

Biophysical properties of heterochromatin in totipotent mouse embryos



KUMULATIVE DISSERTATION
DER FAKULTÄT FÜR BIOLOGIE
DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Vorgelegt von
Manuel Guthmann

München, den 26.10.2022

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

Zweitgutachter: Prof. Dr. Heinrich Leonhardt

Tag der Einreichung: 26/10/2022

Tag der mündlichen Prüfung: 18/07/2023

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.

Manuel Guthmann

München, den 26.10.2022

TABLE OF CONTENTS

LIST OF PUBLISHED MANUSCRIPTS.....	9
LIST OF UNPUBLISHED MANUSCRIPTS.....	9
SUMMARY	11
AIMS.....	13
INTRODUCTION	15
MOUSE PRE-IMPLANTATION DEVELOPMENT	17
METABOLISM STATE DURING PRE-IMPLANTATION DEVELOPMENT....	18
GENOME ORGANISATION IN PRE-IMPLANTATION DEVELOPMENT	18
IN-VITRO MODEL TO STUDY TOTIPOTENCY	21
LIQUID-LIQUID PHASE SEPARATION	22
GENERAL PRINCIPLES OF LIQUID-LIQUID PHASE SEPARATION IN BIOLOGY	22
LIQUID-LIQUID PHASE SEPARATION IN CHROMATIN ORGANISATION.	24
ASSAYS TO STUDY LIQUID-LIQUID PHASE SEPARATION	26
RESULTS	27
PART 1: METABOLIC STATE OF TOTIPOTENT CELLS	27
A DISTINCT METABOLIC STATE ARISES DURING THE EMERGENCE OF 2-CELL-LIKE CELLS.....	29
PART 2: ROLE OF LIQUID-LIQUID PHASE SEPARATION IN HETEROCHROMATIN FORMATION DURING MOUSE PRE- IMPLANTATION DEVELOPMENT	55

EXPRESSION AND PHASE SEPARATION POTENTIAL OF HETEROCHROMATIN PROTEINS DURING EARLY MOUSE DEVELOPMENT	57
A PHASE TRANSITION ACCOMPANIES HETEROCHROMATIN FORMATION IN MOUSE EMBRYOS	75
ABSTRACT	77
MAIN TEXT	77
MATERIAL AND METHODS	86
DISCUSSION	115
PART 1: METABOLIC STATE OF TOTIPOTENT CELLS	117
PART 2: ROLE OF LIQUID-LIQUID PHASE SEPARATION IN HETEROCHROMATIN FORMATION DURING MOUSE PRE- IMPLANTATION DEVELOPMENT	117
EXPRESSION AND PHASE SEPARATION POTENTIAL OF HETEROCHROMATIN PROTEINS DURING EARLY MOUSE DEVELOPMENT	118
A PHASE TRANSITION ACCOMPANIES HETEROCHROMATIN FORMATION IN MOUSE EMBRYOS	120
OUTLOOK.....	123
REFERENCES	124
ACKNOWLEDGMENTS.....	133

LIST OF ABBREVIATIONS

ICM	Inner Cell Mass
TE	Trophectoderm
ZGA	zygotic genome activation
ESCs	Embryonic Stem Cells
iPSCs	induced Pluripotent Stem Cells
H3K9me3	trimethylation of histone H3 at lysine 9
TAD	Topologically Associated Domain
LAD	Lamina Associated Domains
DamID	DNA Adenine Methyltransferase Identification
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
FRAP	Fluorescence Recovery After Photobleaching
HP1	Heterochromatin Protein 1
NLBs	Nucleolar-Like Bodies
TALE	Transcription Activator-Like Effector
2CLCs	2-cell-like cells
LLPS	Liquid-Liquid Phase Separation
RICS	Raster Imaging Correlation Spectroscopy
N&B	Number and Brightness
PRC1	Polycomb-Repressive Complex 1
SAFB	Scaffold Attachment Factor b
FCS	Fluorescent Correlation Spectroscopy
ROS	reactive oxygen species
CAF-1	Chromatin Assembly Factor 1
OCR	Oxygen Consumption Rate
HBP	Hexosamine Biosynthetic Pathway
PPP	Pentose Phosphate Pathway
PMSG	Pregnant Mare Serum Gonadotropin
hCG	human Chorionic Gonadotropin
IDR	Intrinsically Disordered Domain
Mommes	Modifiers of murine metastable epialleles
PLD	Prion-Like Domain
PiCH	Proteomics of isolated Chromatin segments
TGV	TALE-based Genome Visualization
LC	Low Complexity domain

KSOM	K-modified Simplex Optimized Medium
MajSat	Major Satellite
LOESS	Locally Estimated Scatterplot Smoothing
STED	Stimulated Emission Depletion Microscopy
ASO	Antisense Oligo

LIST OF PUBLISHED MANUSCRIPTS

Rodriguez-Terrones, D., Hartlben, G., Gaume, X., Eid, A., **Guthmann, M.**, Iturbide, A., Torres-Padilla, M.E. (2020). A distinct metabolic state arises during the emergence of 2-cell-like cells. EMBO reports. 21: e48354.

Guthmann, M., Burton, A., Torres-Padilla, M.E. (2019). Expression and phase separation potential of heterochromatin proteins during early mouse development. EMBO reports. 20: e47952

LIST OF UNPUBLISHED MANUSCRIPTS

Guthmann, M., Quian, C., Gialdini, I., Nakatani, T., Ettinger, A., Schauer, T., Kukhtevich, I., Schneider, R., Lamb, D.C., Burton, A., Torres-Padilla, M.E. (Unpublished manuscript). A phase transition accompanies heterochromatin formation in mouse embryos.

SUMMARY

Totipotency is an incredibly plastic and transient state during the development of an embryo, which in mouse is restricted to the zygote and 2-cell stage. Indeed, totipotent embryos, are unique in terms of their chromatin architecture, metabolism and transcriptional status. Studying this short window of development is thus very important not only to better understand how normal development at these early stages occurs, but also to obtain conceptual and methodological tools to better manipulate the potency of cells in culture. During my PhD I first studied the metabolism of totipotent cells by measuring the mitochondrial membrane potential of 2-cell embryos and comparing it to pluripotent inner cell mass and differentiated trophoblast of blastocysts. My results show, that although these two embryo stages differ in their oxygen consumption and mitochondria matrix shape, there are no major changes in mitochondrial membrane potential. The main objective of my PhD was the characterization of the biophysical properties of heterochromatin during the process of chromocenter formation that occurs at the 2-cell stage. I first identified a set of core heterochromatic proteins and revealed their higher potential to phase separate based on an *in silico* analysis. Using a wide variety of microscopy techniques, I showed that pericentromeric heterochromatin transitions from a liquid state to a more solid or gel like state during the process of chromocenter formation. Overall, my work contributes to a better understanding of the features that characterise totipotency as well as developing state of the art tools to study the biophysical properties of constitutive heterochromatin at these early stages of development.

AIMS

PART 1: METABOLIC STATE OF TOTIPOTENT CELLS

- Assessing the mitochondrial membrane potential during mouse pre-implantation development

PART 2: ROLE OF BIOPHYSICAL PROPERTIES IN HETEROCHROMATIN FORMATION DURING MOUSE PRE-IMPLANTATION DEVELOPMENT

- Identify in silico new candidate proteins important for the biophysical properties of constitutive heterochromatin.
- Develop methods to study the biophysical properties of constitutive heterochromatin in the mouse pre-implantation embryo.
- Assess the changes of the biophysical properties of constitutive heterochromatin during the process of chromocenter formation.
- Study the functional relevance of the biophysical properties of constitutive heterochromatin.
- Investigate the in vivo role of the candidate proteins identified in silico.

