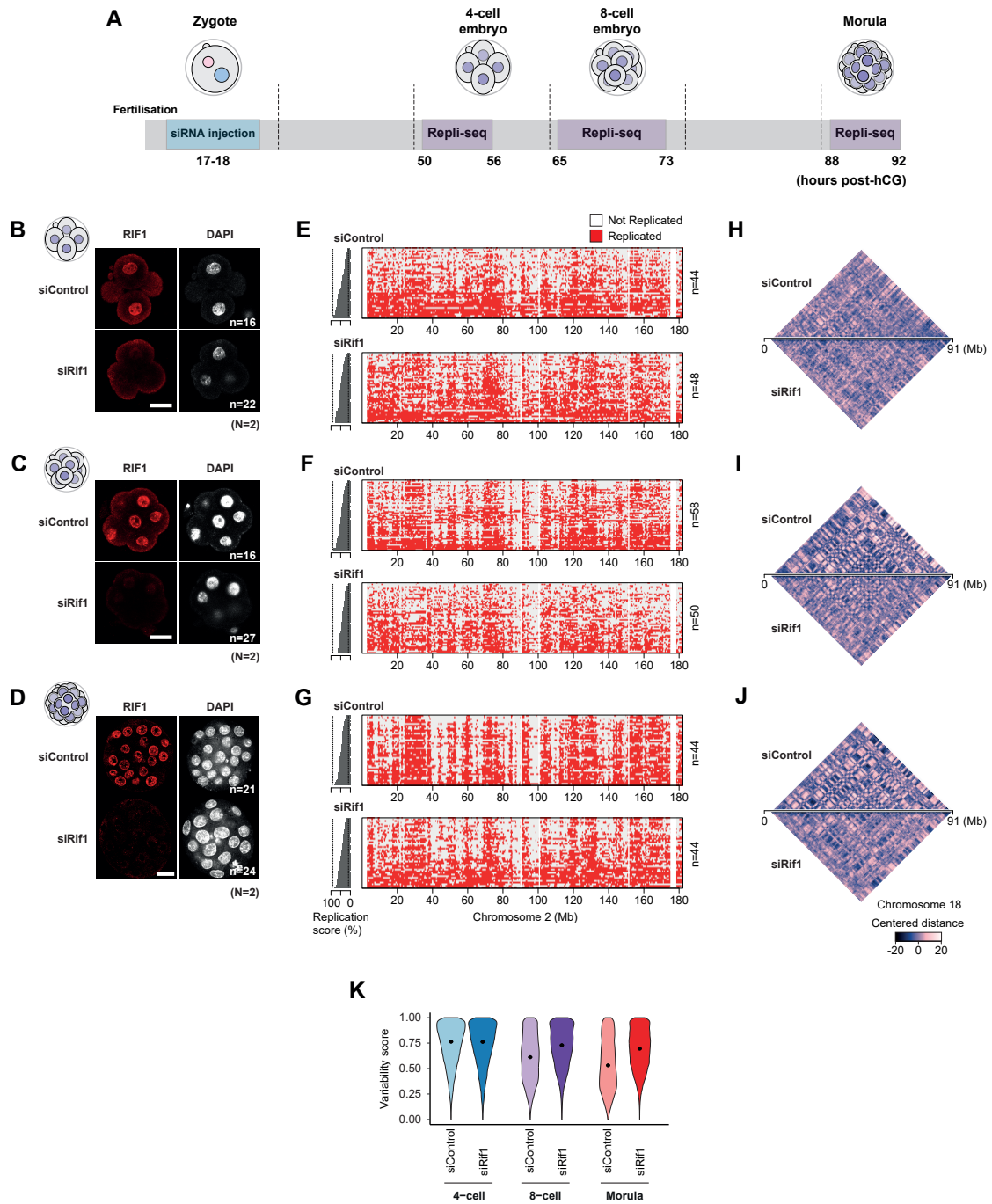
Figure Titles and Legends

Figure 1. RIF1 regulates genome-wide replication timing in mouse embryos.

(A) Experimental timeline to generate replication timing profiles from single-cells in mouse pre-implantation embryos.

(B-D) RIF1 immunostaining at 4-cell (B), 8-cell (C), and morula (D) stages after microinjection of control siRNA or Rif1 siRNA at the zygote stage. n: number of analyzed embryos. N: number of independent experiments. Scale bar, 25 μm. Maximum-intensity projections are shown for representative embryos. Note that 4-cell stage embryos were imaged as 3D mounted, hence the cytoplasmic background is higher compared to the 8-cell and morula stages, albeit nuclear RIF1 signal is clearly depleted upon RIF1 siRNA.

(E-G) Heatmaps of single-cells indicating the replication status based on binarized copy number at 4-cell (E), 8-cell (F), and morula (G) stages after depletion of RIF1. Grey: not replicated; Red: replicated. Cells are ranked by their percentage of replicated genome (replication score), which indicates progress in S-phase and is plotted as a bar plot on the left.

(H-J) Pair-wise Manhattan distance between genomic bins on the binary data over a representative chromosome (chr18) at 4-cell (H), 8-cell (I), and morula (J) stage after depletion of RIF1. Distance values are mean-centered and thus the relative distances are comparable between the samples. Darker colour indicates higher similarity (closer distance).

(K) Variability score at each embryonic stage after depletion of RIF1. The score is 1 when 50% of the cells replicated the genomic bin and it is 0 when either all cells are replicated (100%) or non-replicated (0%). Each violin shows the distribution of scores for all genomic bins. Dots indicate the median.
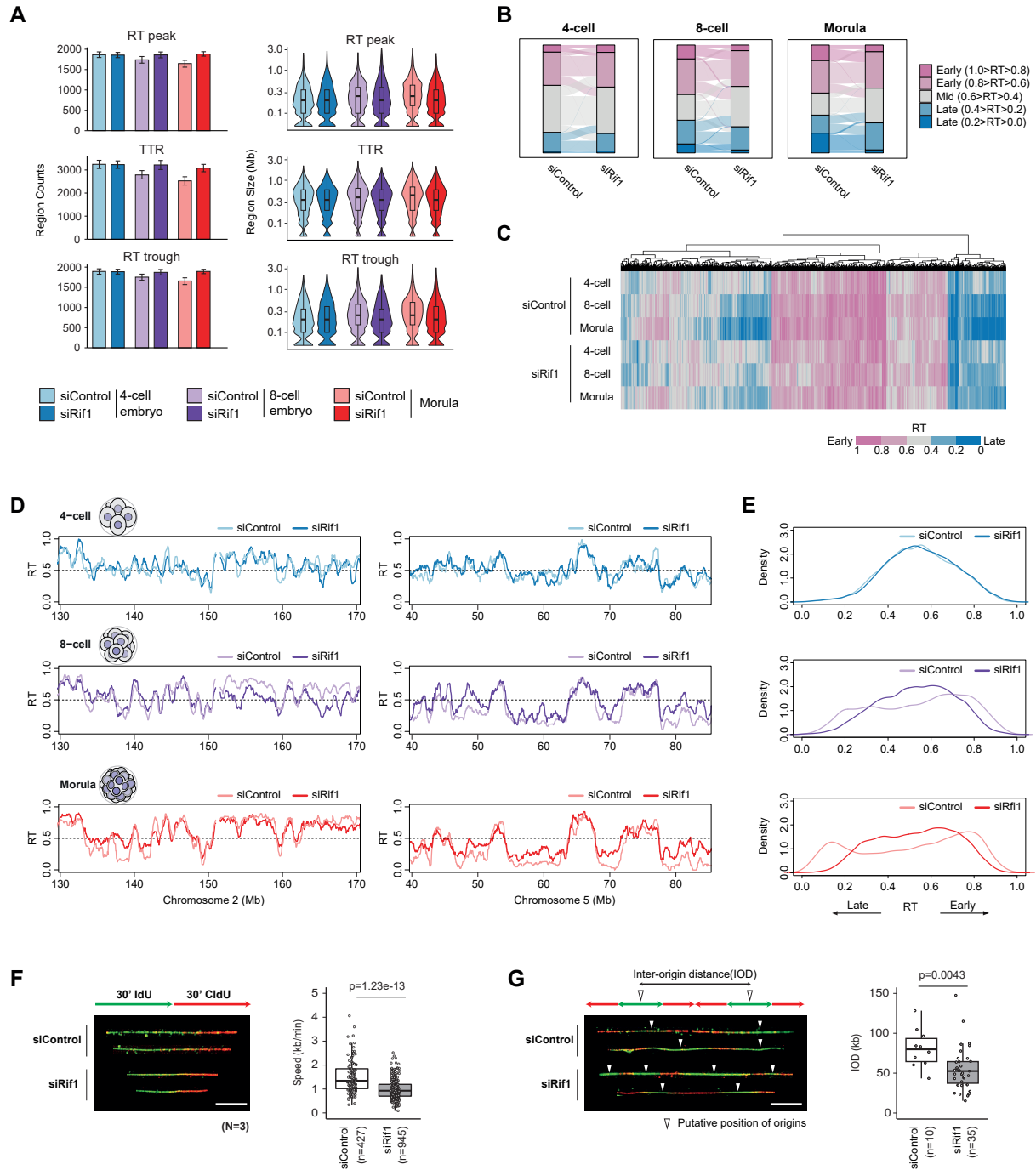
Figure 2. Depletion of RIF1 prevents consolidation of RT during the progression of embryogenesis.

(A) Number (left bar plots) and size (right violin plots) of replication features RT peaks (alternatively known as initiation zones or IZs); timing transition regions, TTRs; RT troughs (alternatively known as termination zones or TZs) at 4-cell, 8-cell, and morula stages after depletion of RIF1. Error bars on bar plots indicate the 95% bootstrap confidence interval. Statistical analysis is shown in Fig. S2A-B. The box plots inside the violin plots show the median and the interquartile range (IQR) and whiskers depict the smallest and largest values within 1.5 × IQR.

(B) Alluvial plot depicting RT changes of all genomic bins showing changes of RT after depletion of RIF1 at the indicated stages. RT values were categorised in 5 groups from the earliest (1.0>RT>0.8) to latest RT (0.2>RT>0.0) across the genome.

(C) Heatmap with hierarchical clustering depicting RT of 50kb genomic bins in RIF1-depleted 4-cell, 8-cell, and morula stages and those from controls (siControl).

(D) Average RT profiles of RIF1-depleted embryos on representative segments of chromosome 2 and 5 at 4-cell, 8-cell, and morula stages overlayed with their controls.

(E) Density plots showing the distribution of replication timing of 50kb genomic bins in RIF1-depleted 4-cell, 8-cell, and morula stages overlayed with those from controls (siControl). Statistical analysis is shown in Fig. S2F.

(F and G) DNA fiber analysis of RIF1-depleted morula stage embryos by sequential labelling of nascent DNA. Quantification results of fork speed (F) and inter-origin distance (IOD) (G) are shown along with representative images. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR. Statistical analysis was performed with a two-sided Wilcoxon's rank-sum test. Scale bar, 15 μm.

In F and G, n and N show the number of fibers analyzed and number of independent experiments, respectively.

19
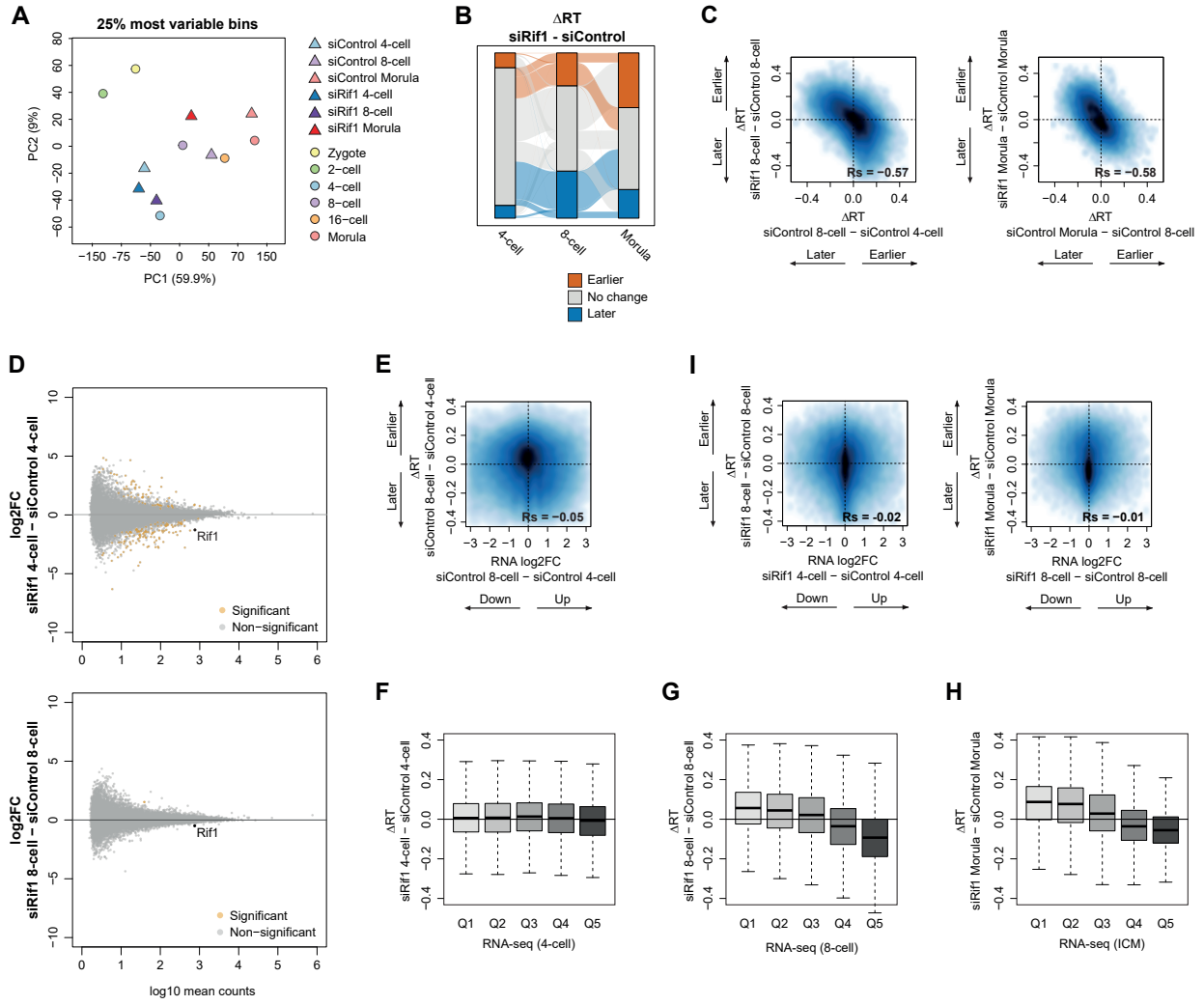
Figure 3. RIF1 depletion prevents changes of RT that would occur between stages without majorly affecting gene expression.

(A) Principal component analysis (PCA) of RT profiles from siControl and siRif1 embryos, compared with publicly available data from wild-type embryonic stages, as indicated. The 25% most variable bins are used.

(B) Alluvial plot showing significantly differential RT bins between stages identified by bootstrap. Genomic bins identified as significantly changing to earlier and later RT upon RIF1 depletion, centered at the 8-cell stage are shown.

(C) Smoothed scatter plots of RT differences (ΔRT) between the same stages of Rif1 and control siRNA microinjected embryos versus RT differences (ΔRT) between different stages of control siRNA injected embryos.

(D) MA plots showing differentially expressed genes between RIF1-depleted (siRif1) and control siRNA injected embryos at 4-cell (n=175 DE genes) and 8-cell stage (n=1 DE gene). P-values were obtained by DESeq2.

(E) Smoothed scatter plot of RT differences (ΔRT) between 8-cell and 4-cell stages of control siRNA injected embryos versus changes in RNA expression (log2FC) between 8-cell and 4-cell stages of control siRNA injected embryos.

(F) Box plots depicting changes in RT (ΔRT) in genomic bins upon RIF1 depletion at the 4-cell stage according to gene expression levels GSE45719,[50] Q5 is the quintile with highest expression and Q1 with the lowest expression values. Box plots show median of ΔRT values and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR.

(G) As in F, but at the 8-cell stage

(H) As in F, but at the morula stage. Note that expression data from morula is not available and thus we used data from ICM.

(I) Smoothed scatter plots of RT differences (ΔRT) between same stages of Rif1 and control siRNA injected embryos versus changes in RNA expression (log2FC) between Rif1 and control siRNA injected embryos.

In C, E, and I, Spearman's correlation (Rs) is indicated.
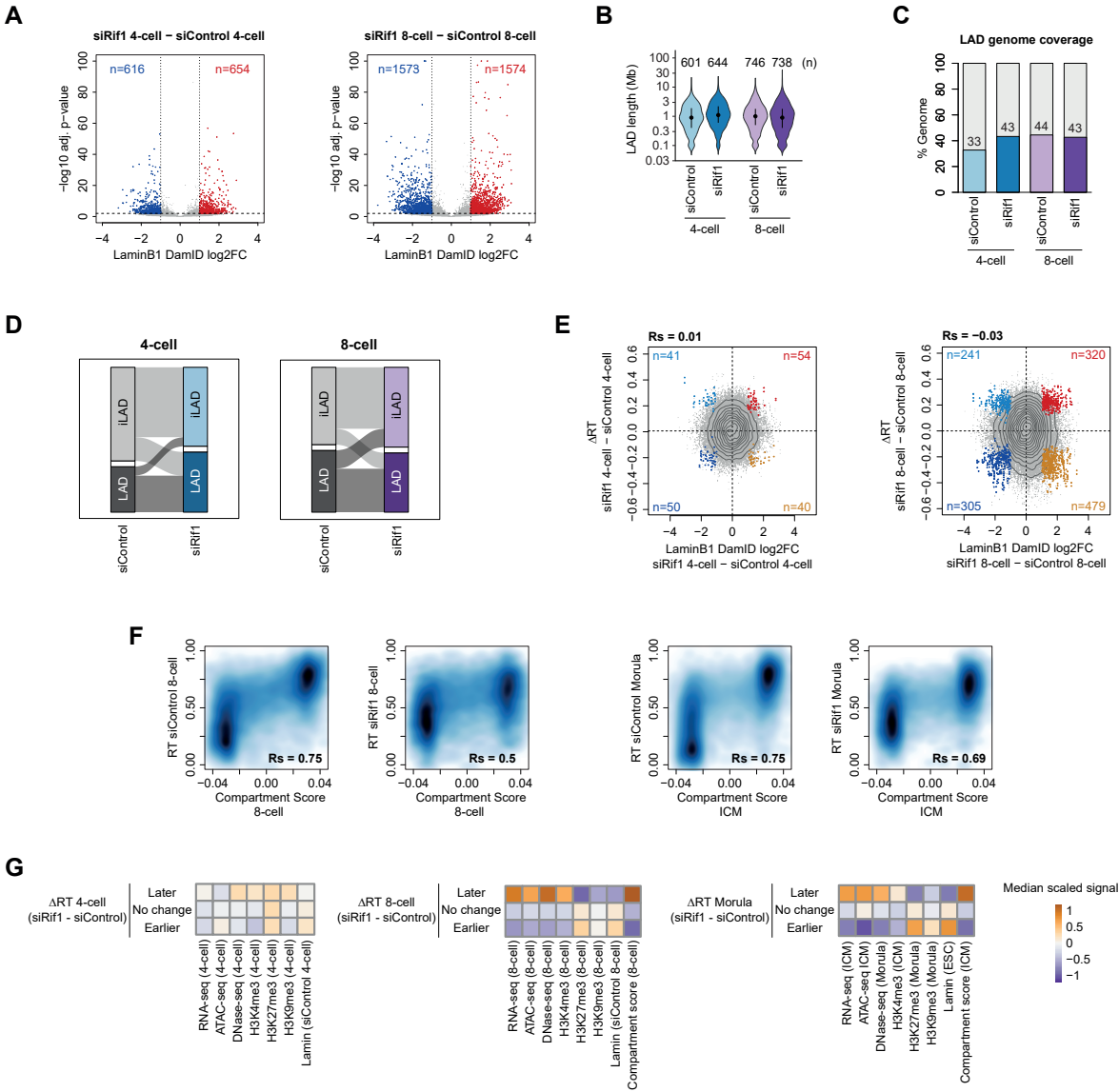
Figure 4. RIF1 controls RT independently of genome-lamina interactions.

(A) Volcano plots showing genomic bins with significantly increased (red) and decreased (blue) lamina interactions upon RIF1 depletion at indicated stages. P-values were calculated by a generalized linear model of the gamma family.

(B) Violin plots depicting the length distribution of LADs at the indicated stage and experimental group. The 25th and 75th percentiles (black lines), median (dots) and the number of LADs (n) are indicated.

(C) Percentage genomic coverage of LADs and inter-LADs (iLADs) after depletion of RIF1 at indicated stages.

(D) Alluvial plots depicting the genomic distribution of LADs and iLADs in control (siControl) and their changes upon RIF1 depletion (siRif1) at the 4- and the 8-cell stage.

(E) Scatter plots of RT differences (ΔRT) between same stages of Rif1 and control siRNA embryos versus changes in Dam LaminB1 OE values (log2FC) between same stages of Rif1 and control siRNA injected embryos. Contour (black) lines indicate the genomic bin density. The number of significantly changed genome bins in each quadrant is indicated (n) and highlighted as coloured dots, respectively. Spearman's correlation (Rs) is indicated.

(F) Smoothed scatter plots showing correlation between RT values of control (siControl) and RIF1-depleted (siRif1) embryos and Hi-C compartment score of wild-type unmanipulated embryos at the indicated stages. Positive compartment scores define A compartment. Spearman's correlation (Rs) is indicated.

(G) Median enrichment of chromatin features in wild-type embryos at differential RT genomic bins between RIF1-depleted (siRif1) and control (siControl) embryos. When data from the same stage as RT is not available, the closest stage data are used for analysis.

Methods

Mouse ethics statement

All experiments were performed under the authorization of the authorities from Upper Bavaria (Tierversuchsantrag von Regierung von Oberbayern). Values for housing temperature, humidity, and light cycle of mouse cage are defined for 20-24 ºC, 45-65%, and 12h dark/12h light, respectively.

Embryo collection and culture

F1 female mice (C57BL/6J × CBA) < 10 weeks of age were super-ovulated by intraperitoneal injection of 10 U of pregnant mare serum gonadotropin (PMSG) followed by 10 U of human chorionic gonadotropin (hCG) 48 hours later, and then mated with DBA/2J male mice. Zygotes were collected from the oviduct and cumulus cells were removed upon brief incubation in M2 media containing hyaluronidase (Sigma Aldrich). Zygotes were placed in drops of KSOM and cultured at 37 ºC with 5% CO_2 as previously described.

Immunostaining

Cells were washed with PBS, fixed for 15 min in 4% PFA in PBS at room temperature and permeabilised with 0.5% Triton X-100 in PBS for 15 min at room temperature. Cells were blocked in 5% normal goat serum in PBS for 1 h at room temperature and incubated overnight at 4ºC with the following primary antibodies: Rif1 UCRIII (1:1000 dilution),[14] γH2AX (ab2251, 1:1000 dilution). After washing with PBS, the cells were incubated with the corresponding secondary antibodies (Anti-mouse Alexa 488; 1:800 dilution). DNA was stained with 1 µg/ml 4',6-diamidino-2-phenylindole (DAPI). Images were acquired on a SP8 confocal laser-scanning microscope (Leica). We set acquisition parameters in order to obtain fluorescence intensity signal in the linear range of the hybrid detectors of the confocal microscope. These detectors have negligible detector noise and linearly amplify incoming photons into photoelectrons, thus enabling counting of measured photons as long as the detector is not saturated. The recovered signal therefore accurately reflects the level of antigen present in the system and quantifications are possible between experimental and control samples since, in addition, we used identical settings for acquisitions.

Quantification of RIF1 fluorescence intensity in 3D

We used an image analysis pipeline with the following software and software libraries (version numbers indicated): Fiji (ImageJ 1.54f),[51] Python (3.12.2), Cellpose (3.0.6),[52,53] Pytorch (2.2.1), pytorch-cuda (11.8), cudnn (8.0), scikit-image (0.22),[54] pandas (2.2.1), R (4.3.1), and ggplot2 (3.4.3). First, we trained a custom Cellpose model using 66 single optical sections in

22

the DAPI channel that were randomly extracted with Fiji for manual annotation. These images were split into a training and test dataset of which the latter contained 20% of all images. The Cellpose 'nuclei' model was then re-trained with these images and a mean object diameter of 75 pixels, as determined from the average size of training masks. The model was trained with default parameters, for 300 epochs. After manually checking the quality of several segmentation masks, we segmented the nuclei of mouse early 4-cell embryos and combined masks per optical section to obtain 3-dimensional volumes using a Cellpose 'stitch threshold' of 0.02. Finally, from the Cellpose masks and the raw intensity images, we extracted the mean intensity value for each cell in both DAPI and RIF1 channels with the scikit-image 'regionprops' module. In 'R', we verified the distribution of volumes of the found objects and filtered out small objects (typically arising from segmentation of polar bodies or noisy regions in the DAPI channel). Plots were subsequently generated with ggplot2 and we performed statistical evaluation in 'R' with the Wilcoxon rank sum test.

DNA fibres

DNA fibres were prepared as described [32] based on [55]. Embryos injected with siRNA for control and for *Rif1* at 17-18 hours post-hCG were cultured until they reached the morula stage and were sequentially pulse-labelled with 25 μM IdU (Sigma) and 50 μM CldU (Sigma) for 30 min each and collected at 89 hours post-hCG. Labelled embryos were lysed and DNA fibres were stretched onto the slide glass by tilting. The fibres were fixed in methanol/acetic acid (3:1), then denatured with 2.5 M HCl for 1 h, neutralised with PBS, and blocked with 1% BSA / 0.1% Tween 20 in PBS. CldU and IdU tracks were detected with anti-BrdU antibodies (described in Key resources table) recognizing CldU and IdU, respectively, and appropriate secondary antibodies (described in Key resources table). Images were acquired on a Leica SP8 confocal microscope using a 40x Plan/Apo NA1.3 oil immersion objective (Leica) at 2048x2048 pixels at an effective pixel size of 142 nm. To calculate fork speed, we used the established conversion 1 μm = 2 kb.[56] Analysis of DNA fibres was performed as described before[32] by two different researchers using a custom image analysis pipeline (https://github.com/IES-HelmholtzZentrumMunchen/dna-fibres-analysis) based in part on implementing a structure reconstruction with a spatially variant morphological closing.[57] We employed masks to select region of interest with sufficient amount of fibres and extracted fibres manually. To detect patterns in the extracted fibres, we used a branches detection strategy. Because the fluorescence channels are not directly comparable in absolute intensity values, we used the logarithm of their point-wise instead. We used regression trees structures in combination with the CART algorithm,[58] and a semi-automated step to verify fibre detection and assignment of patterns. To calculate inter-origin distance (IOD), we manually selected sufficiently long fibre

23

stretches from the DNA fibre dataset in the DNA channel, which encompassed several IdU/CldU boundaries using a custom made Fiji (ImageJ) macro to open the regions of interest in the images and applied the ImageJ 'Straighten' function with a width of 19 pixels to convert bent fibres into approximately 2-dimensional images, where the channel intensities were interpolated along the x-axis. In the stretched fibre images, we then manually selected all identifiable IdU/CldU boundaries. The remaining analysis was performed in 'R'. We first calculated from the x-coordinates of the boundaries all origin positions by averaging between two adjacent boundary points. We then determined the pair-wise difference between origins to obtain the IOD. IOD and box plots were created with the ggplot2 library in 'R'.

Repli-seq

Single-cell Repli-seq in embryos was performed as we previously described [10] based on references [59] and [38]. Briefly, early stage zygotes were collected and micro-injected with 20 μM siRNA at 17-18 hours post-hCG injection (phCG), and then they were cultured until they reached the S-phase at each developmental stage, based on their time after hCG injection. Embryos were collected at different time points at each developmental stage to achieve sampling over the entire S-phase. For each developmental stage, embryos were obtained from several litters and embryos from different litters were collected across different dates to ensure robust data collection. The number of mice used to collect samples for each developmental stage is indicated in the parenthesis, as follows: siControl 4-cell (20); siRif1 4-cell (20); siControl 8-cell (20); siRif1 8-cell (20); siControl Morula (12); siRif1 Morula (12). Zona pellucida was removed by exposure to acid Tyrode and each blastomere was dissociated by gentle pipetting after trypsin treatment. Individual blastomeres were placed into 8-strip PCR tubes containing lysis buffer and extracted DNA was fragmented by heat incubation. Fragmented DNA was tagged by the universal primer (5'-TGTGTTGGGTGTGTTTGGKKKKKKKKKKKNN-3') and amplified with WGA primer sets which have individual barcode. Amplified DNA was purified by the QIAquick 96 PCR Purification Kit (QIAGEN) and concentration was determined by the NanoDrop (Thermo Scientific). Equal amount of DNA from each sample (up to 96 samples) were pooled and 1μg of them was ligated with the Illumina adaptors using the NEBNext Ultra II DNA Library Prep Kit (NEB). Illumina sequences (NEBNext Multiplex Oligos for Illumina; NEB) were added to the adaptor- ligated samples by PCR. Clean up and size selection of the PCR product was done using SPRIselect (Beckman Coulter) and the quality of the library was confirmed by 2100 Bioanalyzer with the High Sensitivity DNA Kit (Agilent).

scRepli-seq read alignment and quality control filtering

24

The summary of sample collection, alignment statistics and data quality is included in Table S1. Sequencing reads were mapped to the GRCm38 genome using bowtie2 (version 2.5.1)[60] with the parameters --local --no-unal --no-mixed --no-discordant. Reads were filtered by mapping quality using samtools with the parameter -q 20. Read duplicates were removed using picard MarkDuplicates (version 3.0.0) with the parameter -REMOVE_DUPLICATES true. Using bedtools intersect (version 2.31.0), reads were counted in consecutive 50kb genomic bins. For each cell, the mean of the bin counts was calculated per chromosome to obtain the between chromosome coefficient of variation (CoV) as the ratio of the standard deviation to the mean (Fig. S1F). Cells with a CoV threshold above 0.15 were filtered out from the analyses. The threshold was set to be able to process a large number of single cells and to accommodate the observed higher coefficient of variation in RIF1-depleted embryos. Cells were added back if the CoV threshold was not passed due to an individual chromosome, which was masked. The CoV filter serves to remove cells with abnormal, aneuploid genotypes. Finally, cells with more than 750,000 mapped reads were used for downstream analyses.

Assignment of replication status

Read counts in consecutive 50kb genomic bins were used in the scRepli-seq bioinformatic pipeline that we previously described.[10] Briefly, bin counts were first RPM (reads per million) normalized. To correct for the mappability bias, for each bin its respective average of all samples within the same condition was calculated. Outlier regions (<5th percentile and >1st percentile) were masked. To correct for low mappability, bins were segmented with the R package copynumber (version 1.38.0, R version 4.2.3)[61] and segments with the highest 95% of values were kept. For each cell, the data were centered by the mean and scaled by the interquartile range and smoothed by a median filter (running window width of 15), followed by segmentation with the R package copynumber. The segmented values were used to fit a two component mixture model to identify replicated and non-replicated genomic bins using the R package mixtools (version 2.0.0).[62] To find a threshold that separates the bins, the intersect of two normal distribution functions were used. If no intersect was found, the center of the means served as threshold, as previously described.[10]

Replication score, bin-bin distance, replication timing value and variability score

Genome-wide replication score was defined as the percentage of replicated genomic bins for each cell. Cells with a replication score greater than 90% and less 10% were removed from the analysis. We ranked the cells by the replication score as a proxy of S-phase progression for visualization on the binary replicated/non-replicated heatmaps (Fig. 1E-G). To assess the relationship between genomic regions we calculated the Manhattan distance between all pairs of genomic bins across cells on the binarized data ('1s' replicated, '0s' non-replicated). The

25

distance matrix was centered by the mean and visualized as heatmap (Fig. 1H-J). Because the RT values are relative (maximum value is always 1 and minimum value is always 0) and the Manhattan distances are centered, the data between two experimental conditions are comparable. To obtain replication timing values and to correct for potential sampling bias of cells, we calculated the fraction of replicated cells in overlapping intervals of the genome-wide replication score with interval size of 35% and increment of 4.33% (e.g. 0-35%, 4.33-39.33% etc.) for each genomic bin. The average of these 16 intervals served as replication timing value that was used for downstream analyses. A higher replication timing value indicates earlier replication timing, as higher proportion of cells replicated the bin. The variability score was calculated as described elsewhere.[10] For the PCA analysis, we used a given percentage of the most variable bins (i.e. to minimise noise and capture the most meaningful signal that reduces dimensionality). We display the data with the 25% most variable bins but we obtained similar trends using a higher number of bins.

Identification of IZs (RT peaks), TTRs and TZs (RT troughs)

IZs, TTRs and TZs were defined based on the replication timing values as described before.[8,10] Briefly, genomic bins were grouped into 15 clusters by their replication timing values using the Mclust function from the mclust (version 6.0.0) R package (R version 4.1.2). The clusters were ranked by their average replication timing. IZs or TZs were defined as consecutive bins with local maxima or local minima of their cluster ranks, respectively, in sliding windows of 21 genomic bins using the rollapply function from zoo R package (version 1.8-10). Regions between IZs and TZs were defined as TTRs.

Analysis of RT heterogeneity

Heterogeneity analysis was performed using the sigmoidal model fitted for each genomic bin as described previously.[10,38,63,6410,37,61,62] Two parameters were calculated from the curve fitting, M-value and T_{width}. The M-value (sometimes also referred to as Trep in the literature[8]) is the replication score (~S-phase time) at which 50% of the cells replicated the given bin. A greater M-value indicates later replication timing. T_{width} is a measure of RT heterogeneity and is defined as the replication score difference (~S-phase time difference) between 25% and 75% of the cells replicated the genomic bin. A higher T_{width} value indicates greater heterogeneity, as the transition from non-replicated to replicated status is larger.

scRepli-seq statistical analysis

For statistical analyses of scRepli-seq, we used a bootstrapping method and calculated 95% confidence intervals to determine statistical significance.[65] We have previously used this method[10] as it avoids the inflation of p-values when n is large due to large number of genomic

26

185

bins (n~49000) and thus we applied bootstrapping to single cells (n~30-70). Namely, we iteratively re-sampled individual cells with replacement 1000x times for each condition and calculated confidence intervals from the bootstrap distribution using the percentile method. The 95% percent confidence interval is the interval between 2.5th and 97.5th percentile of the distribution. When 95% percent confidence intervals do not include zero or two intervals do not overlap, they are significantly different from zero or different from each other, respectively. Using the bootstrap method we called genomic bins that are significantly different between conditions (e.g. siRif1 vs. siControl). We also applied the bootstrap method to judge the significance of the differences in the mean variability score (Fig. S1I), the IQR of the M-values (Fig. S2H) and in the mean of T$_{width}$ (Fig. S2J).

Single embryo RNA-seq and library preparation

20 µM siControl (Dharmacon, D-001210-01) or 20 µM siRif1 (Dharmacon, D-040028-01) were injected into zygotes at 17-18 hphCG injection and embryos were cultured until 63 and 70 hphCG injection, respectively, at which point 4-cell and 8-cell embryos, respectively were collected, washed with PBS, placed in tubes with 1× Clontech lysis buffer (Z5013N) containing ERCC RNA Spike-In Mix (Invitrogen) and flash-frozen in liquid nitrogen. The *Rif1* siRNA used in this study was previously validated against 3 other individual siRNAs.[46] RNA-seq was carried out using the SMART-seq2 protocol[66] and subjected to 150bp paired-end sequencing on a Novaseq 6000 (Illumina) platform. The quality and quantity of the cDNA libraries were verified with the 2100 Bioanalyzer with the High Sensitivity DNA Kit (Agilent). A total of seven siControl and thirteen siRif1 injected 4-cell embryos and eight siControl and twelve siRif1 injected 8-cell embryos derived from two independent experiments were sequenced. Collection developmental timepoints for RNA-seq, which overlap with S-phase, were chosen to enable comparisons across public datasets but, most importantly, because most of the transcriptional changes occur during the course of S-phase due to the short duration of G1 in mouse embryos (only 1-2 hours).

RNA-seq analysis

Sequencing reads (paired) were aligned to the mouse genome (GRCm38, primary assembly) using STAR aligner (version 2.7.6a) with the annotation (GRCm38.101) and ERCC92 (Thermofisher). Reads were counted per gene by the same STAR run by setting quantMode GeneCounts. Downstream data analysis and visualization was done using R (version 4.1.2). Embryos with at least 500,000 genic reads, less than 20% ERCC and mitochondrial reads were considered. Differential expression analysis was performed with DESeq2 (version 1.34.0) and functions from HelpersforDESeq2 package (link: https://github.com/tschauer/HelpersforDESeq2). Genes were filtered for at least one read

counted in at least one fourth of all samples. Differentially expressed genes were called by a cut-off of an adjusted p-value of less than 0.05. Results were visualized as the relationship between the log2 fold change of siRif1 vs. siControl and the log10 mean normalized counts on MA-plots.

DamID-sequencing and analysis

Zygotes were collected and injected with 20 μM siRNA at 17-18 hours post-hCG, followed by culture in KSOM medium until the late 2-cell stage. At 46-48 hours post-hCG, both blastomeres of the 2-cell embryos were injected with 250 ng/μL Tir1, 100 ng/μL membrane-eGFP, and 20 ng/μL AID-Dam-LaminB1 of in vitro transcribed mRNA, and subsequently cultured in medium containing auxin (500 μM). For DamID the 4-cell stage, auxin was removed at 54 hours post-hCG, and late 4-cell embryos were collected at 60–62 hours post-hCG. For DamID in the late 8-cell stage, auxin was washed out from 66 to 72-74 hours post-hCG. Before processing for DamID, the zona pellucida was removed by treatment with 0.5% pronase in M2 at 37 °C for 5 minutes and the polar body was removed by gentle pipetting after trypsin treatment. For each replicate, a pool of 16-24 blastomeres (four to six 4-cell embryos or two to three 8-cell embryos) was collected in 2 μL DamID buffer and stored at -80 °C until processing. All experiments were conducted in three independent biological replicates. Sample processing and library preparation were performed as previously described.[28,67] Libraries were sequenced using Illumina HiSeq2500 platform in 150 bp PE mode but only the first read was used for analysis. The first 6 random bases of reads were discarded by trimmomatic (version 0.39) and reads were demultiplexed by the DamID indexes. The processed reads starting with GATC were then aligned to the GRCm38 using bowtie2 (version 2.5.1) with default settings. Reads with a mapping quality score below 30 were removed using samtools (version 1.17). Duplicated reads were filtered using picard (version 3.0.0). Reads were counted in 100kb consecutive genomic bins using bedtools (version 2.31.0). OE (Observed/Expected) values per bin were calculated similarly as described previously.[68] Briefly, genomic GATC sites were extended to the trimmed read length (123 bp) in both directions using R (version 4.1.2) Biostrings (version 2.62.0) and GenomicRanges (version 1.46.1) packages. GATC reads were processed the same way as DamID reads (observed) and served as expected reads counts. Read counts were normalized by rpkm (reads per kilobase per million) and the smallest non-zero rpkm value (pseudo-count) was added. The final DamID Score was calculated as the ratio of the observed over expected rpkm values. Bins with zero rpkm for both observed and expected values were treated as zero. OE mean signal was obtained by averaging the rpkm values of the biological replicates prior OE value calculation. The OE mean values were used for visualization and LAD calling. To distinguish

28

between LADs and iLADs, a two-state hidden Markov model (HMM)[69] was used on the non-zero OE mean values. For differential analysis between siRif1 and siControl at genomic bins, a generalized linear model of the gamma family with log link was fitted on the replicate OE values using R as described previously (Pal et al., in revision). P-values were obtained on the z-distribution and were corrected for multiple testing by the Benjamini & Hochberg procedure. Significant bins were identified by an adjusted p-value threshold of 0.01 and a log2 fold change threshold of one.

Analysis of public chromatin datasets

Published datasets were obtained from GEO with accession numbers GSE66581 (ATAC-seq[70]), GSE71434 (H3K4me3 ChIP[71]), GSE98149 (H3K9me3 ChIP[72]), GSE73952 (H3K27me3 ChIP[73]) and GSE135457 (DNAse-seq[70]). Paired-end reads were trimmed by cutadapt (version 3.4) with parameters -a CTGTCTCTTATA -A CTGTCTCTTATA -a AGATCGGAAGAGC -A AGATCGGAAGAGC --minimum-length=20. After trimming, reads were mapped to the mouse reference (GRCm38) using bowtie2 (version 2.3.5) with parameters --end-to-end --very-sensitive --no-unal --no-mixed --no-discordant -I 10 -X 500. Reads were filtered by mapping quality by samtools (version 1.3) with parameter -q 12. Read pairs were imported to R (version 4.1.2) using the readGAlignmentPairs function from the GenomicAlignments package (version 1.30.0) and unique fragments were selected. Mitochondrial reads and reads mapped to scaffolds were not considered. Fragments were counted with the countOverlaps function from the GenomicRanges package (version 1.46.1) in 50kb consecutive genomic bins, normalized by the sum of the fragments counts and multiplied by a million. Normalized counts were log2 transformed after adding a pseudo-count of 0.1.

Analysis of publicly available RNA-seq datasets

Published RNA-seq datasets were downloaded from GEO with accession number GSE45719.[50] Data processing, read counting and TPM calculations were carried out as described in.[74] Early blastocyst cells were further divided to ICM and TE cells by hierarchical clustering on the TPM values of selected marker genes.[74] For direct comparison between RT, DamID OE values and RNA-seq, we counted the RNA-seq reads in 100kb consecutive bins similarly to ChIP-seq datasets as described above.

Analysis of public Hi-C dataset

Hi-C compartment coordinates and scores (GSE82185)[30] were analyzed as described.[28]

29

Image analysis of morula and blastocyst stage embryos

For segmentation of cells in morula and blastocysts, we first manually outlined individual embryos in Fiji[51] then segmented cells of each embryo using the DAPI channel with Cellpose (version 3.0.6 and 2.0.5, respectively)[52,53] using a custom-trained model, or the 'nuclei' model, respectively. In both cases, we used a radius of 80 pixels for object detection and a stitch threshold of 0.01 to assemble 2D segmentation masks into 3D objects. For analysis of phosphorylated H2A.X intensities we applied the 'regionprops' module of the 'scikit-image' Python package (version 0.22.0)[54] with the phosphorylated H2A.X channel as 'intensity_image' parameter. For detection of mitotic cells in blastocysts, we trained a pixel classifier in Ilastik (version 1.4.0)[75] on DAPI images. Mitotic cells were defined by applying hysteresis thresholding on the Ilatsik probabilities (low threshold 0.5, high threshold 0.99) and further refined by removing small objects with volumes < 3000 voxels. The post-processed Ilastik masks were then counted to derive the number of mitotic cells per embryo or subtracted from Cellpose masks to quantify the number of interphase cells per embryo. Data were exported to csv files with the Python 'pandas' package. We only considered embryos with more than 9 cells. Data were plotted in 'R' (version 4.3.1) with the 'ggplot2' package (version 3.4.3). Statistical tests were performed by Wilcoxon rank sum exact test, or a generalized linear model with 'poisson' parameter for evaluating count data.