INTRODUCTION

MOUSE PRE-IMPLANTATION DEVELOPMENT

Pre-implantation development starts when the oocyte is fertilised by the sperm to give rise to the 1-cell embryo called the zygote, containing the maternal and paternal genomes in 2 different pronuclei. The mouse embryo will then continue development until it implants in the mother's uterus 4 days after fertilisation at the late blastocyst stage. Pre-implantation development can be separated in 3 stages corresponding to the totipotent zygote and 2-cell (Tarkowski, 1959), the pluripotent 4- and 8-cell and the morula and blastocyst stage with the expression of embryonic markers like Nanog and extra-embryonic factors like Cdx2 (Condic, 2014) (Fig 1). The embryo will undergo compaction at the 8-cell stage before cavitating to form the blastocoel in the blastocyst. Concomitantly the first the lineage segregation takes place to form the pluripotent inner cell mass (ICM) and the differentiated trophectoderm (TE) (Fig 1).

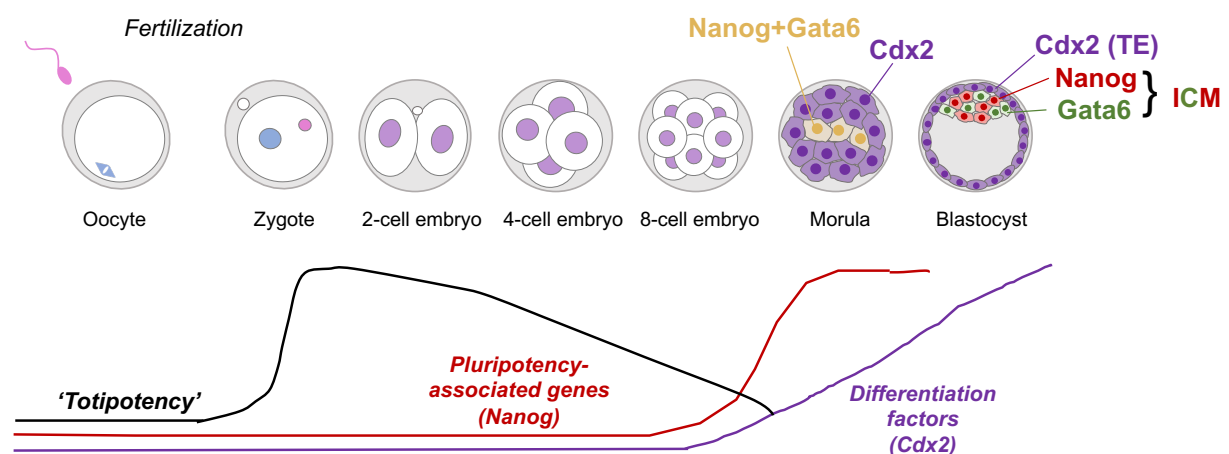


Figure 1. Overview of mouse pre-implantation development with the corresponding pattern of expression of lineage-associated transcription factors.

At fertilization, the embryo is transcriptionally silent and is dependent on proteins and RNA stocked in the oocyte. The genome of the embryo will be activated in two successive waves occurring in the zygote and the 2-cell stage embryo, namely the minor and major zygotic genome activation (ZGA) respectively (Mintz, 1964; Woodland and Graham, 1969). Another important feature of pre-implantation development is the expression of different families of transposable elements at specific stages (Hermant and Torres-Padilla, 2021; Rodriguez-Terrones and Torres-Padilla, 2018). For example, MERV-L (Peaston et al., 2004) and LINE1 (Fadloun et al., 2013) sub-families of transposable elements have a peak of expression at the 2-cell stage. These major changes in gene and transposon expression are accompanied by dramatic changes in epigenetic marks and nuclear architecture which will be described in

detail later in the introduction (Borsos and Torres-Padilla, 2016; Burton and Torres-Padilla, 2014).

METABOLISM STATE DURING PRE-IMPLANTATION DEVELOPMENT

The changes in cellular plasticity during pre-implantation development are accompanied by changes in the cellular metabolism of the embryos (Kaneko, 2016). Indeed, prior to the morula stage, the embryos rely on monocarboxylates such as pyruvate and lactate as their main source of energy through low-level oxidative phosphorylation (Brinster, 1965). On the other hand, at the morula stage the embryos switch to glucose as the major source of energy through glycolysis, which is also characteristic of the blastocyst stage (Houghton et al., 1996). Oxygen consumption also increases at the morula stage (Trimarchi et al., 2000). Interestingly, embryos cultured in vitro from the zygotic but not the 2-cell stage arrest at the morula stage in the absence of glucose in the culture media (Martin and Leese, 1995), indicating that a short uptake of glucose with unknown function is required at the zygotic stage. The ICM and the TE also differ in their metabolism in the blastocyst. The ICM, which has a high proliferation rate, primarily uses aerobic glycolysis, while the TE, which needs high levels of ATP in order to promote blastocoel formation, relies on anaerobic glycolysis (Hewitson and Leese, 1993). Embryonic Stem Cells (ESCs) derived from the ICM and induced pluripotent stem cells (iPSCs) also use aerobic glycolysis while differentiated cells rely on anaerobic glycolysis (Teslaa and Teitell, 2015). Finally, the mitochondria, which are the site of oxidative phosphorylation, are maternally inherited and do not start replicating until the blastocyst stage (St John et al., 2010). In the first part of my PhD, I contributed to the study of the metabolic differences between totipotent and pluripotent cells by comparing the mitochondrial membrane potential between the 2-cell and the blastocyst (Rodriguez-Terrones et al., 2020).

GENOME ORGANISATION IN PRE-IMPLANTATION DEVELOPMENT

Following the fusion of the differentiated oocyte and sperm, the embryo undergoes a massive epigenetic and chromatin organisation remodelling. The oocyte completes the second meiosis and forms the maternal pronucleus while the sperm-derived DNA/chromatin decondenses and exchange its protamines with maternally supplied histones to form the paternal pronucleus. During the first hours following fertilisation, contrary to the maternal chromatin, the paternal genome acquires histone modifications de novo. This creates an epigenetic asymmetry

between the two parental genomes that will slowly equilibrate until reaching the blastocyst stage (Morgan et al., 2005; Puschendorf et al., 2008; Santenard et al., 2010).

One example of the epigenetic asymmetry after fertilisation is the classical heterochromatin mark, trimethylation (me₃) of histone H3 at lysine 9 (H3K9me₃), which is present on the maternal but not on the paternal chromatin immediately after fertilisation. The paternal genome acquires low levels of H3K9me₃ during the zygotic stage although the asymmetry between the paternal and maternal genomes is still clearly visible at the 2-cell stage, in which the maternal and paternal genomes are still spatially segregated inside each nuclei (Burton et al., 2020; Wang et al., 2018a). The two parental genomes also differ in the histone variants they incorporate in their nucleosomes. At fertilisation the paternal genome rapidly incorporates the histone variant H3.3 in a replication-independent manner which is not the case for the maternal genome (Santenard et al., 2010; Torres-Padilla et al., 2006; van der Heijden et al., 2005). The two parental genomes also go through a major wave of DNA demethylation (Fig 2), although this occurs at different rates on the two genomes, which creates another level of epigenetic asymmetry (Mayer et al., 2000; Oswald et al., 2000; Santos et al., 2002).