Data availability

The scRepli-seq, RNA-seq, and DamID-seq data from this study are available from the Gene Expression Omnibus, accession number GSE262791, and can be accessed using the Reviewer's token: atipgkeizvkztuj. The code used for scRepli-seq and DamID analysis can be accessed here: https://ascgitlab.helmholtz-muenchen.de/public_pipelines/repliseq_rif1
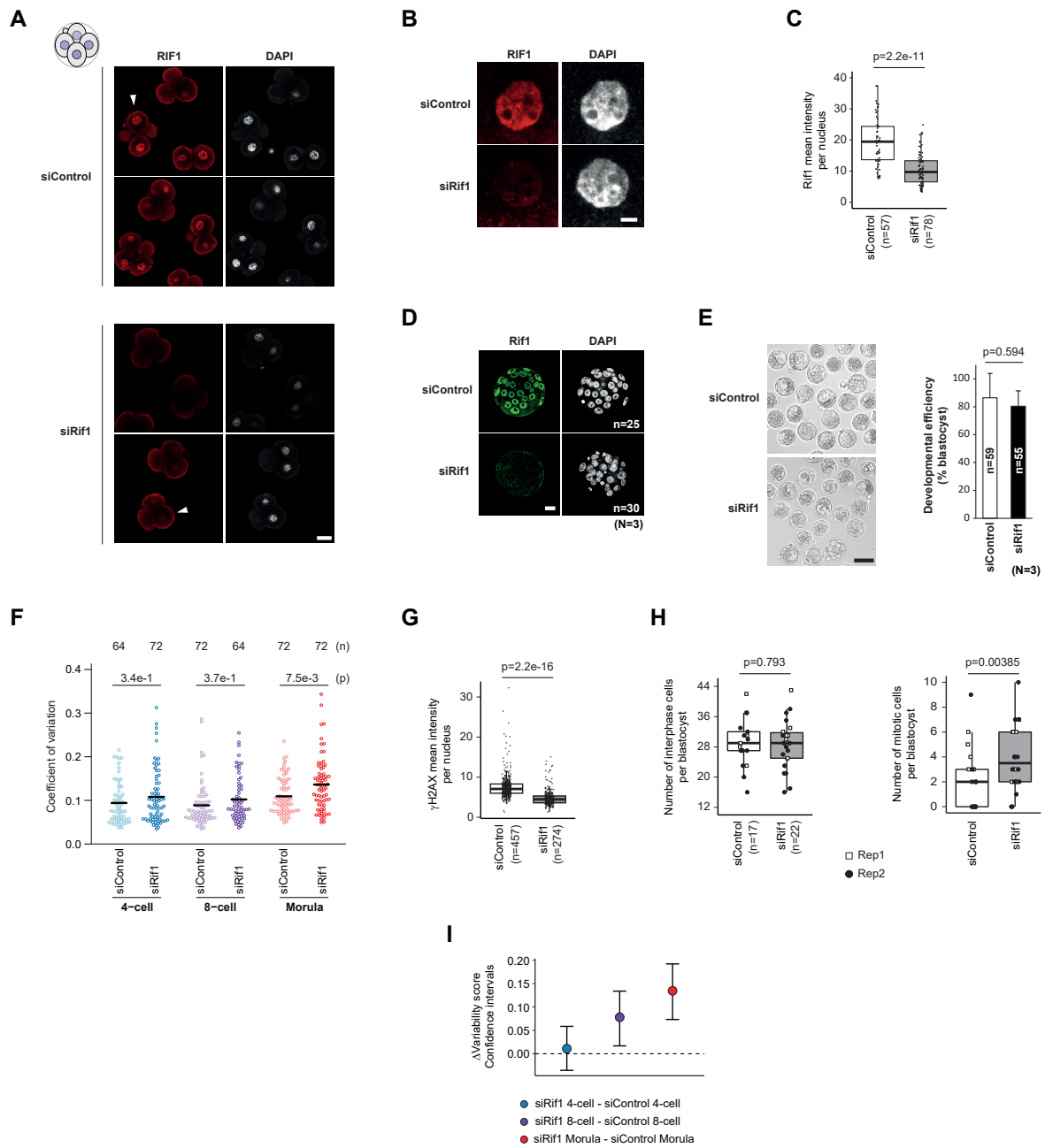
Supplemental Information
Supplemental Figure Legends S1 to S4.
Supplemental Tables S1 to S3

Supplemental Figure Legends
Figure S1. Developmental phenotype after RIF1 depletion.
(A) RIF1 immunostaining in several 4-cell stage control embryos and embryos microinjected with siRNA for RIF1 showing reproducible reduction in nuclear RIF1 protein. The embryo indicated with an arrowhead is the same embryo shown in Figure 1. Shown are single confocal sections. Scale bar, 25 μm.

(B) Representative nuclei of a 4-cell control or RIF1-depleted embryo following RIF1 immunostaining indicating depletion of RIF1 in the nucleus upon siRNA injection. Scale bar, 5 μm.

(C) Quantification of RIF1 levels in 4-cell stage control embryos and embryos microinjected with siRNA for RIF1 based on the mean intensity of fluorescence in each nucleus. Dots represent each nucleus (n).

(D) RIF1 immunostaining at the blastocyst stage after microinjection of control siRNA or siRNA for RIF1. Scale bar, 25 μm.

(E) Developmental progression of embryos upon RIF1 depletion. Zygotes collected at 17-18 h post-hCG were microinjected with siRNA for control or against *Rif1* and cultured until 96 h post-hCG. Representative image of a RIF1-depleted embryo (E) after immunostaining for RIF1. In (F), brightfield images of representative embryos (left) for both groups are shown. On the right, the percentage of embryos reaching the blastocyst stage is indicated; n: number of embryos analyzed from 3 (N) independent experiments. Statistical analyses are by two-sided Student's *t*-test. Mean±SD. Scale bar, 100μm.

(F) Coefficient of variation calculated on the average read counts per chromosome using the scRepli-seq data. The number of cells analyzed in each stage are indicated on top (n). Black lines show mean. P-values were obtained by a linear model and were adjusted for multiple comparisons.

(G) Quantification of the signal intensity of γH2AX at morula stage upon RIF1 depletion. n: number of nuclei analyzed from two independent experiments. n: number of analyzed nuclei.

(H) Quantification of the number of interphase (left) and mitotic (right) cells per blastocyst in control embryos or after depletion of RIF1. White rectangles and black circles indicate the values of each replicate. P-values were calculated by a generalized linear model of the Poisson family. n: number of analyzed embryos. On the right panel, the mean values for replicate 1 are 1.45 mitotic cells per blastocyst in controls and 3.73 in siRif1 embryos (median

for siControl is 0 and for siRif1 is 3 cells). For replicate 2, the mean number of mitotic cells per blastocyst is of 3.83 for control embryos and of 4.57 for siRif1 embryos (median for siControl is 3.5 and for siRif1 is 6 cells).

(I) Statistical analysis of data in Figure 1K. Shown are the pairwise differences in the mean variability score between the siControl and siRif1 embryos at the same stage. Error bars indicate the 95% bootstrap confidence intervals (CIs). CIs are calculated at the 95% confidence level and thus a CI that does not overlap with '0' indicates significance at the 0.05 level.

In C, G and H, box plots show median values and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR.
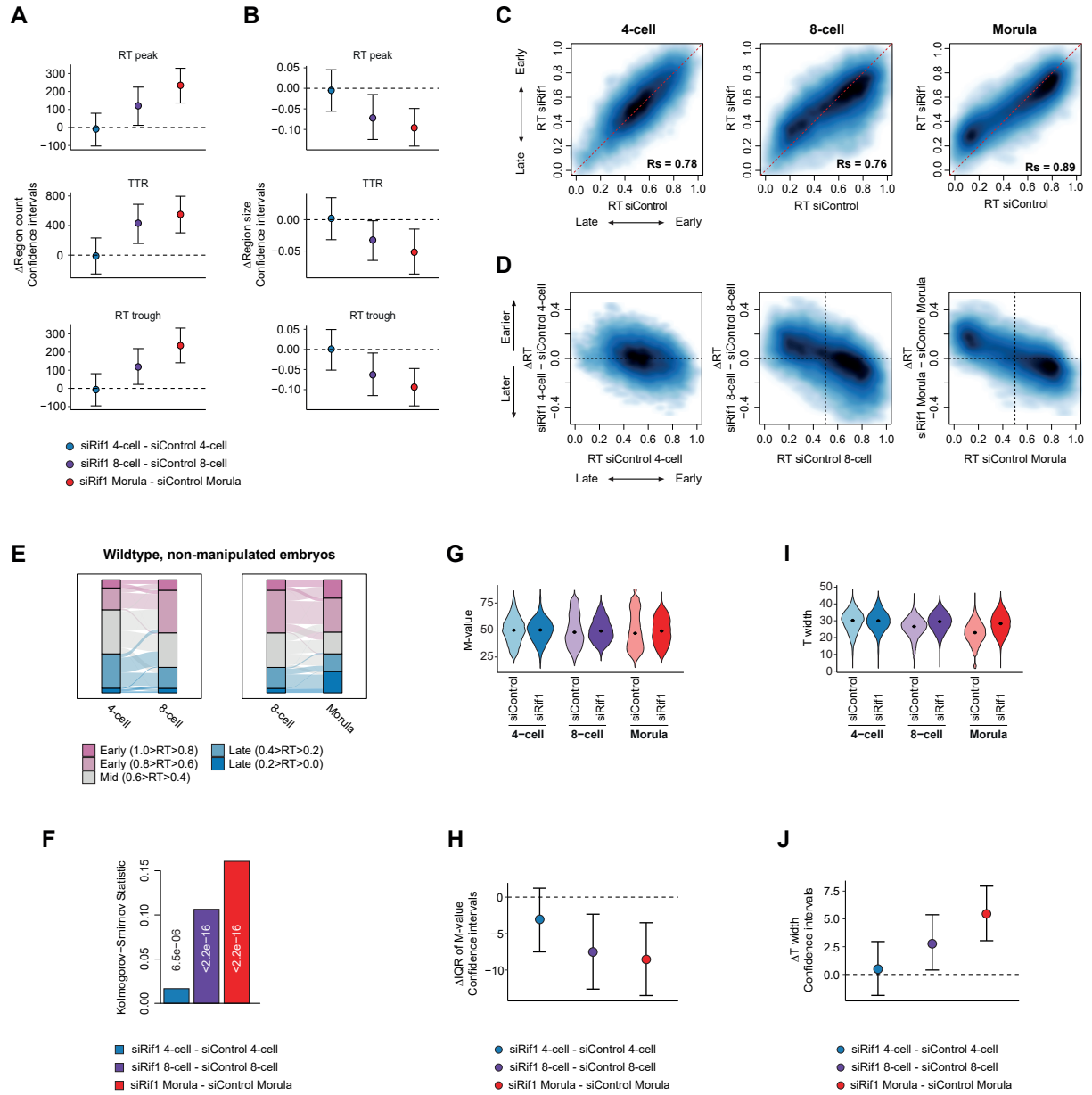
Figure S2. Depletion of RIF1 increases cell-to-cell heterogeneity of the RT program.

(A-B) Analysis of statistical significance of data in Figure 2A for the number (A) and size (B) of RT peaks, TTRs and RT troughs. Pairwise differences in the mean of the RT features in each developmental stage are plotted with error bars indicating the 95% bootstrap confidence intervals (CIs). CIs are calculated at the 95% confidence level and thus a CI that does not overlap with '0' indicates significance at the 0.05 level.

(C) Smoothed scatter plots of RT values in control (siControl) compared to RIF1-depleted embryos (siRif1). Spearman's correlation (Rs) is indicated.

(D) Smoothed scatter plots comparing the RT values in control (siControl) versus the differences in RT (ΔRT) between RIF1-depleted and control embryos at the indicated stages.

(E) Alluvial plot depicting RT changes of all genomic bins showing changes of RT in wild-type, unperturbed embryos (GSE218365) as development proceeds. RT values were categorised in 5 groups from the earliest (1.0>RT>0.8) to latest RT (0.2>RT>0.0) across the genome.

(F) Statistical analysis of Figure 2E. Bar plot showing the Kolmogorov-Smirnov statistic and the corresponding p-values indicating whether two samples (siRif1 vs. siControl) came from the same distribution.

(G) Violin plots showing the *M*-value, which is the replication score at which 50% of the cells replicated a given 50kb bin of the indicated experimental and control groups.

(H) Statistical analysis of S2G. Pairwise differences in the interquartile range (IQR) of M-values in each developmental stage are plotted with error bars indicating the 95% bootstrap confidence intervals.

(I) Violin plot depicting the T_{width}, which is the replication score difference between 25% and 75% of cells replicated the 50kb bin, at the indicated experimental and control groups.

(J) Statistical analysis of S2I. Pairwise differences in the mean of T_{width} in each developmental stage are plotted with error bars indicating the 95% bootstrap confidence intervals.

In G and I, each violin shows the distribution of scores for all genomic bins and dots indicate median.
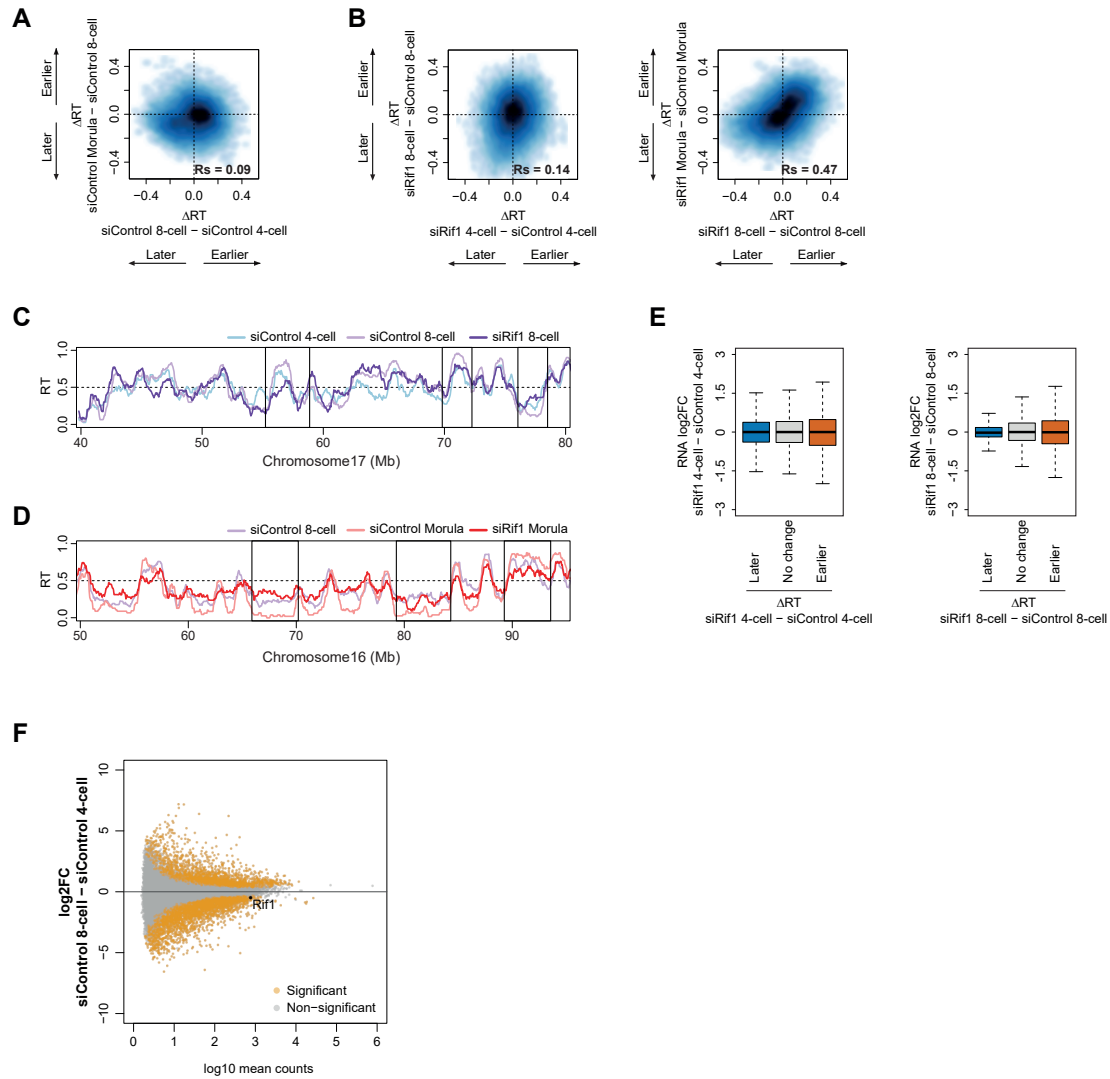
Figure S3. RIF1 regulates replication timing in a stage-specific manner.

(A) Smoothed scatterplot of RT differences (∆RT) between morula and 8-cell stage of control siRNA injected embryos versus RT differences (∆RT) between 8-cell and 4-cell stage of control siRNA injected embryos.

(B) Smoothed scatterplot of RT differences (∆RT) between same stages of RIF1-depleted (siRif1) and control (siControl) embryos versus RT differences (∆RT) between same stages of RIF1-depleted and control injected embryos.

(C and D) Representative RT profiles where regions that change RT during development display an opposite direction in RT change due to RIF1 depletion. 4- to 8-cell stage (C), and 8-cell stage to morula (D) patterns are shown.

(E) Changes in RNA expression (log2FC) between RIF1-depleted (siRif1) and control (siControl) embryos across genomic bins displaying differential RT changes upon RIF1 depletion in 4-cell and 8-cell stage embryos. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR.

(F) MA plot showing differentially expressed genes between control (siControl) embryos at the 8-cell stage compared to control 4-cell stage embryos. P-values were obtained by DESeq2.

In A and B, Spearman's correlation (Rs) is indicated.
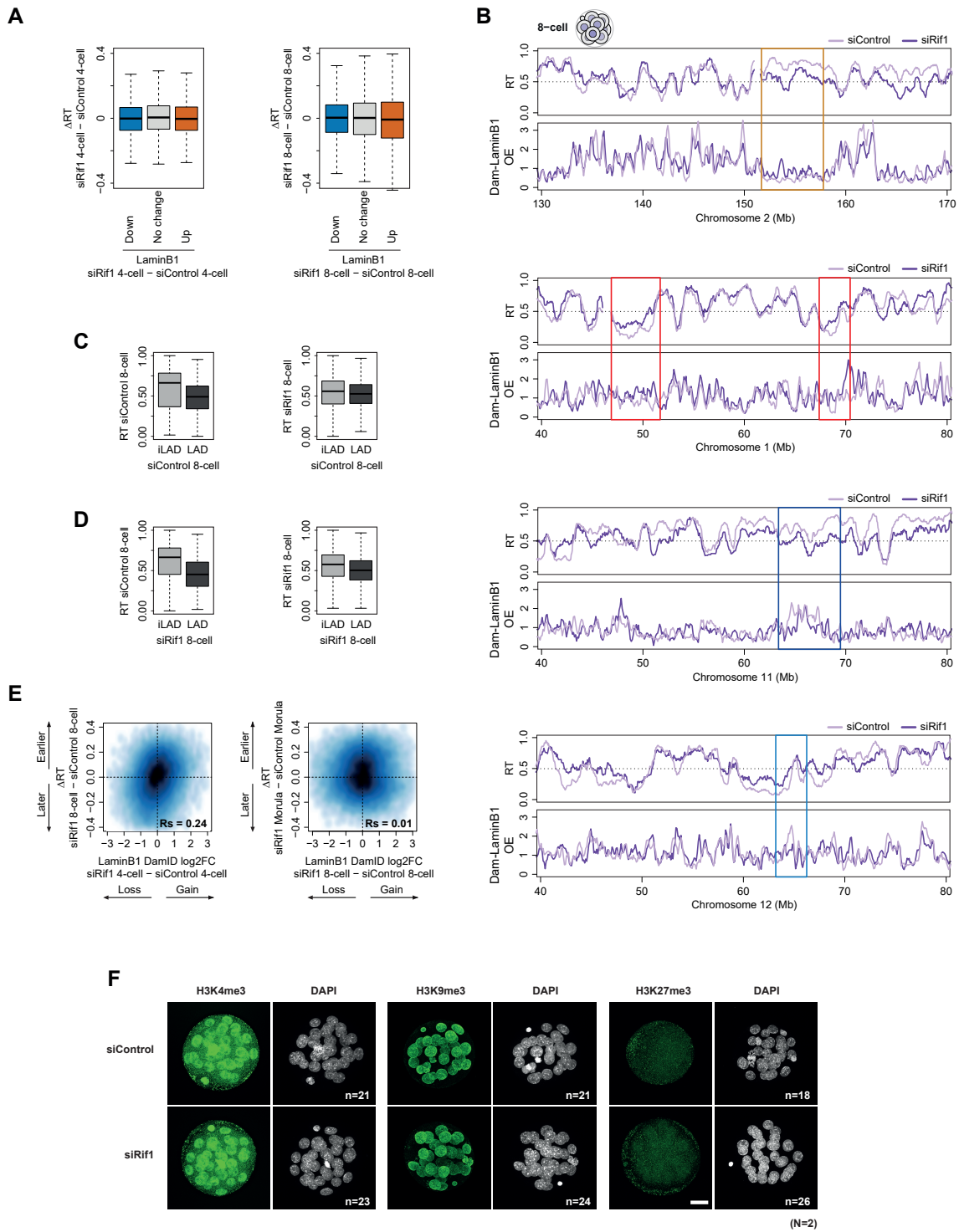
Figure S4. Changes of RT and remodelling of LADs distribution are independent upon RIF1 depletion.

(A) Box plots depicting differences in RT (ΔRT) between RIF1-depleted (siRif1) and control (siControl) embryos at genomic bins with significantly increased (red) and decreased (blue) lamina interactions upon RIF1 depletion at indicated stages. Box plots show median and the interquartile range (IQR), whiskers depict the smallest and largest values within 1.5 ×IQR.

(B) Representative examples of changes in RT and genome-lamina interactions upon RIF1 depletion at the 8-cell stage. For each track example, regions marked by the colored rectangles show different types of behavior. The color code corresponds to the colors in Figure 4D: later replication and gained lamina interaction upon RIF1 depletion (top, yellow); earlier replication and gained lamina interaction upon RIF1 depletion (second row, red); later replication and reduced lamina interaction upon RIF1 loss (third row, darker blue), and earlier replication with reduced lamina interaction upon RIF1 depletion (bottom, lighter blue).

(C and D) Box plots depicting RT values of RIF1-depleted (siRif1) and control (siControl) 8-cell embryos in iLADs and LADs from control (B) and from RIF1-depleted (C) embryos.

(E) Smoothed scatter plots of RT differences (ΔRT) between same stages of RIF1-depleted (siRif1) and control (siControl) embryos versus changes in Dam-LaminB1 OE values (log2FC) at the earlier developmental stage as indicated. Spearman's correlation (Rs) is indicated.

(F) Immunostaining of histone H3 lysine 4 trimethylation (H3K4me3), histone H3 lysine 9 trimethylation (H3K9me3), and histone H3 lysine 27 trimethylation (H3K27me3) in morula stage embryos after depletion of RIF1. Representative maximum intensity projection images are shown. Total number of embryos (n) analyzed in each condition from two independent experiments (N) are shown. Scale bar, 25 μm.

Supplementary Table S1.

Metrics and QC data of single-cell Repli-seq samples.

Overview of sample collection for scRepliseq in siRNA controls and siRNA for RIF1 at the 4-cell, 8-cell and morula stages, as well as mapping statistics and QC including coefficient of variation and whether cells passed QC: cells with high coefficient of variation were removed from the analyses.

Supplementary Table S2.

Differentially expressed genes at 4-cell stage embryo upon RIF1 depletion. Base counts, p values, log2 fold changes and statistics are indicated.

Supplementary Table S3.

Differentially expressed genes at 8-cell stage embryo upon RIF1 depletion. Base counts, p values, log2 fold changes and statistics are indicated.

References

1. Aladjem, M.I., Rodewald, L.W., Lin, C.M., Bowman, S., Cimbora, D.M., Brody, L.L., Epner, E.M., Groudine, M., and Wahl, G.M. (2002). Replication initiation patterns in the beta-globin loci of totipotent and differentiated murine cells: evidence for multiple initiation regions. Mol Cell Biol *22*, 442-452. 10.1128/MCB.22.2.442-452.2002.
2. Ryba, T., Hiratani, I., Sasaki, T., Battaglia, D., Kulik, M., Zhang, J., Dalton, S., and Gilbert, D.M. (2011). Replication timing: a fingerprint for cell identity and pluripotency. PLoS Comput Biol *7*, e1002225. 10.1371/journal.pcbi.1002225.
3. Alvarez, V., Bandau, S., Jiang, H., Rios-Szwed, D., Hukelmann, J., Garcia-Wilson, E., Wiechens, N., Griesser, E., Ten Have, S., Owen-Hughes, T., et al. (2023). Proteomic profiling reveals distinct phases to the restoration of chromatin following DNA replication. Cell Rep *42*, 111996. 10.1016/j.celrep.2023.111996.
4. Alabert, C., Bukowski-Wills, J.C., Lee, S.B., Kustatscher, G., Nakamura, K., de Lima Alves, F., Menard, P., Mejlvang, J., Rappsilber, J., and Groth, A. (2014). Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. Nat Cell Biol *16*, 281-293. 10.1038/ncb2918.
5. Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.W., Lyou, Y., Townes, T.M., Schubeler, D., and Gilbert, D.M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol *6*, e245. 10.1371/journal.pbio.0060245.
6. Fragkos, M., Ganier, O., Coulombe, P., and Mechali, M. (2015). DNA replication origin activation in space and time. Nat Rev Mol Cell Biol *16*, 360-374. 10.1038/nrm4002.
7. Dimitrova, D.S., and Gilbert, D.M. (1999). The spatial position and replication timing of chromosomal domains are both established in early G1 phase. Mol Cell *4*, 983-993. 10.1016/s1097-2765(00)80227-0.
8. Zhao, P.A., Sasaki, T., and Gilbert, D.M. (2020). High-resolution Repli-Seq defines the temporal choreography of initiation, elongation and termination of replication in mammalian cells. Genome Biol *21*, 76. 10.1186/s13059-020-01983-8.
9. Wang, W., Klein, K.N., Proesmans, K., Yang, H., Marchal, C., Zhu, X., Borrman, T., Hastie, A., Weng, Z., Bechhoefer, J., et al. (2021). Genome-wide mapping of human DNA replication by optical replication mapping supports a stochastic model of eukaryotic replication. Mol Cell *81*, 2975-2988 e2976. 10.1016/j.molcel.2021.05.024.
10. Nakatani, T., Schauer, T., Altamirano-Pacheco, L., Klein, K.N., Ettinger, A., Pal, M., Gilbert, D.M., and Torres-Padilla, M.E. (2024). Emergence of replication timing during early mammalian development. Nature *625*, 401-409. 10.1038/s41586-023-06872-1.
11. Xu, S., Wang, N., Zuccaro, M.V., Gerhardt, J., Baslan, T., Koren, A., and Egli, D. (2023). DNA replication in early mammalian embryos is patterned, predisposing lamina-associated regions to fragility. bioRxiv. 10.1101/2023.12.25.573304.
12. Peace, J.M., Ter-Zakarian, A., and Aparicio, O.M. (2014). Rif1 regulates initiation timing of late replication origins throughout the S. cerevisiae genome. PLoS One *9*, e98501. 10.1371/journal.pone.0098501.
13. Malzl, D., Peycheva, M., Rahjouei, A., Gnan, S., Klein, K.N., Nazarova, M., Schoeberl, U.E., Gilbert, D.M., Buonomo, S.C.B., Di Virgilio, M., et al. (2023). RIF1 regulates early replication timing in murine B cells. Nat Commun *14*, 8049. 10.1038/s41467-023-43778-y.
14. Yoshizawa-Sugata, N., Yamazaki, S., Mita-Yoshida, K., Ono, T., Nishito, Y., and Masai, H. (2021). Loss of full-length DNA replication regulator Rif1 in two-cell embryos is associated with zygotic transcriptional activation. J Biol Chem *297*, 101367. 10.1016/j.jbc.2021.101367.
15. Klein, K.N., Zhao, P.A., Lyu, X., Sasaki, T., Bartlett, D.A., Singh, A.M., Tasan, I., Zhang, M., Watts, L.P., Hiraga, S.I., et al. (2021). Replication timing maintains the global epigenetic state in human cells. Science *372*, 371-378. 10.1126/science.aba5545.
16. Foti, R., Gnan, S., Cornacchia, D., Dileep, V., Bulut-Karslioglu, A., Diehl, S., Buness, A., Klein, F.A., Huber, W., Johnstone, E., et al. (2016). Nuclear Architecture Organized by

Rif1 Underpins the Replication-Timing Program. Mol Cell *61*, 260-273. 10.1016/j.molcel.2015.12.001.

17. Buonomo, S.B., Wu, Y., Ferguson, D., and de Lange, T. (2009). Mammalian Rif1 contributes to replication stress survival and homology-directed repair. J Cell Biol *187*, 385-398. 10.1083/jcb.200902039.

18. Chapman, J.R., Barral, P., Vannier, J.B., Borel, V., Steger, M., Tomas-Loba, A., Sartori, A.A., Adams, I.R., Batista, F.D., and Boulton, S.J. (2013). RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. Mol Cell *49*, 858-871. 10.1016/j.molcel.2013.01.002.

19. Daxinger, L., Harten, S.K., Oey, H., Epp, T., Isbel, L., Huang, E., Whitelaw, N., Apedaile, A., Sorolla, A., Yong, J., et al. (2013). An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. Genome Biol *14*, R96. 10.1186/gb-2013-14-9-r96.

20. Enervald, E., Powell, L.M., Boteva, L., Foti, R., Blanes Ruiz, N., Kibar, G., Piszczek, A., Cavaleri, F., Vingron, M., Cerase, A., and Buonomo, S.B.C. (2021). RIF1 and KAP1 differentially regulate the choice of inactive versus active X chromosomes. EMBO J *40*, e105862. 10.15252/embj.2020105862.

21. Masser EA, N.T., Siefert JC, Goins D, SansamCG, Sansam CL (2023). Zebrafish Rif1 impacts zygotic genome activation, replication timing, and sex determination. eLife *12:RP87671*.

22. Siefert, J.C., Georgescu, C., Wren, J.D., Koren, A., and Sansam, C.L. (2017). DNA replication timing during development anticipates transcriptional programs and parallels enhancer activation. Genome Res *27*, 1406-1416. 10.1101/gr.218602.116.

23. Adams, I.R., and McLaren, A. (2004). Identification and characterisation of mRif1: a mouse telomere-associated protein highly expressed in germ cells and embryo-derived pluripotent stem cells. Dev Dyn *229*, 733-744. 10.1002/dvdy.10471.

24. Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., et al. (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402-405. 10.1038/nature13986.

25. Moindrot, B., Audit, B., Klous, P., Baker, A., Thermes, C., de Laat, W., Bouvet, P., Mongelard, F., and Arneodo, A. (2012). 3D chromatin conformation correlates with replication timing and is conserved in resting cells. Nucleic Acids Res *40*, 9470-9481. 10.1093/nar/gks736.

26. Yaffe, E., Farkash-Amar, S., Polten, A., Yakhini, Z., Tanay, A., and Simon, I. (2010). Comparative analysis of DNA replication timing reveals conserved large-scale chromosomal architecture. PLoS Genet *6*, e1001011. 10.1371/journal.pgen.1001011.

27. Chen, N., and Buonomo, S.C.B. (2023). Three-dimensional nuclear organisation and the DNA replication timing program. Curr Opin Struct Biol *83*, 102704. 10.1016/j.sbi.2023.102704.

28. Borsos, M., Perricone, S.M., Schauer, T., Pontabry, J., de Luca, K.L., de Vries, S.S., Ruiz-Morales, E.R., Torres-Padilla, M.E., and Kind, J. (2019). Genome-lamina interactions are established de novo in the early mouse embryo. Nature *569*, 729-733. 10.1038/s41586-019-1233-0.

29. Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., Gao, L., et al. (2017). 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. Cell *170*, 367-381 e320. 10.1016/j.cell.2017.06.029.

30. Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y., et al. (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. Nature *547*, 232-235. 10.1038/nature23263.

31. Pal, M., Altamirano-Pacheco, L., Schauer, T., and Torres-Padilla, M.E. (2023). Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity. Genes Dev *37*, 901-912. 10.1101/gad.350799.123.

32. Nakatani, T., Lin, J., Ji, F., Ettinger, A., Pontabry, J., Tokoro, M., Altamirano-Pacheco, L., Fiorentino, J., Mahammadov, E., Hatano, Y., et al. (2022). DNA replication fork speed

underlies cell fate changes and promotes reprogramming. Nat Genet *54*, 318-327. 10.1038/s41588-022-01023-0.

33. Cornacchia, D., Dileep, V., Quivy, J.P., Foti, R., Tili, F., Santarella-Mellwig, R., Antony, C., Almouzni, G., Gilbert, D.M., and Buonomo, S.B. (2012). Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. EMBO J *31*, 3678-3690. 10.1038/emboj.2012.214.
34. Seller, C.A., and O'Farrell, P.H. (2018). Rif1 prolongs the embryonic S phase at the Drosophila mid-blastula transition. PLoS Biol *16*, e2005687. 10.1371/journal.pbio.2005687.
35. Yamazaki, S., Ishii, A., Kanoh, Y., Oda, M., Nishito, Y., and Masai, H. (2012). Rif1 regulates the replication timing domains on the human genome. EMBO J *31*, 3667-3677. 10.1038/emboj.2012.180.
36. Fernandez-Capetillo, O., Lee, A., Nussenzweig, M., and Nussenzweig, A. (2004). H2AX: the histone guardian of the genome. DNA Repair (Amst) *3*, 959-967. 10.1016/j.dnarep.2004.03.024.
37. Xu, L., and Blackburn, E.H. (2004). Human Rif1 protein binds aberrant telomeres and aligns along anaphase midzone microtubules. J Cell Biol *167*, 819-830. 10.1083/jcb.200408181.
38. Dileep, V., and Gilbert, D.M. (2018). Single-cell replication profiling to measure stochastic variation in mammalian replication timing. Nat Commun *9*, 427. 10.1038/s41467-017-02800-w.
39. Palmerola, K.L., Amrane, S., De Los Angeles, A., Xu, S., Wang, N., de Pinho, J., Zuccaro, M.V., Taglialatela, A., Massey, D.J., Turocy, J., et al. (2022). Replication stress impairs chromosome segregation and preimplantation development in human embryos. Cell *185*, 2988-3007 e2920. 10.1016/j.cell.2022.06.028.
40. Greil, F., Moorman, C., and van Steensel, B. (2006). DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. Methods Enzymol *410*, 342-359. 10.1016/S0076-6879(06)10016-6.
41. Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178-192. 10.1016/j.cell.2013.02.028.
42. Alabert, C., and Groth, A. (2012). Chromatin replication and epigenome maintenance. Nature Reviews Molecular Cell Biology *13*, 153-167. 10.1038/nrm3288.
43. Alver, R.C., Chadha, G.S., Gillespie, P.J., and Blow, J.J. (2017). Reversal of DDK-Mediated MCM Phosphorylation by Rif1-PP1 Regulates Replication Initiation and Replisome Stability Independently of ATR/Chk1. Cell Rep *18*, 2508-2520. 10.1016/j.celrep.2017.02.042.
44. Dan, J., Liu, Y., Liu, N., Chiourea, M., Okuka, M., Wu, T., Ye, X., Mou, C., Wang, L., Wang, L., et al. (2014). Rif1 maintains telomere length homeostasis of ESCs by mediating heterochromatin silencing. Dev Cell *29*, 7-19. 10.1016/j.devcel.2014.03.004.
45. Liu, C., Yu, P., Ren, Z., Yao, F., Wang, L., Hu, G., Li, P., and Zhao, Q. (2023). Rif1 Regulates Self-Renewal and Impedes Mesendodermal Differentiation of Mouse Embryonic Stem Cells. Stem Cell Rev Rep *19*, 1540-1553. 10.1007/s12015-023-10525-1.
46. Rodriguez-Terrones, D., Gaume, X., Ishiuchi, T., Weiss, A., Kopp, A., Kruse, K., Penning, A., Vaquerizas, J.M., Brino, L., and Torres-Padilla, M.E. (2018). A molecular roadmap for the emergence of early-embryonic-like cells in culture. Nat Genet *50*, 106-119. 10.1038/s41588-017-0016-5.
47. Li, P., Wang, L., Bennett, B.D., Wang, J., Li, J., Qin, Y., Takaku, M., Wade, P.A., Wong, J., and Hu, G. (2017). Rif1 promotes a repressive chromatin state to safeguard against endogenous retrovirus activation. Nucleic Acids Res *45*, 12723-12738. 10.1093/nar/gkx884.
48. Halliwell, J.A., Martin-Gonzalez, J., Hashim, A., Dahl, J.A., Hoffmann, E.R., and Lerdrup, M. (2024). Sex-specific DNA-replication in the early mammalian embryo. Nat Commun *15*, 6323. 10.1038/s41467-024-50727-w.

39

49. Xu, S., Wang, N., Zuccaro, M.V., Gerhardt, J., Iyyappan, R., Scatolin, G.N., Jiang, Z., Baslan, T., Koren, A., and Egli, D. (2024). DNA replication in early mammalian embryos is patterned, predisposing lamina-associated regions to fragility. Nat Commun *15*, 5247. 10.1038/s41467-024-49565-7.
50. Deng, Q., Ramskold, D., Reinius, B., and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science *343*, 193-196. 10.1126/science.1245316.
51. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682. 10.1038/nmeth.2019.
52. Stringer, C., and Pachitariu, M. (2024). Cellpose3: one-click image restoration for improved cellular segmentation. bioRxiv, 2024.2002.2010.579780. 10.1101/2024.02.10.579780.
53. Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. Nat Methods *18*, 100-106. 10.1038/s41592-020-01018-x.
54. van der Walt, S., Schonberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., and scikit-image, c. (2014). scikit-image: image processing in Python. PeerJ *2*, e453. 10.7717/peerj.453.
55. Techer, H., Koundrioukoff, S., Azar, D., Wilhelm, T., Carignon, S., Brison, O., Debatisse, M., and Le Tallec, B. (2013). Replication dynamics: biases and robustness of DNA fiber analysis. J Mol Biol *425*, 4845-4855. 10.1016/j.jmb.2013.03.040.
56. Conti, C., Sacca, B., Herrick, J., Lalou, C., Pommier, Y., and Bensimon, A. (2007). Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. Mol Biol Cell *18*, 3059-3067. 10.1091/mbc.e06-08-0689.
57. Tankyevych, O., Talbot, H. & Dokladal, P (2008). Curvilinear morpho-Hessian filter. Internal Symposium on Biomedical imaging: From Nano to Macro (ISBI). 1011-1014.
58. Breiman, L., Friedman, J., Stone, C. & Olshen, R (1984). Classification and Regression Trees. CRC Press.
59. Bartlett, D.A., Dileep, V., Baslan, T., and Gilbert, D.M. (2022). Mapping Replication Timing in Single Mammalian Cells. Curr Protoc *2*, e334. 10.1002/cpz1.334.
60. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359. 10.1038/nmeth.1923.
61. Nilsen, G., Liestol, K., Van Loo, P., Moen Vollan, H.K., Eide, M.B., Rueda, O.M., Chin, S.F., Russell, R., Baumbusch, L.O., Caldas, C., et al. (2012). Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. BMC Genomics *13*, 591. 10.1186/1471-2164-13-591.
62. Benaglia, T. (2009). Mixtools: An R Package for Analyzing Finite Mixture Models.
63. Takada, S., Lis, J.T., Zhou, S., and Tjian, R. (2000). A TRF1:BRF complex directs Drosophila RNA polymerase III transcription. Cell *101*, 459-469. 10.1016/s0092-8674(00)80857-0.
64. Takahashi, S., Miura, H., Shibata, T., Nagao, K., Okumura, K., Ogata, M., Obuse, C., Takebayashi, S.I., and Hiratani, I. (2019). Genome-wide stability of the DNA replication program in single mammalian cells. Nat Genet *51*, 529-540. 10.1038/s41588-019-0347-5.
65. Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics Monograph *38*. doi:10.1137/1.9781611970319.
66. Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc *9*, 171-181. 10.1038/nprot.2014.006.
67. Pal, M., Kind, J., and Torres-Padilla, M.E. (2021). DamID to Map Genome-Protein Interactions in Preimplantation Mouse Embryos. Methods Mol Biol *2214*, 265-282. 10.1007/978-1-0716-0958-3_18.

68. Kind, J., Pagie, L., de Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C.A., Amendola, M., et al. (2015). Genome-wide maps of nuclear lamina interactions in single human cells. Cell *163*, 134-147. 10.1016/j.cell.2015.08.040.

69. Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., and van Steensel, B. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212-224. 10.1016/j.cell.2010.09.009.

70. Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., Xia, W., et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652-657. 10.1038/nature18606.

71. Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature *537*, 553-557. 10.1038/nature19361.

72. Wang, C., Liu, X., Gao, Y., Yang, L., Li, C., Liu, W., Chen, C., Kou, X., Zhao, Y., Chen, J., et al. (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nat Cell Biol *20*, 620-631. 10.1038/s41556-018-0093-4.

73. Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature *537*, 558-562. 10.1038/nature19362.

74. Abe, K., Schauer, T., and Torres-Padilla, M.E. (2022). Distinct patterns of RNA polymerase II and transcriptional elongation characterize mammalian genome activation. Cell Rep *41*, 111865. 10.1016/j.celrep.2022.111865.

75. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. Nat Methods *16*, 1226-1232. 10.1038/s41592-019-0582-9.

41

Part VI
Unpublished manuscript

The establishment of nuclear organization in mouse embryos is orchestrated by multiple epigenetic pathways

The establishment of nuclear organization in mouse embryos is orchestrated by multiple epigenetic pathways

Mrinmoy Pal[1], Tamas Schauer[1], Adam Burton[1], Tsunetoshi Nakatani[1], Federico Pecori[1], Alicia Hernández-Giménez[2], Iliya Nadelson[1], Marc A. Marti-Renom[2,3,4,5] & Maria-Elena Torres-Padilla[1,6] *

[1] *Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München D-81377 München, Germany*

[2] *CNAG, National Center for Genome Analysis, Baldiri Reixac 4, 08028 Barcelona, Spain*

[3] *Centre for Genomic Regulation, The Barcelona Institute for Science and Technology, Carrer del Doctor Aiguader 88, 08003 Barcelona, Spain.*

[4] *ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

[5] *Universitat Pompeu Fabra (UPF), 08002 Barcelona, Spain*

[6] *Faculty of Biology, Ludwig-Maximilians Universität, München, Germany.*

* Correspondence: torres-padilla@helmholtz-muenchen.de

Summary

The folding of the genome in the 3D-nuclear space is a fundamental eukaryotic feature for the regulation of all DNA-related processes. How nuclear organization is first established during development is not understood at the molecular level. The association of the genome with the nuclear lamina into lamina-associated domains (LADs) represents the earliest feature of nuclear organization. Here, we performed a gain-of-function screening to investigate mechanisms affecting LAD establishment *in vivo,* in mouse embryos. We identify chromatin pathways that lead to severe disruption of nuclear architecture in zygotes and 2-cell stage embryos. Remarkably, our data indicate that the initial establishment of LADs in zygotes is dispensable for early development as embryos with disrupted LADs can rebuild nuclear architecture at the 2-cell stage. Our work provides an unprecedented resource for the molecular understanding of nuclear organization and highlights dependencies between chromatin pathways and structural nuclear components that guide genome-lamina interactions at the beginning of development.

Highlights

- Low-input genomics screening *in vivo* generates an unprecedented and resourceful dataset in mouse embryos

- Identification of molecular pathways affecting nuclear architecture after fertilization

- Maternal inheritance, rather than active H3K27 methylation, contributes to LAD formation after fertilization

- LAD boundaries are reorganized based on positional information provided by H3K4me3 and H3K9me3 domains

Keywords

3D genome organization; mouse embryo; epigenetics; nuclear architecture; LAD

1

208

Introduction

How the genome folds into the nucleus of eukaryotic cells is fundamental for all DNA-dependent processes including transcription and replication. The packaging of the chromatin into the 3D space renders specific DNA sequences accessible for the transcription machinery[1], but also plays a role in protecting DNA from damage[2]. Chromosomes undergo long-range intrachromosomal interactions that lead to their folding into highly organized structures in the 3D space such as A-B compartments[3] and topologically associating domains or TADs[4–6]. Such organization is linked to replication and transcription. Compartments of type A largely encompass active chromatin regions, which replicate early during S-phase, and are characterized by higher chromatin accessibility. B compartments, in contrast, comprise heterochromatic regions, replicate later and are largely inaccessible[4,7,8]. Likewise, TADs, which are a smaller unit of organization than compartments, are known to constrain enhancer-promoter interactions enabling a more robust transcriptional regulatory program[6,9–11]. In addition to its folding into TADs and compartments, the genome also organizes with respect to nuclear landmarks and organelles. Amongst these, the partitioning of the genome into regions that associate with the nuclear lamina and those which are more centrally positioned is a major pillar of the 3D genome organization. Lamina-associated domains (LADs) are large genomic regions ranging from 100 kb up to 10 Mb in size that associate with the nuclear lamina, a heterogenous meshwork of intermediate filaments composed of lamins – a structural component of the nuclear envelope[12–15].