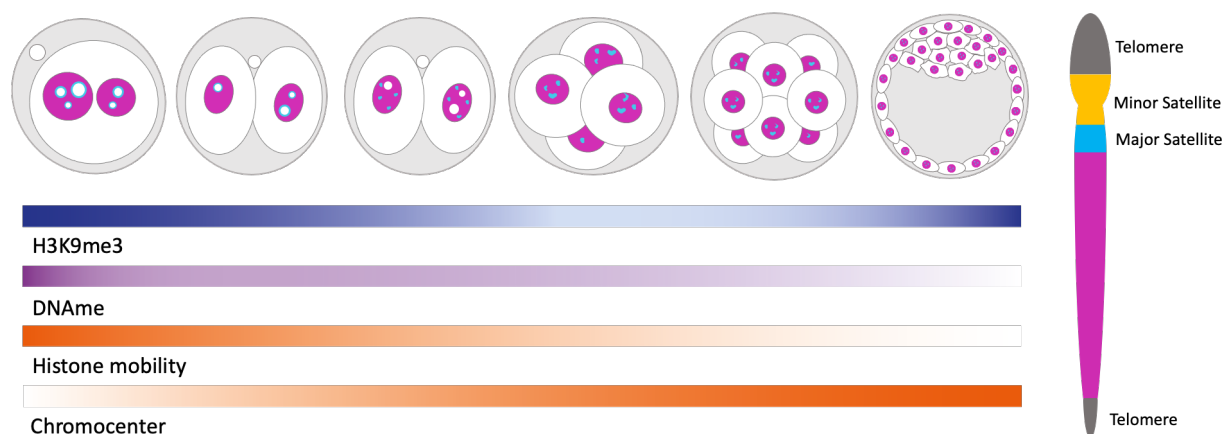


Figure 2. Major epigenetic and chromatin organisation changes occurring during pre-implantation development with the temporal presence of H3K9me₃ and DNA methylation, the histone mobility and the formation of the chromocenters. Major Satellite regions are shown in cyan to visualize their localisation during chromocenter formation. Right: Schematic representation of a chromosome with the Telomeres and Minor and Major Satellite regions highlighted.

In recent years new low-input genomics approaches have enabled major advances in understanding of chromatin architecture in the mouse pre-implantation embryo. First of all, several studies applied genome-wide chromosome conformation technique (Hi-C) to the different stages of the pre-implantation embryo. These studies showed that loop structure, topologically associated domains (TADs) and chromatin compartments are weak in zygotes and become stronger during pre-implantation development (Du et al., 2017; Flyamer et al.,

2017; Ke et al., 2017). Another important feature of chromatin architecture is the position of the genome with respect to the nuclear periphery. Indeed, lamina associated domains (LADs) have been shown in somatic cells to correlate with the heterochromatic B compartment (Guelen et al., 2008). Using a low-input DNA adenine methyltransferase identification (DamID) method, our laboratory has shown that LADs are formed de novo in the zygote before undergoing extensive rearrangements during pre-implantation development (Borsos et al., 2019). The zygotic LADs indeed resemble LADs typically observed in somatic cells, while several of the 2-cell LADs unusually belong to the A compartment. The LADs will continue to rearrange until the blastocyst stage, in which the LADs resemble those found in somatic cells. Finally, genome accessibility and chromatin mobility in the preimplantation embryo have been studied by assay for transposase-accessible chromatin using sequencing (ATAC-seq) (Wu et al., 2016) and fluorescence recovery after photobleaching (FRAP) (Fig 2) (Boskovic et al., 2014; Ooga et al., 2016) respectively. These studies showed that chromatin was highly accessible in 2-cell embryos, becoming more restricted later in development, coinciding with higher chromatin mobility in 2-cell compared to 8-cell embryos.

The zygote and 2-cell stage embryo also differ from somatic cells in the localisation and expression of the DNA that forms the pericentromeric chromatin (Probst et al., 2010; Probst et al., 2007). In somatic cells the major and minor satellite repeats, which constitute the pericentric and centric chromatin respectively, are clustered into chromocenters. These domains, which can be visualised by DAPI staining in mouse cells, are marked by H3K9me3 and the heterochromatin protein 1 α (HP1 α) and are transcriptionally silent. Interestingly, in the mouse embryo, pericentromeric chromatin is initially located around the nucleolar like bodies (NLBs), which are the precursors of the nucleoli in mouse pre-implantation embryos, and will only later form chromocenters between the late 2 and the 4-cell stage (Fig 2). This particular localisation is accompanied with active transcription of the major satellites and a different chromatin state marked by tri-methylation of lysine 27 of the histone variant H3.3 and HP1 β (Santenard et al., 2010). Furthermore, knocking down major satellite transcripts or perturbing major satellite localisation leads to defects in chromocenter formation and proper embryo development (Casanova et al., 2013; Jachowicz et al., 2013). During my PhD, I studied chromocenter formation in the mouse preimplantation embryo as a model to study de novo heterochromatin formation. For this, I implemented several imaging approaches, which relied mainly on the use of a fluorescently labelled transcription activator-like effector (TALE), which are sequence specific DNA-binding proteins, that was designed to bind specifically the major satellite repeats (Miyayari et al., 2013).

IN-VITRO MODEL TO STUDY TOTIPOTENCY

The discovery 10 years ago of the first in-vitro model for totipotency, the 2-cell-like cells (2CLCs), has proven invaluable for the characterisation of this state (Genet and Torres-Padilla, 2020). 2CLCs are a small population of cells arising spontaneously from ESC cultures, which were first identified based on the expression of the 2-cell specific MERV-L subfamily of transposable elements (Macfarlan et al., 2012). This cell population shares several of the features specific to totipotent embryos that were described previously in the introduction (Table 1). For example, their transcriptional profile is more similar to 2-cell embryos, most strikingly showing the up-regulation of 2-cell specific genes like *Duxf3* and *Zscan4* (De Iaco et al., 2017; Hendrickson et al., 2017; Ishiuchi et al., 2015; Macfarlan et al., 2012). Moreover, 2CLCs have several chromatin features in common with totipotent embryos, such as a higher histone mobility, the absence of chromocenters, the global reduction of DNA methylation and an increase in accessibility at MERV-L elements and promoters of genes upregulated in 2-cell embryos (Ishiuchi et al., 2015; Rodriguez-Terrones et al., 2018). Finally, 2CLCs have an expanded cell fate potential compared to ESCs as they can contribute to all lineages of the embryo, including extra-embryonic tissues, in chimera assays (Macfarlan et al., 2012) and they have a higher reprogramming efficiency upon nuclear transfer (Ishiuchi et al., 2015). The study of the differences in metabolism between totipotent and pluripotent cells presented in the first part of this thesis mainly relied on 2CLCs as a model system (Rodriguez-Terrones et al., 2020).

Features		2-cell-like cells	Embryonic stem cells
Transcription	Pluripotency markers	+/-	+++
	2-cell specific genes	+++	+/-
	Repetitive elements	+	-
Chromatin	Histone mobility	++	+
	Chromocenters	-	+
	DNA methylation	+	++
	Accessibility at MERV-L and promoters of 2-cell upregulated genes	+++	-

Table 1. Specific features of 2-cell-like cells compared to embryonic stem cells (adapted from (Genet and Torres-Padilla, 2020)).

LIQUID-LIQUID PHASE SEPARATION

GENERAL PRINCIPLES OF LIQUID-LIQUID PHASE SEPARATION IN BIOLOGY

Eukaryotic cells compartmentalise their different components to store them or to facilitate complex biochemical reactions in space and time. This is done in cells by packaging these components in different organelles where they are needed. We can subclassify these organelles into membrane-bound, which are surrounded by a lipid membrane, or membrane-less, which form via phase separation without any physical barrier (Banani et al., 2017). Several of these membrane-less organelles, also referred to as condensates, have been shown to assemble through liquid-liquid phase separation (LLPS) in which the compartment behaves like a liquid while being separated from its surrounding in the same way oil is separated from water. Based on their physical properties, liquid droplets should be spherical due to surface tension, be able to undergo fusion and fission and return to their original shape after deformation by shear stress. The internal components should be highly mobile but should remain constrained to these droplets (Fig 3). These membrane-less compartments play a role in enriching their internal components compared to the surrounding nucleoplasm or cytoplasm while excluding other factors. On the other hand, these membrane-less compartments are able to buffer the concentration of their internal components by modifying their volume (Fig 3).