In addition to their structural role, lamins associate with nuclear pore components[16–18] and interact with the cytoskeleton[19–21]. The nuclear lamina is thought to protect chromatin from mechanical stress and/or forces through actin and myosin remodeling, which can be transmitted to the nucleus through the LINC complex[22–24]. The latter is formed by KASH and SUN domain proteins at the outer nuclear lamina, which sense and transmit mechanical stimuli to the nucleus by bridging to the inner membrane. Thus, the nuclear lamina enables functional interactions with other components of the nuclear membrane to ensure nucleo-cytoplasmic homeostasis[25,26] in addition to providing a site for chromatin anchoring. The nuclear lamina is primarily constituted of two types of Lamin proteins: B and A/C. Zygotic LaminB1 knock-out mice die at birth due to lung and bone abnormalities[27]. The homozygous mutants cannot breed[27] and thus a potential role of maternally inherited LaminB1 has not been addressed. Similarly, LaminA is dispensable for early development but, as in humans, null LaminA mutant mice develop muscular dystrophies[28].

2

LADs have been identified across all cell types studied so far[14,29,30], except for fully grown oocytes[31]. Across cell types, LADs share distinctive features including high AT content, low gene density and tend to contain functionally repressed chromatin[14,29]. A certain degree of overlap with genomic regions associated with the nucleolus (NADs) and LADs has also been documented[32,33], presumably due to the silencing nature of the surrounding nucleolar environment and to the localization of a pool of lamins to the nucleolus[34,35]. Generally, LADs are late replicating and correspond primarily to B compartments while inter-LADs (iLADs) replicate early during S-phase and correspond to A compartments[8,13,36]. iLADs display higher transcriptional activity than LADs and dislodging of transcriptional units from the nuclear lamina has been observed upon gene activation[37–39]. However, tethering a reporter or endogenous gene to the nuclear periphery does not always result in transcriptional silencing[40,41], indicating that nuclear positioning is not the sole determinant of gene expression.

In some cell types, LAD boundaries are delineated by sharp changes in H3K4me2 and H3K27me3[13,42]. LADs in differentiated and embryonic stem cells are enriched in H3K9me2[13,43,44] and inhibition of the H3K9 methyltransferase Ehmt2 (G9a) reduces their contacts with the nuclear lamina[42,45,46]. Euchromatic factors and histone acetylation have also been proposed to mediate radial segregation of chromatin in *C. elegans*[47]. However, disruption of key components of the nuclear lamina, including lamins themselves, results in a largely unaffected LAD landscape[48], suggesting that once interactions with the nuclear lamina are established, LADs are robust. Thus, the investigation of the pathways that lead to the initial establishment of LADs at the beginning of development is fundamental for our understanding of the mechanisms that direct nuclear organization.

Previous work using DNA adenine methyltransferase identification (DamID) for LaminB1 to map genome-nuclear lamina interactions in oocytes and early mouse embryos revealed that LADs are undetectable in mature oocytes but become rapidly established after fertilization[31]. Dynamic remodeling of LADs occurs after fertilization, in particular prior to the completion of the maternal-to-zygotic transition (MZT) at the 2-cell stage. This reorganization takes place after the first mitosis but also through the progression of the second cell cycle and is characterized by the dislodgement of genomic regions with typical LAD features away from the nuclear lamina[31,49]. In spite of unusual LAD features at the 2-cell stage, the reorganization of LADs and iLADs at the 2-cell stage follows the typical correlation with transcriptional activity: 2-cell stage specific LADs contain genes that are generally silent at the mid- and late 2-cell stages[31,49]. Remarkably, however, around 40% of the genome in mouse zygotes constitutes constant LADs, which are genomic regions that are

LADs in all cell types studied, or constant iLADs. This indicates that while significant remodeling of LADs occurs during development and differentiation, almost half of the genome adopts its 'long-life' positioning immediately after fertilization. We previously showed that the demethylase Kdm5b affects the establishment of LADs in the paternal chromatin right after fertilization[31]. More recently, maternal H3K27me3 has been shown to influence the heterogeneity of genome-nuclear lamina interactions between cells at the 2-cell stage[50]. However, our understanding of the structural and chromatin pathways involved in this process is limited and the factors that regulate the dynamic reshuffling of genome organization in early embryos remain largely unknown.

Here, we set out to investigate the epigenetic and structural components that dictate LAD establishment in mouse embryos and their involvement in the dynamic changes in LADs during the maternal-to-zygotic transition. We performed 50 different perturbations to examine the impact of several molecular pathways and generated 183 DamID libraries of mouse embryos. By performing a mid-scale, low-input genomics-based screening in mouse embryos, we demonstrate that multiple chromatin pathways contribute to the integrity of LAD establishment and that different chromatin pathways can influence the nuclear organization programme at the time of zygotic genome activation. We also show that LAD establishment in zygotes is dispensable for early developmental progression. Instead, mouse embryos show a remarkable capacity to reset nuclear organization at the 2-cell stage, even when this process is prevented in zygotes. Our work presents unprecedented data to enable understanding of nuclear organization and their hierarchies at the beginning of development.

Results

Multiple phenotypes of disrupted nuclear organization emerge upon perturbation of chromatin and nuclear structural pathways after fertilization

To provide an in-depth molecular understanding of the regulation of nuclear organization *in vivo*, we performed a mid-scale screening in mouse embryos to identify determinants of LAD establishment. We aimed to cover a range of molecular pathways, including histone modifiers related to genome organization and LADs in somatic cells; chromatin anchors, and structural proteins of the nuclear envelope including cytoplasmic and nuclear actomyosin (Fig. 1A and Table S1). We devised a 2-step screening strategy with a first screening phase with pooled candidates within similar molecular pathways using gain-of-function or dominant negative approaches (Fig. S1A). The individual pool constituents are shown in Fig. S1A and described in detail in Table S1.

4

As a readout, we mapped LADs using DamID for LaminB1 with our optimized low-input DamID protocol[31,51,52]. For each pool of candidates, we performed mRNA microinjections in hybrid (C57BL/6J×CBA/H × DBA/2J) early zygotes immediately after fertilization and collected DamID libraries using at least three biological replicates at the late zygote stage, prior to the onset of the first embryonic mitosis (Fig. 1B and Table S2). We verified the perturbation of the respective molecular pathways by immunostaining, including global changes in the levels of the expected histone modifications (Fig. S1B). Unsupervised principal component analysis (PCA) revealed that while some samples were close to control zygotes, other candidate pools deviated along the PC1 or the PC2 (Fig. 1C). For example, the pool containing exportin6 and a cortical actin nucleator (Pool B), which lead to changes in nuclear and cortical actin in early embryos[53,54], as well as the pool comprising of histone variant macroH2A and H1 subtypes (Pool M) were located furthest away from control samples on PC2 and PC1, respectively. This suggests that perturbation of some of the candidate proteins in these pools triggered altered genome-nuclear lamina interactions. Additional candidate pools showed spreading along PC1 away from the controls, albeit less pronounced (Fig. 1C). To examine the candidate pools in detail, we called LADs using a two-state hidden Markov model (HMM) based on Dam-LaminB1 methylation levels[51]. We confirmed changes in nuclear positioning of selected LADs and iLADs by 3D-DNA FISH (Fig. S1C). Visual inspection of chromosome tracks revealed a highly similar LAD profile between control samples and our previously mapped LADs in wildtype zygotes[31] (Fig. 1C-D). Globally, some candidate pools had no major differences in their LAD profiles compared to controls including, for example, manipulation of proteins of the nuclear pore such as expression of the dominant negative nucleoporin 98 and Tpr (Pool C) (Fig. 1D). However, a group of pools displayed severely impaired LAD profiles, which included for example the expression of Kdm6a/b H3K27me3 demethylases (Pool F) and H4K20 methyltransferases (Pool G) (Fig. 1D). Manipulation of these pathways resulted in an apparent strong increase in average LAD size and in the proportion of the genome associated with the nuclear lamina, potentially due to the inability of HMM to distinguish between LADs and iLADs when the dynamic range of genome-wide OE values is small (Fig. S1D and S1E). Another group of samples displayed a phenotype with distinguishable LAD and iLADs but at different genomic locations compared to controls. These included for example the subtypes of H1 and histone macroH2A (Pool M) and a histone deacetylase group composed of Hdac1/6 and Sirt1 (Pool L) (Fig. 1D), suggesting that manipulation of histone content in zygotic chromatin and/or of global acetylation results in LAD formation at aberrant genomic regions. Overall, the effects observed on LAD size and number varied widely across all the tested candidate pools (Fig. S1D and S1E).

5

To further characterize the nature and extent of the perturbations, we categorized phenotypes in relation to control LADs in zygotes based on metaplots of LaminB1-DamID scores from all the screening pool samples over control LAD boundaries (Fig. 1E). This revealed four main patterns of nuclear organization, which we classified as: i) unchanged or strengthened LADs (increased Dam-LaminB1 methylation levels within LADs); ii) weakening of LADs (decreased Dam-LaminB1 methylation levels within LADs; iii) collapse of control LADs (similar Dam-LaminB1 methylation levels between control LADs and iLADs), and iv) inversion of LAD architecture (control LADs become iLADs and vice versa) (Fig. 1E). To investigate whether these phenotypes are due to restructuring of LAD boundaries, to the emergence of new LADs, to a change in the strength of interactions with the nuclear lamina, or a combination of these, we determined *de novo* LAD coordinates in perturbed embryos across all our screening pools. We then compared DamID scores in such *de novo* called LAD coordinates against control LAD coordinates (Fig. 1F and Fig. S1F). Pools such as the one expressing Kdm6a/b (Pool F) led to an overall reduction of interactions with the nuclear lamina of control LADs leading to 'weakened' LADs (Fig. 1F). Such reduced interactions with the nuclear lamina were most pronounced for the pools that elicited an inversion of LADs. Although perturbation with these pools (M and N) resulted in a clear partitioning of the genome into regions that associate to the nuclear lamina and regions that do not, higher *de novo* DamID scores are now found in control iLADs and the lower DamID scores correspond to control LADs, explaining the inversion phenotype (Fig. 1F and Fig. S1F). Thus, expression of H1 subtypes and macroH2A, and of H3K9 methyltransferases Ehmt1/2 and Setdb1/2 leads to an inverted LAD architecture. This analysis also confirmed that the phenotype of LAD strengthening is characterized by changes in the DamID methylation levels within existing LADs and not by a major global repositioning of genomic regions (see Pool E; Fig. 1F and S1G). This is potentially due to the reduced nuclear size in zygotes from Pool E (Fig. S1H)[55], and suggests that nuclear size is important for increased strength of genome-nuclear lamina interactions.

Overall, the extent of the phenotypes we observed varied largely, ranging from only a small fraction of the genome affected, for example upon remodeling of nuclear and cortical actin (Pool B) to practically most of the genome affected when either the H3K27me3 demethylases (Kdm6a/b, Pool F) or H3K9me2 demethylases (Kdm3a/b, Pool H) were expressed (Fig. S1G). The alluvial plots also highlight the large-scale interchange of genomic regions between LADs and iLADs caused by the hits that led to an inversion phenotype (Pools M and N; Fig. S1E) and to a lesser degree in the pools that led to a collapse of control LAD architecture (Pools K and L; Fig. S1E) indicating that the collapse is an intermediate phenotype between control LADs and

6

inverted LADs. Indeed, while in Pools K and L (collapse) some conversion of LADs into iLADs occurs, control LADs and iLADs have overall similar median DamID values (Fig. S1F-G and Fig. 1F), in contrast to Pools M and N in which median DamID values are inversed such that control iLADs have now higher DamID values (Fig. 1F and Fig. S1F). Because the two parental genomes establish LADs with different features[31], we asked whether the paternal and maternal allele are equally affected by the candidate pools. In general, most perturbations affect both pronuclei (Fig. S1I). LADs were inverted in both paternal and maternal pronuclei for those pathways that led to LAD inversion (Pool N; Fig. S1J). However, we observed a larger effect on the paternal pronucleus, for example, in zygotes in which we targeted the H4K20 pathway (Pool G) or upon expression of H3K9me2 demethylases (Kdm3a/b, Pool H) (Fig. S1I-J). Thus, our phase I screening recovered multiple phenotypes of disrupted nuclear organization and suggests that several pathways can influence nuclear architecture after fertilization in the zygote.

Integration of nuclear organization phenotypes reveals chromatin features associated with disrupted LAD establishment

Considering the variety of phenotypes elicited, we next asked whether similar pathways affect similar genomic regions. As a first approach, we performed a genome-wide correlation analysis of LaminB1-DamID scores across all the candidate pools, including 14 different perturbation conditions and the controls (Fig. 2A). This analysis revealed two major clusters, which largely coincided with the specific phenotypes of nuclear organization that we described above. The first major cluster (Cluster I) primarily contained zygotes in which genomic interactions with the nuclear lamina remained globally unaffected or became stronger upon manipulation (Fig. 2A). The second major cluster (Cluster II) mainly included the molecular pathways that led to disruption of control LADs (Fig. 2A). Overall, the clusters were defined primarily by pools that contained candidates targeting components of nuclear envelope or cytoskeleton (Cluster I) and pools of heterochromatin manipulation (Cluster II). Cluster II further subdivided into two smaller clusters: Cluster IIa included candidate pools that led to the weakening of LAD interactions with the nuclear lamina, such as Kdm6a/b (Pool F), Kdm3a/b (Pool H), as well as Lsd1 and Kdm7a/c (Pool I); and Cluster IIb included the pathways that led to collapsed (Pools K and L) or inverted control LAD architecture (Pools M and N) (Fig. 2A).

Next, we asked whether similar genomic regions respond in the same way to the perturbation of related molecular pathways. We first extracted the genomic bins (100-kb resolution) that displayed significantly higher and lower LaminB1-DamID scores upon perturbation compared to

7

214

controls. The number of genomic bins affected differed across pooled candidate pathways (Fig. 2B and Table S3). Overall, perturbation of histone content (Pool M) or the expression of H3K9 methyltransferases Ehmt1/2 and Setdb1/2 (Pool N) contained the highest number of genomic bins affected (Fig. 2B). For example, upon expression of the histone H1 subtypes and macroH2A (Pool M), 2418 genomic bins gain interactions with the nuclear lamina. Out of these, only 22 correspond to control LADs while 2396 correspond to iLADs (Fig. 2B), in line with our conclusion of LAD inversion. We also analyzed the degree of overall similarity between affected genomic bins across pools. We observed a high overlap in the genomic bins that changed under conditions in which a LAD inversion or collapse phenotype emerged (top left corner of the heatmap, Fig. 2C). This group consisted of candidate pools K, L, M and N, which affect very different molecular pathways such as histone acetylation (Hdac1/6, Sirt1) in Pool L versus histone content (subtypes of H1/macroH2A) in Pool M (Fig. 2C). This trend was overall similar also for regions that lost interactions with the nuclear lamina (Fig. S2A-B). These observations suggest that the same genomic regions are affected upon LAD disruption, even when they emerge from different molecular perturbations.

We next asked whether our pooled candidate hits affect genomic regions with specific epigenomic features. We analyzed chromatin marks, accessibility, compartments, and RNA Polymerase (Pol) II occupancy in all regions that gained or lost interactions with the nuclear lamina. Zygotic wildtype LADs correspond mostly to B compartment[31], are globally less accessible, tend to be depleted of H3K36me3 and RNA Pol II, but also display lower H3K9me3 levels than iLADs (Fig. 2D). Overall, the perturbations of chromatin pathways led to increased nuclear lamina association of genomic regions in wildtype A compartments, with strong RNA Pol II occupancy and higher accessibility (Fig. 2D-E). In other words, perturbation of all the chromatin pathways that we tested led to increased genome-nuclear lamina interactions of control iLADs. In contrast, the pools in which we disrupted structural nuclear membrane components, anchor proteins, and actomyosin led instead to increased nuclear lamina association of specific regions with no strong compartment score that were overall depleted of RNA Pol II (Fig. 2D). An exception to this was the group of zygotes in which we expressed nuclear actin deficient in polymerization (Pool E), in which regions that increased interactions with the nuclear lamina were B compartment regions depleted from RNA Pol II (Fig. 2D-E). Analysis of genomic regions that lose interactions with the nuclear lamina upon perturbation led to a similar clustering (Fig. S2C). Additionally, we note that all the conditions in which we perturbed chromatin modifiers, maternally marked H3K27me3 regions that lose H3K27me3 methylation after fertilization gained interactions with the nuclear lamina (Fig. 2D).

Overall, our analyses suggest that chromatin pathways such as those regulating H3K27me2/me3 and H3K9me2/me3 affect nuclear organization potentially by restricting the association of genome regions marked by specific chromatin properties from the nuclear lamina.

Identification of pathways that regulate LAD reorganization after the first mitotic division

Having established that perturbing multiple molecular pathways can disrupt LAD formation after fertilization in zygotes, we then searched for factors that alter LAD remodeling during the maternal-to-zygotic transition, which occurs by the late 2-cell stage[31]. We performed a new phase I screening approach at the late 2-cell stage using an auxin degron to temporally control DamID[31](Fig. 3A). As above, we generated at least three independent replicates of LaminB1 DamID per condition for each pool of candidate targets as well as control samples (Fig. S1A and Table S2) and confirmed the perturbation of the respective molecular pathways by immunostaining (Fig. S3A). Note that we could not map LADs in embryos in which we perturbed nuclear export and cortical actin dynamics (Pool B) because they cannot complete cytokinesis and do not reach the 2-cell stage[53]. PCA revealed that in general the replicates grouped together and that most samples did not majorly deviate from the control samples (Fig. 3B). An exception to this were embryos in which we i) expressed the H3K4me2 demethylase Lsd1 together with the H3K9me2/K27me2 demethylases Kdm7a/c (Pool I); ii) the H3K9me2/me3 methyltransferases Ehmt1/2 and Setdb1/2 (Pool N), and iii) the subtypes of histone H1 and histone macroH2A (Pool M), which all grouped away from the rest of the samples (Fig. 3B).

To better understand how the molecular pathways that we perturbed through gain-of-function affect nuclear organization at the late 2-cell stage, we applied HMM to determine LADs, their size, number, and the proportion of the genome associated to the nuclear lamina for each condition. We confirmed changes in nuclear positioning of selected LADs and iLADs by 3D-DNA FISH (Fig. S3B). Visual inspection of DamID scores and the corresponding LADs and iLADs along a representative chromosome for all our candidate pools revealed, in general, more subtle changes in the LAD structure, when compared to the same perturbations in zygotes (Fig. 3C against Fig. 1D). This may suggest that the nuclear organization in zygotes is more sensitive to such global perturbations than that of 2-cell stage embryos. However, the perturbation of specific molecular pathways led to a clear change in LAD architecture. These included, for example, the H3K9me2/me3 'writer' pathway (Ehmt1/2 and Setdb1/2; Pool N in Fig. 3C). The number of LADs was also affected, ranging from 278 (Suv39h1 and Hp1α/γ; Pool J) to 646 (dominant negative Nup98 and Tpr; Pool C) compared to 831 in the controls (Fig. S3C). This was accompanied by

9

changes in the LAD size and in the proportion of the genome associated with the nuclear lamina (Fig. S3C-D).

Next, we categorized phenotypes of nuclear disruption at the 2-cell stage based on metaplot profiles. Piling-up LaminB1-DamID scores over control LAD boundaries led to the identification of two major 2-cell stage LAD phenotypes with different levels of perturbation: i) unaffected-to-weaker LAD/iLAD distinction with a globally preserved wildtype LAD structure (decreased Dam-LaminB1 methylation levels within control LADs), and ii) collapse of control LAD architecture (overall similar Dam-LaminB1 methylation levels between control LADs and iLADs) (Fig. 3D). Most candidate pools belonged to the unaffected-to-weaker group, including those targeting the H4K20 pathway (Pool G), H3K27me3 demethylation (Pool F), histone deacetylation (Pool L), as well as all the pools perturbing nuclear membrane structural components and regulators of nucleo/cytoskeletal dynamics (Pools A, C, D, E) (Fig. 3D, top panel). However, manipulation of the constitutive heterochromatin H3K9me3 pathway (Pool J), expression of the H3K4me2, H3K9me2/K27me2 demethylases Lsd1 and Kdm7a/c (Pool I), H3K9me2/3 methyltransferases Ehmt1/2 and Setdb1/2 (Pool N) and also of histone H1 subtypes/macroH2A (Pool M), led to a profound alteration of DamID values across the control 2-cell stage LAD boundaries (Fig. 3D, bottom panel).

Plotting LaminB1-DamID scores for all genomic regions against *de novo* called LADs, which we calculated using HMM for each of the screening candidate pools, confirmed that the control LAD structure was mostly preserved in the majority of our screening hits (Pools A-H, K, and L; Fig. 3E and Fig. S3E). Amongst the strongest phenotypes at the 2-cell stage, the constitutive heterochromatin H3K9me3 pathway (Pool J) led to a 'flattening' of the LAD structure (Fig. 3C-D) that resulted from a less distinct DamID methylation levels between LADs and iLADs (Fig. 3E). This suggests that ectopic heterochromatin induction[56] leads to flattening of nuclear organization at the 2-cell stage. In contrast, global perturbation of H3K4me2 and H3K9me2/K27me2 with the demethylases Lsd1 and Kdm7a/c (Pool I) leads to ectopic expansion of some LADs beyond their normal genome boundaries (Fig. 3C), resulting in an apparent collapse of control LADs (Fig. 3D) and a reduction in the relative DamID values between control LADs and iLADs (Fig. 3E). This suggests that H3K4me2 and/or H3K9me2/K27me2 may be involved in the correct positioning of LAD boundaries at the 2-cell stage. Remarkably, expression of the H3K9me2/me3 'writers' Ehmt1/2 and Setdb1/2 (Pool N) led to the collapse of control 2-cell stage LADs through yet another process: due to a strong decrease of DamID scores in wildtype LADs (Fig. 3E). However, 2-cell embryos from Pool N still displayed a clear LAD-iLAD structure, with DamID values

10

partitioned into *de novo* called LAD and iLAD coordinates (Pool N) (Fig. 3E). This indicates that LAD boundaries are repositioned upon expression of these H3K9me2/me3 'writers'. The collapse of control LADs upon expression of histone H1 subtypes and macroH2A (Pool M) stemmed from a similar remodeling of LAD boundaries (Fig. 3C-D and Fig. S3E). Interestingly, we detected an enrichment of the tagged histones that we expressed in the nuclear periphery, suggesting that the chromatin regions in which H1 subtypes and macroH2A are incorporated, become repositioned to the periphery (Fig. S3F). In addition, global levels of H3K9me2 are increased in Pool N as expected, but not in Pool M, suggesting different mechanisms of LAD disruption between these two collapse phenotypes (Fig. S3G). This suggests that there is no common histone modifier pathway resulting in a specific type of LAD phenotype. Thus, perturbation of specific chromatin pathways leads to a major reorganization of LADs and their boundaries at the 2-cell stage.

To address whether the pathways that regulate LAD establishment in zygotes also affect LAD reorganization at the 2-cell stage more directly, we compared the phenotypic perturbations in zygotes and in 2-cell stage embryos. First, we combined all LaminB1-DamID scores and their corresponding controls onto a single PCA. Overall, embryos were grouped according to their developmental stage, with 2-cell stage embryos together and separating from zygotes along the PC2, regardless of the perturbation condition (Fig. S3H). An exception to this were the 2-cell stage embryos expressing the H3K4me2 and H3K9me2/K27me2 demethylases Lsd1 and Kdm7a/c (Pool I), which were together with the zygote group (Fig. S3H). We interpret these observations as the expression of these demethylases prevents the rearrangement of LADs that occurs naturally during the maternal-to-zygotic transition, which is also supported by the overall reduced LAD numbers but increased LAD size in these 2-cell embryos as in control zygotes (Fig. S3C and Fig. S1D). Second, to assess whether the pooled candidate hits affect genomic regions in zygotes and 2-cell stage embryos in a similar manner, we compared *de novo* called LADs and iLADs upon perturbation to LADs and iLADs in control zygotes and 2-cell stage embryos (Fig. S3I). Intriguingly, we find that while LAD formation was disrupted in zygotes by several pathways, the affected genomic regions regained their overall LAD/iLAD wildtype structure at the 2-cell stage (Pools F, G, H, K, and L in Fig. S3I). Such a 'recovery' phenotype was striking, for example perturbing histone acetylation led to a collapse of control, zygotic LADs, but only to a minor phenotype at the 2-cell stage (Pool L; Fig. S3I). This was not due to lack of perturbation of the targeted histone modification(s) at the 2-cell stage, as we verified that the respective targeted modifications are affected in embryos in which the LAD architecture is unaffected at the 2-cell

stage (Pools F, G, H, L) (Fig. S3A). Because the expression of the pooled candidates is performed continuously, from the early zygote to the 2-cell stage, we conclude that nuclear organization in zygotes is highly adaptive and that even strong perturbations of LADs in the zygote can be reset to conditions similar to wildtype in the next cell cycle. We find these observations remarkable, as they collectively indicate that, while interactions established after fertilization in the zygote may contribute to the definition of the nuclear organization at the 2-cell stage, embryos at the 2-cell stage can reestablish LADs *de novo* under conditions in which nuclear organization is not correctly initiated in zygotes. This may suggest that the genome content itself or developmental processes such as global remodeling of the chromatin landscape or transcription during zygotic genome activation (ZGA) are sufficient to determine the sites of interaction with the nuclear lamina at the 2-cell stage following first mitosis.

Disrupting specific molecular pathways leads to remodeling of LAD boundaries at the 2-cell stage

Next, we determined the extent of genomic changes at the 2-cell stage and whether those changes are shared amongst similar molecular pathways. Genome-wide correlation analyses of LaminB1-DamID values in all 2-cell stage samples resulted in two primary clusters, which separated 2-cell stage embryos with perturbation of H3K9me2/me3 'writers' Ehmt1/2 and Setdb1/2 as well as the subtypes of histone H1 and macroH2A (Pools N and M, respectively), from the remainder of the samples (Fig. 4A). We next extracted the genomic regions that significantly gained or lost interactions with the nuclear lamina across all our 2-cell samples. Globally, all candidate pools tested led some regions to reposition towards the nuclear lamina (gained Dam-LaminB1 methylation levels) or towards the interior (lost Dam-LaminB1 methylation levels (Fig. 4B and Table S4). However, the number of genomic bins with altered Dam-LaminB1 methylation varied, with most pools leading to only minor changes overall (Pools A-H). As expected, the pool containing the H3K9 methyltransferases Ehmt1/2 and Setdb1/2 showed the strongest phenotype (7555 genomic bins affected in Pool N; Fig. 4B). Also, a large number of genomic bins that are iLADs in control 2-cell stage embryos gained lamina interactions with candidate Pools M and N (2212 and 3507, respectively) (Fig. 4B), suggesting a partial inversion phenotype.

Analysis of the overlap in the number of genomic regions affected revealed 2 major clusters (Fig. 4C). The pools with the strongest phenotype of control LAD collapse (Cluster I: Pools M and N) separated from the main cluster (Cluster II; Fig. 4C). This suggests that globally, similar genomic

12

regions and phenotypic outcomes follow upon perturbation of the H3K9me2/me3 'writers' pathway and of histone subtypes/variants related to chromatin compaction. Within Cluster II, a secondary cluster (Cluster IIa) contained a group of pools with milder phenotypes at the 2-cell stage (Pools A, C, D, E, F, G, H; Fig. 4C). The other secondary cluster (Cluster IIb) contained four more pools targeting different molecular pathways, which comprised embryos with LAD disruption phenotypes of either collapse or unaffected-to-weakened. Cluster IIb included the constitutive H3K9me3 pathway (Pool J), the demethylases Lsd1 and Kdm7a/c (Pool I), the nucleolar interactors and euchromatic 'readers' (Pool K), and the histone deacetylases (Pool L; Fig. 4C). Thus, this analysis reveals some relations across the phenotypes of LAD disruption at the 2-cell stage and the molecular pathways affected, for example, the repressive constitutive H3K9me3 pathway and histone deacetylase pathway (Fig. 4C and Fig. S4A-B).

We observed a clear pattern of chromatin features that characterized the genomic regions affected upon manipulation of the candidate pathways at the 2-cell stage, in particular, for those regions that gained interactions with the nuclear lamina. The strongest association was their location into A or B compartments in control embryos (Fig. 4D and S4C). For example, the molecular pathways that led to an inversion or a collapse of control LADs in zygotes led to increased association to the nuclear lamina of wildtype A compartment regions at the 2-cell stage, regardless of whether the control LAD structure recovered (Pool K and L) or not (Pool M and N) (Fig. 4D). In addition, for the subtypes of histone H1 and macroH2A (Pool M), regions repositioned to the lamina are highly enriched in H3K27ac and RNA Pol2 but depleted in H3K27me3 (Fig. 4D), suggesting that this manipulation of embryonic histone content affects potential regulatory regions, which became embedded into LADs. In contrast, all remainder pathways that we perturbed led to increased interactions with the nuclear lamina of wildtype B compartment regions enriched in H3K27me3 (Fig. 4D). Amongst them, the constitutive H3K9me3 heterochromatin pathway led to increased lamina association of B compartment regions marked by H3K27me3 but also with H3K9me3 in 2-cell stage embryos (Pool J; Fig. 4D), most likely reflecting the nucleation and spreading property of H3K9 methylation[57,58]. Thus, while specific chromatin pathways direct the repositioning of euchromatic A compartment regions towards the nuclear lamina, B compartment regions marked by H3K27me3 and H3K9me3 also gained interactions with the nuclear lamina upon some other perturbations. These results suggest the presence of specific histone modifications and their higher-order organization into compartments are important determinants for the association with the nuclear lamina in early embryos.

13

We also asked whether the phenotypes elicited by manipulating chromatin relate to zygotic genome activation (ZGA). For this, we compared LaminB1-DamID values from all the pools containing chromatin-modifiers with α-amanitin-treated late 2-cell stage embryos[49]. Comparing all genomic regions that significantly change interactions with the nuclear lamina revealed a high overlap of regions affected in α-amanitin-treated embryos with those affected upon expression of subtypes of histone H1 and macroH2A (Pool M) and the H3K9me2/me3 methyltransferases Ehmt1/2 and Setdb1/2 (Pool N) (Table S4; Fig. S4D). Interestingly, we find that *de novo* LAD boundaries that form upon transcriptional inhibition or upon chromatin perturbations that lead to collapse of LADs respect A/B compartment boundaries (Fig. 4E). These observations indicate that changes in association with nuclear lamina remain constrained within A/B compartments, in line with our suggestion above that compartment boundaries provide a primary scaffolding cue on genome organization at the beginning of development.

Because of the strong phenotypic demarcation of our candidate pools by their compartment score (Fig. 4D), we further investigated the relationship with compartments. Specifically, we asked whether LAD regulation by modifiers of H3K27 and H3K9 methylation occurs at the level of compartment boundaries. Co-expression of the H3K27me3 demethylases Kdm6a and Kdm6b led to increased interactions of genomic regions inside the B compartment with the nuclear lamina and a stronger demarcation of DamID score at the compartment boundaries (Pool F; Fig. 4F). Thus, LADs expand inside the B compartment, but not beyond, upon expression of Kdm6a/b. These observations suggest that compartment boundaries limit the expansion of LADs upon removal of H3K27me3. We obtained similar results upon expression of the H3K4me2 and H3K9me2/K27me2 demethylases Lsd1 and Kdm7a/c (Pool I; Fig. 4F), which had even lower average DamID scores in A compartment regions. Expression of Kdm6a and Kdm6b (Pool F) led to the fusion of LADs along regions enriched in H3K27me3 (and H3K9me3) that did not extend beyond regions demarcated by H3K4me3 domains (Fig. 4G). We obtained similar results with Lsd1 and Kdm7a/c expression (Pool I), with larger, more defined LADs overall covering the complete H3K27me3 domains, although the 'merging' phenotype was much stronger compared to Kdm6a/b (Fig. 4G). Thus, we conclude that the interplay between H3K9me3, H3K27me2/3, and H3K4me3 is a major determinant of LAD reorganization during the maternal-to-zygotic transition at the 2-cell stage.

Identification and characterization of individual effectors that regulate the establishment of LADs after fertilization and during MZT

14

221

To provide a better mechanistic understanding of the pathways of our candidate pools, we performed a phase II screening in which we split the pools displaying the most pronounced LAD phenotypes. We chose to focus on the constitutive H3K9me3 heterochromatin hits (Pool J), the H3K9me2/me3 methyltransferases Ehmt1/2 and Setdb1/2 (Pool N) and the H3K4me2 and H3K9me2/K27me2 demethylases Lsd1 and Kdm7a/c (Pool I). In total, we assessed the effects of expressing 10 individual chromatin modifiers and 'readers' on LAD establishment (Suv39h1, Ehmt1, Ehmt2, Setdb1, Setdb2, Lsd1, Kdm7a, Kdm7c, Hp1α and Hp1γ). As above, we conducted three biological replicates to map LADs in zygotes (Fig. 5A and Table S5) and verified that efficient nuclear expression of all proteins individually persists until the 2-cell stage (Fig. S5A). Genome-wide correlation analysis of LaminB1-DamID indicated that zygote samples clustered globally together, indicating a similar phenotype overall, except for zygotes in which we expressed the Kdm7a or Kdm7c demethylases and the Ehmt2 methyltransferase (Cluster I), which clustered separately from the remainder of the perturbations (Cluster II; Fig. 5B). Zygotic LADs were highly disrupted by Kdm7a, Ehmt2 and Kdm7c (Cluster I), as determined with HMM (Fig. 5C). Both the mean LAD size and the proportion of the genome assigned to the nuclear lamina were particularly affected upon expression of Ehmt2 and of Kdm7a (Fig. S5B-C). The changes caused by these two histone modifiers were stronger than those arising upon expression of the H3K4me3 demethylase Kdm5b, which we previously reported[31] (Fig. 5C and Fig. S5B-C).

Plotting LaminB1-DamID scores across the LAD boundaries of control embryos confirmed that the strongest phenotypes were caused by Kdm7a, Kdm7c, Ehmt2 along with Kdm5b, all of which resulted in flattening of wildtype LAD structure (Fig. 5D). While zygotes expressing Kdm7a do not have distinguishable DamID methylation levels between control LADs and iLADs, they do have a clear distinction of DamID methylation between *de novo* called LADs and iLADs (Fig. S5D). This indicates that Kdm7a expression leads to a severe remodeling of zygotic LAD boundaries. In comparison, zygotes expressing Ehmt2 and Kdm7c retain different DamID methylation levels between control LADs and iLADs, but the difference in DamID values is larger between *de novo* LADs and iLADs (Fig. S5D). This suggests that upon Ehmt2 and Kdm7c expression, the control LAD and iLAD structure is partially retained but some LAD boundaries are affected. Expression of all other chromatin proteins either did not affect overall wildtype LAD structure (Hp1α, Hp1γ and Setdb2) or only caused a weakening of genome-lamina interactions of control LADs (Suv39h1, Lsd1, Setdb1 and Ehmt1) (Fig. 5D and Fig. S5D). Overall, we conclude that methylation of H3K9 and H3K27 are key factors of LAD boundaries and their structural organization in zygotes. Additionally, our findings suggest that the correct amount of H3K9me in

15

zygotic chromatin is a critical determinant of correct genome-lamina scaffolding establishment since the expression of an H3K9me2 'writer' (Ehmt2) as well as 'eraser' (Kdm7a/c) perturbs LADs.

Next, we determined the effect of expressing the same 10 individual chromatin effectors in LAD restructuring during the first embryonic mitosis by mapping LADs at the late 2-cell stage (Fig. 5E and Table S5). A genome-wide correlation analysis indicated that Ehmt2 displayed the strongest phenotype of LAD disruption in 2-cell stage embryos, along with Kdm7a and Kdm7c (Fig. 5F-H). Ehmt2 expression also induced changes in LAD size, number, and genome percentage at the lamina (Fig. S5E-F). These observations are interesting, as Pool N included Ehmt1, Ehmt2, Setdb1 and Setdb2. Yet, Setdb2 alone did not seem to elicit a major LAD phenotype (Fig. 5F-H), suggesting that it is the other H3K9 methyltransferases that disrupt LAD integrity. Lsd1, Setdb1, Suv39h1, and Ehmt1 led to an overall reduction in the difference between DamID values of control LADs and iLADs (Fig. 5H) that was due to increased interactions of control iLADs with the nuclear lamina and reduced interactions of control LADs (Fig. 5H and Fig. S5G). Additionally, Ehmt2 expression prevented the natural reorganization of LADs/iLADs that occurs between the zygote and the 2-cell stage, in which a large proportion of iLADs remained instead associated with the lamina as determined by HMM (Fig. S5H). Overall, we conclude that methyltransferase activities towards H3K9me2/me3 – presumably at euchromatic regions – as well as demethylation of H3K9me2 and/or H3K27me2 affect both, the initial establishment of nuclear organization as well as the reorganization of LADs during the 2-cell stage.

Manipulation of pathways rather than individual effectors interferes with LAD establishment and reorganization

To further understand the effect of specific chromatin proteins in LAD organization and establishment, we next examined the genomic regions affected by individual candidates and their associated chromatin features. We analyzed all differentially Dam-LaminB1 methylated regions across each of the 10 individual proteins that we manipulated, both in zygotes and in 2-cell stage embryos. (Fig. S6A and Table S6). Interestingly, while expression of Hp1α or Hp1γ only displayed a few genomic bins that changed interactions with the nuclear lamina in zygotes (n=31 and 29, respectively), both proteins had a stronger phenotypic defect at the 2-cell stage overall, and in particular a higher proportion of the genome that repositioned towards or away from the nuclear lamina (n=1778 and 1909; Fig. S6A). We interpret this to suggest that heterochromatin 'readers' may have a stronger influence in positioning the genome in the 2-cell stage. To investigate this possibility, we analyzed the chromatin marks of the genomic regions affected by the 10 individual

16

hits. Excepting for Hp1γ, all chromatin modifiers led to increased nuclear lamina association of strong A compartment regions in zygote (Fig. 6A). Expression of Suv39h1 and Setdb2 similarly affected A compartment regions without a particular enrichment in active histone modifications but marked by H3K9me3 (Fig. 6A). Kdm7a and Kdm7c also affected regions of similar epigenetic marking belonging to A compartments and that are enriched in H3K27me3 in the oocyte, but not in the fertilized zygote (Fig. 6A). Analysis of the genomic regions that reduced and/or lost interactions with the nuclear lamina revealed a roughly opposite pattern (Fig. S6B). Namely, in zygotes, most hits affected B compartment regions, except for Hp1α/γ and Setdb2 (Fig. S6B).

A similar analysis at the 2-cell stage indicated that Setdb1, Ehmt1, Ehmt2, Suv39h1, and Hp1α increased interactions with the nuclear lamina of A compartment regions enriched in active chromatin marks (Fig. 6B). Since all these proteins mediate and/or read H3K9 di- or tri-methylation, this suggests that ectopic and/or spreading of heterochromatin promotes aberrant interactions with the nuclear lamina. Kdm7a and Kdm7c, also involved in heterochromatin regulation, displayed a similar pattern (Fig. 6B). In contrast, Setdb2 and Hp1γ led to repositioning of regions marked by H3K9me3 and H3K27me3 (Fig. 6B). Overall, all 10 hits except for Setdb2 led to a decrease in the DamID methylation of B compartment regions depleted in H3K27ac (Fig. S6C). Thus, the features of embryonic chromatin associated with LAD reorganization upon manipulation of these histone modification pathways differ between the zygote and the 2-cell stage.

Because our phase I pathway manipulation generally led to stronger phenotypes than the phase II, we next compared the pooled candidates to their individual effectors, with a specific focus on the H3K9 and H3K27 methylation pathways. Direct comparison of the effect of the constitutive heterochromatin pool (Pool J), which contains Suv39h1, Hp1α, and Hp1γ, on zygotic LADs indicated that the strong disruption of the nuclear organization by the complete pool was mostly recapitulated upon expression of Suv39h1 alone (Fig. 6C). In contrast, at the 2-cell stage, the Hp1α or Hp1γ could individually weaken the control LADs to a comparable level of the pooled perturbation (Fig. 6D). To better understand the regulation of LADs by these H3K9 pathways, we computed H3K9me3 levels within genomic regions affected upon expression of the constitutive heterochromatin pool and its individual components. We find that, in zygotes, the regions that increase interactions with the nuclear lamina have higher levels of H3K9me3 compared to non-affected regions or regions that move towards the nuclear interior in both, the complete Pool J or upon expression of each of its individual hits (Fig. S6D). In contrast, in 2-cell stage embryos, regions that reposition towards the nuclear lamina with the complete Pool J or with Hp1α or Hp1γ

17

individually have higher levels of H3K9me3 than those affected by Suv39h1 alone (Fig. S6E). This is particularly evident for Hp1γ, whose expression leads to larger LADs at the 2-cell stage over H3K9me3-marked regions (Fig. S6F). We interpret these observations as the 'readers' and/or the presence of H3K9me3 pre-modified chromatin at the 2-cell stage are more relevant for the regulation of LADs. This is in line with earlier observations indicating that the levels of H3K9me3 are low after fertilization and increase progressively to define a more mature, repressive chromatin as development proceeds[56,59].

Similarly, we analyzed the H3K9me2/me3 'writers' pathway (Pool N), which contains Ehmt1/2 and Setdb1/2 and led to an inversion of LADs in zygotes and collapse of control LADs in 2-cell via boundary remodeling. None of these methyltransferases alone recapitulate the extent of LAD disruption of the pool (Fig. 6E-F). This indicates that the combined action of these H3K9 di- and tri-methyltransferases are required for the pool phenotype (Pool N).

Lastly, we explored the individual roles of the H3K4me2 and H3K9me2/K27me2 demethylases Lsd1 and Kdm7a/c, which, when expressed together lead to a disruption and weakening of LADs in zygotes and formation of broad, ectopic LADs in 2-cell stage embryos (Pool I). We find that, individually, only Kdm7a or Kdm7c but not Lsd1 affect wildtype LAD structure in both zygotes and the 2-cell stage (Fig. S6G-I). This suggests that in combination with H3K4me2 depletion, removal of H3K9me2/K27me2 can regulate LAD reorganization differently. In addition, our data suggest that the interplay between H3K9 and/or H3K27 methylation may act as a major determinant of LAD (re)positioning at the beginning of development.

LAD boundaries reorganize based on positional information of H3K9me3 and H3K4me3 domains

We next aimed to gain mechanistic insights on LAD positioning in embryos. For this, we performed histone modification profiling under conditions in which LAD boundaries are displaced. We focused on the constitutive heterochromatin pool (Pool J), in which LADs expand beyond control LAD boundaries (Fig. 3C and Fig. S6F). We first asked if LAD expansion is linked to the acquisition of H3K9me3 at those genomic regions. CUT&RUN for H3K9me3 (Fig. S6J) indicated that upon combined expression of Suv39h1, Hp1α and Hp1γ, broad H3K9me3 domains are formed that reposition towards the nuclear lamina (Fig. 6G), also visible by immunostaining (Fig. 6H). Interestingly, profiling H3K9me3 in 2-cell stage embryos upon expression of Suv39h1 alone revealed that while H3K9me3 expands into broad domains within B compartments beyond control

18

225

2-cell LADs (Fig. 6G, 6I), these domains do not necessarily relocate to the nuclear lamina (Fig. 6G, 6J). Thus, H3K9me3 alone is not sufficient to drive stable interactions with the nuclear lamina. To further discern these observations, we next asked whether other histone modifications contribute to defining LAD boundaries. Specifically, because our findings above suggest that the interplay of H3K9me3, H3K27me2/3, and H3K4me3 may be a major determinant for LAD reorganization at the 2-cell stage and the potential role of euchromatin in counteracting heterochromatin spreading, we profiled H3K4me3 in embryos from Pool J using CUT&Tag (Fig. S6K). H3K4me3 levels are largely unchanged upon expression of the constitutive heterochromatin pool (Fig. S6L). However, we find that ectopic H3K9me3 domains are delimited by regions enriched with H3K4me3 (Fig. 6K), suggesting that H3K4me3 resists the spreading of H3K9me3 and the anchoring to the nuclear lamina. Remarkably, LAD boundaries are repositioned precisely at those sites demarcated by boundaries of H3K4me3 and H3K9me3, whereby H3K9me3 is enriched inside LADs and H3K4me3 just outside (Fig. 6L). We further addressed directly whether H3K4me3 can 'resist' lamina anchoring by performing CUT&Tag for H3K4me3 in embryos in which LADs become expanded, namely upon expression of the demethylases Lsd1 and Kdm7a/c (Pool I) (Fig. 6M and Fig. S6K). We find that these *de novo* broad ectopic LADs are also delimited by H3K4me3 domains (Fig. 6N), which are largely unchanged (Fig. 6M and Fig. S6M). We did not observe global changes in opposing modifications to those targeted by our pathways. For example, expression of the H3K9me3 methyltransferases and readers (Pool J) leads to increased H3K9me3 levels (Fig. 6H and Fig. S3A), H3K4me3 levels are globally unchanged (Fig. S6M) and expression of the H3K9me2 and H3K27me2 demethylases (Pool I) leads to a global reduction of H3K9me2 and H3K27me2 levels (Fig. S3A), but H3K4me3 remains largely unchanged (Fig. 6M and Fig. S6M).

In summary, our results indicate that *de novo* LAD boundaries can form at places in which H3K4me3 and H3K9me3 'oppose' each other. Overall, our data suggest that anchoring of broad, ectopic H3K9me3 domains to the nuclear lamina is promoted by the HP1 proteins and H3K4me3 resists spreading of H3K9me3 domains, preventing anchoring at the nuclear lamina. Thus, we conclude that embryos can reshuffle LADs based on positional information of H3K9me3 and H3K4me3 domains (Fig. 6O).

Developmental consequences associated with LAD disruption

Finally, we sought to address potential developmental relevance of LADs. In particular, we investigated whether inheritance versus establishment of *de novo* chromatin marks after

19

fertilization make different contributions to LAD establishment. We first focused on H3K27 methylation because of the known role of inherited maternal H3K27me3 in directing imprinting after fertilization[60]. Our phase I screening indicates that the combined expression of the H3K27me3 demethylases Kdm6a (Utx) and Kdm6b (Jmjd3) leads to a strong disruption of nuclear organization (Pool F; Fig. 1E). Similarly to the combined expression of Kdm6a/b, individual expression of Kdm7a and Kdm7c strongly affected genomic regions enriched in H3K27me3 in the oocyte, but not in the fertilized zygote (Fig. 2D and Fig. 6A). Thus, we asked whether zygotic LADs are regulated by maternally deposited H3K27me3 and/or by *de novo* methylation of H3K27 catalyzed after fertilization. To address this, we incubated embryos immediately after fertilization with the specific Ezh2 inhibitor GSK343[61] and performed LaminB1-DamID in zygotes (Fig. 7A). While incubation with GSK343 is expected to prevent *de novo* methylation by Ezh2, the catalytic subunit of polycomb repressive complex 2 (PRC2), after fertilization, expression of Kdm6a and Kdm6b is expected to also demethylate the maternally inherited H3K27me3, which is the main source of H3K27me3 in early zygotes[62]. Indeed, the majority of H3K27me3 signal in zygotes was strongly reduced, and became practically undetectable, upon Kdm6a/b expression while Ezh2 inhibition by GSK343 resulted in a lesser reduction of H3K27me3 (Fig. 7A).

Plotting DamID scores over wildtype LAD boundaries indicated that Ezh2 inhibition does not majorly affect zygotic LADs/iLADs (Fig. 7B). Genome-wide correlation analysis of DamID values confirmed that GSK343-treated embryos are similar to controls (Fig. S7A). In contrast, zygotes expressing Kdm6a/b correlate strikingly less well with both, controls as well as zygotes treated with the GSK343 inhibitor (Fig. S7B). Thus, expression of Kdm6a/b, but not GSK343 treatment, severely affect LAD architecture in zygotes (Fig. 7C, see also Fig. 1D). The combined expression of Kdm6a and Kdm6b led to overall weaker contacts of LADs with the nuclear lamina that expanded beyond the H3K27me3 regions into H3K4me3-enriched domains (Fig. 7D). We conclude that maternally deposited H3K27me3 contributes to LAD establishment after fertilization. Continued expression of Kdm6a/b until the 2-cell stage had a much less pronounced effect (Fig. S7C), with altered interactions with the nuclear lamina that remained contained within B compartment regions flanked by H3K4me3 domains (Fig. 7E), suggesting that while inherited H3K27me3 may be important to set up the LAD landscape in zygotes, it may not play a major role at the 2-cell stage. These observations could be potentially explained by the fact that maternal (oocyte) H3K27me3 levels demarcate zygotic LAD boundaries, but neither zygotic H3K27me3 nor 2-cell H3K27me3 demarcates the LAD boundaries of their corresponding stage (Fig. S7D).

20

227

Thus, we conclude that inheritance rather than active methylation of H3K27 contributes primarily to LAD formation in zygotes after fertilization.

We also analyzed the developmental competence of embryos expressing Kdm6a/b and found that they develop to the blastocyst stage at a comparable rate to controls (Fig. 7F). We find these observations remarkable, as they posit that, despite the major disruption of LADs in zygotes, embryos can 'rebuild' nuclear architecture and can continue successful preimplantation development, implying that the initial establishment of LADs in zygotes is dispensable for early development. We further performed developmental assays in representative perturbations in which embryos failed to restore LAD organization at the 2-cell stage following zygotic disruption (weakening for Pools I, J; inversion for Pools M, N) and assessed developmental progression of these embryos. These experiments indicate that developmental competence is affected in all cases in which both zygotic and 2-cell stage LADs are disrupted, with a reduced percentage of embryos reaching the blastocyst stage compared to controls (Fig. S7E). We conclude that when both zygotic and 2-cell stage LADs are affected, development is perturbed, but in conditions in which zygotic LADs are affected and their organization is rebuilt at the 2-cell stage, development can progress. In addition, our data suggest that correct LAD establishment in zygotes is dispensable for early developmental progression.

Lastly, we asked whether the developmental consequences associated with LAD disruption that we report are also reflected in transcriptional changes during zygotic genome activation (ZGA), for which the major wave occurs at the late 2-cell stage[63–65]. For this, we focused on pathways that result in different LAD phenotypes, namely the constitutive heterochromatic pathway containing Suv39h1, Hp1α and Hp1γ (Pool J) that results in weakening and collapse of control LADs in zygotes and 2-cell stage embryos, respectively, and the histone variant pathway with H1 subtypes and macroH2A (Pool M), which leads to an inversion and collapse of control LADs phenotype. We performed single-embryo RNAseq at the late 2-cell stage (Fig. S7F). RNAseq analyses revealed that both perturbations are associated with changes in gene expression, with 3,311 and 6,551 up- and 2,558 and 5,042 down-regulated genes in Pool J and Pool M, respectively (p.adj <0.05; Table S7). Notably, up-regulated genes include maternal transcripts (Fig. S7G). Differentially regulated genes also comprise major ZGA genes (Fig. 7G), indicating an impairment of a timely maternal-to-zygotic transition characterized by the accumulation of maternally deposited transcripts and inability to efficiently undergo ZGA. Interestingly, the stronger effects on transcription in Pool M (subtypes of histone H1 and variant macroH2A)

21

corresponds to the repositioning of gene-rich iLADs towards the nuclear lamina (Fig. 7H). LAD expansion into gene-poor, lowly expressed genomic regions in Pool J elicits a comparatively weaker transcriptional phenotype (Fig. 7H). Since previous work indicates that H3K9me3 alone is not repressive[56], the above findings can be separated, at least in apart, from effects of H3K9me3 alone. Although genomic regions that gain lamina interaction overlap with down-regulated major ZGA genes, in both Pool J and Pool M (Fig. 7I-J), this relationship is not always respected and thus repositioning to and away from the lamina does not necessarily imply changes in gene expression. Furthermore, our data suggests that the extent of changes in gene expression upon LAD perturbation may relate to the intrinsic properties of those genomic regions that relocate in the nuclear space, for example gene density. While it is likely that the changes in gene expression result from multifactorial effects involving chromatin modifications and nuclear positioning, which are impossible to disentangle as they depend on each other, overall, our data suggest that successful ZGA is associated with the correct nuclear organization with respect to the nuclear lamina in early embryos.

Discussion

In this study, we generated an unprecedented catalogue of LAD disruption phenotypes (Fig. S7H) and identified a collection of molecular pathways that disrupt nuclear architecture *in vivo*, in early mouse embryos. We find zygotic LADs to be labile and highly sensitive to changes in heterochromatin-associated histone modifications. For example, while expression of H3K9me3 demethylase, and thus global depletion of H3K9me3, does not affect LAD establishment[31], depositing or demethylating H3K9me2 through the expression of the corresponding Kdm3a/b or Ehmt2 enzymes, leads to a severely disrupted wildtype LAD structure. This suggests that the specific balance of H3K9 di and/or tri-methylation contributes to LADs architecture in zygotes.

We find that expression of Hp1α or Hp1γ does not affect LADs in zygotes but, consistent with the timing of heterochromatin maturation throughout the 2-cell stage, expression of Hp1α as well as Hp1γ leads to substantial remodeling of genome-lamina interactions at the 2-cell stage (Fig.7K; left panel). Overall, we identify four major pathways, primarily heterochromatic, which perturb the establishment of LADs in zygotes and their reorganization at the 2-cell stage: i) the constitutive H3K9me3 pathway (Pool J), ii) a H3K4me2 and H3K9me2/K27me2 pathway regulated by demethylases Lsd1 and Kdm7a/c (Pool I), iii) the H3K9me2/me3 'writers' Ehmt1/2 and Setdb1/2 (Pool N), and iv) the subtypes of H1 and macroH2A (Pool M) (Fig. S7H).

22

Our work also provides a potential framework for how LAD spreading is regulated. Our data indicate a model whereby H3K4me3 delimits the spreading of H3K9me3 anchoring to the nuclear lamina, thereby determining the position of LAD boundaries (Fig. 6O). However, over-expression of the H3K9me3 pathway, with both the 'writer' and the 'readers' overwrites the characteristic fragmented LADs, leading to a more canonical LAD structure in 2-cell stage embryos. Thus, our observations also explain the unusual and unique LAD fragmentation observed in wild-type 2-cell stage embryos, which we propose is due to the non-canonical distribution of H3K4me3 and the lack of the canonical H3K9me3 pathway at these stages. This also implies that H3K4me3 domains in early embryos contribute to maintaining robust nuclear organization and can counteract lamina anchoring. This builds on findings documenting a role for Kdm5b in LAD regulation in zygotes[31] and extends our understanding of the interplay between the chromatin landscape and nuclear organization. Additionally, our data point towards a possible role for H3K27me2 in LAD organization in early embryos by potentially determining the position of the LAD boundaries, as the expression of the H3K27me2 demethylase Kdm7a led to disrupted LAD boundaries. On the other hand, H3K27me3 removal through the action of the demethylases Kdm6a/b resulted in an expansion of genome-lamina contacts inside the B compartment in 2-cell stage embryos. A similar observation has been made using EZH2 inhibitor in human leukemia cells, in which the authors suggested that H3K27me3 may repel association to the nuclear lamina within B compartments[66]. Accordingly, recent findings in embryos from crosses in which EED was maternally knocked out indicate an antagonizing role for H3K27me3 in genome-nuclear lamina interactions, specifically in regards to cell-to-cell variability of LADs at the 2-cell stage[50]. Thus, methylation of H3K27 plays a role in the robustness and the definition of LAD boundaries in early embryos (Fig. 7K, middle). Interestingly, our work using a chemical inhibitor for EZH2 allowed us to further separate the contribution of inherited versus *de novo* H3K27me3 and suggest that demethylation of inherited H3K27me3 contributes to a most drastic LAD phenotype. Indeed, we find that H3K27me3 is enriched in oocytes just outside future zygotic LAD boundaries, and active demethylation of H3K27 rather than EZH2i inhibition leads to disruption of zygotic LADs. This is particularly interesting considering that oocytes do not have detectable LADs, and thus these results suggest that maternal chromatin would carry a 'programming' mark to reset nuclear organization in embryos. Maternally inherited non-canonical H3K4me3 domains could serve a similar purpose, by imparting a 'stop' signal for spreading of H3K9me3 during heterochromatin establishment[67].

23

Interestingly, targeting DNA anchors, the components of nuclear pore complex, and nuclear/ cortical cytoskeleton does not majorly perturb LAD establishment or rearrangement in early embryos. In agreement, very few structural proteins were identified as LAD regulators in a recent genome-wide screening in human cells[68]. In the case of BAF, these findings are in line with work in human epithelial cells, in which BAF knockdown does not affect genome-nuclear lamina interactions[69]. However, they also differ from work in human fibroblasts in which TPR has been proposed to repel heterochromatin[70]. Decompaction of chromatin has been shown to be sufficient to relocate specific loci towards the nuclear interior[38]. Theoretical simulations suggest that any interaction that densifies chromatin might guide preferential peripheral localization of condensed chromatin[71]. The association of the nuclear lamina of initially less dense, A compartment regions upon expression of histone deacetylases and the subtypes of histone H1/macroH2A could reflect a favoured repositioning due to increased compaction by histone deacetylation or the presence of H1 subtypes and macroH2A[72–76] (Fig. 7K; right panel). This is in accord with the strong enrichment of the ectopically expressed histones in the nuclear periphery, suggesting that the chromatin regions in which H1 subtypes and macroH2A are incorporated, become repositioned to the lamina. Overall, this suggests that the heterochromatic state can play a role in nuclear organization in early embryos by promoting or impeding the relocation of specific regions towards the periphery.

Our results also indicate that the correct establishment of LADs in zygotes is dispensable for early embryonic development. Despite severe LAD disruption in zygotes upon perturbation of several candidates, we observed that genome-lamina interactions were successfully established *de novo* after the first mitosis. This may suggest that nuclear organization in 2-cell embryos is more resilient, compared to the more labile nuclear-lamina contacts in zygotes. A potential explanation for this could be linked to the fact that the major wave of ZGA, which is essential for preimplantation development, takes place at the 2-cell stage and that transcriptional activity is highest just outside of LAD boundaries at this stage[49]. In addition, preventing ZGA leads to a drastic remodeling of late 2-cell stage LAD boundaries[49]. Therefore, we speculate that ZGA itself might provide robustness in organizing the nuclear architecture at the 2-cell stage embryos.

Early mammalian embryos undergo a gradual consolidation of compartment strength, lack mature TAD architecture[77–80], and are subject to global changes in chromatin remodeling[81,82]. These processes coincide temporally with transcriptional activation of the embryonic genome[63–65]. Generally LADs and B-compartments correlate well, excepting at the 2-cell stage, in which this relationship is weaker[31]. Interestingly, while both compartment score and replication timing are

24

largely insensitive to transcription inhibition[83], LAD boundaries are heavily disrupted[49], suggesting that LADs and compartments may not necessarily respond to the same mechanisms in early embryos. Although compartments consolidate only gradually, we often find compartment boundaries to correlate with the LAD phenotypes that we describe. Whether this correlation depends on genetic and/or epigenetic features remains to be investigated.

Lastly, while overall genome-lamina associations are robust and reports of global alterations in lamina interactions are scarce, cellular transformation can be accompanied by a distinct nuclear organization. In senescent cells, H3K9me2/3 enriched regions detach from the lamina and form senescence-associated heterochromatin domains[84,85]. A global collapse of genome-lamina interactions also occurs in oncogene-induced senescent human fibroblasts, where constitutive LADs lose lamina contacts and aberrant genome-nuclear lamina contacts emerge[86]. Our work in early embryos shows that very drastic reorganization of the genome can emerge upon perturbation of given pathways and their identification can shed light on potential chromatin-based mechanisms at play in these cell types. The findings that LAD structuring is more sensitive to manipulation of several proteins within a pathway may reflect robustness associated to chromatin anchoring at the nuclear lamina during development, whereby manipulation of a single modulator is less likely to fully perturb nuclear organization at fertilization. The contribution of several individual modulators with different specificities and binding to distinct genomic regions may also provide an additional molecular basis for a combined, stronger effect on LADs.

In summary, we show that a complex interplay of chromatin modifications influences the scaffolding of genome-lamina interactions post-fertilization and during maternal-to-zygotic transition. Specific chromatin-states are prone to altered lamina interactions under different perturbation conditions and this behaviour depends upon the developmental stage. Our work lays the ground for further investigation of embryonic chromatin and the fundamental process of genome organization during early development.

Limitations of the study

The presence of maternally produced proteins in the oocyte that are inherited by the zygote and 2-cell stage embryos is a bottleneck for studying gene function at these earliest stages. Indeed, genetic depletion is only achieved by conditional knockout approaches that must be performed in the germline, typically prior to oocyte growth. These conditional strategies often lead to defects in germline development itself. In this context, the gain-of-function strategy of our screening enabled

25

us to successfully interrogate multiple molecular pathways for a mid-scale screening coupled with low-input genomics for the first time *in vivo*. Although the use of dominant negative constructs and demethylases targeting chromatin modifications allowed us to perform functional loss-of-function perturbations, further research is required to dissect the exact mechanistic processes leading to the LAD disruption phenotypes as chromatin modifying enzymes often function on multiple target histone marks and could also have non-histone targets. Likewise, potential crosstalk between histone modifications cannot be ruled out, as it is known that modifications of specific residues depend on others, particularly those for which e.g., methylation is processive. Lastly, live imaging approaches could, in the future, complement our understanding of the dynamics of nuclear lamina-genome interaction upon perturbation, thereby providing a more in-depth interpretation of the underlying mechanisms. Importantly, our datasets could also be used for further exploration of the effects in the two parental genomes by others in the field. Our work also provides a toolbox to investigate how chromatin processes such as transcription and replication are mechanistically linked to nuclear organization in early embryos and delineate their interdependencies.
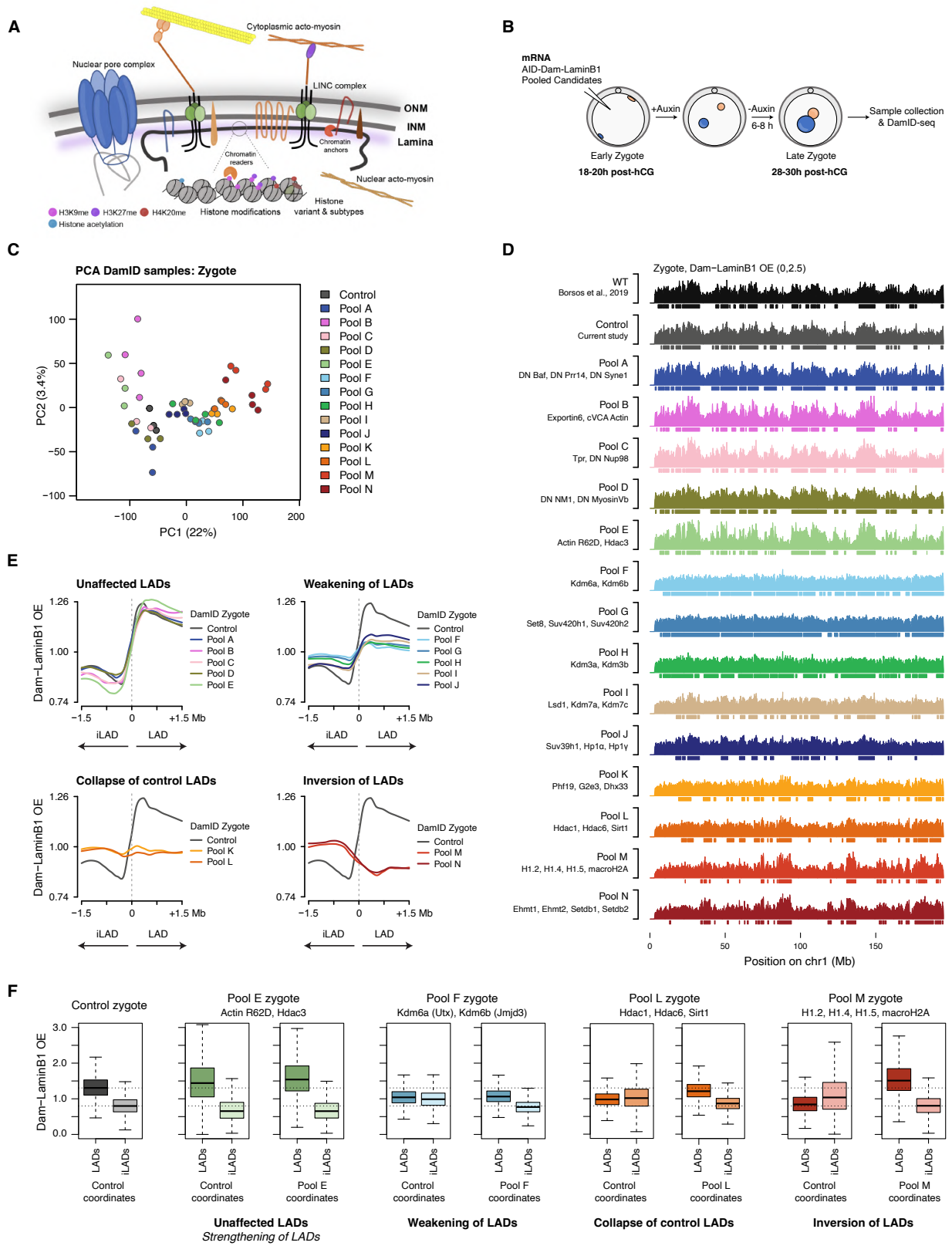
26

Figure Legends

Figure 1. Multiple phenotypes of disrupted nuclear organization emerge upon perturbation of chromatin and structural pathways after fertilization

(A) Schematic showing the association of chromatin with the nuclear envelope and indicating different molecular pathways that were targeted by screening in this study to identify mechanisms underlying the establishment of lamina-associated domains (LADs) in early embryos. ONM: Outer nuclear membrane; IMM: Inner nuclear membrane.

(B) Experimental design of embryo manipulation and sample collection for low-input LaminB1 DamID. Early zygotes collected immediately after fertilization (18-20h post-hCG injection) were microinjected with pooled candidates (screening phase I) and processed for DamID at the late zygote stage (28-30h post-hCG injection). Experiments were performed in at least three biological replicates. +Auxin: no methylation; -Auxin: GATC methylation by Dam-LaminB1 under the control of auxin-inducible degron (AID).

(C) Principal component analysis (PCA) of zygotic DamID samples from phase I screening. Each data point represents a biological replicate for the corresponding manipulation indicated by the colour code. The percentage of variance explained by PC1 and PC2 is indicated in axis labels.

(D) Observed over Expected (OE) Dam-LaminB1 mean values calculated from biological replicates visualized on chromosome 1. Boxes below the tracks represent LADs determined using a two-state hidden Markov model (HMM). Previously published zygotic DamID data from GSE112551 is reanalyzed with the same pipeline and shown as WT for comparison. OE values were calculated in consecutive 100-kb genomic bins. Candidates belonging to each pool are shown on the left-hand side of the panel.

(E) Average OE values over zygotic LAD boundaries of control embryos. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot. The 1.5 Mb region towards the right-hand side depicts LAD. iLAD: inter-LAD.

(F) Box plots of Dam-LaminB1 OE mean values in control and *de novo* called LAD and iLAD regions. Box plots show the median, and the interquartile range (IQR), and whiskers depict the smallest and largest values within 1.5 ×IQR. The horizontal dotted lines indicate the median OE values from the control zygotic LADs (upper line) and iLADs (lower line).

See also Figure S1 and Tables S1 and S2.

27

A — Correlation of zygotic DamID samples

B — Volcano plots: Pool A through Pool N vs Control zygote (log2 fold change vs −log10 adj. p-value)

C — Overlap of differentially Dam-LaminB1 methylated regions in zygote

D — Chromatin features of regions with increased DamID scores in zygote

E — LAD strengthening (Pool E, Actin R62D, Hdac3); LAD weakening (Pool F, Kdm6a, Kdm6b); LAD inversion (Pool M, H1.2, H1.4, H1.5, macroH2A)
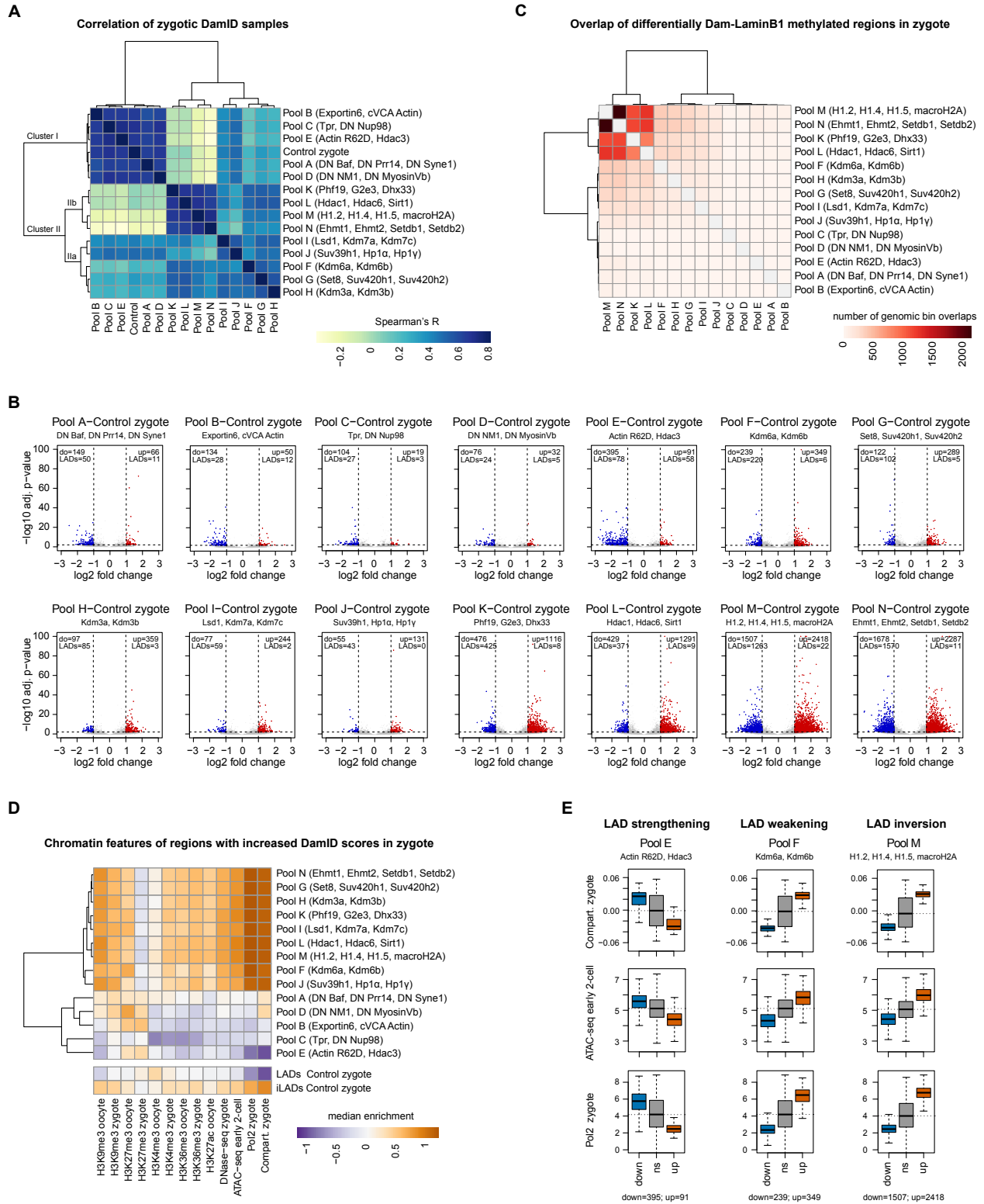
Figure 2. Integration of nuclear organization phenotypes reveals chromatin features associated with disrupted LAD establishment

(A) Hierarchical clustering and genome-wide correlation of Dam-LaminB1 OE mean values between the zygotic samples from phase I screening using Spearman's R.

(B) Volcano plots showing genomic regions that show differential LaminB1 OE values in comparison to control zygotes. 'Up' indicates the number of 100-kb genomic bins that have significantly higher OE values (log2 fold change >1 and adjusted p-value <0.01; red dots) and 'do'(down) shows the number of genomic bins that show significantly reduced OE value (log2 fold change <-1 and adj. p-value <0.01; blue dots). The number of 'up' or 'do' genomic bins that belong to LADs in control zygotes is indicated below.

(C) Heatmap showing overlap of all genomic regions that show differential lamina interactions ('up' + 'do') in zygote with respect to controls.

(D) Enrichment of wildtype chromatin features in genomic regions that show increased OE values (marked as 'up' in volcano plots) compared to controls in zygote. Chromatin feature enrichment in zygotic LADs and inter-LADs (iLADs) in control embryos is shown below for comparison. Positive compartment scores (Compart.) define A compartment.

(E) Compartment score (Compart.), chromatin accessibility (log2 transformed), and RNA polymerase II occupancy (log2 transformed) in genomic regions that gain ('up') or lose ('down') lamina interactions in zygote. Box plots show the median, and the interquartile range (IQR), and whiskers depict the smallest and largest values within 1.5 ×IQR. ns: all genomic bins with non-significant changes in OE value compared to control. The horizontal dotted lines indicate the median signal in the 'ns' genomic regions for the corresponding analysis.
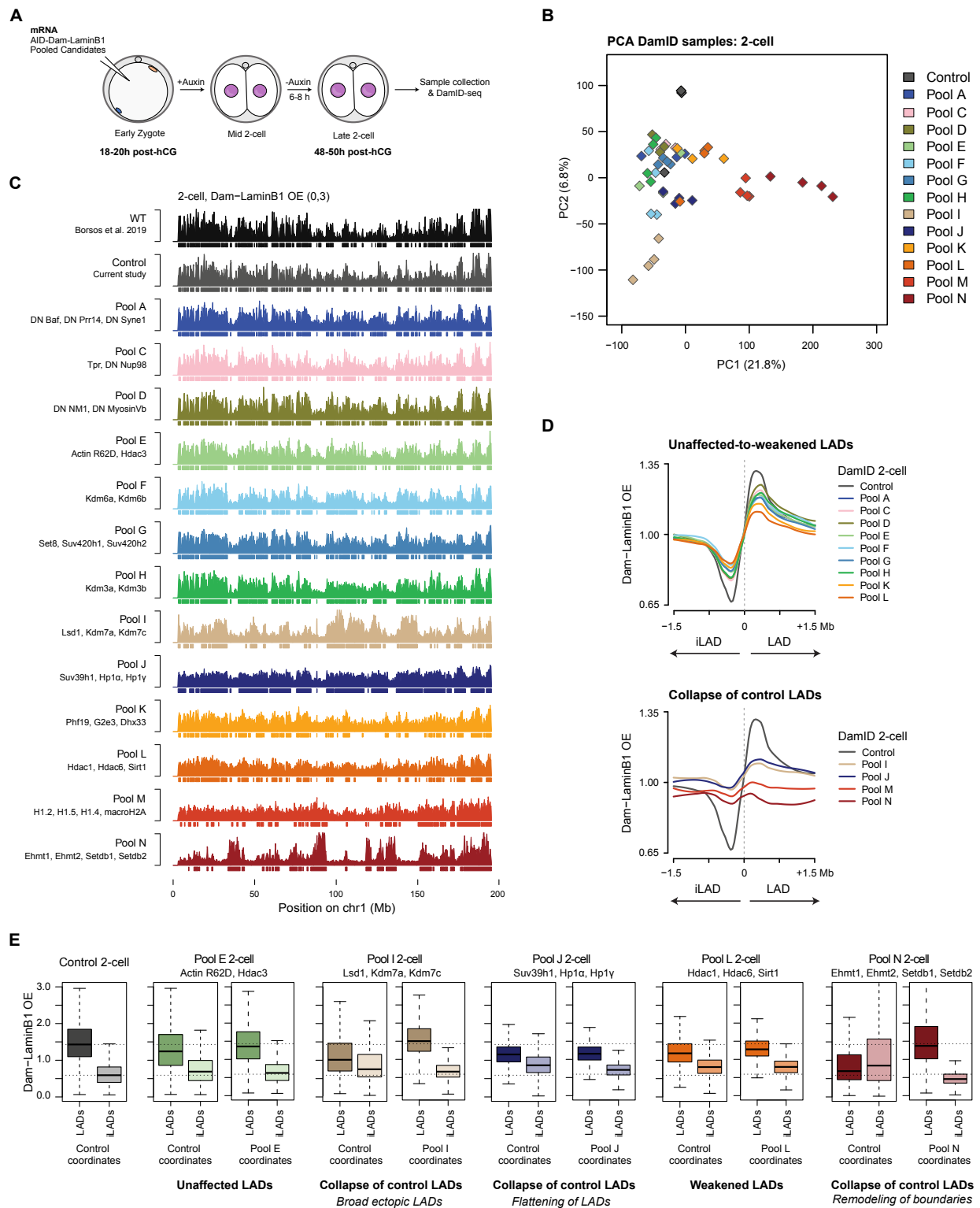
See also Figure S2 and Table S3.

A

mRNA
AID-Dam-LaminB1
Pooled Candidates

Early Zygote
18-20h post-hCG

+Auxin

Mid 2-cell

-Auxin
6-8 h

Late 2-cell
48-50h post-hCG

Sample collection
& DamID-seq

B

PCA DamID samples: 2-cell

PC2 (6.8%)

PC1 (21.8%)

Control
Pool A
Pool C
Pool D
Pool E
Pool F
Pool G
Pool H
Pool I
Pool J
Pool K
Pool L
Pool M
Pool N

C

2-cell, Dam−LaminB1 OE (0,3)

WT
Borsos et al. 2019

Control
Current study

Pool A
DN Baf, DN Prr14, DN Syne1

Pool C
Tpr, DN Nup98

Pool D
DN NM1, DN MyosinVb

Pool E
Actin R62D, Hdac3

Pool F
Kdm6a, Kdm6b

Pool G
Set8, Suv420h1, Suv420h2

Pool H
Kdm3a, Kdm3b

Pool I
Lsd1, Kdm7a, Kdm7c

Pool J
Suv39h1, Hp1α, Hp1γ

Pool K
Phf19, G2e3, Dhx33

Pool L
Hdac1, Hdac6, Sirt1

Pool M
H1.2, H1.5, H1.4, macroH2A

Pool N
Ehmt1, Ehmt2, Setdb1, Setdb2

Position on chr1 (Mb)

D

Unaffected-to-weakened LADs

Dam−LaminB1 OE

DamID 2-cell
Control
Pool A
Pool C
Pool D
Pool E
Pool F
Pool G
Pool H
Pool K
Pool L

iLAD   LAD

Collapse of control LADs

Dam−LaminB1 OE

DamID 2-cell
Control
Pool I
Pool J
Pool M
Pool N

iLAD   LAD

E

Control 2-cell

Pool E 2-cell
Actin R62D, Hdac3

Pool I 2-cell
Lsd1, Kdm7a, Kdm7c

Pool J 2-cell
Suv39h1, Hp1α, Hp1γ

Pool L 2-cell
Hdac1, Hdac6, Sirt1

Pool N 2-cell
Ehmt1, Ehmt2, Setdb1, Setdb2

Dam−LaminB1 OE

Control
coordinates

Control
coordinates   Pool E
coordinates

Control
coordinates   Pool I
coordinates

Control
coordinates   Pool J
coordinates

Control
coordinates   Pool L
coordinates

Control
coordinates   Pool N
coordinates

Unaffected LADs

Collapse of control LADs
Broad ectopic LADs

Collapse of control LADs
Flattening of LADs

Weakened LADs

Collapse of control LADs
Remodeling of boundaries

Figure 3. Identification of pathways that regulate LAD reorganization after the first mitotic division

(A) Experimental design of embryo manipulation and sample collection for low-input LaminB1 DamID in late 2-cell embryos (48-50h post-hCG injection) for phase I of the screening performed with pooled candidates. Experiments were performed in at least three biological replicates. +Auxin: no methylation; -Auxin: GATC methylation by Dam-LaminB1 under the control of auxin inducible degron. Microinjections are performed immediately after fertilization (18-20h post-hCG) and therefore candidates are expressed from zygote stage.

(B) Principal component analysis (PCA) of 2-cell DamID samples from phase I screening. Each data point represents a biological replicate for the corresponding manipulation indicated by the colour code. The percentage of variance explained by PC1 and PC2 is indicated in axis labels.

(C) 2-cell Dam-LaminB1 OE mean values visualized on chromosome 1. Boxes below the tracks represent LADs called by a two-state HMM. Previously published 2-cell DamID data from GSE112551 is reanalyzed with the same pipeline and shown as WT for comparison. Candidates belonging to each pool are shown on the left-hand side of the panel.

(D) Average Dam-LaminB1 OE value over LAD boundaries of control 2-cell embryos. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot. The 1.5 Mb region towards the right-hand side indicates LAD. iLAD: inter-LAD.

(E) Box plots of Dam-LaminB1 OE mean values in control and *de novo* called LAD and iLAD regions. Box plots show the median, and the interquartile range (IQR), and whiskers depict the smallest and largest values within 1.5 ×IQR. The horizontal dotted lines indicate the median OE values from the control 2-cell LADs (upper line) and iLADs (lower line).

See also Figure S3 and Table S2.

A **Correlation of 2-cell DamID samples**

C **Overlap of differentially Dam-LaminB1 methylated regions in 2-cell embryos**

B

D **Chromatin features of regions with increased DamID scores in 2-cell**
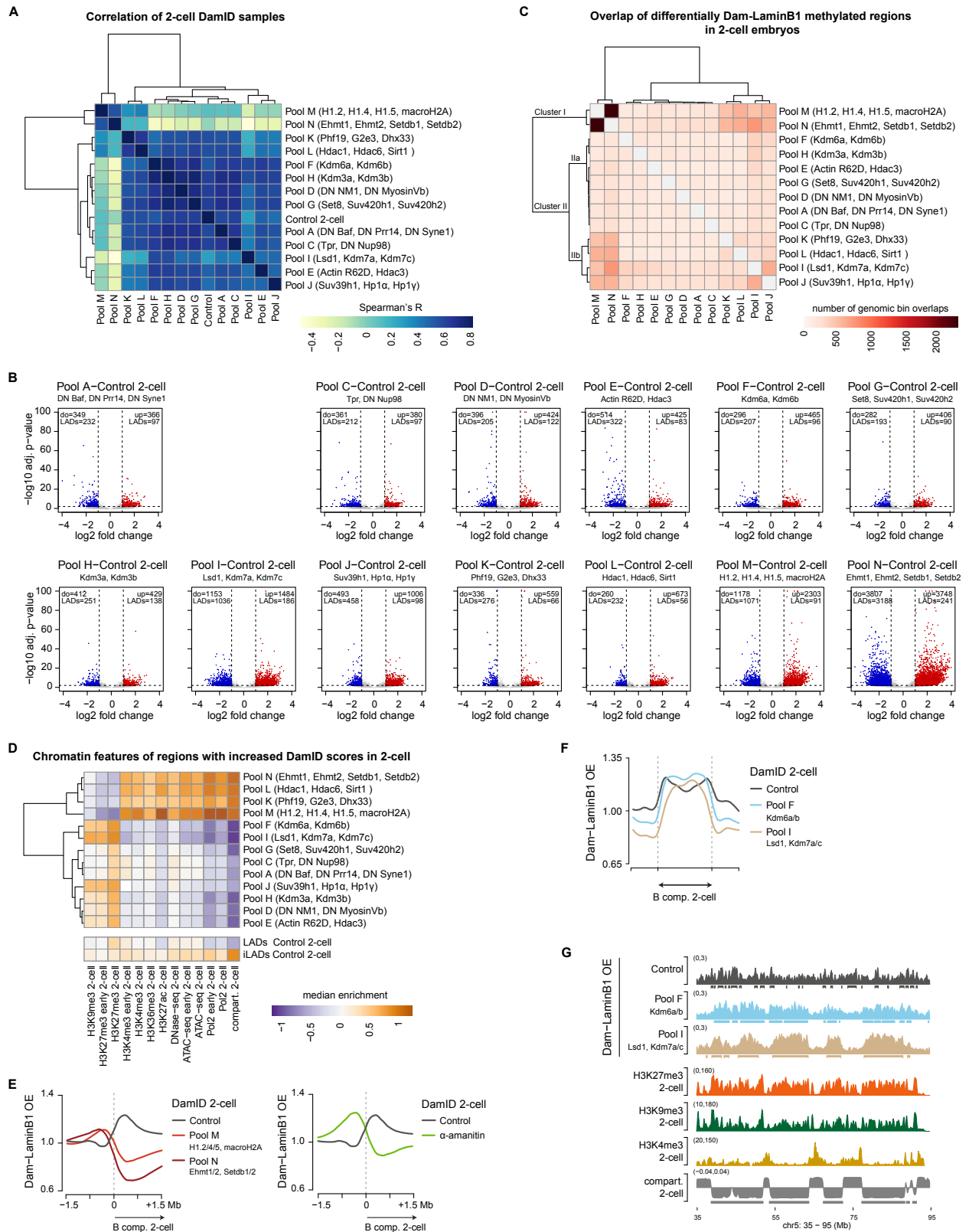

E

F

G

240

Figure 4. Disrupting specific molecular pathways leads to remodeling of LAD boundaries at the 2-cell stage with altered concordance with compartments

(A) Hierarchical clustering and genome-wide correlation of OE mean values between the 2-cell DamID samples from phase I screening using Spearman's R.

(B) Volcano plots showing genomic regions that show differential Dam-LaminB1 OE values in comparison to control 2-cell embryos. 'Up' indicates the number of 100-kb genomic bins that have significantly higher OE values (log2 fold change >1 and adjusted p-value <0.01; red dots) and 'do'(down) shows the number of genomic bins that show significantly reduced OE value (log2 fold change <-1 and adj. p-value <0.01; blue dots). The number of 'up' or 'do' genomic bins that belong to LADs in control 2-cell stage embryos is indicated below.

(C) Heatmap showing overlap of all genomic regions that show differential lamina interactions ('up' + 'do') with respect to controls in 2-cell DamID samples.

(D) Enrichment heatmap for wildtype chromatin features in genomic regions that show increased Dam-LaminB1 OE values (marked as 'up' in volcano plots) compared to controls in 2-cell stage embryos. Chromatin feature enrichment in control 2-cell LADs and inter-LADs (iLADs) is shown below for comparison. Positive compartment scores (Compart.) define A compartment.

(E) Metaplots of average Dam-LaminB1 OE value over A/B compartment boundaries of control 2-cell embryos. The 1.5 Mb region at the right from the dotted line indicate B compartments in wildtype embryos. DamID data from α-amanitin-treated 2-cell embryos was analyzed from GSE241483.

(F) Metaplot showing average Dam-LaminB1 OE value over scaled B compartment regions in 2-cell embryos.