Since the discovery that P granules form through LLPS more than a decade ago (Brangwynne et al., 2009), an increasing number of membrane-less compartments have been shown to have similar biophysical properties (Boeynaems et al., 2018). Among others, we can cite the nucleolus (Brangwynne et al., 2011), the stress granules (Molliex et al., 2015), and the paraspeckles (Hennig et al., 2015). LLPS has also been implicated in transcription initiation through the condensation of the C-terminal repeat of RNA polymerase II, transcription factors and the Mediator complex (Cho et al., 2018; Henninger et al., 2021; Sabari et al., 2018). It is noteworthy that membrane-less compartments can also form through liquid-gel phase separation or polymer-polymer phase separation, behaving as a gel or solid respectively and can transition between these two states (Frank and Rippe, 2020).

Properties		Assays
Sphericity		Droplet formation Assay in vitro
Fusion/Fission		Live imaging in vivo
High internal diffusion and preferential internal mixing		FRAP or Single particle tracking
Higher internal concentration ([I])		FCS methods (RICS, N&B)
Boundary constraining diffusion		FCS methods (RICS, N&B)
Concentration buffering		Imaging of compartment volume upon increase in concentration
Exclusion of non-interacting factors		Exclusion of eg dextran in live imaging
Dependence on weak hydrophobic interactions		Treatment with 1,6-hexanediol
Sensitivity to concentration, temperature, pH, salt levels,...		Generation of a phase diagram in vitro (can be done with optoDroplet system in vivo)

Figure 3. Table depicting the properties of LLPS and how they can be studied in vitro or in vivo. (a, b) The droplets should be spherical and be able to fusion and fission (arrow in cartoon) which can be visualised with a droplet formation assay in vitro or live imaging in vivo. (c) The internal components of these droplets should have a high diffusion and preferential internal mixing as these components are constrained to the condensate. This can be visualised by half-FRAP experiments (depicted in cartoon) or by single particle tracking. (d, e) FCS methods are used to measure the concentration of the components inside versus outside the condensate and the diffusion at the boundary. In LLPS, the concentration should be higher inside (depicted in top of cartoon) while the diffusion should decrease at the boundary of the condensate (depicted in the bottom of cartoon). (f) Liquid condensates should buffer the concentration of their internal components by changing their volume when the concentration varies (depicted in cartoon). This can be tested using microscopy by measuring the dependence of the volume of the domains to different protein concentrations. (g) The condensates should also exclude non-interacting factors (depicted in cartoon) which can be visualised in vivo by the exclusion of compounds like fluorescent dextrans. (h) Some condensates depend on weak hydrophobic interactions for their nucleation. In that case, they should be sensitive to treatment with 1,6-hexanediol (depicted in the cartoon). (i) Nucleation of condensates should be sensitive to scaffold concentration, but also to temperature, pH, salt levels and other factors. To test this, it is useful to generate phase diagrams by visualising the dependence of droplet formation to these parameters in vitro (depicted in cartoon).

LLPS is driven by different types of weak multivalent interaction between proteins and/or nucleic acids that are called scaffolds. These scaffolds can interact with other proteins, called clients, that are thus recruited to the condensate. Proteins involved in LLPS often contain multiple copies of interaction domains or contain intrinsically disorder regions (IDRs), which have been the most extensively studied presently (Banani et al., 2016; Wheeler and Hyman, 2018). These IDRs can be simplified as polymers containing stickers that mediate the interactions and which are separated from each other by spacers (Wang et al., 2018b). These

stickers can drive LLPS through a variety of interactions which have been extensively studied in recent years, namely electrostatic, hydrophobic (Fig 3), cation and Pi-Pi interactions (Kato et al., 2012; Martin et al., 2020; Pak et al., 2016; Vernon et al., 2018; Wang et al., 2018b). Building on this knowledge, several algorithms have been put forward to predict the phase separation propensity of proteins based on their amino acid sequence (van Mierlo et al., 2021; Vernon et al., 2018).

LIQUID-LIQUID PHASE SEPARATION IN CHROMATIN ORGANISATION

In recent years several studies focussed on the biophysical properties of chromatin in vitro and in cell culture. The aim of these studies was to determine if LLPS was a general property of chromatin which would drive the compartmentalisation of the genome. The first study, performed in vitro, showed that nucleosomal arrays can form liquid droplets under physiological conditions and that the length of the linker DNA, the incorporation of histone H1 and histone acetylation were key regulators of this process (Gibson et al., 2019). More recently, these conclusions have been challenged by another study that suggests that chromatin actually behaves as a solid using both nucleosomal arrays and fragmented native chromatin (Strickfaden et al., 2020). Finally, another study showed that the biophysical properties of chromatin in vitro depend on the length of the DNA fragment used in the assay with long DNA fragments behaving more solid-like while short ones had liquid-like properties (Muzzopappa et al., 2021). In vivo, the biophysical properties of chromatin are also debated. To assess the biophysical properties of chromatin in vivo, Strickfaden and colleagues performed FRAP experiments on fluorescently labelled DNA in living cells. Their results showed no recovery of chromatin both in eu- and heterochromatin suggesting that it behaves as a solid (Strickfaden et al., 2020). On the other hand, a recent study measured the viscoelastic properties of chromatin by applying magnetic forces inside the nucleus of interphase cells. Their results show rapid displacement of chromatin showing its high fluidity, suggesting liquid-like properties (Keizer et al., 2022). Hence the biophysical properties of chromatin in vivo are currently unclear. These differences could be explained by several factors such as cell cycle, culture medium and genomic chromatin compartment analysed.

Chromatin in vivo is however more complex than in vitro assembled nucleosomal arrays as there is a multitude of histone post-translational modifications and interacting molecules which can impact the biophysical properties of the different domains. For example, a subunit of the Polycomb-repressive complex 1 (PRC1) called Cbx2, has been shown to drive phase

separation of facultative heterochromatin in vitro and in vivo (Eeftens et al., 2021; Plys et al., 2019; Tatavosian et al., 2019). Similarly, several studies focused on the biophysical properties of constitutive heterochromatin with a particular interest in HP1 α . The first two studies on the biophysical properties of heterochromatin suggested that HP1 α could both phase separate in vitro and in the drosophila embryo in vivo (Larson et al., 2017; Strom et al., 2017). These HP1 α condensates have been shown to be highly dynamic while being able to compact chromatin in vitro. The hinge of HP1 α is necessary and sufficient to drive phase separation while the N and C terminal extensions play a role in regulating droplet formation (Keenen et al., 2021). Droplet formation of HP1 α in vitro is also regulated in vitro by phosphorylation of the its N-terminal extension, the presence of the other HP1 paralogs, namely HP1 β and HP1 γ , and the presence of major satellite RNA (Keenen et al., 2021; Larson et al., 2017; Novo et al., 2022). On the other hand, HP1 β can phase separate in the presence of histones, while binding to H3K9me3 increased droplet formation at physiological concentrations (Qin et al., 2021).

The biophysical properties of HP1 α in vivo are still under debate, however, as another study suggested that HP1 α , even if it can phase separate in vitro, does not show hallmarks of LLPS at pericentric heterochromatin of mouse fibroblasts. The authors from this work suggested that heterochromatin instead adopts a collapsed polymer globule state which is percolated by the nucleoplasm (Erdel et al., 2020). Interestingly, these biophysical properties have never been studied in the context of de novo heterochromatin formation. During my PhD I took advantage of the unique process of chromocenter formation during mouse pre-implantation development to study the biophysical properties of heterochromatin in vivo and in a dynamic system.