(G) Dam-LaminB1 OE value, histone modification enrichment and compartment score from 2-cell stage embryos calculated from public datasets and visualized on part of chromosome 5. Boxes below the OE value tracks represent called LADs and wildtype 2-cell B compartment regions are indicated below the compartment score track.
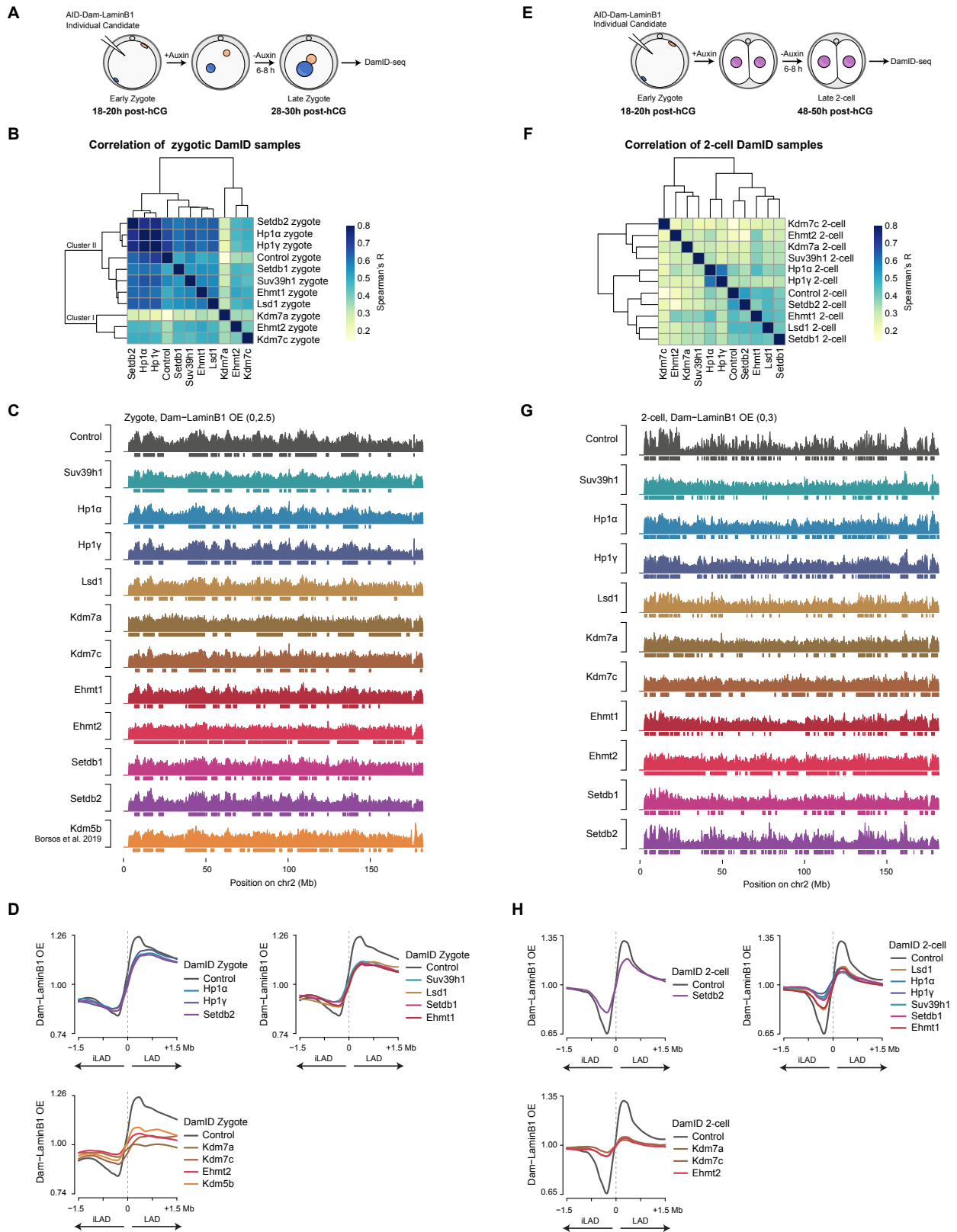
See also Figure S4 and Table S4.

Figure 5. Identification and characterization of individual effectors that regulate the establishment of LADs after fertilization and their dynamics during MZT

(A) Experimental design of embryo manipulation and sample collection for DamID in zygotes for phase II of the screening performed for individual candidates. Experiments were performed in at least three biological replicates. +Auxin: no methylation; -Auxin: GATC methylation by Dam-LaminB1 construct under the control of auxin-inducible degron (AID).

(B) Hierarchical clustering and genome-wide correlation of Dam-LaminB1 OE mean values between the zygotic samples using Spearman's R.

(C) Zygotic Dam-LaminB1 OE mean values visualized on chromosome 2. Boxes below the tracks represent LADs called by a two-state HMM. DamID data from Kdm5b overexpressed zygotes is analyzed from GSE112551.

(D) Metaplot of Dam-LaminB1 OE value over control zygotic LAD boundaries. Zero and the dotted line represent the position of the LAD/iLAD boundary. The 1.5 Mb region towards the right-hand side indicates LAD. iLAD: inter-LAD.

(E) Experimental design of 2-cell LaminB1 DamID for phase II of the screening performed for individual candidates. Experiments were performed in at least three biological replicates. Microinjections are performed immediately after fertilization (18-20h post-hCG) and therefore candidates are expressed from zygote stage.

(F) Correlation of genome-wide OE mean values between the 2-cell DamID values using Spearman's R.

(G) Dam-LaminB1 OE mean values at the 2-cell stage visualized on chromosome 2. Boxes below the tracks represent LADs.

(H) Average OE value plotted over LAD boundaries of control 2-cell embryos. Zero and the dotted line represent the position of the LAD/iLAD boundary.
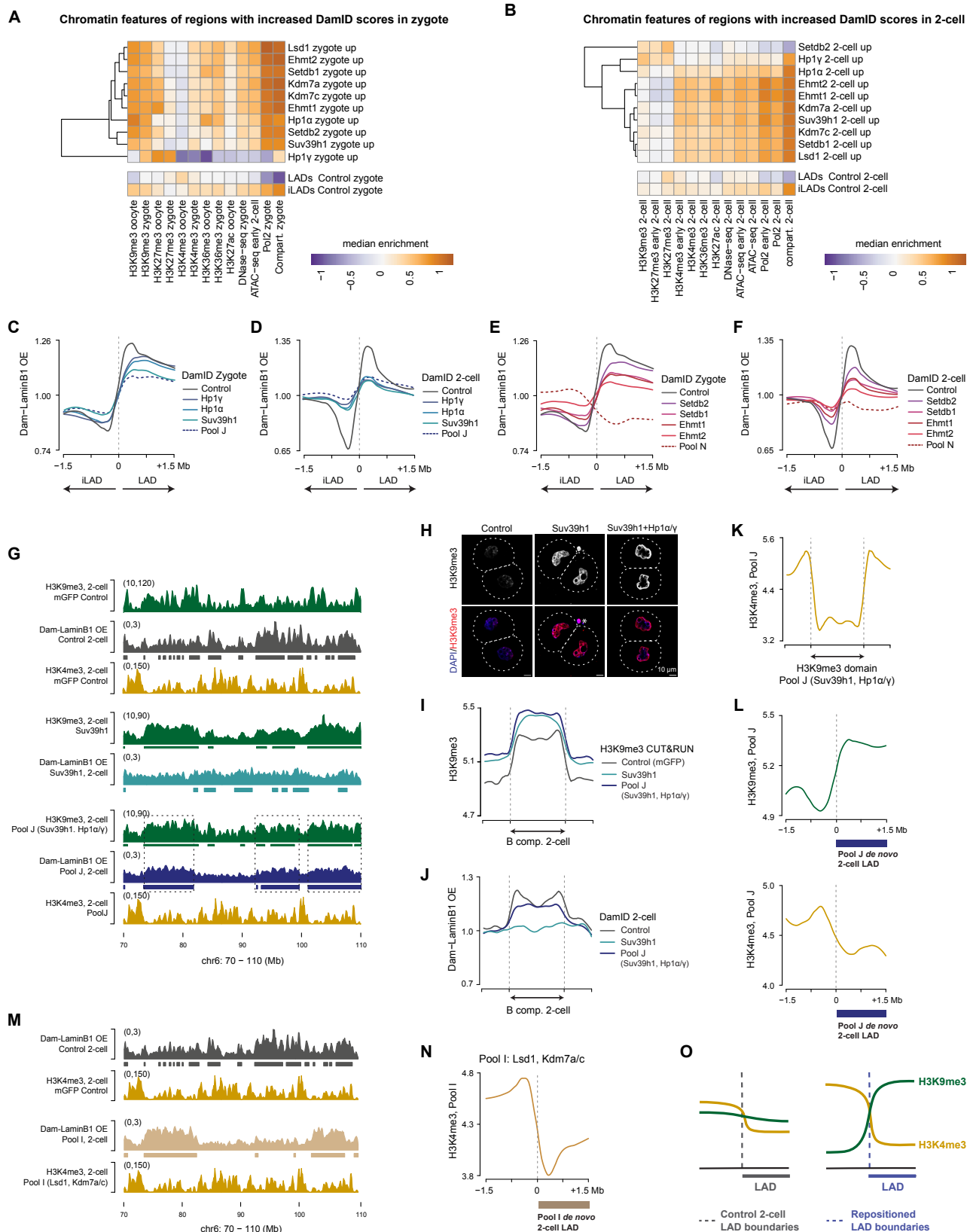
See also Figure S5 and Table S5.

A Chromatin features of regions with increased DamID scores in zygote

B Chromatin features of regions with increased DamID scores in 2-cell

Figure 6. Manipulation of pathways rather than individual effectors interferes with LAD establishment and reorganization

(A, B) Enrichment heatmap for wildtype chromatin features in genomic regions that show increased Dam-LaminB1 OE values (marked as 'up' in volcano plots) compared to controls in zygote (A) and 2-cell stage (B) embryos. Chromatin feature enrichment in control LADs and inter-LADs (iLADs) is shown below for comparison. Positive compartment scores (Compart.) define A compartment.

(C, D, E and F) Average Dam-LaminB1 OE value over control LAD boundaries across different DamID samples for zygote (C and E) and 2-cell (D and F) stage embryos comparing individual effectors (solid lines; phase II) to respective candidate pools (dotted lines; phase I).

(G, M) Dam-LaminB1 OE values and histone modification enrichment in 2-cell stage embryos visualized on part of chromosome 6. Boxes under the OE value tracks indicate LADs. Mean H3K9me3 and H3K4me3 enrichment was computed from CUT&RUN or CUT&Tag replicates, respectively. H3K9me3 domains established by HMM are shown as thick lines under the corresponding tracks in green. The dotted rectangles in G point to some examples of broad H3K9me3 domains that correspond to LADs in Pool J (Suv39h1, Hp1α and Hp1γ).

(H) Representative single confocal planes from H3K9me3 immunostaining in late 2-cell stage embryos (48-50h post-hCG injection). DAPI stains for DNA. Asterisks indicate the polar bodies. Dashed lines roughly demarcate the contour of the embryos. Scale bars, 10 µm. N = 3.

(I, J) Metaplots showing average H3K9me3 enrichment (CUT&RUN, log2 transformed; panel I) or Dam-LaminB1 OE values (panel J) over scaled wildtype B compartment regions in 2-cell embryos.

(K) Metaplot showing average H3K4me3 enrichment (CUT&Tag, log2 transformed) over scaled broad H3K9me3 domains as called using two-state HMM on mean H3K9me3 CUT&RUN enrichment data from 2-cell embryos from Pool J.

(L) Average enrichment of the indicated histone modifications (log2 transformed) over *de novo* called 2-cell LAD boundaries in Pool J.

(N) Average enrichment of H3K4me3 (CUT&Tag, log2 transformed) over *de novo* called LAD boundaries in 2-cell stage embryos from Pool I.

(O) Cartoon model depicting LAD boundary remodeling upon perturbation based on positional enrichment of H3K9me3 and H3K4me3 domains.
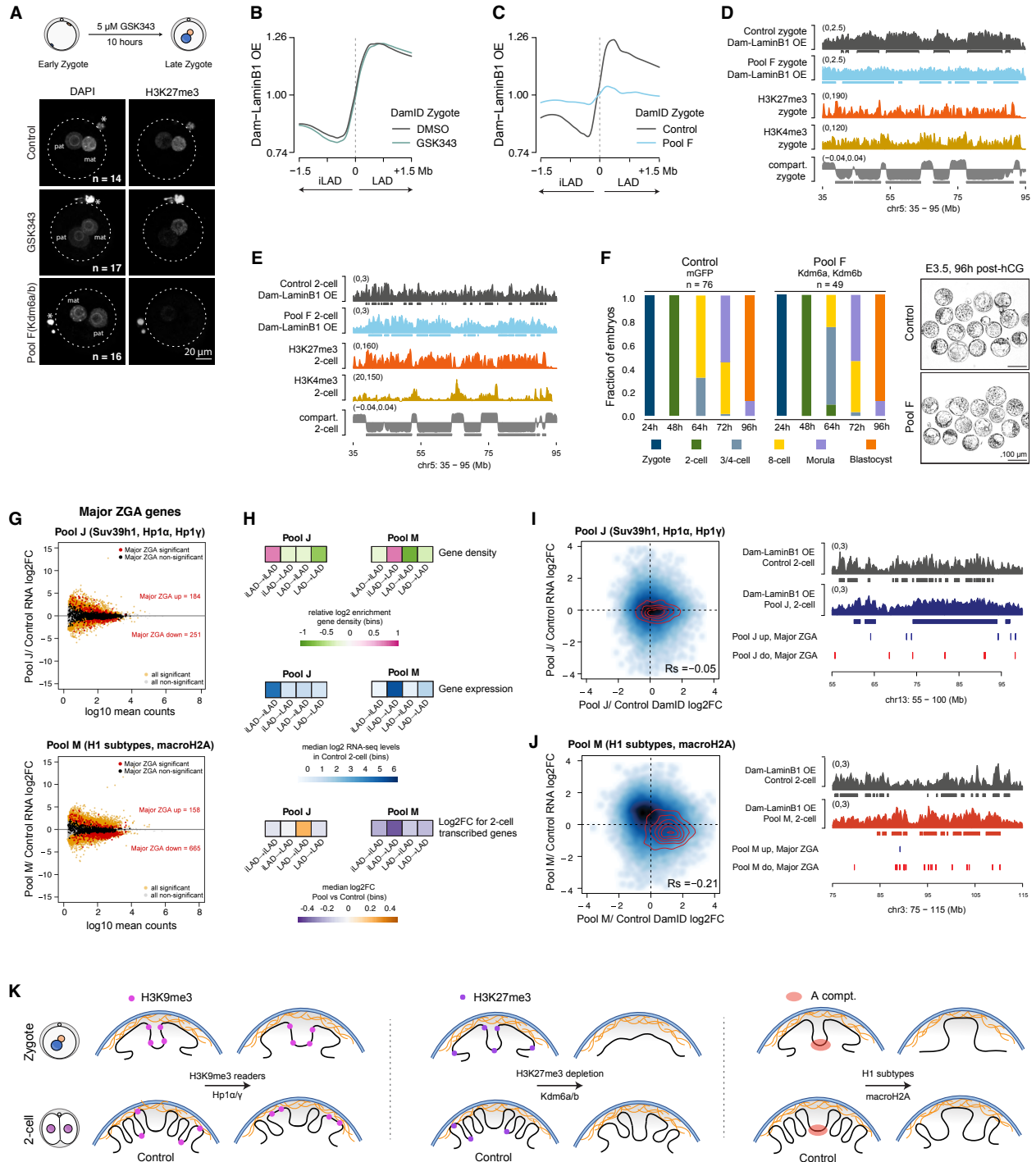
See also Figure S6 and Table S6.

32

Figure 7. Developmental consequences associated with LAD disruption

(A) Representative maximum intensity projections of H3K27me3 immunostaining. Zygotes were isolated immediately after fertilization (18-20h post-hCG injection), treated for 10 hours with GSK343 and collected prior to the first mitosis, at 28-30h post-hCG. A cartoon representation of the treatment is shown above. Note that the fluorescence signal of H3K27me3 is practically undetectable upon expression of Kdm6a/6b (Pool F). DAPI stains for DNA. Asterisks indicate the polar bodies. Dashed lines roughly demarcate the contour of the embryos. Scale bar, 20 μm. N = 3. mat: maternal pronucleus; pat: paternal pronucleus.

(B, C) Average Dam-LaminB1 OE value over control zygotic LAD boundaries. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot.

(D, E) Dam-LaminB1 OE value, histone modification enrichment and compartment score calculated from public datasets and visualized on a region of chromosome 5 for zygote (D) or 2-ell stage embryos (E). Boxes below the OE value tracks represent LADs. Wildtype B compartments are indicated below the compartment score tracks.

(F) Developmental progression of embryos microinjected with mRNA for mGFP only (control) and embryos coexpressing H3K27me3 demethylases Kdm6a and Kdm6b (Pool F). On the x-axis, h indicates hours post-hCG injection. n = total number of embryos analyzed from at least three independent experiments. Representative images from the developmental assay are shown to the right. Images were captured at embryonic day 3.5 (E3.5) which corresponds to 96h post-hCG. Scale bars, 100 μm.

(G) MA plots of log2-fold change in transcript abundance (RNA-seq counts) in 2-cell stage embryos against mean RNA-seq counts (log10 transformed). Differentially expressed genes are labeled in orange (adj. p-value <0.05), non-differential ones in gray. Differentially expressed major ZGA genes (as per DBTMEE classification) are marked in red, non-differential major ZGA genes are shown in black.

(H) Enrichment heatmap for gene density (top), transcript abundance (RNA-seq counts) in control late 2-cell embryos (middle) and log2-fold change in gene expression (bottom) for genes transcribed at the 2-cell stage (sum of the DBTMEE categories: major ZGA, 2-cell transient and MGA but excluding maternal RNA and minor ZGA) for groups of genomic regions that reorganize between LAD and iLADs in Pool J or Pool M perturbations compared to control 2-cell embryos. For example, the 'iLAD→LAD' category includes genomic regions that are iLADs in control embryos but become LADs as per *de novo* calling for the indicated Pool.

(I, J) Smoothed scatter plots of genome-wide changes in transcript abundance (RNA-seq counts; log2FC) versus changes in LaminB1 DamID score (log2FC) between Pool J (panel I) or Pool M

(panel J) and control late 2-cell stage embryos. Spearman's correlation (Rs) is indicated. The red lines on top of the scatter plot demarcate the contour of genomic bin density (n = 902) that contain major ZGA genes but not maternal transcripts (as per DBTMEE database). To the right, position of significant (adj. p-value <0.05) up- (log2FC >0; in blue) or down- (log2FC <0; in red) regulated major ZGA genes are shown below the tracks of Dam-LaminB1 OE values on representative parts of chromosome 13 (panel I) or chromosome 3 (panel J).

(K) Graphical summary of LAD reorganization resulting from perturbing H3K27, H3K9 methylation pathways as well as histone content in zygote (top row) and 2-cell embryos (bottom row). Enrichment of the indicated histone modifications is shown by dots. Blue lines depict the nuclear envelope, orange mesh the nuclear lamina. A compartments in control embryos are illustrated with a red cloud.

See also Figure S7 and Table S7.

34

Methods

Embryo collection, culture, and manipulation

All experiments were approved by the Government of Upper Bavaria. Mice housed in Helmholtz Zentrum München were maintained and bred in accordance with institutional guidelines. To obtain embryos, 5-8-week-old F1 (C57BL/6J × CBA/H) female mice were mated with DBA/2J males. To induce ovulation, females were injected with 10 IU pregnant mare serum gonadotropin (PMSG) (Ceva) and then 46-48 h later with human chorionic gonadotrophin (hCG) (MSD Animal Health). Collected embryos were cultured in KSOM drops under paraffin oil (Sigma) at 37 °C with 5% CO2 as previously described. For DamID, early zygotes (18 h post-hCG) were isolated and injected with 250 ng/µL Tir1, 50 ng/µL membrane-eGFP and 10 ng/µL AID-Dam-LaminB1 along with mRNA encoding candidate protein(s) and cultured in auxin (500 µM)-containing medium. For mapping LADs in the zygotic stage, auxin was removed from 22 h and late zygotes were collected at 28–30 h post-hCG. For DamID in the late 2-cell stage, auxin was washed out from 42 to 48-50 h post-hCG and embryos were cultured in KSOM. All candidate cDNAs were subcloned into the pRN3P vector containing identical 5' and 3'UTR and a consensus KOZAK, to ensure efficient and equivalent expression (excepting for Kdm6b and Tpr, which were already obtained in pcDNA, suitable for in vitro transcription). The mRNA concentration of candidate proteins was decided based on the size of the ORF and was chosen based on earlier titration experiments[31,56,83,87,88]. The concentration of mRNA used was calculated to ensure a molarity equivalence range (0.8-1.5 µM) across all candidates, both for the individually microinjected candidates or within the pools, to achieve a similar equimolar expression (Table S1). For the EZH2 inhibition experiment in zygote, embryos were treated with 0.01% DMSO (as control) or 5 µM GSK343 (Selleckchem, #S7164) from 18 h to 28 h post-hCG. To monitor developmental effects, microinjected embryos were scored daily after microinjection, up until day 4. As control, we used embryos injected with mRNA for membrane-eGFP only. To validate for the expression of candidate Pools and individual proteins, we performed immunostaining against HA, the targeted histone modification, or monitored fluorescent of fusion proteins (e.g., mCherry-DN Syne1). For the Pools containing structural nuclear membrane components or exportin and actin dominant negative constructs (Pool B and Pool E), validation of expression was inferred from the expected published phenotype, namely lack of cell division[53] and reduced pronuclear size, respectively (Fig. S1H).

DamID sample processing and library preparation

Zona pellucida was removed by treatment with 0.5% pronase in M2 at 37 °C for 5 minutes with visual inspection. Polar bodies were separated from the embryos by gentle pipetting after a short

35

trypsin treatment (up to 1 minute with visual inspection) and discarded. For each replicate, a pool of 10-20 blastomeres (10 to 20 zygotes or 5 to 10 2-cell embryos) were collected in 2 µL DamID buffer (10 mM TRIS acetate pH 7.5, 10 mM magnesium acetate, 50 mM potassium acetate) and stored at -80 °C until processing. All experiments were performed in at least three independent biological replicates. Sample processing and library preparation were done as described[31,52].

DamID sequencing and analysis

Samples were sequenced using Illumina HiSeq4000 or HiSeq2500 platforms in 150 bp PE mode but only read1 was used for analysis. For read pre-processing, the first 6 random bases were discarded using trimmomatic (version 0.39). Subsequently, reads were demultiplexed according to DamID indexes using fastx barcode splitter and the additional 15 bp of adaptors were trimmed using trimmomatic. Pre-processed reads starting with GATC were then mapped to the GRCm38 using bowtie2 (version 2.5.1) with default parameters. Reads aligning to the genome with a quality score below 30 were discarded using samtools (version 1.17). Duplicates were removed using picard (version 3.0.0) to obtain unique GATC reads. Reads were counted in 100-kb consecutive genomic bins using bedtools (version 2.31.0). The computation of OE (Observed/Expected) values per bin was carried out similarly as described[51]. Briefly, to obtain the expected number of reads, all genomic GATC sites were extended to the trimmed read length (123 bp) in both directions using R (version 4.1.2) Biostrings (version 2.62.0) and GenomicRanges (version 1.46.1). Extended GATC reads were processed the same way as reads obtained by DamID (observed). Read counts were normalized by rpkm (reads per kilobase per million) and a pseudo-count was added (smallest non-zero rpkm value). Finally, the observed over expected rpkm ratio was calculated. Bins with zero rpkm for both observed and expected values were treated as zero. For obtaining the OE mean signal, rpkm values of at least 3 replicates were averaged prior OE value calculation. The OE mean signal was used for data visualization and LAD calling. To distinguish LAD domains from inter-LADs, a two-state hidden Markov model (HMM)[89] was applied to non-zero OE mean values. For differential testing between treatment and control conditions, a generalized linear model of the gamma family with log link was fitted on the replicate OE values for each genomic bin using R. P-values were calculated based on the z-distribution and were adjusted for multiple comparisons by the Benjamini & Hochberg method. Significant bins were called by an adjusted p-value cutoff of 0.01 and a log2 fold change cutoff of 1.

Allelic analysis of DamID dataset

Allelic analysis was performed using SNPsplit (version 0.6.0). SNP annotation for GRCm38 genome was downloaded from https://ftp.ebi.ac.uk/pub/databases/mousegenomes/REL-1505-

. Genome files were prepared using SNPsplit_genome_preparation with the parameters --strain CBA_J --strain2 DBA_2J. SNPsplit was applied on the DamID alignment files using 1,708,377 DBA/2J-specific SNPs. To obtain allelic OE values, the splitting was also carried out on the genomic GATC reads, which served as allelic expected read counts. For downstream analysis, only those genomic bins were considered that contained more than 30 allele specific genomic GATC reads and thus the allelic analysis is limited to 4,254 100-kb bins. Allelic OE values were visualized at LADs/iLADs using our previously published maternal and paternal LAD coordinates[31].

RNA-seq library preparation and sequencing

The SMART-seq+5' protocol was adapted from the Smart-seq2 protocol as described by Oomen et al. (in preparation)[90]. Late 2-cell stage embryos (membrane-eGFP control, Pool J, and Pool M) were collected in the same lysis buffer, stored at -80 °C until use. The lysis buffer was prepared by diluting Clontech 10× lysis buffer (635015) to 1× in H2O, supplemented with ERCC RNA spike-ins (diluted to 1:581,000), and aliquoted into PCR tubes (5.8 µL per tube). The embryos were washed three times in PBS, transferred to tubes containing the lysis buffer, snap-frozen in liquid nitrogen, and stored at -80 °C until further processing. RNA was extracted using AMPure RNA magnetic beads (Beckman Coulter), washed with 80% ethanol, and resuspended in 1 µL of dNTP mix (ThermoFisher, R0192), 1 µL of oligo-dT30V (10 µM, Sigma, 5'-AAGCAGTGGTATCAACGCAGAGTACT30V-3'), and 1 µL of nuclease-free water containing 5% RNase inhibitor (Takara, 2313A). The samples were incubated for 3 minutes at 72 °C and kept on ice until further processing. The reverse transcription solution was prepared by mixing 2 µL of Superscript II 5× RT buffer (Thermo-Fisher, 18064014), 1.6 µL of 40% PEG-8000 (Sigma), 0.5 µL of DTT, 0.25 µL of RNase inhibitors (Takara, 2313A), 0.1 µL of 100 µM TSO (TIB MolBiol, 5'-AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3'), 0.06 µL of 1M MgCl2 (Sigma, M1028), 2 µL of 5M Betaine (Sigma, B0300-1VL), and 0.5 µL of Superscript II RT. A total of 7 µL of this reverse transcription mix was added to the 3 µL of the annealed RNA mix, and the mixture was incubated for 90 minutes at 42 °C, followed by 15 minutes at 70 °C. Preamplification of the resulting cDNA was performed using KAPA HiFi ReadyMix (KM2605) for 14 cycles with ISPCR primers (10 µM, Sigma, 5′-AAGCAGTGGTATCAACGCAGAGT-3′), and the product was purified using Agencourt AMPure XP beads (Beckman Coulter). For each sample, 2.5 µL of 120 µg/µL cDNA was used for tagmentation, which was carried out using the Nextera XT kit (Illumina, 15032354). The preamplified cDNA was mixed with 5 µL of tagment DNA buffer and 2.5 µL of Amplicon Tagment Mix, and the reaction was incubated at 55 °C for 5 minutes. The tagmentation reaction was stopped with 2.5 µL of NT buffer, and the samples were incubated at room

temperature for 5 minutes. The tagmented DNA was then amplified for 12 cycles using the standard i5 and i7 Nextera Unique Double Indexes along with a tailed i7 index, which contains an overhang enabling the capture of the 5' end of the transcripts. The libraries were sequenced in 150 bp paired-end mode using the Illumina NovaSeq6000 platform.

RNA-seq analysis

RNA-seq read pairs were aligned to the mouse reference (GRCm38 primary assembly genome) using STAR (version 2.7.6a) and the gencode annotation (vM20). ERCC spike-in sequences and annotations were obtained from https://assets.thermofisher.com/TFS-Assets/LSG/manuals/ERCC92.zip. ERCC, human SUV39H1 and membrane-eGFP sequences were added to the reference genome prior alignment. STAR parameters were set to --outFilterMultimapNmax 100 --winAnchorMultimapNmax 100. Reads were counted at genes and transposable elements (TE) using TEcount with the parameters --mode multi --stranded no. TE annotation was obtained from https://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/mm10_rmsk_TE.gtf.gz. Only samples that met our quality criteria of at least 500 thousand read counts, less than 15 percent mitochondrial and ERCC reads, respectively, were included in the analysis. Genes and TEs with at least one read detected in one fourth of the samples were considered. Differential expression analysis was performed using DESeq2 (version 1.26.0) in R (version 3.6.3). Read counts were normalized by the default DESeq2 method. Results were visualized on MA-plots for which genes were colored by significance (adjusted p-value < 0.05) and labeled according to DBTMEE[91] gene clusters. DBTMEE data were obtained from the table cluster_gene.tsv at the link https://dbtmee.hgc.jp/download/data/tables.tar.gz. For principal component analysis (PCA), normalized counts were log2 transformed after adding a pseudo-count of 1. To generate RNA-seq genome coverages, STAR alignments were filtered for uniquely mapped reads by samtools (version 1.16.1) with the parameter -q 255. Read pairs were counted in 100-kb consecutive bins (same bin size as for DamID profiles), normalized by the sum of the counts, and multiplied by a million. Replicates were averaged for downstream analysis. The bin-based log2 fold change between Pool vs. Control of the RNA-seq was directly compared to the log2 fold change of the DamID data.

CUT&RUN library generation and sequencing

CUT&RUN for H3K9me3 was conducted following the published protocol[92] with modifications for embryos. 60 to 80 late 2-cell stage embryos (48 h post-hCG) with intact zona pellucida were washed three times in a buffer containing 20 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 0.5 mM

38

Spermidine, 0.1% BSA, and 1× Roche protease inhibitor cocktail. The embryos were then transferred to antibody buffer (1:100 dilution of anti-H3K9me3 antibody (Abcam, ab8898 or Millipore, 17-625) in wash buffer containing 0.05% Triton-X and 2 mM EDTA pH 8.0). After overnight incubation with the antibody at 4 °C, embryos were washed in Triton wash buffer (wash buffer containing 0.05% Triton-X) and incubated with pAG-MNase (1:20 or 1:200 dilution; EpiCypher, 15-1016) for 1 hour at room temperature. The embryos were then washed in Triton wash buffer and transferred to a drop of ice-cold calcium isolation buffer (2 mM CaCl2 in wash buffer) and incubated at 4 °C for 30 minutes or 2 hours for targeted chromatin digestion. An equal volume of 2× EGTA-STOP buffer (340 mM NaCl, 20 mM EDTA pH 8.0, 10 mM EGTA pH 8.0, 50 µg/mL Glycogen, 50 µg/mL RNaseA, 0.05% Triton-X) was added to stop the reaction. The embryos were then incubated at 37 °C for 30 minutes to release digested chromatin fragments, and the supernatant was carefully collected. DNA extraction was performed using the QIAquick kit, and the purified CUT&RUN DNA was stored at -20 °C. Illumina library preparation was performed as previously described[93] with 15 or 18 PCR cycles. Libraries were then sequenced in 150 bp paired-end mode on the NovaSeq6000 platform.

CUT&Tag library generation and sequencing

CUT&Tag for H3K4me3 was performed as previously described[94], with modifications for embryos. Briefly, 60 to 80 late 2-cell embryos with intact zona pellucida were collected 48 h post-hCG, permeabilized with a Triton-X-containing buffer, and incubated overnight at 4 °C with primary antibody (anti-H3K4me3: EpiCypher, 13-0041; 1:100 dilution). This was followed by a 30-minute incubation at room temperature with a secondary antibody (Guinea Pig anti-Rabbit IgG, AntibodiesOnline, ABIN101961). After incubation with the pA-Tn5 adaptor complex (Diagenode, C01070001) for 1 hour at room temperature, tagmentation was carried out for 1 hour at 37 °C in a MgCl2-containing buffer. DNA was then extracted by incubating with a 0.1% SDS-containing buffer for 1 hour at 58 °C. SDS was neutralized with Triton-X, and the PCR reaction was conducted directly. After 18 cycles of PCR amplification using NEBNext High-Fidelity 2X PCR Master Mix (M0541), the PCR product was cleaned and size-selected using AMPure XP (Beckman Coulter). Samples were sequenced using the Illumina NovaSeq6000 platform in 150 bp paired-end mode.

Analysis of CUT&RUN and CUT&Tag data

Paired-end reads were trimmed by cutadapt (version 3.4) with parameters -a CTGTCTCTTATA -A CTGTCTCTTATA -a AGATCGGAAGAGC -A AGATCGGAAGAGC --minimum-length=20. After trimming, reads were aligned to the mouse reference (GRCm38) using bowtie2 (version 2.3.5)

with parameters --end-to-end --very-sensitive --no-unal --no-mixed --no-discordant -I 10 -X 500. Reads were filtered by mapping quality score using samtools (version 1.3) with parameter -q 12. Read pairs were read into R (version 4.1.2) using the readGAlignmentPairs function from the GenomicAlignment package (version 1.30) and were filtered for unique fragments. Fragments aligned to the mitochondrial genome or small scaffolds were not considered in the analysis. Fragments were counted in 100-kb consecutive genomic bins (same bin size as for DamID profiles), normalized by the sum of the fragment counts, and multiplied by a million. For chromosomal tracks, replicates were averaged, and normalized counts were visualized along the genomic coordinates. To call broad H3K9me3 domains, a two-state hidden Markov model (HMM)[89] was applied to the normalized counts. For other subsequent analyses, normalized counts were log2 transformed after adding a pseudo-count of 0.1.

Analysis of public chromatin datasets

Published datasets were downloaded from GEO with accession numbers GSE66581, GSE101571[95,96] (ATAC-seq), GSE71434[97] (H3K4me3 ChIP), GSE112834[98] (H3K36me3 ChIP), GSE72784[99] (H3K27ac ChIP), GSE98149[59] (H3K9me3 ChIP), GSE76687[62] (H3K27me3 ChIP) and GSE135457[100] (Pol2 Stacc-seq), GSE76642[101] (DNaseI-seq). Chromatin datasets were processed and analyzed as the CUT&RUN and CUT&Tag data. For heatmap visualizations, log2 normalized counts were scaled (centered to the genome-wide mean and divided by the standard deviation), and the median of genomic bins with significantly increased or decreased DamID OE values was taken.

Hi-C data analysis

Hi-C compartment coordinates and scores were obtained from GEO with accession GSE82185[77] and analyzed as previously described[31].

Immunofluorescence

Embryos were fixed in 4% PFA for 20 min at room temperature and permeabilized in PBS containing 0.5% Triton-X for 20 min. Embryos were kept in blocking buffer (3% BSA in PBS) for 4-5 h and then incubated overnight in primary antibody (H3K27me3, Millipore, 07-449, 1:250; H4K20me3: Millipore, 07-463, 1:250; H3K9me2: Abcam, ab1220, 1:250 or Active Motif, 39239, 1:250; H3K9me3: Abcam, ab8898, 1:250 or Active Motif 39286, 1:100; H3K9ac: Abcam, ab4441, 1:250; H3K4me3: Abcam, ab8580, 1:250 or Diagenode, C15410003, 1:250; H3K27me2: Abcam, ab24684, 1:250; HA-tag: Roche, 11867423001, 1:500) diluted in blocking buffer. After overnight incubation, samples were washed three times in PBS and stained with secondary antibodies conjugated with (Alexa Fluor 555 or Alexa Fluor 647) in blocking buffer for 2-3 h. After washing

40

three times in PBS, embryos were mounted in Vectashield (Vector Laboratories) containing DAPI. Confocal imaging was performed using a 40× oil objective in a Leica SP8 confocal microscope. Experiments were performed on 2 independent experimental days, with each replicate including embryos from several mice (20 mice for 4 experimental conditions), which we randomly assign to experimental groups. For quantification of pronuclear size, the PN size was determined by the DAPI area on maximum-intensity projection images. Relative pronuclear size was calculated by normalizing the size of each pronuclei over the mean size of control paternal pronucleus, which was set to 1.

DNA FISH

DNA FISH was performed as previously described[31,102]. In brief, BACs (Table S1) were obtained from BACPAC (https://bacpacresources.org/home.htm), purified with NucleoBond BAC 100 kit (Macherey-Nagel) and nick-translated with dUTPs conjugated to ATTO550, ATTO594 or ATTO647N (Jena Bioscience). To combine nuclear lamina staining with DNA FISH, we performed immunostaining with an anti-LaminB1 antibody (Santa Cruz, sc-6216, 1:100) as described above, followed by postfixation in 2% PFA for 10 min at 37 °C. Next, embryos were briefly permeabilized (0.5% Triton-X 100, 0.02% RNAseA and 1 mg/mL PVP in PBS) for 10 min at room temperature and treated with HCl solution (0.1N HCl, 0.7 Triton-X 100 and 1 mg/mL PVP in water) for 1 min and equilibrated into prehybridization buffer (10% dextran sulfate, 2× SSC, 0.5 mM EDTA, 50% formamide, 1 mg/mL PVP, 0.05% TritonX, 0.5 mg/mL BSA) at 55 °C for one hour. Embryos were incubated in hybridization buffer containing 1 µg/µL mouse Cot-1 DNA, denatured at 83 °C for 10 min and blocked at 37 °C for one hour. Lastly, embryos were transferred into drops of hybridization buffer containing a mixture of probes, each at 250 ng/µL which were previously denatured at 83 °C for 10 min under mineral oil. After overnight hybridization at 37 °C on a dry-bath, embryos were washed once in 2× SSC, 0.1% Triton-X 100, 1 mg/mL PVP at room temperature followed by washing three times for 10 min in 0.2× SSC, 0.01% Triton-X 100, 1 mg/mL PVP at 52 °C and mounted in Vectashield containing DAPI (Biozol) on a high precision (170 µm +/- 5 µm) glass bottom dish (Ibidi) to preserve 3D structure.

DNA FISH image analysis

Microscopy images were acquired on a Leica SP8 point scanning confocal equipped with a Plan Apochromat 100×/1.4 NA oil objective at a voxel size of 0.025 × 0.025 × 0.3 µm (x, y, z). For distance analysis, DNA FISH foci and LaminB1 signals were segmented from unadjusted raw images using Ilastik[103] and analyzed using a custom Python script. The script segments nuclear boundaries (LaminB1 signal), identifies and filters FISH signals, and calculates distances between

41

FISH signal centroids, the nuclear surface and the center of the nucleus (defined as the inner volume encapsulated by LaminB1). These distance measurements were then used to compute the distance ratio (*dRatio*) by taking the shortest distance from the DNA FISH centroid to LaminB1 and dividing it by the total distance. The total distance is the sum of two parts: the distance from the center of the nucleus to the DNA FISH signal, and the shortest distance between the DNA FISH signal and LaminB1. Thus, a *dRatio* close to 1 indicates proximity to the center of the nucleus, while a *dRatio* close to 0 indicates proximity to the nuclear periphery. dRatio was compared to the log2 mean DamID OE values of the overlapping genomic bins. Representative images are shown as a maximum intensity projection of 2 to 3 z-stacks (0.3 µm), in which noise was smoothed with a gaussian blur on LaminB1 and the FISH signal (sigma 1.0 and 2.0 radius, respectively) uniformly across all images equally and contrast/brightnes adjusted and thus fluorescence intensity is not comparable across the images presented in the Figures.

Data availability

All datasets generated in this study were deposited in GEO under SuperSeries Accession: GSE278721; token: kzopgkgklpeztsz.

LaminB1-DamID datasets are accessible at GSE244496; token: sbsfmgugldmzjar

H3K9me3 CUT&RUN datasets are accessible at GSE278718; token: cxetwomwzbizhyt

H3K4me3 CUT&Tag datasets are accessible at GSE278719; token: avyxaiqordgbhwn

Single-embryo RNA-seq datasets are accessible at GSE278720; token: axshmcwivzsnnwf

Code availability

Custom code used in this work is available at https://ascgitlab.helmholtz-muenchen.de/public_pipelines.

42

A Phase-I Screening

Pool	Brief description	Candidates
A	Chromatin anchors to NE, LINC complex	DN Baf, DN Prr14, DN Syne1
B	Nuclear export, Cortical actin dynamics	Exportin6, cVCA Actin
C	Nuclear basket, Nucleoporins	Tpr, DN Nup98
D	Nuclear and cortical myosin	DN NM1 (Nuclear myosin 1), DN MyosinVb
E	Nuclear actin polymerization mutant, Deacetylase	Actin R62D, Hdac3
F	H3K27me3 demethylases	Kdm6a (Utx), Kdm6b (Jmjd3)
G	H4K20 HMTs	Set8, Suv420h1, Suv420h2
H	H3K9me1/me2 demethylases	Kdm3a, Kdm3b
I	H3K4me2, H3K9me2/K27me2 demethylases	Lsd1 (Kdm1a), Kdm7a, Kdm7c (Phf2)
J	H3K9me3 HMT and Heterochromatin readers	Suv39h1, Hp1α (Cbx5), Hp1γ (Cbx3)
K	Nucleolar binder, Euchromatin reader	Phf19, G2e3, Dhx33
L	Histone deacetylases	Hdac1, Hdac6, Sirt1
M	H1 subtypes and H2A variant	H1.2, H1.4, H1.5, macroH2A
N	H3K9me2/me3 HMTs	Ehmt1 (GLP), Ehmt2 (G9a), Setdb1, Setdb2

DN Baf: G25E Baf DNA-binding mutant (Haraguchi et al., 2001)
DN Prr14: HP1α binding mutant of Prr14 (Yang et al., 2015)
DN Syne1: KASH domain of Nesprin1α tagged with mCherry (Lombardi et al., 2011)
DN Nup98: fragment (amino acid 1-504) of Nup98 (Liang et al., 2013)
DN NM1: Mutant with impaired motor activity and actin-binding (Ye et al., 2008)
DN Myosin Vb: Tail fragment of MyosinVb (Provance et al., 2008)
Actin R62D: Nuclear-targeted (NLS) polymerisation-defective actin (Posern et al., 2002)
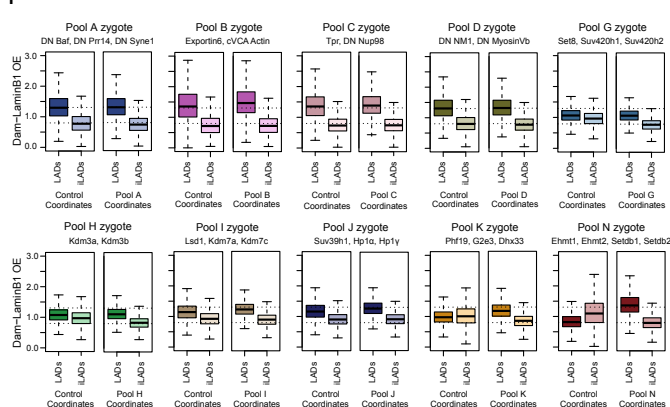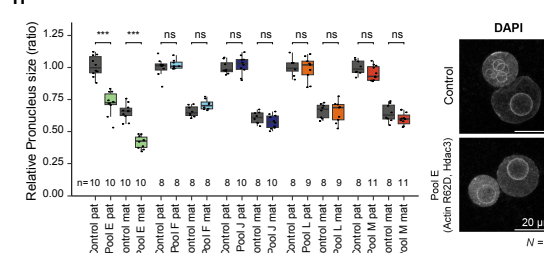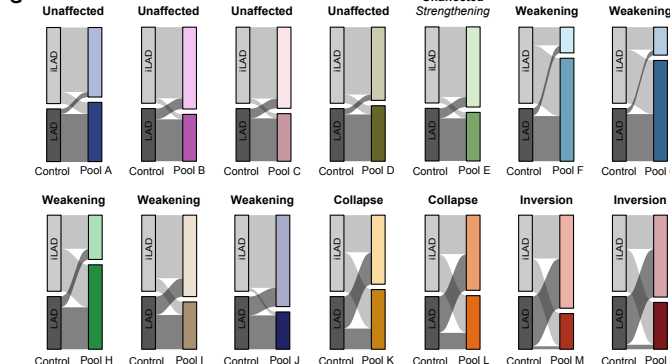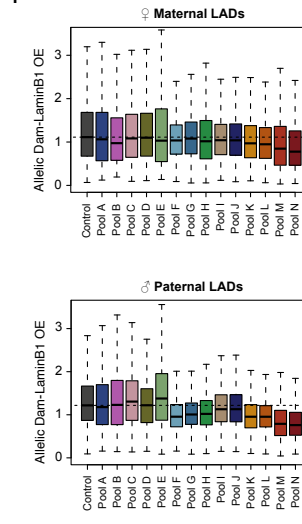**NE - Nuclear envelope; DN - Dominant negative; HMT - Histone methyltransferase**
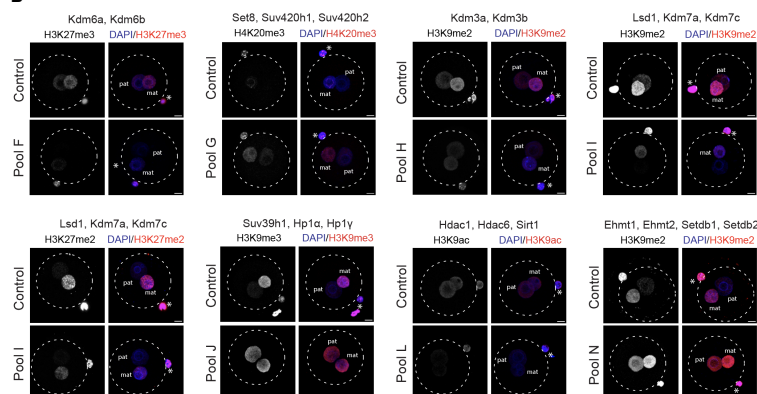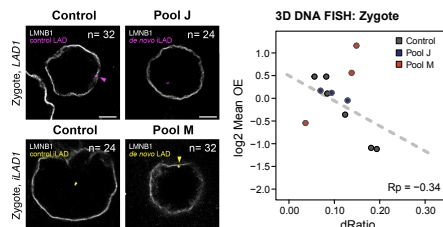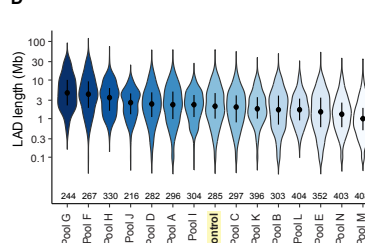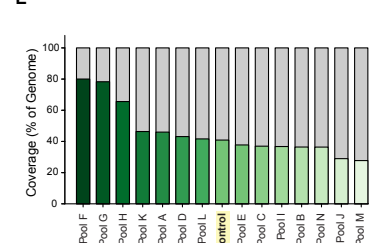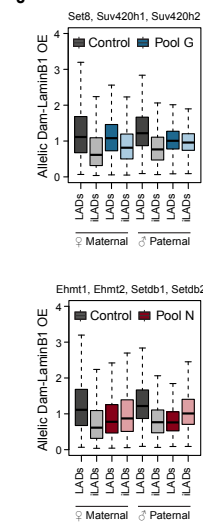
Supplemental Information

Supplemental Figure Legends S1 to S7.
Supplemental Tables S1 to S7.

Supplemental Figure Legends

Figure S1 related to Figure 1

(A) Brief description of the candidates investigated in this study and pooling of those for phase I of screening. Information about the dominant negative (DN) constructs is shown below with references[104–110].

(B) Representative maximum intensity projections of confocal images from immunostainings of the indicated histone modifications in late zygotes (28-30h post-hCG injection) from control or experimental groups. Dashed lines roughly demarcate the plasma membrane. DNA was counterstained with DAPI. On the merged images, asterisks indicate the polar body. Scale bars, 10 μm. N$\geq 2; n \geq 10$. mat: maternal pronucleus; pat: paternal pronucleus.

(C) (*Left*) Representative confocal images of immuno-3D FISH in late zygotes for LaminB1 (LMNB1) and genomic regions within LADs or iLADs as indicated. Regions corresponding to LADs and iLADs in control embryos are shown in purple and yellow, respectively. Arrowheads point to DNA-FISH spots localizing at nuclear lamina. Scale bars, 5 μm. n = number of DNA FISH spots analyzed. Data derive from two biologically independent experiments. (*Right*) Correlation between DamID values and distance measurements from DNA FISH of all indicated datasets (control, Pool J and Pool M). The y-axis is the log2 transformed mean OE values for genomic loci corresponding to selected LADs and iLADs derived from the DamID replicates. The x-axis indicates the average distance ratio (*dRatio*) of the individual FISH probes determined from at least 24 measurements ($24 \leq n \leq 44$). A *dRatio* close to 1 indicates proximity to the center of the nucleus while a *dRatio* close to 0 indicates proximity to the nuclear periphery. Pearson's correlation (Rp) is indicated. Note the overall negative correlation between DamID values and distance to the nuclear periphery, as expected.

(D) Distribution of zygotic LAD length. Violin plots show the 25th and 75th percentiles (black lines) and median (circles). n indicates the number of LADs called, shown below violin plots. The candidate pools are arranged in a descending order based on median LAD size and control is highlighted in yellow.

43

(E) Percentage genomic coverage of LADs and iLADs. The candidate pools are arranged in a descending order of LAD coverage and control is highlighted in yellow.

(F) Box plots of Dam-LaminB1 OE mean values in control and *de novo* called LAD and iLAD regions. The horizontal dotted lines indicate the median OE values from the control zygotic LADs (upper line) and iLADs (lower line).

(G) Alluvial plot showing zygotic LAD reorganization upon perturbations performed with respect to the control.

(H) Quantification of pronuclear area across several Pools in phase I screening. n = number of pronuclei analyzed from two independent experiments (N = 2). Pairwise Wilcoxon rank-sum test was performed and *** indicates p <0.001, where p is adjusted p-value for multiple comparisons. Representative maximum intensity projections of DAPI staining in late zygotes are shown for Pool E and control. Scale bars, 20 µm; ns: non-significant (p >0.05); mat: maternal pronucleus; pat: paternal pronucleus.

(I, J) Box plots of allelic (paternal and maternal) Dam-LaminB1 OE mean values from hybrid zygotes (C57BL/6J×CBA/H female × DBA/2J male) in control LAD and iLAD regions. Only genomic bins containing more than 30 allelic GATC reads were analyzed. The number of 100-kb genomic bins analyzed are: n = 1,835 for maternal LADs; 2,419 for maternal iLADs; 1,446 for paternal LADs; and 2,808 for paternal iLADs. The horizontal dotted lines in panel I indicate the median allelic OE values from the control maternal or paternal LADs. The allelic LAD and iLAD coordinates used for this analysis were extracted from GSE112551.
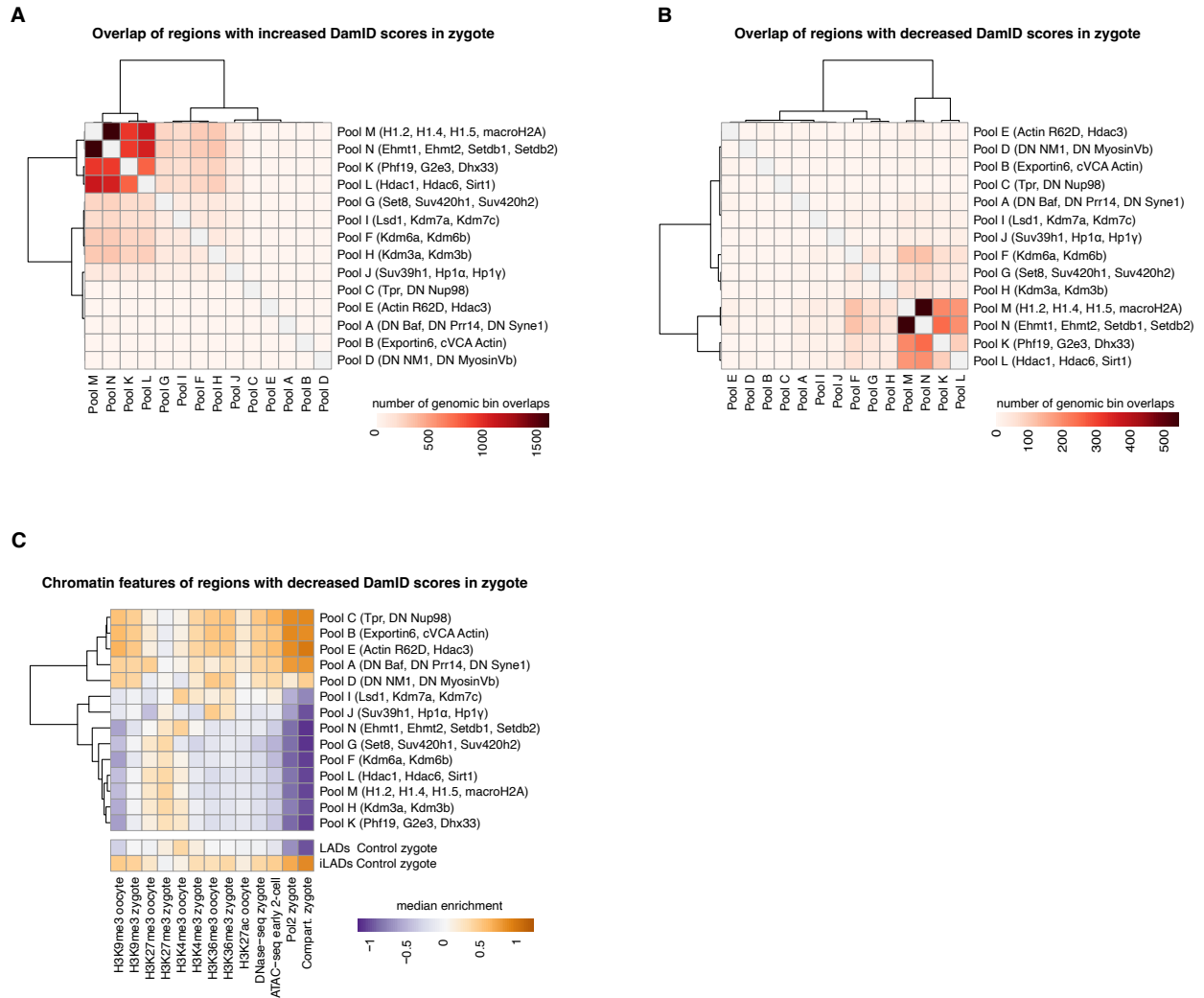
See also Figure 1 and Tables S1 and S2.

A

Overlap of regions with increased DamID scores in zygote

B

Overlap of regions with decreased DamID scores in zygote

C

Chromatin features of regions with decreased DamID scores in zygote

Figure S2 related to Figure 2

(A) Heatmap showing overlap of genomic regions that show increased Dam-LaminB1 OE values ('up') in zygote with respect to control.

(B) Heatmap showing overlap of genomic regions that show significantly reduced Dam-LaminB1 OE values ('down') in zygote with respect to control.

(C) Heatmap showing enrichment of chromatin features in genomic regions that lose lamina interactions ('down') compared to controls in zygote. Chromatin feature enrichment in LADs and inter-LADs (iLADs) of control zygotes is shown below for comparison. Positive compartment scores (Compart.) define A compartment.

See also Figure 2 and Table S3.
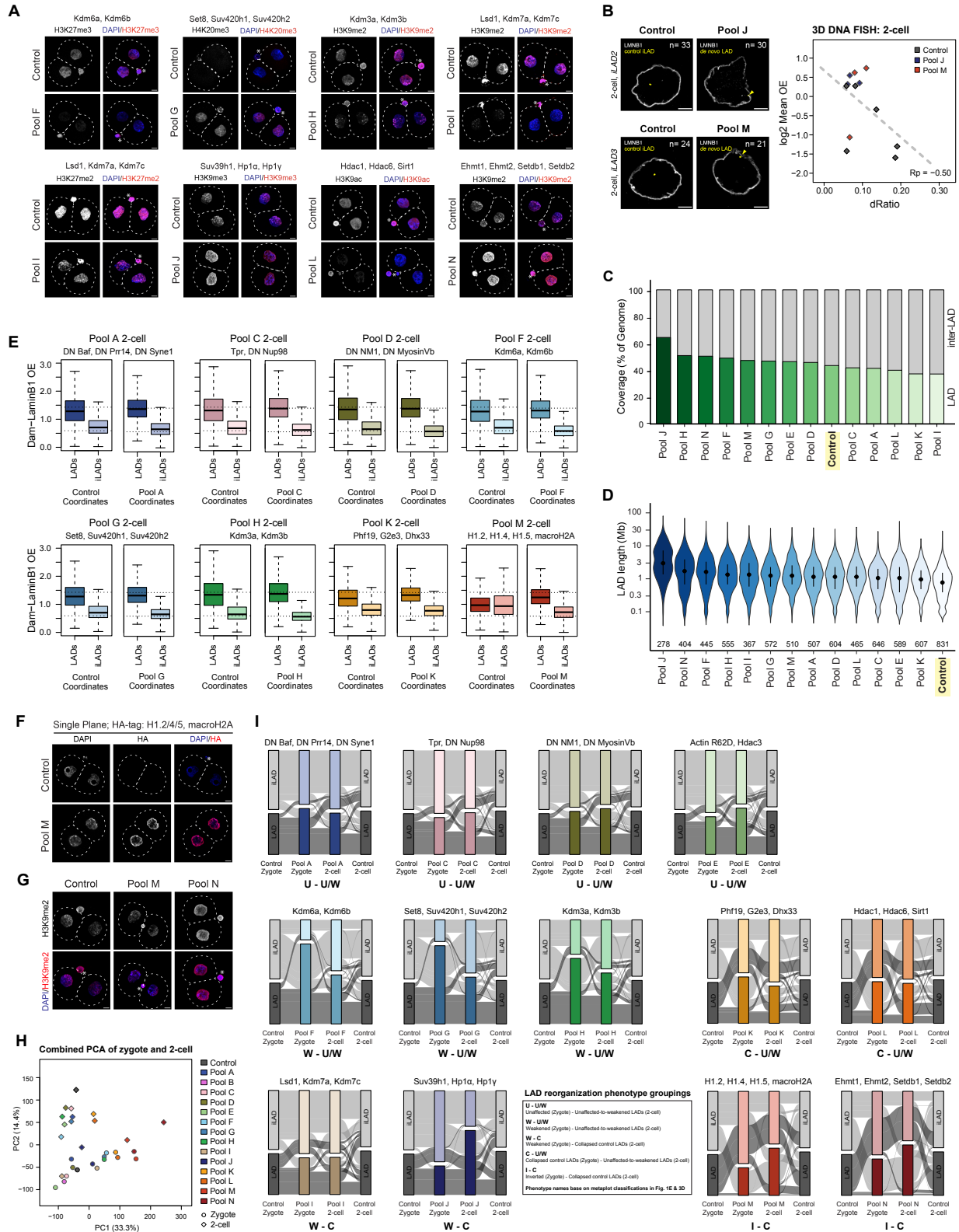
Figure S3 related to Figure 3

(A) Representative confocal images showing maximum intensity projections from immunostainings of the indicated histone modifications in control and experimental late 2-cell stage embryos (48-50h post-hCG injection). Dashed lines roughly demarcate the contour of the embryos. DNA was counterstained with DAPI. On the merged images, asterisks indicate the polar body. Scale bars, 10 µm. $N \geq 2; n \geq 10$.

(B) (*Left*) Representative confocal images of immuno-3D FISH in late 2-cell stage embryos for LaminB1 (LMNB1) and genomic regions within LADs or iLADs as indicated. Regions corresponding to iLADs in control embryos are shown in yellow. Arrowheads point to DNA-FISH spots localizing at nuclear lamina. Scale bars, 5 µm. n = number of DNA FISH spots analyzed. Data derive from two biologically independent experiments. (*Right)* Correlation between DamID values and distance measurements from DNA FISH of all indicated datasets (control, Pool J and Pool M). The y-axis is the log2 transformed mean OE values for genomic loci corresponding to selected LADs and iLADs derived from the DamID replicates. The x-axis indicates the average distance ratio (*dRatio*) of the individual FISH probes determined from at least 17 measurements ($17 \leq n \leq 33$). A *dRatio* close to 1 indicates proximity to the center of the nucleus while a *dRatio* close to 0 indicates proximity to the nuclear periphery. Pearson's correlation (Rp) is indicated. Note the overall negative correlation between DamID values and distance to the nuclear periphery, as expected.

(C) Distribution of LAD length in 2-cell embryos. Violin plots show the 25th and 75th percentiles (black lines) and median (circles). n indicates the number of LADs called, shown below violin plots. The candidate pools are arranged in a descending order based on median LAD size and control is highlighted in yellow.

(D) Percentage genomic coverage of LADs and iLADs. The candidate pools are arranged in a descending order of LAD coverage in 2-cell embryos and control is highlighted in yellow.

(E) Box plots of Dam-LaminB1 OE mean values in control and *de novo* called LAD and iLAD regions. The horizontal dotted lines indicate the median OE values from the control 2-cell LADs (upper line) and iLADs (lower line).

(F) Representative single confocal images from immunostaining using an HA-antibody in late 2-cell stage embryos. In Pool M, all histones (H1.2, H1.4, H1.5 and macroH2A) contain an N-terminal HA-tag. Scale bars, 10 µm. $N \geq 2; n \geq 10$.

(G) Representative confocal images from maximum intensity projections from H3K9me2 immunostainings in late 2-cell stage embryos in control or embryos from Pool M and Pool N. Scale bars, 10 µm. $N \geq 2; n \geq 10$.

46

(H) Combined Principal component analysis (PCA) of all zygotic and 2-cell DamID samples from phase I screening. Each data point represents the mean of the biological replicates for the corresponding manipulation indicated by the colour code. The percentage of variance explained by PC1 and PC2 is indicated in axis labels.

(I) Alluvial plots showing reorganization of genomic regions between LAD and iLAD during the maternal-to-zygotic transition and how that is affected upon perturbations with respect to the control zygote and 2-cell embryos. LAD reorganization phenotype groupings (U-U/W, W-U/W, W-C, C-U/W, and I-C) are explained below the plots.

See also Figure 3 and Table S2.

A

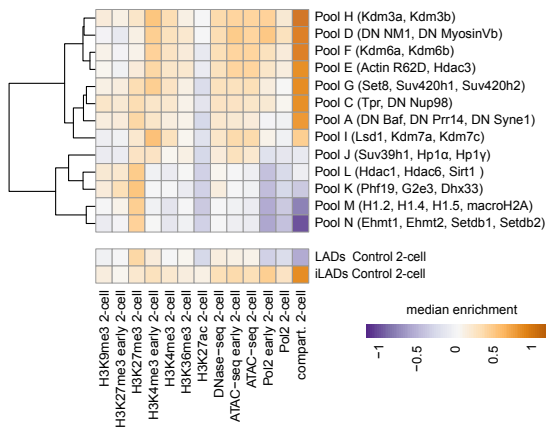Overlap of regions with increased DamID scores in 2-cell

B
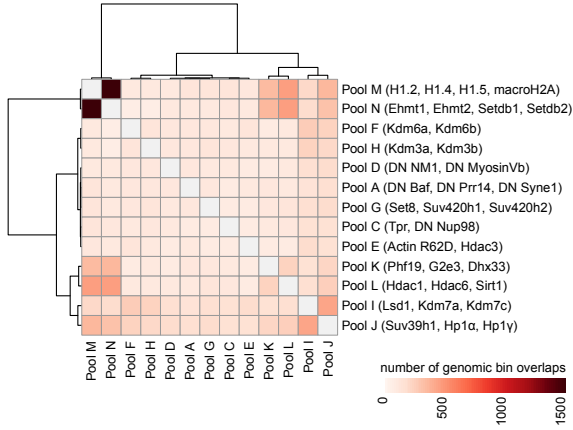
Overlap of regions with decreased DamID scores in 2-cell

C
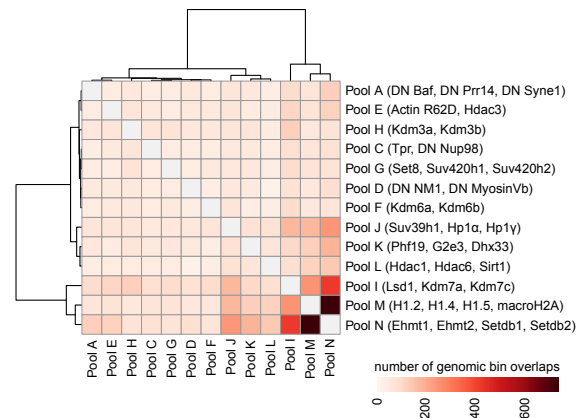
Chromatin features of regions with decreased DamID scores in 2-cell

D

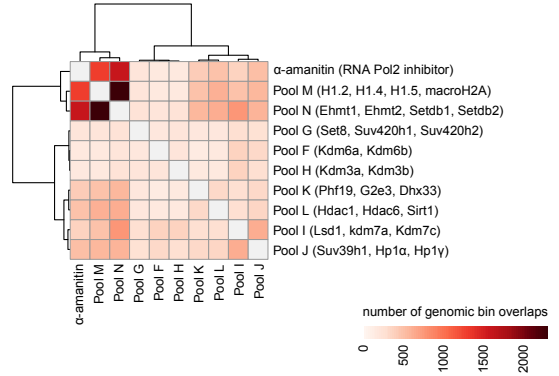Overlap of differentially Dam-LaminB1 methylated regions in 2-cell

Figure S4 related to Figure 4

(A) Heatmap showing overlap of genomic regions that show significantly increased Dam-LaminB1 OE values ('up') in phase I screening samples with respect to control at the 2-cell stage.

(B) Heatmap showing overlap of genomic regions that show significantly reduced Dam-LaminB1 OE values ('down') with respect to control in 2-cell embryo samples.

(C) Enrichment of chromatin features in genomic regions that lose lamina interactions ('down') compared to controls in 2-cell stage embryos. Chromatin feature enrichment in control 2-cell LADs and inter-LADs (iLADs) is shown below for comparison. Positive compartment scores (Compart.) define A compartment.

(D) Heatmap showing overlap of all genomic bins that show differential lamina interactions ('up' + 'down') in 2-cell stage embryos with respect to controls comparing α-amanitin treatment to DamID samples from screening phase I where different chromatin pathways are targeted with pooled candidates.
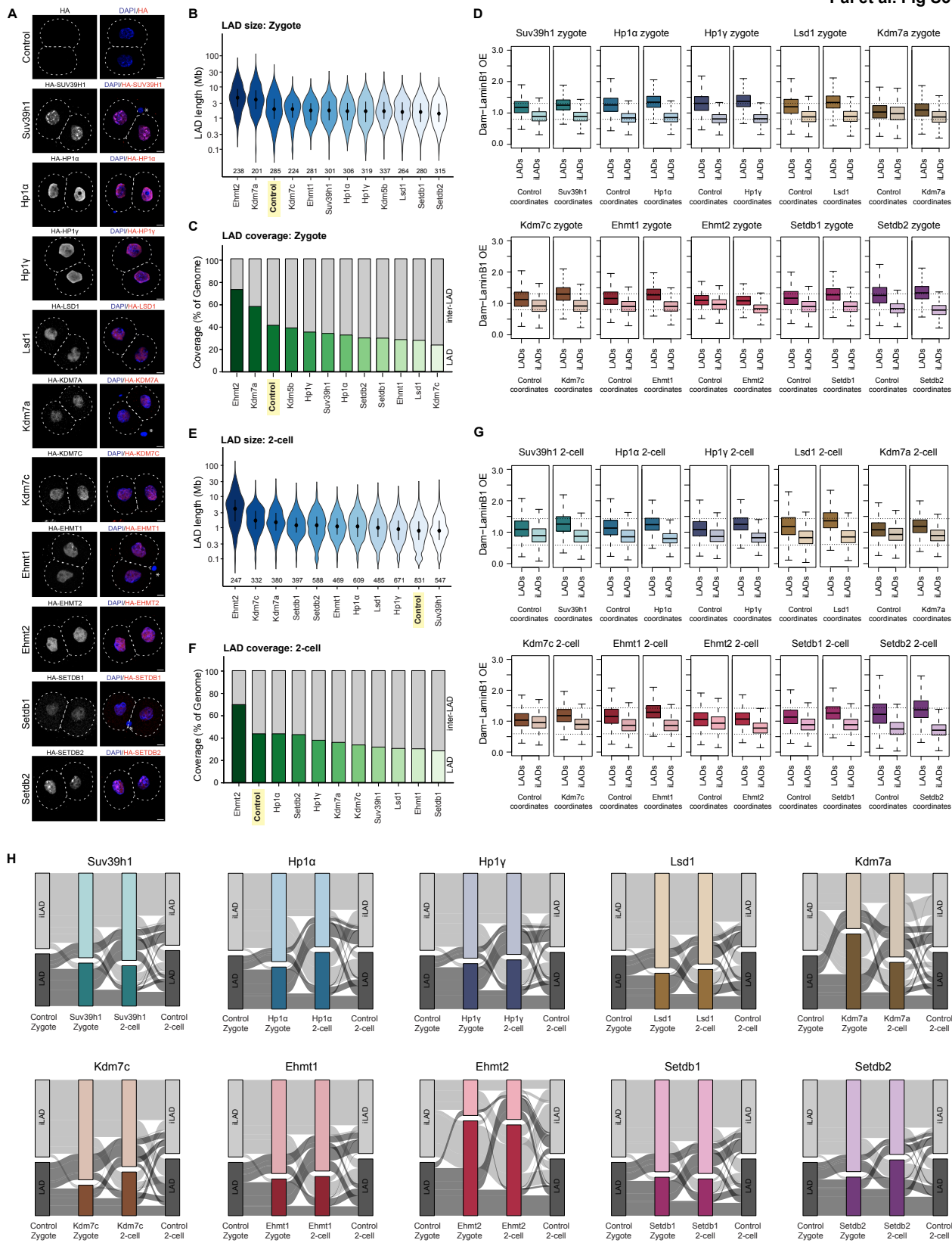
See also Figure 4 and Tables S4.

Figure S5 related to Figure 5

(A) Representative confocal images showing maximum intensity projections from immunostaining using an HA-antibody in late 2-cell stage embryos (48-50h post-hCG injection). All candidates expressed in the phase II screening contain N-terminal HA-tag. Dashed lines roughly demarcate the cell membrane. DNA was counterstained with DAPI. On the merged images, asterisks indicate the polar body. Scale bars, 10 μm. N$\geq 2; n \geq 10$.

(B) Distribution of zygotic LAD length in phase II screening samples. Violin plots show the 25th and 75th percentiles (black lines) and median (circles). n indicates the number of LADs called, shown below violin plots. The candidates are arranged in a descending order based on median LAD size and control is highlighted in yellow.

(C) Percentage genomic coverage of LADs and iLADs in zygotes. The candidates are arranged in a descending order of LAD coverage and control is highlighted in yellow.

(D) Box plots of Dam-LaminB1 OE mean values in control and *de novo* called LAD and iLAD regions for zygotes. Box plots show the median, and the interquartile range (IQR), and whiskers depict the smallest and largest values within 1.5 ×IQR. The horizontal dotted lines indicate the median OE values from the control zygotic LADs (upper line) and iLADs (lower line).

(E) Distribution of LAD length in 2-cell embryos. The candidates are arranged in a descending order based on median LAD size and control is highlighted in yellow.

(F) Percentage genomic coverage of LADs and iLADs in 2-cell embryos. The candidates are arranged in a descending order of LAD coverage and control is highlighted in yellow.

(G) Box plots of Dam-LaminB1 OE mean values in control and *de novo* called LAD and iLAD regions for 2-cell embryos. The horizontal dotted lines indicate the median OE values from the control 2-cell LADs (upper line) and iLADs (lower line).

(H) Alluvial plots showing reorganization of genomic regions between LAD and iLAD during the maternal-to-zygotic transition and how that is affected upon candidate expression with respect to the control zygote and 2-cell embryos.
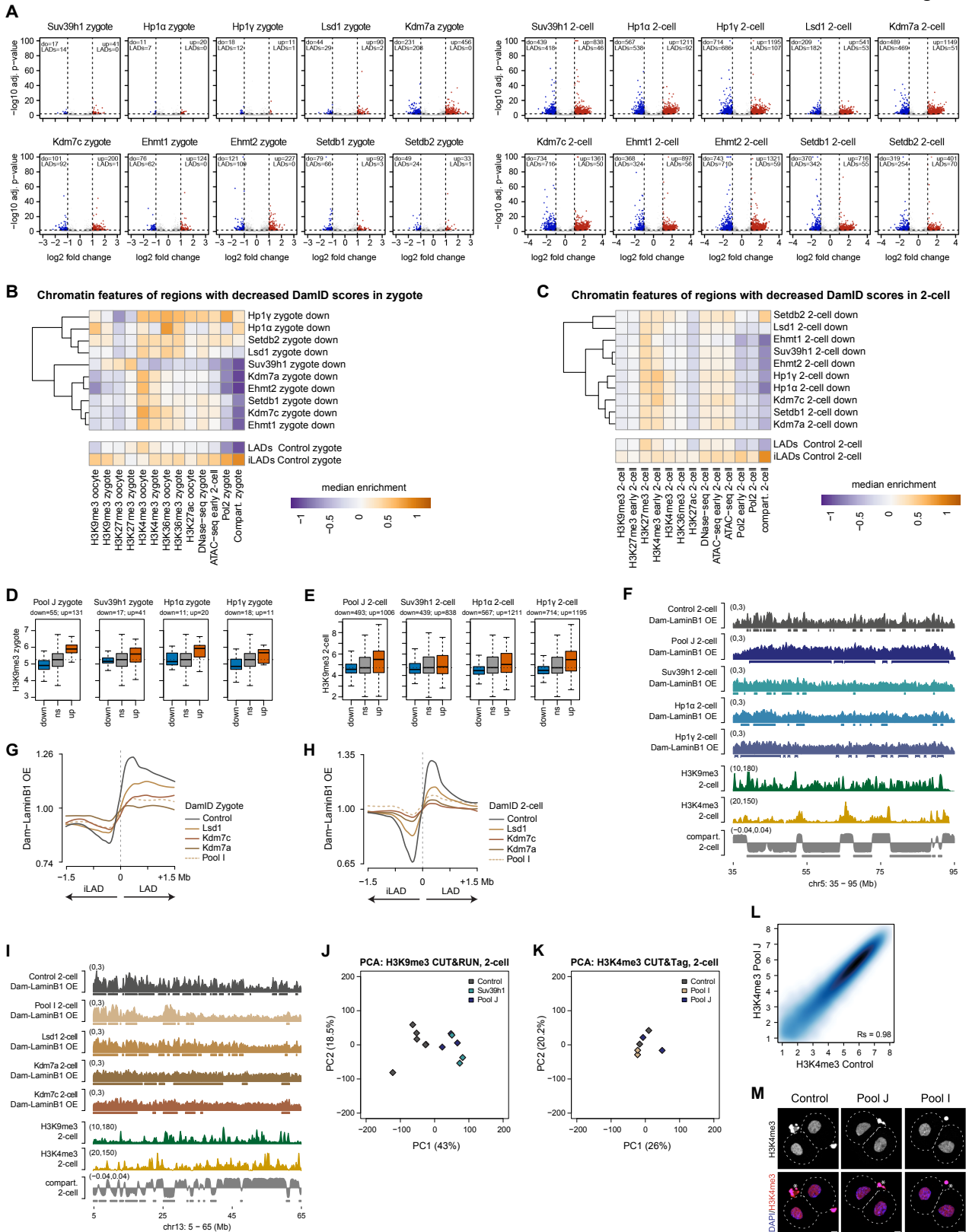
See also Figure 5 and Table S5.

Figure S6 related to Figure 6

(A) Volcano plots of genomic regions that show differential Dam-LaminB1 OE values in comparison to control embryos for corresponding developmental stages (left: zygote, right: 2-cell stage embryo). 'Up' indicates the number of 100-kb genomic bins that have significantly higher OE values (log2 fold change >1 and adjusted p-value <0.01; red dots) and 'do'(down) shows the number of genomic bins that show significantly reduced OE value (log2 fold change <-1 and adj. p-value <0.01; blue dots). The number of 'up' or 'do' genomic bins that belong to LADs in control embryos is indicated below.

(B) Heatmap showing enrichment of chromatin features in genomic regions that lose lamina interactions ('down' in volcano plots) compared to controls in zygotic DamID samples when candidate effectors are expressed. Chromatin feature enrichment in zygotic LADs and inter-LADs (iLADs) in control embryos is shown below for comparison. Positive compartment scores (Compart.) define A compartment.

(C) Enrichment of chromatin features in genomic regions that lose lamina interactions ('down' in volcano plots) compared to controls in 2-cell stage samples.

(D, E) Box plots showing log2 transformed H3K9me3 enrichment in genomic regions that gain ('up') or lose ('down') OE value compared to controls in zygote (E) and 2-cell stage (F) embryos. Box plots show the median, and the interquartile range (IQR), and whiskers depict the smallest and largest values within 1.5 ×IQR. Number of 100-kb genomic bins analyzed is indicated. ns: genomic regions with non-significant changes in lamina association. The horizontal dotted lines indicate the median signal in the 'ns' genomic regions for the corresponding analysis. The H3K9me3 data is publicly available data and derive from control zygote and 2-cell stage embryos.

(F, I) Dam-LaminB1 OE value, histone modification enrichment and compartment score from 2-cell stage embryos calculated from publicly available datasets. Boxes under the OE value tracks represent called LADs and wildtype 2-cell B compartment regions are indicated below the compartment score tracks.

(G, H) Average Dam-LaminB1 OE value over control LAD boundaries across different DamID samples for zygote (G) and 2-cell (H) stage embryos comparing individual effectors (Kdm7a, Kdm7c, Lsd1; solid lines) to candidate pools (Pool I; dotted line).

(J) Principal component analysis (PCA) on all H3K9me3 CUT&RUN samples and replicates from late 2-cell stage. CUT&RUN was performed in at least two independent biological replicates. Data points are colored based on the experimental conditions as indicated.

50

(K) PCA on H3K4me3 CUT&Tag samples. CUT&Tag was performed at the late 2-cell stage in two independent biological replicates. Data points are colored based on the experimental conditions as indicated.

(L) Smoothed scatter plot of genome-wide H3K4me3 enrichment values (log2 transformed). Spearman's correlation (Rs) is indicated.

(M) Representative confocal images from maximum intensity projections from immunostaining against H3K4me3 in late 2-cell stage embryos (48-50h post-hCG injection). Dashed lines roughly demarcate the cell membrane. DNA was counterstained with DAPI. On the merged images, asterisks indicate the polar body. Scale bars, 10 μm. N$\geq 2; n \geq 10$.

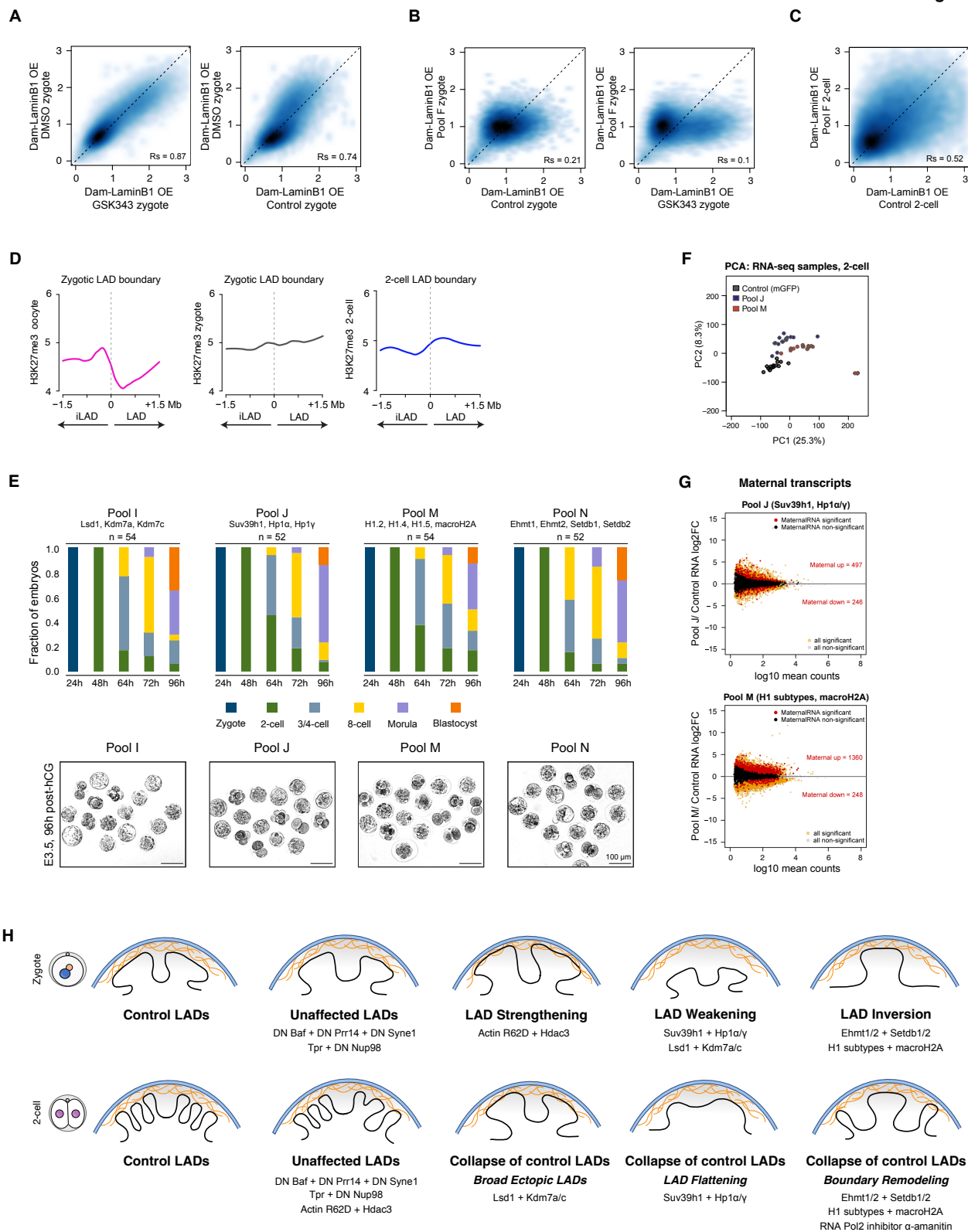See also Figure 6 and Table S6.

Figure S7 related to Figure 7

(A, B, C) Smoothed scatter plot of genome-wide Dam-LaminB1 OE mean values. Spearman's correlation (Rs) is indicated.

(D) Average H3K27me3 enrichment (log2 transformed) signal from publicly available datasets over LAD boundaries. Zero and the dotted line represent the position of the LAD/iLAD boundary in the metaplot. The 1.5 Mb region towards the right-hand side indicates LAD.

(E) Developmental progression (in percentage) of embryos from the indicated Pools. Embryos were microinjected with mRNAs for each Pool at the zygote stage and developmental progression was monitored daily. Plotted is the percentage of embryos at the indicated stage corresponding to the time post-hCG injection in hours (h; x-axis). Representative images after 3.5 days in culture (E3.5; corresponds to 96 h post-hCG). n = number of embryos analyzed. Data derive from at least three independent experiments. Scale bars, 100 μm. Developmental data for control (mGFP-injected) embryos is shown in Fig. 7F.

(F) Principal component analysis (PCA) of late 2-cell single-embryo RNA-seq read counts (genes and transposable elements combined) across experimental and control samples as indicated. Embryos (n = 16) were collected from two independent biological experiments.

(G) MA plots of log2-fold change in transcript abundance (RNA-seq counts) for 2-cell stage embryos against mean RNA-seq counts (log10 transformed). Differentially expressed genes are labeled in orange (adj. p-value <0.05), non-differential genes in gray. Differentially expressed (adj. p-value <0.05) maternal transcripts (as per DBTMEE classification) are marked in red, non-differential ones are in black.

(H) Summary of LAD disruption phenotypes in zygote (top row) and 2-cell stage embryos (bottom row). Blue lines depict the nuclear envelope and orange mesh the nuclear lamina. DN: Dominant negative construct. The relative strength of interactions with the nuclear lamina is represented by the distance between the nuclear lamina and the LADs.

See also Figure 7 and Table S7.

52

273

Supplemental Table S1. Summary of screening targets

List of protein candidates included in the screening, together with their rationale, expression patterns (RPKM) in early mouse embryos(GSE38495[111], GSE45719[112]) and the experimental design.

Supplemental Table S2. Sequencing metrics for phase I screening in zygotes and 2-cell embryos

Replicates, raw sequencing reads, filtered mapped reads and unique GATC counts per sample.

Supplemental Table S3. Differentially Dam-LaminB1 methylated regions determined by generalized linear model in zygote phase I screening samples

The columns represent information about the 100-kb genomic bins, fitted log2 OE value in the control and for each experimental condition log2 fold-change in OE values, p-values and adjusted p-values compared to control as calculated by the GLM.

Supplemental Table S4. Differentially Dam-LaminB1 methylated regions determined by generalized linear model in 2-cell phase I screening samples

The columns represent information about the 100-kb genomic bins, fitted log2 OE value in the control and for each experimental condition log2 fold-change in OE values, p-values and adjusted p-values compared to control as calculated by the GLM.

Supplemental Table S5. Sequencing metrics for phase II screening samples

Replicates, raw sequencing reads, filtered mapped reads and unique GATC counts per sample.

Supplemental Table S6. Differentially Dam-LaminB1 methylated regions determined by generalized linear model in phase II samples

The columns represent information about the 100-kb genomic bins, fitted log2 OE value in the control and for each experimental condition log2 fold-change in OE values, p-values and adjusted p-values compared to control as calculated by the GLM.

Supplemental Table S7. Differentially expressed genes

Differential gene expression analysis by DESeq2 with the category of genes based on the DBTMEE database.

53

References

1. Gorkin, D.U., Leung, D., and Ren, B. (2014). The 3D Genome in Transcriptional Regulation and Pluripotency. Cell Stem Cell *14*, 762–775. https://doi.org/10.1016/j.stem.2014.05.017.

2. Hauer, M.H., and Gasser, S.M. (2017). Chromatin and nucleosome dynamics in DNA damage and repair. Genes Dev. *31*, 2204–2221. https://doi.org/10.1101/gad.307702.117.

3. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long range interactions reveals folding principles of the human genome. Science *326*, 289–293. https://doi.org/10.1126/science.1181369.

4. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380. https://doi.org/10.1038/nature11082.

5. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. Cell *148*, 458–472. https://doi.org/10.1016/j.cell.2012.01.010.

6. Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385. https://doi.org/10.1038/nature11049.

7. Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., and Gilbert, D.M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome Res. *20*, 761–770. https://doi.org/10.1101/gr.099655.109.

8. Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., et al. (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402–405. https://doi.org/10.1038/nature13986.

9. Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012–1025. https://doi.org/10.1016/j.cell.2015.04.004.

10. Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.-L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., et al. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science *351*, 1454–1458. https://doi.org/10.1126/science.aad9024.

11. Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature *529*, 110–114. https://doi.org/10.1038/nature16490.

54

12. Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M., and van Steensel, B. (2006). Characterization of the Drosophila melanogaster genome at the nuclear lamina. Nat. Genet. *38*, 1005–1014. https://doi.org/10.1038/ng1852.

13. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951. https://doi.org/10.1038/nature06947.

14. van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780–791. https://doi.org/10.1016/j.cell.2017.04.022.

15. Briand, N., and Collas, P. (2020). Lamina-associated domains: peripheral matters and internal affairs. Genome Biol. *21*, 85. https://doi.org/10.1186/s13059-020-02003-5.

16. Aaronson, R.P., and Blobel, G. (1975). Isolation of nuclear pore complexes in association with a lamina. Proc. Natl. Acad. Sci. *72*, 1007–1011. https://doi.org/10.1073/pnas.72.3.1007.

17. Dwyer, N., and Blobel, G. (1976). A modified procedure for the isolation of a pore complex-lamina fraction from rat liver nuclei. J. Cell Biol. *70*, 581–591. https://doi.org/10.1083/jcb.70.3.581.

18. Xie, W., Chojnowski, A., Boudier, T., Lim, J.S.Y., Ahmed, S., Ser, Z., Stewart, C., and Burke, B. (2016). A-type Lamins Form Distinct Filamentous Networks with Differential Nuclear Pore Complex Associations. Curr. Biol. *26*, 2651–2658. https://doi.org/10.1016/j.cub.2016.07.049.

19. Georgatos, S.D., Weber, K., Geisler, N., and Blobel, G. (1987). Binding of two desmin derivatives to the plasma membrane and the nuclear envelope of avian erythrocytes: evidence for a conserved site-specificity in intermediate filament-membrane interactions. Proc. Natl. Acad. Sci. U. S. A. *84*, 6780–6784.

20. Padmakumar, V.C., Libotte, T., Lu, W., Zaim, H., Abraham, S., Noegel, A.A., Gotzmann, J., Foisner, R., and Karakesisoglou, I. (2005). The inner nuclear membrane protein Sun1 mediates the anchorage of Nesprin-2 to the nuclear envelope. J. Cell Sci. *118*, 3419–3430. https://doi.org/10.1242/jcs.02471.