Finally, several additional proteins have been shown to play a role in the biophysical properties of heterochromatin. Scaffold attachment factor b (SAFB), for example, interacts with major satellite RNAs to drive its phase separation in vitro and to stabilise pericentromeric heterochromatin in vivo (Huo et al., 2020). Additionally, SUV39H1, a histone methyltransferase, also plays a role in the biophysical properties of constitutive heterochromatin by interacting with HP1 proteins and increasing the number of multivalent interactions with H3K9me3 via their chromodomains (Wang et al., 2019). During my PhD I performed an in silico analysis to find new potential candidate proteins that could play a role in heterochromatin establishment though LLPS. I also analysed the expression of these heterochromatic proteins during mouse pre-implantation development both at the RNA and protein level (Guthmann et al., 2019).

ASSAYS TO STUDY LIQUID-LIQUID PHASE SEPARATION

In vitro, LLPS can be assessed by visualising droplet formation under a microscope. The droplets formed should be spherical, be able to fuse and should have a high recovery rate after photobleaching (Fig 3). By changing the concentration of the protein of interest and a second parameter of the assay, for example the temperature, the salt levels or the pH, it is possible to plot a phase diagram and calculate a saturation concentration above which droplets start to form (Fig 3). Most of the studies in the field of LLPS start with *in silico* and *in vitro* assays to determine the domains and even amino acids of the protein of interest necessary for droplet formation (Alberti et al., 2019). The effect of mutating these amino acids on the localisation of the protein and membrane-less organelle stability can then be assessed in cells. On the other hand, studying the liquid state of a compartment of interest directly *in vivo* is experimentally more challenging due to the limited number of methods available and complexity of the system. Most of the methods used are microscopy-based and rely on the fluorescent tagging of the domain or the protein of interest. First, as for droplet formation assays *in vitro*, the compartment should have a spherical shape and be able to fuse and fission, which can be visualised by live-imaging (Fig 3). Second, FRAP and fluorescent correlation spectroscopy (FCS) methods enable the measurement of the dynamics of the internal components of the compartment of interest. Indeed, the diffusion of these components should be high inside the condensate, different from the surrounding nucleoplasm or cytoplasm and low at the boundary of the condensate due to an energetic barrier constraining diffusion (Fig 3) (Frank and Rippe, 2020; McSwiggen et al., 2019). FCS is also used to measure the concentration and enrichment of the internal components of the condensate compared to the surrounding nucleoplasm or cytoplasm (Fig 3) (Alberti et al., 2019). Third, The break-down of the condensate by treatment with 1,6-hexanediol, a drug that disrupts weak hydrophobic interactions, has also been extensively used to prove LLPS even if it is important to keep in mind that this drug can have multiple side effects and LLPS is not exclusively dependent on weak hydrophobic interactions (Fig 3) (Kroschwald et al., 2017). Fourth, it is possible to artificially to induce the formation of condensates of a protein of interest in cells, for example with the optoDroplet system and study the biophysical properties of these droplets *in vivo*. This technique works by overexpressing a fusion construct between the protein of interest and the photolyase homology region of Cry2, a protein that self-associates upon blue light exposure (Shin et al., 2017).

RESULTS

PART 1: METABOLIC STATE OF TOTIPOTENT CELLS

A DISTINCT METABOLIC STATE ARISES DURING THE EMERGENCE OF 2-CELL-LIKE CELLS

Statement of contribution

I hereby state that my contribution to the publication:

Rodriguez-Terrones, D., Hartlben, G., Gaume, X., Eid, A., **Guthmann, M.**, Iturbide, A., Torres-Padilla, M.E. (2020). A distinct metabolic state arises during the emergence of 2-cell-like cells. EMBO reports. 21: e48354.

consisted in measuring the mitochondrial membrane potential in mouse 2-cell and blastocyst embryos by JC-1 staining.

Manuel Guthmann
München, September 23rd, 2022

Confirmation of contribution

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

Prof. Dr. ~~Mařia-Elena~~ Torres-Padilla
München, September 23rd, 2022

A distinct metabolic state arises during the emergence of 2-cell-like cells

Diego Rodriguez-Terrones¹ , Götz Hartleben², Xavier Gaume¹, André Eid¹ , Manuel Guthmann¹, Ane Iturbide¹ & Maria-Elena Torres-Padilla^{1,3,*} 

Abstract

Pluripotent stem cells are thought of as a surrogate of early developmental stages that sustain the capacity to generate all cell types in the body, thereby constituting an invaluable tool to address the mechanisms underlying cellular plasticity. In the mouse, cells resembling totipotent 2-cell-stage embryos (2-cell-like cells) arise at a very low frequency in embryonic stem cell (ESC) cultures. However, the extent to which these early-embryonic-like cells recapitulate the molecular features of the early embryo is unclear. Here, we have undertaken a characterization of some of the metabolic features of early-embryonic-like cells in culture. Our data indicate that early-embryonic-like cells exhibit decreased glycolytic and respiratory activity, lower levels of reactive oxygen species and increased glucose uptake, suggesting a shift of the metabolic programme during 2-cell-like cell reprogramming. Accordingly, we find that 2-cell-like cells can be induced by defined metabolites. Thus, in addition to their transcriptional and chromatin features, 2-cell-like cells recapitulate some of the metabolic features of their *in vivo* counterpart. Altogether, our work underscores a distinct metabolic state of early-embryonic-like cells and identifies compounds that can induce their emergence *in vitro*.

Keywords 2-cell-like cells; metabolism; pluripotency; reprogramming; totipotency

Subject Categories Metabolism; Stem Cells & Regenerative Medicine

DOI 10.15252/embr.201948354 | Received 26 April 2019 | Revised 13 November 2019 | Accepted 15 November 2019 | Published online 18 December 2019

EMBO Reports (2020) 21: e48354

Introduction

The metabolic state of a cell is a key feature of cellular identity and has been linked to cellular plasticity. Shifts in metabolic pathways and reactive oxygen species (ROS) have been involved in reprogramming cell fate [1]. Likewise, the Warburg effect, whereby aerobic glycolysis becomes predominant over oxidative phosphorylation, is a well-known feature of cancer cells which is thought to

satisfy the altered metabolic demands that arise upon cellular transformation [2]. The importance of cellular metabolism during changes in cell fate is therefore beginning to emerge, particularly because it opens up the possibility to manipulate cell fate through inducing changes in metabolic programmes.

Pluripotent embryonic stem cells (ESCs) derived from the inner cell mass of the mouse blastocyst can self-renew indefinitely, provided appropriate culture conditions [3]. Mouse ESCs are pluripotent, since they have the capacity to generate all the cells in the body, including the germline, when transplanted into blastocysts to form chimera. ESC cultures are heterogeneous and are known to harbour different cellular states which vary depending on the culture conditions used [4–6]. Namely, when grown in serum and LIF, ESCs fluctuate between a naïve state, which is considered reminiscent of the pre-implantation epiblast, and a primed state, closer in nature to the post-implantation epiblast. The latter has a limited capacity to contribute to chimeras and the germline, compared to naïve ESCs [7]. These two cell populations recapitulate several molecular features of their *in vivo* counterparts, including their DNA methylation profiles [8], the expression of pluripotency markers [9] and their metabolic state [10]. Whereas naïve pluripotent stem cells rely on a mixture of glycolytic and aerobic metabolism, primed pluripotent stem cells rely almost exclusively on glycolysis to satisfy their energetic demands. In other words, naïve mouse ESCs respire more than the more primed EpiSCs [10]. Thus, there appears to be a link between the maintenance and loss of pluripotency, and the state of cellular metabolism.

In addition to the aforementioned heterogeneities of naïve and primed ESCs, cells resembling the blastomeres of the 2-cell-stage embryo have been documented to arise spontaneously in these cultures [11]. These “2-cell-like cells” constitute ~0.5% of the mouse ESC culture and display transcriptional and chromatin accessibility profiles highly similar to those in the 2-cell-stage embryo [11–13], as well as greater histone mobility [14] and dispersed chromocentres [15], all of which are molecular features characteristic of the 2-cell-stage embryo. In addition, 2-cell-like cells display expanded cellular potency and higher reprogrammability upon somatic cell nuclear transfer [11,15], underscoring their broader plasticity. Two-cell-like cells emerge from cells that express the

¹ Institute of Epigenetics and Stem Cells, Helmholtz Zentrum München, München, Germany

² Institute for Diabetes and Cancer, German Center for Diabetes Research, Helmholtz Zentrum München, Neuherberg, Germany

³ Faculty of Biology, Ludwig-Maximilians Universität, München, Germany

*Corresponding author. Tel: +49(0) 89 3187 3317; E-mail: torres-padilla@helmholtz-muenchen.de

transcription factor Zscan4 (Zscan4⁺ cells) [16], which are yet another subpopulation of ESC cultures constituting approximately 5% of the cell population [17,18]. Early-embryonic-like cells (Zscan4⁺ and 2-cell-like cells) can be induced in culture through the modulation of specific chromatin pathways, including the chromatin assembly factor 1 (CAF-1) [15] and the non-canonical polycomb repressive complex PRC1.6 [16,19], as well as the transcription factors Dux and Dppa2/4 [12,20–22].

Pre-implantation mouse embryos up to the 8-cell stage rely exclusively on monocarboxylates such as pyruvate and lactate to satisfy their bioenergetic needs [23–25]. This contrasts to morula and blastocyst-stage embryos, which rely on glucose to produce energy through a combination of glycolysis and oxidative phosphorylation [23,24]. Thus, there is a switch in central carbon metabolism as development proceeds, when the embryo transits from a totipotent, to a more restricted, pluripotent stage. Stem cells maintained *in vitro* may recapitulate some of their counterparts *in vivo*. However, it is unclear whether the different cellular heterogeneities in ESCs also reflect changes in metabolic pathways. In particular, whether 2-cell-like cells recapitulate some of the metabolic characteristics of 2-cell-stage embryos has not been investigated.

Here, we set out to investigate whether 2-cell-like cells display different metabolic features, compared to ESCs. We show that 2-cell-like cells display lower glycolytic and respiratory activity. Notably, this metabolic shift occurs in concert with a marked change in mitochondrial morphology, a significant reduction in ROS levels and a considerable increase in glucose uptake, suggesting a remodelling of metabolic activity upon 2-cell-like cells emergence. Importantly, Zscan4⁺ cells display mostly intermediate metabolic features, between 2-cell-like cells and ESCs, suggesting gradual metabolic reprogramming during the transition from ESCs to 2-cell-like cells. Finally, by carrying out a small-scale metabolite screen, we identified three compounds that promote the spontaneous

emergence of early-embryonic-like cells in a dose-dependent fashion. Overall, our data indicate that 2-cell-like cells transition into an overall “quiet” metabolic state and identify specific metabolites that induce them in culture.

Results and Discussion

We first interrogated our previously reported RNA-seq datasets [15] for changes in the expression levels of genes involved in metabolic regulation (Fig 1A). We compared expression levels between ESCs and 2-cell-like cells obtained through three distinct means, namely: spontaneously arising 2-cell-like cells (endogenous 2-cell-like cells) and CAF-1 knockdown-induced 2-cell-like cells obtained upon depletion of either of the two main subunits of CAF-1 (p60 and p150) [15]. We determined expression changes for most major central carbon metabolism enzymes and regulators, which we broadly classified into four groups: those involved in glycolysis, the TCA cycle, electron transport or glutamine metabolism (Fig 1A). Globally, while glycolytic enzymes displayed a tendency to be downregulated, we did not detect major changes in the expression of TCA cycle enzymes themselves (Fig 1A). In addition, several genes whose activity would be predicted to promote metabolic flux into the TCA cycle were upregulated in 2-cell-like cells, while others whose activity is known to strongly inhibit TCA cycle flux, including PDK and LDH, were downregulated (Fig 1A). Our analysis revealed that overall, 2-cell-like cells display marked differences in the expression levels of several enzymes and regulators involved in central carbon metabolism. These findings suggest potential changes in the metabolic activity of 2-cell-like cells compared to ESCs. Notwithstanding, because metabolic flux cannot be robustly predicted based on gene expression data alone—mainly due to the fact that metabolic changes are primarily regulated through

Figure 1. Early-embryonic-like cells exhibit decreased mitochondrial respiration.

- A Heatmap showing changes in RNA expression levels for various enzymes and regulators of central carbon metabolism in endogenous and in CAF1-knockdown-induced (through p60 or p150 KD) 2-cell-like cells. Fold-changes relative to ESCs were calculated based on bulk RNA-seq data [15].
- B Schematic representation of 2-cell-like cell emergence from ES cells, which transit through the intermediate Zscan4⁺ state before becoming 2-cell-like cells. Reporter constructs used to identify the three cell populations are shown.
- C Oxygen consumption rate of ES (blue line) and 2-cell-like cells (green line). Assay medium was formulated to recapitulate standard ES cell culture conditions and contained glucose, L-glutamine and pyruvate. Basal, maximal (FCCP-induced) and non-mitochondrial (rotenone and antimycin A-mediated) respiratory rates are indicated. A representative graph of three independent biological replicates performed on the Seahorse extracellular flux analyser is shown. Due to the low number of 2-cell-like cells available, compared to ESCs, one technical replicate of the former was analysed per biological replicate, while three or more technical replicates were performed for the latter. Accordingly, mean ± s.d. of technical replicates is shown for ESCs.
- D Basal oxygen consumption rate of ES (blue), Zscan4⁺ (red) and 2-cell-like cells (green) across three independent biological replicates performed on the Seahorse extracellular flux analyser. Assay medium was formulated to recapitulate standard ES cell culture conditions and contained glucose, L-glutamine and pyruvate. Boxes indicate the range between the first and third quartile, the band specifies the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the measurements obtained in each of the individual technical replicates.
- E Oxygen consumption rate of ES (blue line), Zscan4⁺ (red line) and 2-cell-like cells (green line) in glucose-free media and upon acute injection of sodium pyruvate or sodium L-lactate. Note that L-glutamine—but not glucose or pyruvate—was initially present in the assay medium. Maximal (FCCP-induced) and non-mitochondrial (rotenone and antimycin A-mediated) respiratory rates following pyruvate or lactate treatment are also indicated. A graph including data from three independent biological replicates is presented, and the mean ± s.d. of technical replicates is shown.
- F Representative electron micrographs of mitochondria from ES (*n* = 49 sections) and 2-cell-like cells (*n* = 57 sections) generated across two independent technical and biological replicates. Scale bar, 1 μm. Representative images from two independent biological and technical replicates are shown. 162 and 99 mitochondria were analysed for 2-cell-like and ES cells, respectively.
- G Representative single section of CellROX-DeepRed fluorescence in ES and 2-cell-like cells (green arrow) obtained using live-cell microscopy. Scale bar, 10 μm. Representative images from three independent biological replicates are shown.
- H FACS-assisted quantification of CellROX-DeepRed fluorescence intensity in ES and 2-cell-like cells. Measurements were obtained from two independent biological replicates. ****P* < 0.005; Mann–Whitney *U* test.

Source data are available online for this figure.