21. Ho, C.Y., Jaalouk, D.E., Vartiainen, M.K., and Lammerding, J. (2013). Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics. Nature *497*, 507–511. https://doi.org/10.1038/nature12105.

22. Chang, W., Worman, H.J., and Gundersen, G.G. (2015). Accessorizing and anchoring the LINC complex for multifunctionality. J. Cell Biol. *208*, 11–22. https://doi.org/10.1083/jcb.201409047.

23. Osmanagic-Myers, S., Dechat, T., and Foisner, R. (2015). Lamins at the crossroads of mechanosignaling. Genes Dev. *29*, 225–237. https://doi.org/10.1101/gad.255968.114.

24. Kirby, T.J., and Lammerding, J. (2018). Emerging views of the nucleus as a cellular mechanosensor. Nat. Cell Biol. *20*, 373–381. https://doi.org/10.1038/s41556-018-0038-y.

55

25. Burke, B., and Stewart, C.L. (2013). The nuclear lamins: flexibility in function. Nat. Rev. Mol. Cell Biol. *14*, 13–24. https://doi.org/10.1038/nrm3488.

26. Miroshnikova, Y.A., and Wickström, S.A. (2022). Mechanical Forces in Nuclear Organization. Cold Spring Harb. Perspect. Biol. *14,* a039685. https://doi.org/10.1101/cshperspect.a039685.

27. Vergnes, L., Péterfy, M., Bergo, M.O., Young, S.G., and Reue, K. (2004). Lamin B1 is required for mouse development and nuclear integrity. Proc. Natl. Acad. Sci. U. S. A. *101*, 10428–10433. https://doi.org/10.1073/pnas.0401424101.

28. Sullivan, T., Escalante-Alcalde, D., Bhatt, H., Anver, M., Bhat, N., Nagashima, K., Stewart, C.L., and Burke, B. (1999). Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. J. Cell Biol. *147*, 913–920. https://doi.org/10.1083/jcb.147.5.913.

29. Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.-B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and van Steensel, B. (2013). Constitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res. *23*, 270–280. https://doi.org/10.1101/gr.141028.112.

30. Shah, P.P., Keough, K.C., Gjoni, K., Santini, G.T., Abdill, R.J., Wickramasinghe, N.M., Dundes, C.E., Karnay, A., Chen, A., Salomon, R.E.A., et al. (2023). An atlas of lamina-associated chromatin across twelve human cell types reveals an intermediate chromatin subtype. Genome Biol. *24*, 16. https://doi.org/10.1186/s13059-023-02849-5.

31. Borsos, M., Perricone, S.M., Schauer, T., Pontabry, J., Luca, K.L. de, Vries, S.S. de, Ruiz-Morales, E.R., Torres-Padilla, M.-E., and Kind, J. (2019). Genome–lamina interactions are established de novo in the early mouse embryo. Nature *569*, 729–733. https://doi.org/10.1038/s41586-019-1233-0.

32. Bizhanova, A., Yan, A., Yu, J., Zhu, L.J., and Kaufman, P.D. (2020). Distinct features of nucleolus-associated domains in mouse embryonic stem cells. Chromosoma *129*, 121–139. https://doi.org/10.1007/s00412-020-00734-9.

33. Bersaglieri, C., Kresoja-Rakic, J., Gupta, S., Bär, D., Kuzyakiv, R., Panatta, M., and Santoro, R. (2022). Genome-wide maps of nucleolus interactions reveal distinct layers of repressive chromatin domains. Nat. Commun. *13*, 1483. https://doi.org/10.1038/s41467-022-29146-2.

34. Martin, C., Chen, S., Maya-Mendoza, A., Lovric, J., Sims, P.F.G., and Jackson, D.A. (2009). Lamin B1 maintains the functional plasticity of nucleoli. J. Cell Sci. *122*, 1551–1562. https://doi.org/10.1242/jcs.046284.

35. Pochukalina, G.N., Ilicheva, N.V., Podgornaya, O.I., and Voronin, A.P. (2016). Nucleolus-like body of mouse oocytes contains lamin A and B and TRF2 but not actin and topo II. Mol. Cytogenet. *9*, 50. https://doi.org/10.1186/s13039-016-0259-3.

36. Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W.M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M., et al. (2010). Molecular maps of the reorganization of genome – nuclear lamina interactions during differentiation. Mol. Cell *38*, 603–613. https://doi.org/10.1016/j.molcel.2010.03.016.

56

37. Tumbar, T., and Belmont, A.S. (2001). Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. Nat. Cell Biol. *3*, 134–139. https://doi.org/10.1038/35055033.

38. Therizols, P., Illingworth, R.S., Courilleau, C., Boyle, S., Wood, A.J., and Bickmore, W.A. (2014). Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science *346*, 1238–1242. https://doi.org/10.1126/science.1259587.

39. Brueckner, L., Zhao, P.A., van Schaik, T., Leemans, C., Sima, J., Peric-Hupkes, D., Gilbert, D.M., and van Steensel, B. (2020). Local rewiring of genome-nuclear lamina interactions by transcription. EMBO J. *39*, e103159. https://doi.org/10.15252/embj.2019103159.

40. Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. PLOS Genet. *4*, e1000039. https://doi.org/10.1371/journal.pgen.1000039.

41. Jachowicz, J.W., Santenard, A., Bender, A., Muller, J., and Torres-Padilla, M.-E. (2013). Heterochromatin establishment at pericentromeres depends on nuclear position. Genes Dev. *27*, 2427–2432. https://doi.org/10.1101/gad.224550.113.

42. Harr, J.C., Luperchio, T.R., Wong, X., Cohen, E., Wheelan, S.J., and Reddy, K.L. (2015). Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. J. Cell Biol. *208*, 33–52. https://doi.org/10.1083/jcb.201405110.

43. Yokochi, T., Poduch, K., Ryba, T., Lu, J., Hiratani, I., Tachibana, M., Shinkai, Y., and Gilbert, D.M. (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. Proc. Natl. Acad. Sci. *106*, 19363–19368. https://doi.org/10.1073/pnas.0906142106.

44. Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat. Genet. *41*, 246–250. https://doi.org/10.1038/ng.297.

45. Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. Cell *153*, 178–192. https://doi.org/10.1016/j.cell.2013.02.028.

46. Bian, Q., Khanna, N., Alvikas, J., and Belmont, A.S. (2013). β-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. J. Cell Biol. *203*, 767–783. https://doi.org/10.1083/jcb.201305027.

47. Cabianca, D.S., Muñoz-Jiménez, C., Kalck, V., Gaidatzis, D., Padeken, J., Seeber, A., Askjaer, P., and Gasser, S.M. (2019). Active chromatin marks drive spatial sequestration of heterochromatin in C. elegans nuclei. Nature *569*, 734–739. https://doi.org/10.1038/s41586-019-1243-y.

48. Amendola, M., and Steensel, B. van (2015). Nuclear lamins are not required for lamina-associated domain organization in mouse embryonic stem cells. EMBO Rep. *16*, 610. https://doi.org/10.15252/embr.201439789.

49. Pal, M., Altamirano-Pacheco, L., Schauer, T., and Torres-Padilla, M.-E. (2023). Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity. Genes Dev. https://doi.org/10.1101/gad.350799.123.

50. Guerreiro, I., Rang, F.J., Kawamura, Y.K., Kroon-Veenboer, C., Korving, J., Groenveld, F.C., van Beek, R.E., Lochs, S.J.A., Boele, E., Peters, A.H.M.F., et al. (2024). Antagonism between H3K27me3 and genome–lamina association drives atypical spatial genome organization in the totipotent embryo. Nat. Genet., 1–10. https://doi.org/10.1038/s41588-024-01902-8.

51. Kind, J., Pagie, L., de Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C.A., Amendola, M., et al. (2015). Genome-wide maps of nuclear lamina interactions in single human cells. Cell *163*, 134–147. https://doi.org/10.1016/j.cell.2015.08.040.

52. Pal, M., Kind, J., and Torres-Padilla, M.-E. (2021). DamID to Map Genome-Protein Interactions in Preimplantation Mouse Embryos. In Epigenetic Reprogramming During Mouse Embryogenesis: Methods and Protocols Methods in Molecular Biology., K. Ancelin and M. Borensztein, eds. (Springer US), pp. 265–282. https://doi.org/10.1007/978-1-0716-0958-3_18.

53. Chaigne, A., Campillo, C., Voituriez, R., Gov, N.S., Sykes, C., Verlhac, M.-H., and Terret, M.-E. (2016). F-actin mechanics control spindle centring in the mouse zygote. Nat. Commun. *7*, 10253. https://doi.org/10.1038/ncomms10253.

54. Baarlink, C., Plessner, M., Sherrard, A., Morita, K., Misu, S., Virant, D., Kleinschnitz, E.-M., Harniman, R., Alibhai, D., Baumeister, S., et al. (2017). A transient pool of nuclear F-actin at mitotic exit controls chromatin organization. Nat. Cell Biol. *19*, 1389–1399. https://doi.org/10.1038/ncb3641.

55. Okuno, T., Li, W.Y., Hatano, Y., Takasu, A., Sakamoto, Y., Yamamoto, M., Ikeda, Z., Shindo, T., Plessner, M., Morita, K., et al. (2020). Zygotic Nuclear F-Actin Safeguards Embryonic Development. Cell Rep. *31*, 107824. https://doi.org/10.1016/j.celrep.2020.107824.

56. Burton, A., Brochard, V., Galan, C., Ruiz-Morales, E.R., Rovira, Q., Rodriguez-Terrones, D., Kruse, K., Le Gras, S., Udayakumar, V.S., Chin, H.G., et al. (2020). Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. Nat. Cell Biol. *22*, 767–778. https://doi.org/10.1038/s41556-020-0536-6.

57. Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I.S. (2002). Establishment and Maintenance of a Heterochromatin Domain. Science *297*, 2232–2237. https://doi.org/10.1126/science.1076466.

58. Obersriebnig, M.J., Pallesen, E.M.H., Sneppen, K., Trusina, A., and Thon, G. (2016). Nucleation and spreading of a heterochromatic domain in fission yeast. Nat. Commun. 7, 11518. https://doi.org/10.1038/ncomms11518.

59. Wang, C., Liu, X., Gao, Y., Yang, L., Li, C., Liu, W., Chen, C., Kou, X., Zhao, Y., Chen, J., et al. (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian

embryo development. Nat. Cell Biol. *20*, 620–631. https://doi.org/10.1038/s41556-018-0093-4.

60. Inoue, A., Jiang, L., Lu, F., Suzuki, T., and Zhang, Y. (2017). Maternal H3K27me3 controls DNA methylation-independent imprinting. Nature *547*, 419–424. https://doi.org/10.1038/nature23262.

61. Verma, S.K., Tian, X., LaFrance, L.V., Duquenne, C., Suarez, D.P., Newlander, K.A., Romeril, S.P., Burgess, J.L., Grant, S.W., Brackley, J.A., et al. (2012). Identification of Potent, Selective, Cell-Active Inhibitors of the Histone Lysine Methyltransferase EZH2. ACS Med. Chem. Lett. *3*, 1091–1096. https://doi.org/10.1021/ml3003346.

62. Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016). Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. Mol. Cell *63*, 1066–1079. https://doi.org/10.1016/j.molcel.2016.08.032.

63. Zeng, F., and Schultz, R.M. (2005). RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. Dev. Biol. *283*, 40–57. https://doi.org/10.1016/j.ydbio.2005.03.038.

64. Abe, K., Funaya, S., Tsukioka, D., Kawamura, M., Suzuki, Y., Suzuki, M.G., Schultz, R.M., and Aoki, F. (2018). Minor zygotic gene activation is essential for mouse preimplantation development. Proc. Natl. Acad. Sci. *115*, E6780–E6788. https://doi.org/10.1073/pnas.1804309115.

65. Schulz, K.N., and Harrison, M.M. (2019). Mechanisms regulating zygotic genome activation. Nat. Rev. Genet. *20*, 221–234. https://doi.org/10.1038/s41576-018-0087-x.

66. Siegenfeld, A.P., Roseman, S.A., Roh, H., Lue, N.Z., Wagen, C.C., Zhou, E., Johnstone, S.E., Aryee, M.J., and Liau, B.B. (2022). Polycomb-lamina antagonism partitions heterochromatin at the nuclear periphery. Nat. Commun. *13*, 4199. https://doi.org/10.1038/s41467-022-31857-5.

67. Sankar, A., Lerdrup, M., Manaf, A., Johansen, J.V., Gonzalez, J.M., Borup, R., Blanshard, R., Klungland, A., Hansen, K., Andersen, C.Y., et al. (2020). KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3 domains from H3K9me3 invasion in oocytes. Nat. Cell Biol. *22*, 380–388. https://doi.org/10.1038/s41556-020-0494-z.

68. Manzo, S.G., Mazouzi, A., Leemans, C., van Schaik, T., Neyazi, N., van Ruiten, M.S., Rowland, B.D., Brummelkamp, T.R., and van Steensel, B. (2024). Chromatin protein complexes involved in gene repression in lamina-associated domains. EMBO J., 1–28. https://doi.org/10.1038/s44318-024-00214-1.

69. Kind, J., and van Steensel, B. (2014). Stochastic genome-nuclear lamina interactions. Nucleus *5*, 124–130. https://doi.org/10.4161/nucl.28825.

70. Boumendil, C., Hari, P., Olsen, K.C.F., Acosta, J.C., and Bickmore, W.A. (2019). Nuclear pore density controls heterochromatin reorganization during senescence. Genes Dev. *33*, 144–149. https://doi.org/10.1101/gad.321117.118.

71. MacPherson, Q., Beltran, B., and Spakowitz, A.J. (2020). Chromatin Compaction Leads to a Preference for Peripheral Heterochromatin. Biophys. J. *118*, 1479–1488. https://doi.org/10.1016/j.bpj.2020.01.034.

72. Thrower, D.A., and Bloom, K. (2001). Dicentric chromosome stretching during anaphase reveals roles of Sir2/Ku in chromatin compaction in budding yeast. Mol. Biol. Cell *12*, 2800–2812. https://doi.org/10.1091/mbc.12.9.2800.

73. Gunjan, A., Alexander, B.T., Sittman, D.B., and Brown, D.T. (1999). Effects of H1 Histone Variant Overexpression on Chromatin Structure *. J. Biol. Chem. *274*, 37950–37956. https://doi.org/10.1074/jbc.274.53.37950.

74. Prendergast, L., and Reinberg, D. (2021). The missing linker: emerging trends for H1 variant-specific functions. Genes Dev. *35*, 40–58. https://doi.org/10.1101/gad.344531.120.

75. Chakravarthy, S., Patel, A., and Bowman, G.D. (2012). The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome. Nucleic Acids Res. *40*, 8285–8295. https://doi.org/10.1093/nar/gks645.

76. Douet, J., Corujo, D., Malinverni, R., Renauld, J., Sansoni, V., Posavec Marjanović, M., Cantariño, N., Valero, V., Mongelard, F., Bouvet, P., et al. (2017). MacroH2A histone variants maintain nuclear organization and heterochromatin architecture. J. Cell Sci. *130*, 1570–1582. https://doi.org/10.1242/jcs.199216.

77. Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y., et al. (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. Nature *547*, 232–235. https://doi.org/10.1038/nature23263.

78. Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., Gao, L., et al. (2017). 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. Cell *170*, 367-381.e20. https://doi.org/10.1016/j.cell.2017.06.029.

79. Flyamer, I.M., Gassler, J., Imakaev, M., Brandão, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110–114. https://doi.org/10.1038/nature21711.

80. Collombet, S., Ranisavljevic, N., Nagano, T., Varnai, C., Shisode, T., Leung, W., Piolot, T., Galupa, R., Borensztein, M., Servant, N., et al. (2020). Parental-to-embryo switch of chromosome organization in early embryogenesis. Nature *580*, 142–146. https://doi.org/10.1038/s41586-020-2125-z.

81. Burton, A., and Torres-Padilla, M.-E. (2014). Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. Nat. Rev. Mol. Cell Biol. *15*, 723–735. https://doi.org/10.1038/nrm3885.

82. Xia, W., and Xie, W. (2020). Rebooting the Epigenomes during Mammalian Early Embryogenesis. Stem Cell Rep. *15*, 1158–1175. https://doi.org/10.1016/j.stemcr.2020.09.005.

60

83. Nakatani, T., Schauer, T., Altamirano-Pacheco, L., Klein, K.N., Ettinger, A., Pal, M., Gilbert, D.M., and Torres-Padilla, M.-E. (2024). Emergence of replication timing during early mammalian development. Nature *625*, 401–409. https://doi.org/10.1038/s41586-023-06872-1.

84. Chandra, T., Kirschner, K., Thuret, J.-Y., Pope, B.D., Ryba, T., Newman, S., Ahmed, K., Samarajiwa, S.A., Salama, R., Carroll, T., et al. (2012). Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation. Mol. Cell *47*, 203–214. https://doi.org/10.1016/j.molcel.2012.06.010.

85. Sati, S., Bonev, B., Szabo, Q., Jost, D., Bensadoun, P., Serra, F., Loubiere, V., Papadopoulos, G.L., Rivera-Mulia, J.-C., Fritsch, L., et al. (2020). 4D Genome Rewiring during Oncogene-Induced and Replicative Senescence. Mol. Cell *78*, 522-538.e9. https://doi.org/10.1016/j.molcel.2020.03.007.

86. Lenain, C., de Graaf, C.A., Pagie, L., Visser, N.L., de Haas, M., de Vries, S.S., Peric-Hupkes, D., van Steensel, B., and Peeper, D.S. (2017). Massive reshaping of genome-nuclear lamina interactions during oncogene-induced senescence. Genome Res. *27*, 1634–1644. https://doi.org/10.1101/gr.225763.117.

87. Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature *445*, 214–218. https://doi.org/10.1038/nature05458.

88. Burton, A., Muller, J., Tu, S., Padilla-Longoria, P., Guccione, E., and Torres-Padilla, M.-E. (2013). Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. Cell Rep. *5*, 687–701. https://doi.org/10.1016/j.celrep.2013.09.044.

89. Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells. Cell *143*, 212–224. https://doi.org/10.1016/j.cell.2010.09.009.

90. Oomen, M.E., Rodriguez-Terrones, D., Simmet, K., Zakhartchenko, V., Mottes, L., Kurome, M., Noll, C., Nakatani, T., Mourra-Diaz, C.M., Aksoy, I., et al. Article in preparation: Mapping transcription initiation across mammalian species using Smart-seq+5' reveals regulatory principles of embryonic genome activation.

91. Park, S.-J., Shirahige, K., Ohsugi, M., and Nakai, K. (2015). DBTMEE: a database of transcriptome in mouse early embryos. Nucleic Acids Res. *43*, D771–D776. https://doi.org/10.1093/nar/gku1001.

92. Skene, P.J., Henikoff, J.G., and Henikoff, S. (2018). Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat. Protoc. *13*, 1006–1019. https://doi.org/10.1038/nprot.2018.015.

93. Hainer, S.J., and Fazzio, T.G. (2019). High-Resolution Chromatin Profiling Using CUT&RUN. Curr. Protoc. Mol. Biol. *126*, e85. https://doi.org/10.1002/cpmb.85.

94. Kaya-Okur, H.S., Janssens, D.H., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2020). Efficient low-cost chromatin profiling with CUT&Tag. Nat. Protoc. *15*, 3264–3283. https://doi.org/10.1038/s41596-020-0373-x.

95. Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., Xia, W., et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652–657. https://doi.org/10.1038/nature18606.

96. Wu, J., Xu, J., Liu, B., Yao, G., Wang, P., Lin, Z., Huang, B., Wang, X., Li, T., Shi, S., et al. (2018). Chromatin analysis in human early development reveals epigenetic transition during ZGA. Nature *557*, 256–260. https://doi.org/10.1038/s41586-018-0080-8.

97. Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature *537*, 553–557. https://doi.org/10.1038/nature19361.

98. Xu, Q., Xiang, Y., Wang, Q., Wang, L., Brind'Amour, J., Bogutz, A.B., Zhang, Y., Zhang, B., Yu, G., Xia, W., et al. (2019). SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. Nat. Genet. *51*, 844–856. https://doi.org/10.1038/s41588-019-0398-7.

99. Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A.Y., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature *537*, 548–552. https://doi.org/10.1038/nature19360.

100. Liu, B., Xu, Q., Wang, Q., Feng, S., Lai, F., Wang, P., Zheng, F., Xiang, Y., Wu, J., Nie, J., et al. (2020). The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. Nature *587*, 139–144. https://doi.org/10.1038/s41586-020-2847-y.

101. Lu, F., Liu, Y., Inoue, A., Suzuki, T., Zhao, K., and Zhang, Y. (2016). Establishing Chromatin Regulatory Landscape during Mouse Preimplantation Development. Cell *165*, 1375–1388. https://doi.org/10.1016/j.cell.2016.05.050.

102. Miyanari, Y., Ziegler-Birling, C., and Torres-Padilla, M.-E. (2013). Live visualization of chromatin dynamics with fluorescent TALEs. Nat. Struct. Mol. Biol. *20*, 1321–1324. https://doi.org/10.1038/nsmb.2680.

103. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. Nat. Methods *16*, 1226–1232. https://doi.org/10.1038/s41592-019-0582-9.

104. Haraguchi, T., Koujin, T., Segura-Totten, M., Lee, K.K., Matsuoka, Y., Yoneda, Y., Wilson, K.L., and Hiraoka, Y. (2001). BAF is required for emerin assembly into the reforming nuclear envelope. J. Cell Sci. *114*, 4575–4585. https://doi.org/10.1242/jcs.114.24.4575.

105. Yang, M., and Yuan, Z.-M. (2015). A novel role of PRR14 in the regulation of skeletal myogenesis. Cell Death Dis. *6*, e1734–e1734. https://doi.org/10.1038/cddis.2015.103.

106. Lombardi, M.L., Jaalouk, D.E., Shanahan, C.M., Burke, B., Roux, K.J., and Lammerding, J. (2011). The Interaction between Nesprins and Sun Proteins at the Nuclear Envelope Is

Critical for Force Transmission between the Nucleus and Cytoskeleton *. J. Biol. Chem. *286*, 26743–26753. https://doi.org/10.1074/jbc.M111.233700.

107. Liang, Y., Franks, T.M., Marchetto, M.C., Gage, F.H., and Hetzer, M.W. (2013). Dynamic Association of NUP98 with the Human Genome. PLOS Genet. *9*, e1003308. https://doi.org/10.1371/journal.pgen.1003308.

108. Ye, J., Zhao, J., Hoffmann-Rohrer, U., and Grummt, I. (2008). Nuclear myosin I acts in concert with polymeric actin to drive RNA polymerase I transcription. Genes Dev. *22*, 322–330. https://doi.org/10.1101/gad.455908.

109. Provance, D.W., Addison, E.J., Wood, P.R., Chen, D.Z., Silan, C.M., and Mercer, J.A. (2008). Myosin-Vb functions as a dynamic tether for peripheral endocytic compartments during transferrin trafficking. BMC Cell Biol. *9*, 44. https://doi.org/10.1186/1471-2121-9-44.

110. Posern, G., Sotiropoulos, A., and Treisman, R. (2002). Mutant Actins Demonstrate a Role for Unpolymerized Actin in Control of Transcription by Serum Response Factor. Mol. Biol. Cell *13*, 4167–4178. https://doi.org/10.1091/mbc.02-05-0068.

111. Ramsköld, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat. Biotechnol. *30*, 777–782. https://doi.org/10.1038/nbt.2282.

112. Deng, Q., Ramsköld, D., Reinius, B., and Sandberg, R. (2014). Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. Science *343*, 193–196. https://doi.org/10.1126/science.1245316.

Discussion

During PhD work, my primary goal was to interrogate the molecular mechanisms underlying the establishment of the epigenome in early mouse embryos, with a particular focus on genome-lamina association. Additionally, across multiple collaborative projects, I investigated the interplay of 3D nuclear organization with embryonic transcription and DNA replication at several stages of mammalian preimplantation development. Our work has generated a catalogue of perturbations in embryonic chromatin and uncovered fundamental insights into the complex interdependencies between the chromatin landscape, the radial organization of the genome and various DNA-related processes. Below, I discuss the key findings from my PhD thesis in detail and highlight the potential implications and future research directions that arise from our work.

Part I: DamID to map genome-protein interactions in preimplantation mouse embryos

This methods chapter by Pal et al. 2021 focuses on optimizing a protocol for using low-input DNA adenine methyltransferase identification (DamID) to map genome-protein interactions in early-stage mouse embryos. DamID does not require specific antibodies, making it a versatile tool for mapping genome-protein interactions at the single-cell level (Steensel and Henikoff 2000; Kind et al. 2015; Borsos et al. 2019). This study outlines the steps required for embryo manipulation, DNA methylation analysis, and library preparation for sequencing. Although the protocol described here is optimized for LaminB1 DamID in early mouse embryos, it can be adapted to other fusion proteins or to different stages of germline development, such as growing oocytes (*see* **Part II** below; Liu et al. 2024) . This flexibility opens up possibilities for investigating the dynamics of nuclear organization and chromatin behaviour during early mammalian development. For instance, this low-input DamID technique could be employed to study how embryonic chromatin interacts with other nuclear compartments such as the nuclear pore complex or with specific transcription factors, as previously done in other cell types, albeit using bulk samples (Jacinto et al. 2015; Cheetham et al. 2018; Tosti et al. 2018; Tyagi et al. 2023).

Part II: Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators

H3K27ac in gene deserts mark oocyte-specific putative enhancers

Liu et al. 2024 revealed that fully grown oocytes (FGOs) possess numerous unique enhancers in gene deserts, regions typically lacking genes. These enhancers are linked to oocyte-specific genes and transposable element activation, indicating a unique epigenetic regulatory landscape in oocytes compared to somatic cells. In contrast to earlier studies, which suggested that enhancers are inactive in oocytes and zygotes (Majumder et al. 1997; Lawinger et al. 1999), this work demonstrated that the putative enhancers in these cells are bidirectionally transcribed and can drive reporter activity. This was evidenced by the prevalent H3K27ac marks in gene deserts in FGOs, indicating active enhancer regions. Motif analysis of these putative enhancers led to the identification of TCF3 and TCF12 as key activators of oocyte genes and folliculogenesis.

Intriguingly, in mouse oocytes and preimplantation embryos, many enhancers are also marked by H3K4me3, which is a typical promoter marker of differentiated and pluripotent stem cells (Shilatifard 2012). H3K4 demethylases are reported to convert H3K4me3 to H3K4me1 at enhancers to prevent overactivation in mouse ESCs (Kidder et al. 2014; Shen et al. 2016). Therefore, the low expression of KDM5A/B/C enzymes in fully grown oocytes may contribute to the presence of H3K4me3 on enhancers (Shao et al. 2014). Additionally, global DNA hypomethylation in mammalian oocytes and early embryos (Wang et al. 2014) might explain the presence of H3K4me3 at putative enhancers, consistent with the H3K4me1-H3K4me3 seesaw model previously proposed (Sharifi-Zarchi et al. 2017), in which DNA methylation levels differentiate enhancers from promoters. Further research is needed to understand the role of H3K4me3 at enhancers in oocytes and early embryos.

Growing oocytes lack detectable LADs

Since gene deserts are usually heterochromatic regions associated with the nuclear lamina (NL), we were intrigued to investigate the radial organization of the genome

during oocyte growth and maturation. We had previously reported the absence of detectable LADs in FGOs and shown that such structures are established *de novo* after fertilization (Borsos et al. 2019). In the current study, we performed LaminB1 DamID (*see* **Part I** above; Pal et al. 2021) in growing oocytes at postnatal day 10 (GO-P10) and found that LADs are already undetectable at this stage of growing oocytes across all autosomes. This implies that during oocyte growth and maturation when gene desert regions start acquiring the H3K27ac mark, those regions have already lost their NL association.

These findings underscore a unique regulatory interplay between chromatin organization and gene regulation during germ cell development. The absence of LADs in growing oocytes (GOs) and FGOs may facilitate a transcription-permissive environment crucial for the expression of genes essential for oocyte maturation and folliculogenesis. Several open questions arise from this work: Is the loss of LADs causal for the activation of oocyte-specific genes? When and how are LADs lost in autosomes in the female germline? Future research should aim to unravel the mechanisms driving the loss of LADs in oocytes, investigate the precise timing of these events, and determine their direct role in gene activation.

Part III: Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity

LAD reorganization during MZT is gradual and dynamic

Pal et al. 2023 investigated the temporal reorganization of LADs during maternal-to-zygotic transition (MZT). We showed that LADs at the 2-cell stage mature gradually during the complete cell cycle with evolving molecular and genomic features suggesting a dynamic rearrangement of genome-lamina contacts during interphase progression. Although in cultured human cells it was recently reported that LADs evolve during the cell cycle (van Schaik et al. 2020), the changes in molecular features of LADs following the mitosis of the mouse zygote represent distinct biology of the early embryos. This dynamic repositioning of LADs and iLADs during the 2-cell stage

correlates with the transcriptional activity of genes and transposable elements contained therein. Specifically, the enrichment of MERVL elements outside the LAD boundaries underlies a unique feature of the early 2-cell iLADs. Since MERVL elements are highly and transiently expressed at the early 2-cell stage (Ishiuchi et al. 2015; Kruse et al. 2019; Liu et al. 2020; Sakashita et al. 2023), the repositioning of MERVL-containing LADs into iLADs at this stage aligns with their transcriptional activation just before major ZGA.

Inhibition of ZGA leads to atypical features of lamina-associated chromatin

Next, we interrogated the role of RNA polymerase II-mediated transcription at zygotic genome activation (ZGA) in regulating such rearrangement of the nuclear organization during MZT. The use of transcriptional inhibitors, DRB and α-amanitin, resulted in globally altered genome-lamina interactions in 2-cell embryos, causing major zygotic genome activation genes to relocate to the nuclear lamina. Interestingly, Pol II inhibition in early embryos leads to the rearrangement of H3K4me3 and, perhaps more globally, of regions marked by active histone modifications towards the nuclear periphery. This finding is particularly intriguing given that most previous work has identified interactions between repressive chromatin states and the nuclear periphery (Guelen et al. 2008; Harr et al. 2015). However, since broad H3K4me3 domains in early embryos have been proposed to be repressive (Dahl et al. 2016; Zhang et al. 2016), these data raise the possibility that non-canonical H3K4me3 could link genomic regions to the nuclear periphery. Nevertheless, it remains to be explored how, under conditions of ZGA inhibition, chromatin marked with active modifications becomes anchored to the nuclear periphery.

Differential impact of Pol II inhibition on TADs/compartments versus LADs

Interestingly, transcriptional inhibition does not affect the consolidation of TADs in mouse embryos (Du et al. 2017; Ke et al. 2017). Therefore, our results showing the complete reorganization of LAD boundaries in 2-cell embryos upon Pol II inhibition indicate that the contributions of zygotic genome activation to different pillars of nuclear organization may vary. Although compartment scores remain largely unchanged upon α-amanitin treatment in 2-cell embryos (Du et al. 2017; Ke et al.

2017), we demonstrated that A compartment regions gain lamina interactions. These findings challenge the current understanding of the interplay between nuclear organization and the onset of embryonic transcription in mouse embryos and indicate a more complex relationship than previously speculated based on data from somatic cells (also *see* **Part IV** below; Nakatani et al. 2024).

In summary, our findings suggest that dynamic LAD rearrangement is an integral component of the maternal-to-zygotic transition. Furthermore, the naturally evolving molecular characteristics of embryonic LADs are dependent on RNA polymerase II activity during ZGA. However, it remains to be determined whether complete transcriptional inhibition in cultured somatic cells also affects LADs to the same degree as in the early embryos. Lastly, understanding additional molecular pathways that govern this dynamic rearrangement of genome-lamina interaction in the early embryos will require more experimental efforts (*see* **Part VI** below; Pal et al. under review).

Part IV: Emergence of replication timing during early mammalian development

Gradual consolidation of RT program in preimplantation embryos

Using single-cell Repli-seq, Nakatani et al. 2024 investigated the emergence and progression of the replication timing (RT) program during preimplantation development in mouse embryos. We showed that early stages, such as zygotes and 2-cell embryos, exhibit a less defined RT pattern with high variability, suggesting a less coordinated replication program. As development advances beyond the 4-cell stage, the RT profile becomes progressively more defined, with a clearer separation into early and late replication domains.

Limited role of chromatin marks on RT consolidation

We also investigated if and how histone modifications could influence the consolidation of replication timing. Following the 2-cell stage, as RT becomes more defined with the onset of zygotic gene activation (ZGA), specific changes in chromatin

marks were noted. H3K36me3, which usually marks transcriptional elongation, became enriched at RT peaks from the 8-cell stage onwards, suggesting its role in defining RT peaks. However, H3K4me3 levels were stable but slightly higher at RT peaks compared to troughs. Expression of KDM5B, which globally removes H3K4me3, did not significantly alter RT profiles or affect the replication timing of major ZGA genes. Our results indicate that while histone modifications are associated with RT features, their direct impact on RT consolidation is limited.

RNA Pol II activity during ZGA contributes to the precision of the RT program

Treatment with α-amanitin, which inhibits transcription and causes degradation of total RNA polymerase II (Nguyen et al. 1996; Bensaude 2011; Liu et al. 2020), led to moderate RT changes. This included delaying DNA replication in genomic regions associated with major ZGA genes and causing an overall fragmented RT pattern. These findings highlight that global transcription driven by RNA Pol II is crucial for fine-tuning RT initiation and termination sites in early embryos. However, treatment with DRB, which specifically inhibits CDK9 activity and thereby transcriptional elongation (Dubois et al. 1994; Liu et al. 2020), resulted in milder RT alterations and had little impact on ZGA genes. While these two inhibitors, α-amanitin and DRB, have globally similar phenotypes regarding LAD remodelling (*see* **Part III** above; Pal et al. 2023), it is indeed striking that the extent and effects on RT changes differ. Taken together, our results would suggest that RNA Pol II's role is more significant than the act of transcriptional elongation itself in defining RT precision. However, further investigation is required to explore other potential contributing factors.

Organization into LAD/iLAD precedes partitioning of early and late replication

We also investigated the dependency between nuclear architecture and the establishment of the RT program. We found that A compartments exhibited earlier RT profiles across all the stages compared to B compartments. Although the distinction between early and late RT values was less pronounced in zygotes, it became more defined as development progressed. Inhibition of zygotic genome activation (ZGA) with α-amanitin nullified RT differences between A and B compartments while preserving the compartment scores (Du et al. 2017; Ke et al. 2017), indicating that

ZGA affects RT but not the compartment structure itself. Additionally, even though LADs are established immediately after fertilization, replication time segregation between LADs and iLADs becomes apparent only at later developmental stages. Therefore, our work concludes that the organization of the genome into LADs and iLADs precedes and possibly influences the partitioning of early and late replication dynamics, highlighting the temporal relationship between nuclear architecture and RT establishment.

As the chromatin undergoes extensive remodelling following fertilization, we speculate that the heterogeneous RT program in zygotes and 2-cell embryos might stem from greater plasticity in chromatin structure during these stages. However, identifying the molecular regulators of RT consolidation *in vivo* during development remains a crucial theme for further investigation (*see* **Part V** below; Nakatani et al. under review).

Following our publication, three other groups have investigated replication timing establishment in early mouse embryos. Takahashi and colleagues (Takahashi et al. 2024) observed a clear replication timing profile beginning at the 4-cell stage, with no defined pattern at the zygote or 2-cell stage. They noted a strong correlation between late replication and the B compartment, suggesting that replication timing aligns with nuclear compartmentalization. Halliwell and colleagues (Halliwell et al. 2024) found that a replication timing program begins at the 2-cell stage, with no detectable patterns at the zygotic stage. They examined parental differences in replication timing, showing that late-replicating regions in both parental genomes were associated with LADs. Early replication in the maternal genome correlated with H3K27me3, while no such correlation was observed in the paternal genome. Finally, Xu and colleagues (Xu et al. 2024) found that DNA replication timing patterns are clearly defined by the zygote stage. Late-replicating regions correlated with LADs and the B compartment, while early-replicating regions aligned with the A compartment. These patterns were evident at the zygotic stage in both maternal and paternal genomes. Collectively, these four studies on mouse embryos differed slightly in how early RT patterns were observed, ranging from the zygote to the 4-cell stage. The differences between the four studies might stem from differences in sample collection timing or other technical factors. Nevertheless, a clear consensus has emerged on the correlation of late replication with nuclear organization.

Part V: RIF1 regulates the consolidation of replication timing in early mouse embryos independently of changes in nuclear organization towards the nuclear lamina

RIF1 depletion results in a less coordinated RT program

In this manuscript (Nakatani et al. under review), we generated genome-wide replication time (RT) profiles of RIF1-depleted 4-cell, 8-cell, and morula stage embryos. Our analysis reveals that embryos depleted of RIF1 display a less defined replication pattern, particularly at the 8-cell and morula stages, indicating a less coordinated RT program post-4-cell stage. Further examination of replication features indicated that the developmental consolidation of RT is disrupted as RIF1-depleted embryos retained a heterogeneous and less well-segregated replication timing profile. Additionally, replication fork speed was slower in RIF1-depleted embryos, suggesting more origins of replication fire upon loss of RIF1. Overall, the findings suggest that RIF1 depleted preimplantation embryos show DNA replication features characteristic of a more totipotent-like state (*see* **Part IV** above; Nakatani et al. 2024). This is particularly notable, as knockdown studies (Li et al. 2017; Rodriguez-Terrones et al. 2018) in mouse ESCs also showed that RIF1 depletion promotes efficient reprogramming to totipotent-like 2-cell-like cells (2CLCs), which exhibit a similarly slow fork speed (Nakatani et al. 2022).

Lamina association and RT changes are uncoupled upon RIF1 depletion

Mapping LADs in RIF1-depleted embryos using LaminB1 DamID (*see* **Part I** above; Pal et al. 2021) revealed altered genome-lamina interactions and changes in LAD/iLAD boundaries at both the 4-cell and 8-cell stages. However, comparing global RT differences between control and RIF1-depleted embryos against differences in lamina association indicated no correlation between changes in RT and nuclear positioning. This observation was consistent at both the 4-cell and 8-cell stages, suggesting that RIF1's regulation of RT is independent of changes in lamina interactions. Interestingly, such changes in genome-lamina interactions were associated with minimal changes in gene expression in RIF1-depleted embryos. These findings are particularly important as they add to the observations that the two

pillars of the epigenome - the radial positioning towards the nuclear lamina and RT - can be disentangled from each other and from gene transcription.

Instead, the data indicated that the organization of chromatin into A and B compartments had a more consistent relationship with RT changes upon RIF1 depletion. Genomic regions that shift towards earlier RT had a strong B compartment score while those shifting towards later RT had a strong A compartment score. This would suggest chromatin compartments to be a stronger determinant for RT regulation than lamina association during early development.

Previous studies show that RIF1 regulates the replication timing of heterochromatin following zygotic genome activation (ZGA) in *Drosophila* (Seller and O'Farrell 2018). In *zebrafish* embryos, RIF1 loss primarily affected DNA replication timing post-gastrulation (Masser et al. 2023). Extending this, our work identifies RIF1 as a key regulator of the progressive consolidation of RT at the beginning of mammalian development. Mouse oocytes and 2-cell embryos lack functional RIF1 protein, and full-length protein is detected only at a later developmental stage (Yoshizawa-Sugata et al. 2021). Therefore, it is tempting to speculate that the absence of RIF1 is the reason for a less consolidated RT program. Whether the expression of full-length RIF1 in the early embryos can rescue the heterogeneous replication timing remains to be investigated.

Part VI: The establishment of nuclear organization in mouse embryos is orchestrated by multiple epigenetic pathways

In this study (Pal et al. under review), I performed a perturbation screen to identify molecular pathways that regulate the establishment of LADs in early mouse embryos, using low-input LaminB1 DamID as a readout (*see* **Part I** above; Pal et al. 2021). Our findings reveal various chromatin modifications that direct genome-lamina interactions after fertilization and throughout the maternal-to-zygotic transition. This work not only offers an unprecedented resource for the molecular understanding of nuclear organization, but also provides critical insights into the chromatin-based dependencies

of epigenome establishment. The main findings from this manuscript are discussed below.

Nuclear actin dynamics fine-tunes the strength of genome-lamina interactions

Interestingly, targeting chromatin anchors, components of the nuclear pore complex, and the nuclear/cortical cytoskeleton did not significantly perturb LAD establishment or rearrangement in early embryos. Consistently, a recent genome-wide screen in human cells identified only a few structural proteins as LAD regulators (Manzo et al. 2024). In the case of BAF, our findings are in line with work in human epithelial cells, in which BAF knockdown does not affect genome-nuclear lamina interactions (Kind and van Steensel 2014). On the other hand, while TPR has been proposed to repel heterochromatin from the nuclear periphery in human fibroblasts (Boumendil et al. 2019), we did not observe altered nuclear organization in early embryos upon TPR expression. These observations highlight potential differences or similarities between LAD establishment in embryos and LAD maintenance in somatic cells. Intriguingly, when we perturbed the nuclear pool of F-actin, we observed an increase in lamina association of wildtype LAD regions. In 2-cell embryos, targeting nuclear acto-myosin led to broader LADs, with B-compartment regions gaining lamina association. Our findings suggest that actomyosin mediated molecular forces play a role in fine-tuning genome-lamina interactions during early development, influencing the strength of these interactions or possibly contributing to occasional detachment of genomic regions within broad inactive compartments.

The lack of a constitutive heterochromatin pathway enables establishment of a unique nuclear organization

We find the zygotic LADs to be labile and susceptible to changes in heterochromatin-associated chromatin modifications. For example, while global depletion of H3K9me3 did not affect LAD establishment (Borsos et al. 2019), removing as well as depositing H3K9me2 in zygotic chromatin leads to the collapse of wildtype LAD structure in zygotes. HP1 proteins, which are important for the establishment of global nuclear organization during embryonic development in *Drosophila* (Zenk et al. 2021), are lowly expressed in early mouse embryos (Deng et al. 2014; Leonard et al. 2015; Gao et al. 2017). Expression of HP1 proteins (HP1α/γ) does not affect LADs in zygote but

consistent with the timing of heterochromatin maturation, in 2-cell stage embryos, remodels lamina interaction of chromatin. Co-expression of HP1s together with Suv39h1 leads to a more severe LAD flattening as lamina-anchoring of broad ectopic H3K9me3 domains is promoted by the HP1 proteins.

Decompaction of chromatin has been shown to be sufficient to relocate specific loci towards the nuclear interior (Therizols et al. 2014). Theoretical simulations suggest that any interaction that densifies chromatin might guide preferential peripheral localization of condensed chromatin (MacPherson et al. 2020). The association of the nuclear lamina of initially less dense, A compartment regions upon expression of histone deacetylases and the subtypes of histone H1/macroH2A could reflect a favoured repositioning due to increased compaction by histone deacetylation (Thrower and Bloom 2001; Vaquero et al. 2004) or the presence of H1 subtypes and macroH2A (Gunjan et al. 1999; Prendergast and Reinberg 2021; Chakravarthy et al. 2012; Douet et al. 2017). Considering this, specifically in early embryos, where the absence of dense heterochromatin at the nuclear periphery is reported (Ahmed et al. 2010), lamina localization of accessibly genomic regions could be envisioned upon enforced heterochromatinization by histone deacetylation or H1 subtype expression. In the 2-cell stage, a similar 'inversion' phenotype becomes apparent upon global transcriptional inhibition which could mediate similar compaction of gene-rich A compartment region (*see* **Part III** above; Pal et al. 2023).

Overall, our observations could explain the unusual and unique LAD fragmentation observed in wildtype 2-cell stage embryos, which we propose is due to the lack of the canonical heterochromatin pathway i.e., low global levels of H3K9me3 and HP1 reader proteins and absence of chromatin-compacting histone variants at these stages.