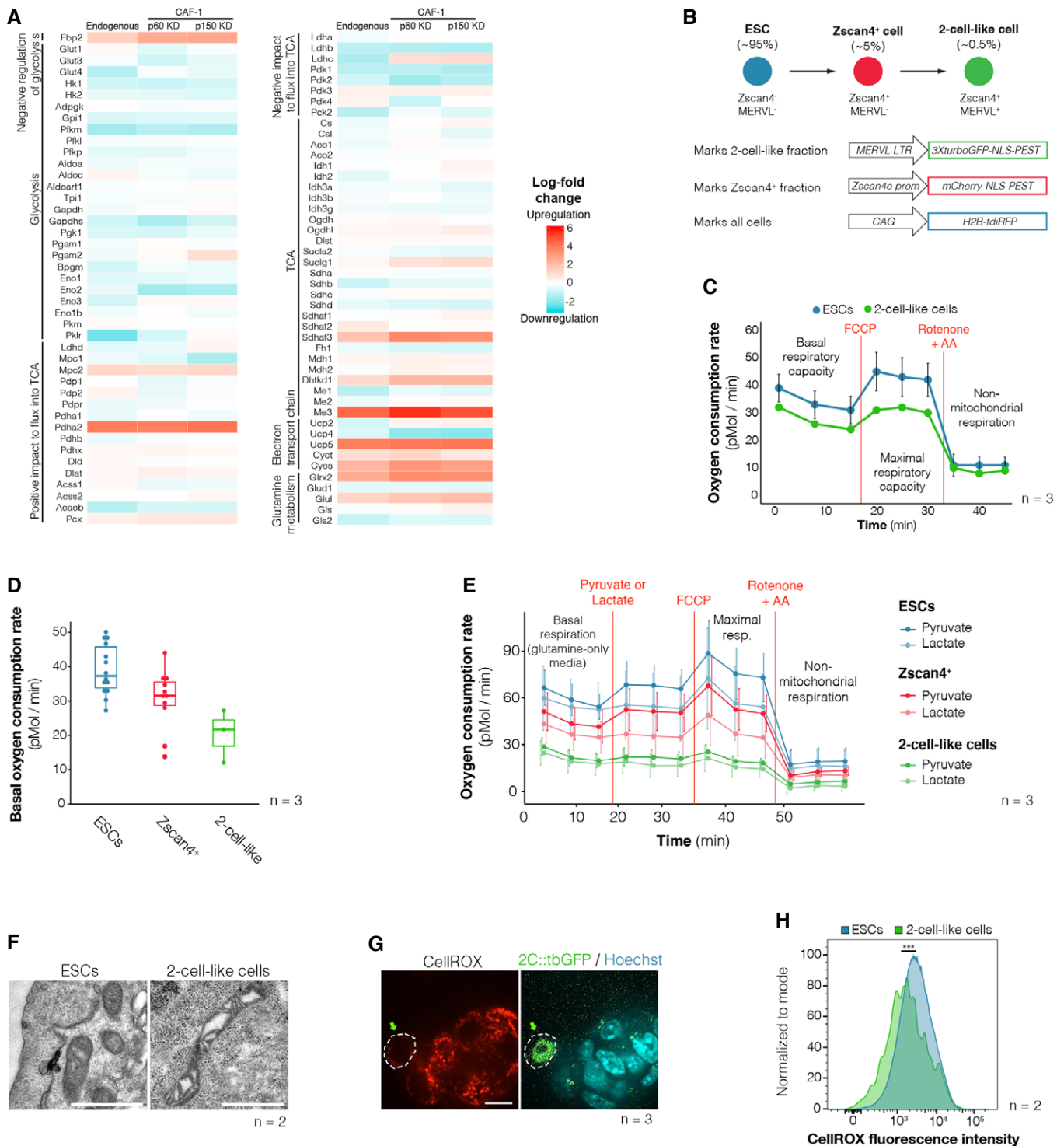


Figure 1.

modulation of enzymatic activity, for example through post-translational modifications and changes in substrate concentration—we set out to investigate the respiratory capacity of 2-cell-like cells directly.

Because of the limiting amounts of 2-cell-like cells available through cell sorting, we were unable to perform metabolomic profiling of these cells. Instead, to directly assess whether early-embryonic-like cells exhibit an overall distinct pattern of metabolic

activity than ESCs, we first measured oxygen consumption in 2-cell-like cells. We used the Seahorse extracellular flux analyser to measure the oxygen consumption rate (OCR) in live cells and determine basal, maximal and non-mitochondrial respiration. We optimized conditions for low cell numbers, which we set at 50,000 cells per well. Using a previously described reporter cell line (Fig 1B and Tables 1 and 2), we FACS-sorted equal numbers of ESCs (Zscan4c::

Table 1. Reporter cell lines used in this study.

Cell line	Green channel	Red channel	Far-red channel	Described in
tbg4	2C::3XturboGFP-NLS-PEST			Ishiuchi et al [15]
tbg4-12	2C::3XturboGFP-NLS-PEST	CAG::NLS-tdTomato (constitutive)		Rodriguez-Terrones et al [16]
tbg4ZH	2C::3XturboGFP-NLS-PEST	Zscan4c::mCherry-NLS-PEST	CAG::H2B-tdiRFP (constitutive)	Rodriguez-Terrones et al [16]
2C-EGFP	2C::EGFP			Ishiuchi et al [15]
Rex1-Zscan4	Rex1::EGFP-PEST (knock-in)	Zscan4c::tdTomato-PEST		Rodriguez-Terrones et al [16]
2C-tdTomato		2C::tdTomato	CAG::H2B-tdiRFP (constitutive)	Ishiuchi et al [15]

mCherry⁻, 2C::tbGFP⁻) and 2-cell-like cells (Zscan4c::mCherry⁺, 2C::tbGFP⁺) and profiled them on the Seahorse analyser (Fig EV1A and B). We used medium containing 25 mM glucose, 1 mM pyruvate and 2 mM glutamine, which is equivalent to the standard concentrations in ESC culture medium. Two-cell-like cells displayed a lower basal oxygen consumption rate compared to ESCs (Fig 1C), indicating a decrease in mitochondrial respiration. When challenged with FCCP, which uncouples the proton gradient from oxidative phosphorylation in mitochondria to reveal maximal respiratory capacity, ESCs augmented their oxygen consumption rate (Figs 1C, and EV2A and B). However, this was not the case for 2-cell-like cells, which remained at similar OCR levels compared to basal conditions (Fig 1C, and EV2A and B). We observed no differences in OCR between these two cell types after rotenone addition, suggesting similar levels of extra-mitochondrial oxygen consumption rates (Fig EV2A and B). Interestingly, Zscan4⁺ cells (Zscan4c::mCherry⁺, 2C::tbGFP⁻) displayed intermediate levels of basal and maximal respiratory capacity, compared to ESCs and 2-cell-like cells (Figs 1D, and EV2A and B), in agreement with their intermediate nature during the transition of ESCs to the 2-cell-like state [16]. Of note, oligomycin treatment was not tolerated by 2-cell-like cells in this experimental set-up, which prevented us from determining the levels of ATP-linked respiration. Altogether, these results indicate that 2-cell-like cells display lower cellular respiratory capacity than ESCs and that in basal conditions, 2-cell-like cells respire at maximal capacity.