H3K4me3 contributes to the robustness of nuclear organization in early embryos

Our study proposes a potential mechanism for regulating LAD expansion. We suggest that H3K4me3 restricts the spread of H3K9me3 and its association with the nuclear lamina, thus defining LAD boundaries. This indicates that in early embryos, H3K4me3 domains play a critical role in maintaining stable nuclear organization and counteracting lamina anchoring. Therefore, maternally inherited broad non-canonical

H3K4me3 domains could counteract H3K9me3 spreading and act as a 'stop' signal during heterochromatin formation (Sankar et al. 2020). This builds on findings documenting a role for KDM5B in LAD regulation in zygotes (Borsos et al. 2019) and extends our understanding of the interplay between the chromatin landscape and nuclear organization.

Maternal bookmarking for LAD establishment and PRC2-lamina antagonism

H3K27me3 removal through the action of the demethylases resulted in an expansion of genome-lamina contacts inside the B compartment in 2-cell stage embryos. A similar observation has been made using EZH2 (functional enzymatic component of the PRC2) inhibitor in human leukemia cells (Siegenfeld et al. 2022), in which the authors suggested that H3K27me3 may repel association to the nuclear lamina within B compartments. Accordingly, recent findings in embryos from crosses in which EED (an essential component of PRC2 that deposits H3K27me3) was maternally knocked out indicate an antagonizing role for H3K27me3 in genome-nuclear lamina interactions, specifically in regard to cell-to-cell variability of LADs at the 2-cell stage (Guerreiro et al. 2024). Thus, methylation of H3K27 plays a role in the robustness and the definition of LAD boundaries in early embryos. Interestingly, our work using a chemical inhibitor for EZH2 allowed us to further separate the contribution of inherited versus *de novo* H3K27me3 and suggest that demethylation of inherited H3K27me3 contributes to a most drastic LAD phenotype. Indeed, we find that H3K27me3 is enriched in oocytes just outside future zygotic LAD boundaries, and active demethylation of H3K27 rather than EZH2 inhibition leads to disruption of zygotic LADs. This is particularly interesting considering that growing (*see* **Part II** above; Liu et al. 2024) as well as mature oocytes (Borsos et al. 2019) do not have detectable LADs and thus these results suggest that maternal chromatin would carry a 'programming' mark to reset nuclear organization in embryos.

Compartment boundaries link LAD disruption phenotypes

LADs and B-compartments generally show a strong correlation, with the exception of the 2-cell stage (Borsos et al. 2019), where this relationship weakens. Interestingly, while both compartment scores (Du et al. 2017; Ke et al. 2017) and replication timing (*see* **Part IV** above; Nakatani et al. 2024) remain largely unaffected by transcription

298

inhibition, LAD boundaries are significantly disrupted (*see* **Part III** above; Pal et al. 2023), indicating that LADs and compartments might be regulated by distinct mechanisms in early embryos. Although compartment organization gradually consolidates, we frequently observe compartment boundaries correlating with the LAD phenotypes we describe. Across different perturbations, LAD 'spreading' remains constrained within B compartment regions. Additionally, LAD inversion phenotypes that include increased lamina-association of wildtype A compartment regions tend to respect the A/B compartment boundaries. These observations suggest that the changes with regard to the nuclear lamina remain constrained within A/B compartments, indicating that compartment boundaries may provide a primary 'scaffolding cue' on genome organization at the beginning of development. Whether this correlation depends on genetic or epigenetic features remains to be investigated.

Zygotic LAD establishment is dispensable as embryos can rebuild nuclear organization in the 2-cell stage

Mouse embryos demonstrate a remarkable ability to reset nuclear organization at the 2-cell stage. Although LAD formation can be significantly affected in zygotes through various pathways, 2-cell stage embryos often recover their normal LAD structure. For instance, despite global H3K27me3 depletion causing a collapse of zygotic LADs, only minor effects on nuclear organization were observed at the 2-cell stage. Moreover, these embryos can develop to the blastocyst stage at a rate comparable to that of control embryos. This suggests that zygotic LAD establishment is dispensable for subsequent preimplantation development, as zygotes exhibit highly adaptive nuclear organization. The *de novo* reconstruction of LAD architecture in 2-cell embryos is likely influenced by chromatin remodeling or transcription during major ZGA. The absence of a defined replication timing program (*see* **Part IV** above; Nakatani et al. 2024) attenuated DNA damage response (Kermi et al. 2019), and lack of regulated long-range transcriptional control (Hamamoto et al. 2014; Abe et al. 2015; Aoki 2022) in zygotes may render the radial positioning of the genome inessential.

2-cell LAD disruption is associated with impaired preimplantation development

Transcriptome analysis indicates that the extent of changes in gene expression upon LAD perturbation relates to the intrinsic properties of the genomic regions that relocate

within the nuclear space, such as gene density. We observe that changes in transcript abundance correlate with their positioning relative to the nuclear lamina-regions gaining interactions with the lamina tend to have lower transcript counts on average. However, our data also show that repositioning to or away from the lamina does not necessarily imply changes in gene expression. These findings are consistent with our previous work in embryos, which shows that while a global relationship between transcript levels and inter-LADs emerges at the 2-cell stage, association with the nuclear lamina is not strictly linked to gene silencing, as observed in other models (Kumaran and Spector 2008; Therizols et al. 2014). Transcriptomic changes in LAD disrupted 2-cell embryos indicate an impaired maternal-to-zygotic transition, including accumulation of maternal transcripts and failure to efficiently undergo major zygotic genome activation. Our work further demonstrates that continuous LAD disruption until the 2-cell stage results in significantly decreased developmental competence. However, it remains to be disentangled whether the developmental defects are solely due to LAD disruption or stem from additional effects of chromatin manipulation.

Implications for development and disease biology

Disruption of nuclear lamina components, including lamins, results in a largely unaltered LAD landscape (Amendola and Steensel 2015), suggesting that once interactions with the nuclear lamina are established, LADs are robust. Indeed, reports of global alterations in genome-lamina association have been limited, with exceptions like mouse oocytes (Borsos et al. 2019; Liu et al. 2024; *see* **Part II** above). Murine rod photoreceptors exhibit a unique 'inverted' nuclear organization, with heterochromatin occupying the nucleus center (Solovei et al. 2009, 2013). In senescent cells, H3K9me2/3-enriched regions detach from the lamina to form senescence-associated heterochromatin domains (Chandra et al. 2012; Sati et al. 2020). A global collapse of genome-lamina interactions also occurs in oncogene-induced senescent human fibroblasts, where constitutive LADs lose lamina contacts and aberrant genome-lamina contacts emerge (Lenain et al. 2017). However, the mechanisms underlying such alterations remain poorly understood. Our work in early embryos shows that drastic genome reorganization can occur upon chromatin perturbation and understanding these molecular pathways may shed light on mechanisms of epigenome alteration in the context of development, senescence and disease.

Concluding remarks

Early embryonic development provides a unique model to study epigenome reprogramming following the fertilization of gametes. During my PhD, by performing perturbation experiments, I identified key molecular pathways that orchestrate the establishment and dynamic reprogramming of the epigenome with a particular focus on the radial 3D organization of chromatin. This work not only provides valuable insights into the complex interplay between embryonic chromatin and DNA-related processes but also provides tools for future studies to dissect the function of the epigenome in regulating developmental plasticity and cell fate. Despite disruptions, redundancy and dispensability in molecular effectors underscore the remarkable robustness and resilience of the early developmental program. Future efforts to uncover mechanisms for reprogramming the epigenome in eggs and embryos hold promising applications for advancing reproductive biology and regenerative medicine.

References

Abe K, Yamamoto R, Franke V, Cao M, Suzuki Y, Suzuki MG, Vlahovicek K, Svoboda P, Schultz RM, Aoki F. 2015. The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3′ processing. *EMBO J* **34**: 1523–1537.

Ahmed K, Dehghani H, Rugg-Gunn P, Fussner E, Rossant J, Bazett-Jones DP. 2010. Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo. *PLOS ONE* **5**: e10531.

Amendola M, Steensel B van. 2015. Nuclear lamins are not required for lamina-associated domain organization in mouse embryonic stem cells. *EMBO Rep* **16**: 610.

Aoki F. 2022. Zygotic gene activation in mice: profile and regulation. *J Reprod Dev* **68**: 79–84.

Bensaude O. 2011. Inhibiting eukaryotic transcription. *Transcription* **2**: 103–108.

Borsos M, Perricone SM, Schauer T, Pontabry J, de Luca KL, de Vries SS, Ruiz-Morales ER, Torres-Padilla M-E, Kind J. 2019. Genome–lamina interactions are established de novo in the early mouse embryo. *Nature* **569**: 729–733.

Boumendil C, Hari P, Olsen KCF, Acosta JC, Bickmore WA. 2019. Nuclear pore density controls heterochromatin reorganization during senescence. *Genes Dev* **33**: 144–149.

Chakravarthy S, Patel A, Bowman GD. 2012. The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome. *Nucleic Acids Res* **40**: 8285–8295.

Chandra T, Kirschner K, Thuret J-Y, Pope BD, Ryba T, Newman S, Ahmed K, Samarajiwa SA, Salama R, Carroll T, et al. 2012. Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation. *Mol Cell* **47**: 203–214.

Cheetham SW, Gruhn WH, van den Ameele J, Krautz R, Southall TD, Kobayashi T, Surani MA, Brand AH. 2018. Targeted DamID reveals differential binding of mammalian pluripotency factors. *Dev Camb Engl* **145**: dev170209.

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY, et al. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**: 548–552.

Deng Q, Ramsköld D, Reinius B, Sandberg R. 2014. Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. *Science* **343**: 193–196.

Douet J, Corujo D, Malinverni R, Renauld J, Sansoni V, Posavec Marjanović M, Cantariño N, Valero V, Mongelard F, Bouvet P, et al. 2017. MacroH2A histone

variants maintain nuclear organization and heterochromatin architecture. *J Cell Sci* **130**: 1570–1582.

Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, et al. 2017. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**: 232–235.

Dubois M-F, Bellier S, Seo S-J, Bensaude O. 1994. Phosphorylation of the RNA polymerase II largest subunit during heat shock and inhibition of transcription in hela cells. *J Cell Physiol* **158**: 417–426.

Gao Y, Liu X, Tang B, Li C, Kou Z, Li L, Liu W, Wu Y, Kou X, Li J, et al. 2017. Protein Expression Landscape of Mouse Embryos during Pre-implantation Development. *Cell Rep* **21**: 3957–3969.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948–951.

Guerreiro I, Rang FJ, Kawamura YK, Kroon-Veenboer C, Korving J, Groenveld FC, van Beek RE, Lochs SJA, Boele E, Peters AHMF, et al. 2024. Antagonism between H3K27me3 and genome–lamina association drives atypical spatial genome organization in the totipotent embryo. *Nat Genet* 1–10.

Gunjan A, Alexander BT, Sittman DB, Brown DT. 1999. Effects of H1 Histone Variant Overexpression on Chromatin Structure *. *J Biol Chem* **274**: 37950–37956.

Halliwell JA, Martin-Gonzalez J, Hashim A, Dahl JA, Hoffmann ER, Lerdrup M. 2024. Sex-specific DNA-replication in the early mammalian embryo. *Nat Commun* **15**: 6323.

Hamamoto G, Suzuki T, Suzuki MG, Aoki F. 2014. Regulation of transketolase like 1 gene expression in the murine one-cell stage embryos. *PloS One* **9**: e82087.

Harr JC, Luperchio TR, Wong X, Cohen E, Wheelan SJ, Reddy KL. 2015. Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biol* **208**: 33–52.

Ishiuchi T, Enriquez-Gasca R, Mizutani E, Bošković A, Ziegler-Birling C, Rodriguez-Terrones D, Wakayama T, Vaquerizas JM, Torres-Padilla M-E. 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat Struct Mol Biol* **22**: 662–671.

Jacinto FV, Benner C, Hetzer MW. 2015. The nucleoporin Nup153 regulates embryonic stem cell pluripotency through gene silencing. *Genes Dev* **29**: 1224–1238.

Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, et al. 2017. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell* **170**: 367-381.e20.

Kermi C, Aze A, Maiorano D. 2019. Preserving Genome Integrity during the Early Embryonic DNA Replication Cycles. *Genes* **10**: 398.

Kidder BL, Hu G, Zhao K. 2014. KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biol* **15**: R32.

Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, de Graaf CA, Amendola M, et al. 2015. Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell* **163**: 134–147.

Kind J, van Steensel B. 2014. Stochastic genome-nuclear lamina interactions. *Nucleus* **5**: 124–130.

Kruse K, Díaz N, Enriquez-Gasca R, Gaume X, Torres-Padilla M-E, Vaquerizas JM. 2019. Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. 523712. https://www.biorxiv.org/content/10.1101/523712v1 (Accessed May 11, 2023).

Kumaran RI, Spector DL. 2008. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol* **180**: 51–65.

Lawinger P, Rastelli L, Zhao Z, Majumder S. 1999. Lack of Enhancer Function in Mammals Is Unique to Oocytes and Fertilized Eggs*. *J Biol Chem* **274**: 8002–8011.

Lenain C, de Graaf CA, Pagie L, Visser NL, de Haas M, de Vries SS, Peric-Hupkes D, van Steensel B, Peeper DS. 2017. Massive reshaping of genome-nuclear lamina interactions during oncogene-induced senescence. *Genome Res* **27**: 1634–1644.

Leonard PH, Grzenda A, Mathison A, Morbeck DE, Fredrickson JR, de Assuncao TM, Christensen T, Salisbury J, Calvo E, Iovanna J, et al. 2015. The Aurora A-HP1γ pathway regulates gene expression and mitosis in cells from the sperm lineage. *BMC Dev Biol* **15**: 23.

Li P, Wang L, Bennett BD, Wang J, Li J, Qin Y, Takaku M, Wade PA, Wong J, Hu G. 2017. Rif1 promotes a repressive chromatin state to safeguard against endogenous retrovirus activation. *Nucleic Acids Res* **45**: 12723–12738.

Liu B, He Y, Wu X, Lin Z, Ma J, Qiu Y, Xiang Y, Kong F, Lai F, Pal M. 2024. Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators. *Nat Cell Biol* 1–13.

Liu B, Xu Q, Wang Q, Feng S, Lai F, Wang P, Zheng F, Xiang Y, Wu J, Nie J, et al. 2020. The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature* **587**: 139–144.

MacPherson Q, Beltran B, Spakowitz AJ. 2020. Chromatin Compaction Leads to a Preference for Peripheral Heterochromatin. *Biophys J* **118**: 1479–1488.

Majumder S, Zhao Z, Kaneko K, DePamphilis ML. 1997. Developmental acquisition of enhancer function requires a unique coactivator activity. *EMBO J* **16**: 1721–1731.

Manzo SG, Mazouzi A, Leemans C, van Schaik T, Neyazi N, van Ruiten MS, Rowland BD, Brummelkamp TR, van Steensel B. 2024. Chromatin protein complexes involved in gene repression in lamina-associated domains. *EMBO J* 1–28.

Masser EA, Noble TD, Siefert JC, Goins D, Sansam CG, Sansam CL. 2023. Zebrafish Rif1 impacts zygotic genome activation, replication timing, and sex determination. *eLife* **12**. https://elifesciences.org/reviewed-preprints/87671 (Accessed November 5, 2024).

Nakatani T, Lin J, Ji F, Ettinger A, Pontabry J, Tokoro M, Altamirano-Pacheco L, Fiorentino J, Mahammadov E, Hatano Y, et al. 2022. DNA replication fork speed underlies cell fate changes and promotes reprogramming. *Nat Genet* **54**: 318–327.

Nakatani T, Schauer T, Altamirano-Pacheco L, Klein KN, Ettinger A, Pal M, Gilbert DM, Torres-Padilla M-E. 2024. Emergence of replication timing during early mammalian development. *Nature* **625**: 401–409.

Nguyen VT, Giannoni F, Dubois M-F, Seo S-J, Vigneron M, Kédinger C, Bensaude O. 1996. In Vivo Degradation of RNA Polymerase II Largest Subunit Triggered by α-Amanitin. *Nucleic Acids Res* **24**: 2924–2929.

Pal M, Altamirano-Pacheco L, Schauer T, Torres-Padilla M-E. 2023. Reorganization of lamina-associated domains in early mouse embryos is regulated by RNA polymerase II activity. *Genes Dev* **37**: 901-912.

Pal M, Kind J, Torres-Padilla M-E. 2021. DamID to Map Genome-Protein Interactions in Preimplantation Mouse Embryos. In *Epigenetic Reprogramming During Mouse Embryogenesis: Methods and Protocols* (eds. K. Ancelin and M. Borensztein), *Methods in Molecular Biology*, pp. 265–282, Springer US, New York, NY.

Prendergast L, Reinberg D. 2021. The missing linker: emerging trends for H1 variant-specific functions. *Genes Dev* **35**: 40–58.

Rodriguez-Terrones D, Gaume X, Ishiuchi T, Weiss A, Kopp A, Kruse K, Penning A, Vaquerizas JM, Brino L, Torres-Padilla M-E. 2018. A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat Genet* **50**: 106–119.

Sakashita A, Kitano T, Ishizu H, Guo Y, Masuda H, Ariura M, Murano K, Siomi H. 2023. Transcription of MERVL retrotransposons is required for preimplantation embryo development. *Nat Genet* **55**: 484–495.

Sankar A, Lerdrup M, Manaf A, Johansen JV, Gonzalez JM, Borup R, Blanshard R, Klungland A, Hansen K, Andersen CY, et al. 2020. KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3 domains from H3K9me3 invasion in oocytes. *Nat Cell Biol* **22**: 380–388.

Sati S, Bonev B, Szabo Q, Jost D, Bensadoun P, Serra F, Loubiere V, Papadopoulos GL, Rivera-Mulia J-C, Fritsch L, et al. 2020. 4D Genome Rewiring during Oncogene-Induced and Replicative Senescence. *Mol Cell* **78**: 522-538.e9.

Seller CA, O'Farrell PH. 2018. Rif1 prolongs the embryonic S phase at the Drosophila mid-blastula transition. *PLoS Biol* **16**: e2005687.

Shao G-B, Chen J-C, Zhang L-P, Huang P, Lu H-Y, Jin J, Gong A-H, Sang J-R. 2014. Dynamic patterns of histone H3 lysine 4 methyltransferases and demethylases during mouse preimplantation development. *In Vitro Cell Dev Biol Anim* **50**: 603–613.

Sharifi-Zarchi A, Gerovska D, Adachi K, Totonchi M, Pezeshk H, Taft RJ, Schöler HR, Chitsaz H, Sadeghi M, Baharvand H, et al. 2017. DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. *BMC Genomics* **18**: 964.

Shen H, Xu W, Guo R, Rong B, Gu L, Wang Z, He C, Zheng L, Hu X, Hu Z, et al. 2016. Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. *Cell* **165**: 331–342.

Shilatifard A. 2012. The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. *Annu Rev Biochem* **81**: 65–95.

Siegenfeld AP, Roseman SA, Roh H, Lue NZ, Wagen CC, Zhou E, Johnstone SE, Aryee MJ, Liau BB. 2022. Polycomb-lamina antagonism partitions heterochromatin at the nuclear periphery. *Nat Commun* **13**: 4199.

Solovei I, Kreysing M, Lanctôt C, Kösem S, Peichl L, Cremer T, Guck J, Joffe B. 2009. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell* **137**: 356–368.

Solovei I, Wang AS, Thanisch K, Schmidt CS, Krebs S, Zwerger M, Cohen TV, Devys D, Foisner R, Peichl L, et al. 2013. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152**: 584–598.

Steensel B van, Henikoff S. 2000. Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nat Biotechnol* **18**: 424–428.

Takahashi S, Kyogoku H, Hayakawa T, Miura H, Oji A, Kondo Y, Takebayashi S, Kitajima TS, Hiratani I. 2024. Embryonic genome instability upon DNA replication timing program emergence. *Nature* **633**: 686–694.

Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. 2014. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* **346**: 1238–1242.

Thrower DA, Bloom K. 2001. Dicentric chromosome stretching during anaphase reveals roles of Sir2/Ku in chromatin compaction in budding yeast. *Mol Biol Cell* **12**: 2800–2812.

Tosti L, Ashmore J, Tan BSN, Carbone B, Mistri TK, Wilson V, Tomlinson SR, Kaji K. 2018. Mapping transcription factor occupancy using minimal numbers of cells in vitro and in vivo. *Genome Res* **28**: 592–605.

Tyagi S, Capitanio JS, Xu J, Chen F, Sharma R, Huang J, Hetzer MW. 2023. High-precision mapping of nuclear pore-chromatin interactions reveals new principles of genome organization at the nuclear envelope. *eLife* **12**. https://elifesciences.org/reviewed-preprints/87462 (Accessed November 6, 2024).

van Schaik T, Vos M, Peric-Hupkes D, HN Celie P, van Steensel B. 2020. Cell cycle dynamics of lamina-associated DNA. *EMBO Rep* **21**: e50636.

Vaquero A, Scher M, Lee D, Erdjument-Bromage H, Tempst P, Reinberg D. 2004. Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. *Mol Cell* **16**: 93–105.

Wang L, Zhang J, Duan J, Gao X, Zhu W, Lu X, Yang L, Zhang J, Li G, Ci W, et al. 2014. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**: 979–991.

Xu S, Wang N, Zuccaro MV, Gerhardt J, Iyyappan R, Scatolin GN, Jiang Z, Baslan T, Koren A, Egli D. 2024. DNA replication in early mammalian embryos is patterned, predisposing lamina-associated regions to fragility. *Nat Commun* **15**: 5247.

Yoshizawa-Sugata N, Yamazaki S, Mita-Yoshida K, Ono T, Nishito Y, Masai H. 2021. Loss of full-length DNA replication regulator Rif1 in two-cell embryos is associated with zygotic transcriptional activation. *J Biol Chem* **297**: 101367.

Zenk F, Zhan Y, Kos P, Löser E, Atinbayeva N, Schächtle M, Tiana G, Giorgetti L, Iovino N. 2021. HP1 drives de novo 3D genome reorganization in early Drosophila embryos. *Nature* **593**: 289–293.

Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q, et al. 2016. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**: 553–557.

Copyright statements

This Agreement between Mrinmoy Pal ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5911750962655
License date	Nov 18, 2024
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	DamID to Map Genome-Protein Interactions in Preimplantation Mouse EmbryosEmbryos
Licensed Content Author	Mrinmoy Pal, Jop Kind, Maria-Elena Torres-Padilla
Licensed Content Date	Jan 1, 2021
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	yes
Title of new work	PhD Student
Institution name	Institute of Epigenetics and Stem Cells, Helmholtz Munich
Expected presentation date	Dec 2024
The Requesting Person / Organization to Appear on the License	Mrinmoy Pal
Requestor Location	Mrinmoy Pal Feodor-Lynen-Strasse 21 Munich, 81377 Germany
Billing Type	Invoice

Terms and Conditions

Springer Nature Customer Service Centre GmbH Terms and Conditions

The following terms and conditions ("Terms and Conditions") together with the terms specified in your [RightsLink] constitute the License ("License") between you as Licensee and Springer Nature Customer Service Centre GmbH as Licensor. By clicking 'accept' and completing the transaction for your use of the material ("Licensed Material"), you confirm your acceptance of and obligation to be bound by these Terms and Conditions.

1. Grant and Scope of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, non-sublicensable, revocable, world-wide License to reproduce, distribute, communicate to the public, make available, broadcast, electronically transmit or create derivative works using the Licensed Material for the purpose(s) specified in your RightsLink Licence Details only. Licenses are granted for the specific use requested in the order and for no other use, subject to these Terms and Conditions. You acknowledge and agree that the rights granted to you under this License do not include the right to modify, edit, translate, include in collective works, or create derivative works of the Licensed Material in whole or in part unless expressly stated in your RightsLink Licence Details. You may use the Licensed Material only as permitted under this Agreement and will not reproduce, distribute, display, perform, or otherwise use or exploit any Licensed Material in any way, in whole or in part, except as expressly permitted by this License.

1. 2. You may only use the Licensed Content in the manner and to the extent permitted by these Terms and Conditions, by your RightsLink Licence Details and by any applicable laws.

1. 3. A separate license may be required for any additional use of the Licensed Material, e.g. where a license has been purchased for print use only, separate permission must be obtained for electronic re-use. Similarly, a License is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the License.

1. 4. Any content within the Licensed Material that is owned by third parties is expressly excluded from the License.

1. 5. Rights for additional reuses such as custom editions, computer/mobile applications, film or TV reuses and/or any other derivative rights requests require additional permission and may be subject to an additional fee. Please apply to journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

2. Reservation of Rights

Licensor reserves all rights not expressly granted to you under this License. You acknowledge and agree that nothing in this License limits or restricts Licensor's rights in or use of the Licensed Material in any way. Neither this License, nor any act, omission, or statement by Licensor or you, conveys any ownership right to you in any Licensed Material, or to any element or portion thereof. As between Licensor and you, Licensor owns and retains all right, title, and interest in and to the Licensed Material subject to the license granted in Section 1.1. Your permission to use the Licensed Material is expressly conditioned on you not impairing Licensor's or the applicable copyright owner's rights in the Licensed Material in any way.

3. Restrictions on use

3. 1. Minor editing privileges are allowed for adaptations for stylistic purposes or formatting purposes provided such alterations do not alter the original meaning or intention of the Licensed Material and the new figure(s) are still accurate and representative of the Licensed Material. Any other changes including but not limited to, cropping, adapting, and/or omitting material that affect the meaning, intention or moral rights of the author(s) are strictly prohibited.

3. 2. You must not use any Licensed Material as part of any design or trademark.

3. 3. Licensed Material may be used in Open Access Publications (OAP), but any such reuse must include a clear acknowledgment of this permission visible at the same time as the figures/tables/illustration or abstract and which must indicate that

the Licensed Material is not part of the governing OA license but has been reproduced with permission. This may be indicated according to any standard referencing system but must include at a minimum 'Book/Journal title, Author, Journal Name (if applicable), Volume (if applicable), Publisher, Year, reproduced with permission from SNCSC'.

4. STM Permission Guidelines

4. 1. An alternative scope of license may apply to signatories of the STM Permissions Guidelines ("STM PG") as amended from time to time and made available at https://www.stm-assoc.org/intellectual-property/permissions/permissions-guidelines/.

4. 2. For content reuse requests that qualify for permission under the STM PG, and which may be updated from time to time, the STM PG supersede the terms and conditions contained in this License.

4. 3. If a License has been granted under the STM PG, but the STM PG no longer apply at the time of publication, further permission must be sought from the Rightsholder. Contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

5. Duration of License

5. 1. Unless otherwise indicated on your License, a License is valid from the date of purchase ("License Date") until the end of the relevant period in the below table:

Reuse in a medical communications project	Reuse up to distribution or time period indicated in License
Reuse in a dissertation/thesis	Lifetime of thesis
Reuse in a journal/magazine	Lifetime of journal/magazine
Reuse in a book/textbook	Lifetime of edition
Reuse on a website	1 year unless otherwise specified in the License
Reuse in a presentation/slide kit/poster	Lifetime of presentation/slide kit/poster. Note: publication whether electronic or in print of presentation/slide kit/poster may require further permission.
Reuse in conference proceedings	Lifetime of conference proceedings
Reuse in an annual report	Lifetime of annual report
Reuse in training/CME materials	Reuse up to distribution or time period indicated in License
Reuse in newsmedia	Lifetime of newsmedia
Reuse in coursepack/classroom materials	Reuse up to distribution and/or time period indicated in license

6. Acknowledgement

6. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract and must be hyperlinked to the journal/book's homepage.

6. 2. Acknowledgement may be provided according to any standard referencing system and at a minimum should include "Author, Article/Book Title, Journal name/Book imprint, volume, page number, year, Springer Nature".

7. Reuse in a dissertation or thesis

7. 1. Where 'reuse in a dissertation/thesis' has been selected, the following terms apply: Print rights of the Version of Record are provided for; electronic rights for use only on institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/) and only up to what is required by the awarding institution.

7. 2. For theses published under an ISBN or ISSN, separate permission is required. Please contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

7. 3. Authors must properly cite the published manuscript in their thesis according to current citation standards and include the following acknowledgement: '*Reproduced with permission from Springer Nature*'.

8. License Fee

You must pay the fee set forth in the License Agreement (the "License Fees"). All amounts payable by you under this License are exclusive of any sales, use, withholding, value added or similar taxes, government fees or levies or other

assessments. Collection and/or remittance of such taxes to the relevant tax authority shall be the responsibility of the party who has the legal obligation to do so.

9. Warranty

9. 1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. **You are solely responsible for ensuring that the material you wish to license is original to the Licensor and does not carry the copyright of another entity or third party (as credited in the published version).** If the credit line on any part of the Licensed Material indicates that it was reprinted or adapted with permission from another source, then you should seek additional permission from that source to reuse the material.

9. 2. EXCEPT FOR THE EXPRESS WARRANTY STATED HEREIN AND TO THE EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR PROVIDES THE LICENSED MATERIAL "AS IS" AND MAKES NO OTHER REPRESENTATION OR WARRANTY. LICENSOR EXPRESSLY DISCLAIMS ANY LIABILITY FOR ANY CLAIM ARISING FROM OR OUT OF THE CONTENT, INCLUDING BUT NOT LIMITED TO ANY ERRORS, INACCURACIES, OMISSIONS, OR DEFECTS CONTAINED THEREIN, AND ANY IMPLIED OR EXPRESS WARRANTY AS TO MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, PUNITIVE, OR EXEMPLARY DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE LICENSED MATERIAL REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION APPLIES NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

10. Termination and Cancellation

10. 1. The License and all rights granted hereunder will continue until the end of the applicable period shown in Clause 5.1 above. Thereafter, this license will be terminated and all rights granted hereunder will cease.

10. 2. Licensor reserves the right to terminate the License in the event that payment is not received in full or if you breach the terms of this License.

11. General

11. 1. The License and the rights and obligations of the parties hereto shall be construed, interpreted and determined in accordance with the laws of the Federal Republic of Germany without reference to the stipulations of the CISG (United Nations Convention on Contracts for the International Sale of Goods) or to Germany's choice-of-law principle.

11. 2. The parties acknowledge and agree that any controversies and disputes arising out of this License shall be decided exclusively by the courts of or having jurisdiction for Heidelberg, Germany, as far as legally permissible.

11. 3. This License is solely for Licensor's and Licensee's benefit. It is not for the benefit of any other person or entity.

Questions? For questions on Copyright Clearance Center accounts or website issues please contact springernaturesupport@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777. For questions on Springer Nature licensing please visit https://www.springernature.com/gp/partners/rights-permissions-third-party-distribution

Other Conditions:

Version 1.4 - Dec 2022

Questions? customercare@copyright.com.

Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators

Author: Bofeng Liu et al

Publication: Nature Cell Biology

Publisher: Springer Nature

Date: Jun 5, 2024

Copyright © 2024, The Author(s), under exclusive licence to Springer Nature Limited

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select no to the question "Are you the Author of this Springer Nature content?".
Ownership of copyright in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:
To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).
The author and any academic institution, where they work, at the time may reproduce the contribution for the purpose of course teaching.
To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.
To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website.
Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read Springer Nature's online author reuse guidelines.

For full paper portion: Authors of original research papers published by Springer Nature are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v1.0

BACK   CLOSE WINDOW

Genes & Development

Warranties

Acknowledgements

I am grateful to the reviewers of my thesis for taking the time to read this work. I hope they find excitement in the science that has kept me motivated and curious over the past years. I would also like to thank all the members of my TAC, Heinrich, Melina, and Irina, for their advice and support throughout my PhD.

I am incredibly grateful to Maria-Elena for taking me on this PhD journey. You taught me so much, mentored me, believed in me, and, most importantly, challenged me. These years in your lab have immensely shaped me, and I am thankful for your time, patience, commitment, and support. Thank you for assembling such an incredible team of scientists and creating an excellent atmosphere for discussion and growth.

I am also profoundly grateful to my past mentors, who guided my scientific journey. I would like to especially thank my undergraduate internship supervisors: Ashwin (IISc Bangalore), Anusha and Munia Ma'am (CSIR IGIB, Delhi), Simon, Oliver, and Steve (UMC Göttingen), Akhila, Saurabh, Manu, and Mouli (IISER Pune). Each of you taught me with patience and kept my curiosity alive. I am also thankful for the excellent teachers and guides I had during my studies, who repeatedly made me fall in love with biology. A special mention goes to Sanjeev, who supervised me during my master's thesis. Your teaching was always inspiring, and your mentorship and support have been invaluable.

I want to thank Tamas, who was always there to create beautiful plots from my data. I am deeply grateful for the support you provided throughout my projects. You taught me much about analysis, experimental design, and scientific rigor. Thank you, Tsune, for your technical mastery and the invaluable discussions about microinjections, embryos, and oocytes. I am immensely grateful for your trust in my experimental skills and for collaborating with me across different projects. I also thank Adam, Fede, and Iliya for helping with revision experiments. Thank you, Marc and Alicia, for helping us with the analysis for the screening manuscript. A special thanks to Luis, who patiently answered my endless questions as I started working on bioinformatics.

I am thankful to everyone in the Torres-Padilla lab, including Antoine, Clara, Fede, Iliya, Marlies, Tsune, Tamas, Adam, Yicong, Jiezhen, Yuki, Andreas, Pilar, and Marga. I am also grateful to our past lab members: Manuel, Marion, Yung-Li, Camille, Ane, Ken, Jiangwei, Natasha, Amelie, Ksenia, Elias, Luis, Hiromi, Mich, Melissa, Lorenza, and my student Mythili. I am genuinely grateful to have shared my PhD journey with so many curious, passionate, and kind people. It has always been an outstanding balance of science, fun, and mutual respect. Thank you for tolerating my attempted puns and the lively and often intense discussions in lab meetings or over lunch and coffee. Thanks for the beer clubs, beach volleyball sessions, ice cream outings, Christmas market visits, and hikes. It has been a wonderful time, and I will always cherish these memories.

I would also like to extend my thanks to everyone in the institute, from Stephan, Eva, Antonio, and Nico to past and present lab members, including Matthias, Henning, Anna, Elisabeth, Maxime, Marcel, Manuel, Atiqa, Clare, Ioannis, Elizabeth, Augusto, Xanthoula, Tomas, Meghana, Ana, Jonathan, Elmir, Mayra, Gabriele, Marco, Veronica, Kim, Martin, Wasif, David, Andres, Gizem, Jasmine, Juana, Marie-Sophie, Thomas, and Laura. Life at IES has been an adventure. It has been a rollercoaster ride, from enduring a global pandemic to surviving a cyberattack. I will never forget the karaoke sessions, Secret Santa exchanges, barbecues, post-seminar beer gatherings, movie nights, cooking sessions, and, last but not least, the IES Olympics. The warm, welcoming, and vibrant atmosphere at IES made this journey truly unforgettable.

My special thanks to the ChromDesign family: Livia, Gianni, Arun, Mike, Tina, Pia, Carla, Antonia, Blanka, Kourosh, Nathalie, Alicia, Luciano, Marc, Genevieve, Edith, Kristian, Luca, Stefan, Giacomo, Saveiro, Francesc, Jonas, and David. It has been wonderful to be part of this PhD network. From endless courses to intense discussions, from parties and fun to catch-up Zoom calls, I am grateful to have connected with and learned from all of you. I am sure we will keep in touch and may our paths cross again and again.

I am also grateful to my past and present flatmates, especially Raju and Prateek, who have been integral to my life. You made Munich feel like home, and I knew I could always count on a smile, even after the most stressful weeks at work. My Munich circle: Meghana, Vibha, Viju, Avani, Muhunden, Pavan, and Mahak - you were instrumental in keeping me sane and functional. My heartfelt thanks to all of you for caring, sharing, and being there for me.

I would also like to thank all my friends, my school buddies, my IISER Kolkata mates (Pathak, Simha, Vinay, Shruti, Varun, Krishna, Prajjwal, Sesha, Nishchhal, Debanjana, Sreyam, Som), and the Pune gang (Mukul, Abhishek, Rini, Sudipta, Sneha, Ankitha, Deepak, Saurabh, Sharma), for staying in touch and checking in from across different time zones. You made me realize that our bonds are stronger than the distance, and I am forever grateful for your continued friendship.

Munich has been an incredible city to call home during this time. I am grateful for the beautiful sunsets, the fall colors, the snow, the spring blossoms, the summer sunshine, and the Alps. This city has been a source of energy, solace, and inspiration. Thank you, Munich, for everything!

Finally, I want to thank my family. Dada, you have always inspired me and are probably the reason I chose a career in research in the first place. Maa and Baba, your unwavering love and support have been my greatest strength. Even though it has been hard to be so far from you, this journey has deepened my appreciation for all the little joys and comforts you have always provided. *This thesis is dedicated to both of you.*

Curriculum vitae

MRINMOY PAL

LinkedIn   ORCID   Google scholar
mrinmoy.pal@helmholtz-muenchen.de
mrinmoy.pal.1995@gmail.com

Curriculum vitae

EDUCATION

PhD | Ludwig Maximilian University, Munich   2019-Current
Supervisor: Maria-Elena Torres-Padilla | IES, Helmholtz Munich

BS MS Dual Degree | IISER Kolkata   2013-2018
Major: Biological Sciences | CGPA: 9.86

FELLOWSHIPS

Marie-Curie Doctoral Fellowship | ChromDesign Innovative Training Network   2019-2022

DAAD WISE | University Medical Centre, Göttingen   2016

KVPY | Department of Science & Technology, Government of India   2013-2018

PUBLICATIONS

From PhD work (2019-Current)

Pal, Mrinmoy, Tamas Schauer, Adam Burton, Tsunetoshi Nakatani, Federico Pecori, Alicia Hernández-Giménez, Iliya Nadelson, Marc A. Marti-Renom, Maria-Elena Torres-Padilla. "The establishment of nuclear organization in mouse embryos is orchestrated by multiple epigenetic pathways." ***Under review***

Nakatani, Tsunetoshi, Tamas Schauer, Mrinmoy Pal, Andreas Ettinger, Luis Altamirano-Pacheco, Julia Zorn, David M. Gilbert, Maria-Elena Torres-Padilla. "RIF1 regulates the consolidation of replication timing in early mouse embryos independently of changes in nuclear organization towards the nuclear lamina." ***Under review***

Liu, Bofeng, Yuanlin He, Xiaotong Wu, Zili Lin, Jing Ma, Yuexin Qiu, Yunlong Xiang, Feng Kong, Fangnong Lai, Mrinmoy Pal, Peizhe Wang, Jia Ming, Bingjie Zhang, Qiujun Wang, Jingyi Wu, Weikun Xia, Jie Na, Maria-Elena Torres-Padilla, Jing Li, Wei Xie. "Mapping putative enhancers in mammalian oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators." ***Nat Cell Biol.*** 2024 Jun; 26(6):962-974.

Nakatani, Tsunetoshi, Tamas Schauer, Luis Altamirano-Pacheco, Kyle N. Klein, Andreas Ettinger, Mrinmoy Pal, David M. Gilbert, Maria-Elena Torres-Padilla. "Emergence of replication timing during early mammalian development." ***Nature.*** 2024 Jan; 625(7994):401-409.

Pal, Mrinmoy, Luis Altamirano-Pacheco, Tamas Schauer, Maria-Elena Torres-Padilla. "Reorganization of Lamina Associated Domains in early mouse embryos is regulated by RNA Polymerase II activity." ***Genes & Development.*** 2023 Oct 1; 37(19-20):901-912.

Pal, Mrinmoy, Jop Kind, and Maria-Elena Torres-Padilla. "DamID to map genome-protein interactions in preimplantation mouse embryos." ***Methods in Molecular Biology***. 2021; 2214, 265–282.

From undergraduate work (2013-2018)

Gungi, Akhila, Shagnik Saha, Mrinmoy Pal, and Sanjeev Galande. "H4K20me1 plays a dual role in transcriptional regulation of regeneration and axis patterning in Hydra." *Life Science Alliance*. 6, no. 5 (2023).

Unni, Manu, Puli Chandramouli Reddy, Mrinmoy Pal, Irit Sagi, and Sanjeev Galande. "Identification of components of the hippo pathway in Hydra and potential role of YAP in cell division and differentiation." *Frontiers in Genetics.* 12 (2021): 676182.

Pradhan, Saurabh J., Puli Chandramouli Reddy, Michael Smutny, Ankita Sharma, Keisuke Sako, Meghana S. Oak, Rini Shah, Mrinmoy Pal, Ojas Deshpande, Greg Dsilva, Yin Tang, Rakesh Mishra, Girish Deshpande, Antonio J Giraldez, Mahendra Sonawane, Carl-Philipp Heisenberg, Sanjeev Galande. "Satb2 acts as a gatekeeper for major developmental transitions during early vertebrate embryogenesis." *Nature Communications*. 12, no. 1 (2021): 6094.

Aditya, Anusha, Sabyasachi Chattopadhyay, Nidhi Gupta, Shamshad Alam, Archana Palillam Veedu, Mrinmoy Pal, Archana Singh, Deenan Santhiya, Kausar M. Ansari, and Munia Ganguli. "ZnO nanoparticles modified with an amphipathic peptide show improved photoprotection in skin." *ACS applied materials & interfaces.* 11, no. 1 (2018): 56-72.

Basu, Srijoni, Chandra Bose, Nupur Ojha, Nabajit Das, Jagaree Das, Mrinmoy Pal, and Sukant Khurana. "Evolution of bacterial and fungal growth media." *Bioinformation.* 11, no. 4 (2015): 182.

SCIENTIFIC TALKS & POSTERS

Chromatin Dynamics Symposium | Platform Talk | LMU, Munich — Oct 2024

Epigenetics and Chromatin | Poster | Cold Spring Harbor Laboratory — Sep 2024

NIH 4DN Scientific Webinar | Virtual Talk — Aug 2023

Chromatin Day | Platform Talk | LMU, Munich — July 2023

NIH 4DN Annual Meeting 2022 | Poster | San Diego — Dec 2022

EpIC Conference | Platform Talk | Granada — Oct 2022

8th Asian Science Camp | Poster | NTU, Singapore — Aug 2014

HONOURS & AWARDS

Selected attendee of 68th Lindau Nobel Laureate Meeting (2018).

Awarded IISER Kolkata Chairman's Gold Medal for Overall Excellence for highest all-round performance in curricular and extracurricular activities and leadership (2018).

Recipient of Director's Gold Medal for 1st rank in Department of Biological Sciences, IISER Kolkata (2018).

Qualified National Eligibility Test (NET) with All India Rank 14 in Life Sciences (June 2017).

RESEARCH & WORK EXPERIENCE

Intern | ChromDesign Secondment | EMBO Press Editorial Board, Heidelberg   Sept 2022

Research Assistant | Prof. Sanjeev Galande | IISER Pune   May- Dec 2018

DAAD Summer Research Fellow | Prof. Steven Johnsen | UMC Göttingen   May-July 2016

KVPY Summer Research Intern | Dr. Munia Ganguli | CSIR IGIB, Delhi   May-July 2015

KVPY Summer Research Intern | Prof. Umesh Varshney | IISc Bangalore   May-July 2014

TEACHING AND LEADERSHIP

Instructor at Helmholtz Summer School on Chromatin biology (2024).

Selected as DAAD Young Ambassador for DAAD India (2022-2023).

Member of The Epigenetics PhD Board at Helmholtz Munich (2021-2023).

Core committee member of IISER Kolkata Alumni Association (2019-2023).

Organizing committee member of EpIC Conference in Granada (2022).

Part of organizing team at Happygene2Sys: virtual meeting of the Epigene2Sys (2020).

Selected by Ministry of Youth Affairs and Sports, Government of India to attend 19th World Festival of Youth and Students at Sochi, Russia as a member of Indian Delegation (2017).

Founding member of Student Affairs Council (SAC), IISER Kolkata (2015-2016).

REFEREES

Prof. Dr. Maria-Elena Torres-Padilla | PhD Supervisor
Director of the Institute of Epigenetics and Stem Cells, Helmholtz Munich
torres-padilla@helmholtz-munich.de

Prof. Dr. Sanjeev Galande | Undergraduate Supervisor
Dean, School of Natural Sciences
Head of the Center of Excellence in Epigenetics, Shiv Nadar Institute of Eminence
sanjeev.galande@snu.edu.in

Prof. Dr. Marc A. Marti-Renom
Structural Genomics Group Leader at Centre de Regulació Genòmica (CRG), Barcelona
martirenom@crg.eu

Dr. Stephan Hamperl
Group Leader at the Institute of Epigenetics and Stem Cells, Helmholtz Munich
stephan.hamperl@helmholtz-munich.de