The lower respiratory capacity of Zscan4⁺ and 2-cell-like cells compared to ESCs prompted us to investigate whether this decrease could be attributed to different substrate preferences in Zscan4⁺ and 2-cell-like cells. Because 2-cell-stage embryos rely on monocarboxylates such as pyruvate or lactate, we next measured the respiratory response of ES, Zscan4⁺ and 2-cell-like cells to acute supplementation of these two metabolites. For these experiments, we used medium without glucose and pyruvate, but containing L-glutamine to sustain a basal level of respiration. ES and Zscan4⁺ cells increased their oxygen consumption upon pyruvate supplementation, but 2-cell-like cells did not (Figs 1E and EV2C). None of the three cell types increased their oxygen consumption rate upon lactate supplementation (Fig EV2D). These observations may reflect the inability of 2-cell-like cells to take up exogenous pyruvate and/or the fact that 2-cell-like cells are already respiring at maximal capacity.

The observation that maximal respiratory capacity decreases in 2-cell-like cells raised the possibility that mitochondrial architecture might change upon reprogramming to the 2-cell-like state. To test this hypothesis, we examined mitochondrial morphology in ESCs and 2-cell-like cells by electron microscopy (Figs 1F, and

EV3A and B). Two-cell-like cells contained a larger proportion of elongated mitochondria in comparison with ESCs (25% of mitochondria were longer than 1.5 μ m in 2-cell-like cells, versus only 9% in ES cells, $n = 162$ and 99 mitochondria, respectively). Instead of the more developed cristae typical of serum/LIF-grown ESCs [10,26], 2-cell-like cells exhibited mitochondria with a matrix that was electron poor and tended to exhibit irregularly folded cristae (Fig 1F), in agreement with their overall lower maximal respiratory capacity. Mitochondria with irregularly folded cristae have been associated with lower oxygen consumption [27,28]. The electron micrographs also suggested increased vacuolization in the cytoplasm of 2-cell-like cells, which prompted us to measure autophagy. We found that 2-cell-like cells display slightly higher levels of autophagic vesicles, as measured by Cyto-ID fluorescence (Fig EV3C–E). However, this difference was not statistically significant. The analysis of additional autophagic markers is necessary to address the biological relevance of these changes. In agreement with previous reports [28], electron micrographs of pre-implantation embryos showed that mitochondria in the zygote and 2-cell-stage embryo also possess an electron-poor matrix with concentrically organized cristae around it (Fig EV4A and B), as opposed to the more opaque matrix observed already at the 8-cell stage (Fig EV4C) or the transverse cristae observed in blastocysts [28] and ES cells (Figs 1F and EV3A). While the mitochondria of 2-cell-like cells also display increased electron-poor matrix volume, they may not fully recapitulate the mitochondrial morphology of the 2-cell embryo. Further studies will be needed to assess the biological relevance of these changes. Indeed, despite the known differences in mitochondrial morphology between 2-cell embryos and blastocysts, we did not detect changes in mitochondrial membrane potential, as assayed using JC-1 staining (Fig EV4D and E).

As an additional indicator of mitochondrial activity, we measured levels of reactive oxygen species (ROS) in 2-cell-like cells, since altered ROS levels are often indicative of altered respiration [29,30]. For this, we incubated the 2C::tbGFP cell line with CellROX, a ROS-sensitive fluorescent probe which detects the oxidative species HO \cdot and \cdot O $_2^-$, and measured fluorescence intensity in 2-cell-like (2C::tbGFP⁺) and ES (2C::tbGFP⁻) cells by direct visualization using confocal microscopy (Fig 1G). CellROX staining was heterogeneous in mouse ESCs, but 2-cell-like cells clearly displayed an overall lower reactivity to CellROX, indicating lower ROS accumulation (Fig 1G). FACS analysis confirmed these results quantitatively, indicating reduced ROS levels in 2-cell-like cells (Figs 1H and EV1C), in line with the reduced respiratory activity of these cells.

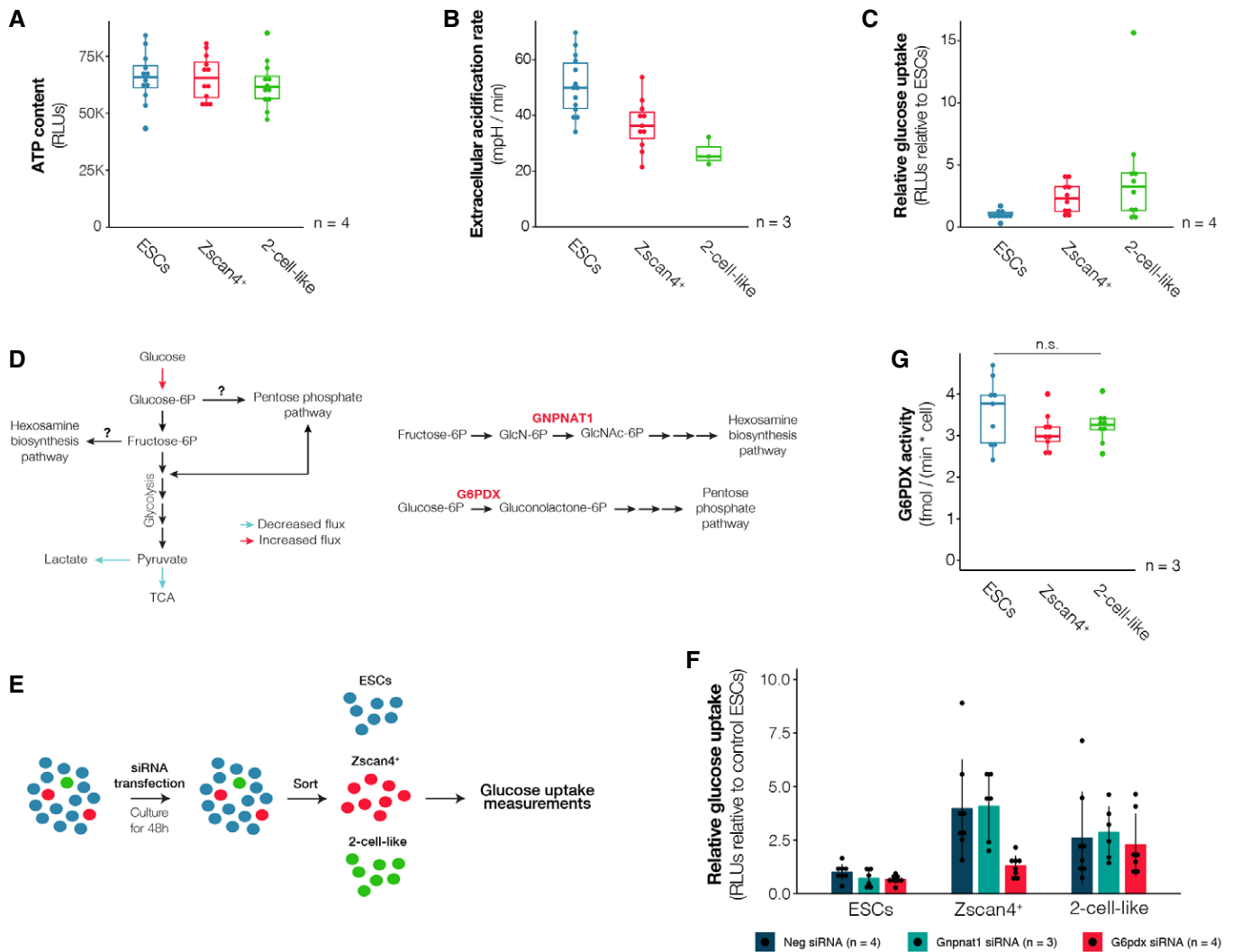


Figure 2. Increased glucose uptake supports higher flux into the pentose phosphate pathway in Zscan4⁺ cells.

- A** ATP content in ES (blue), Zscan4⁺ (red) and 2-cell-like cells (green) across four independent biological replicates.
- B** Extracellular acidification rate of ES (blue), Zscan4⁺ (red) and 2-cell-like cells (green) across three independent biological replicates performed on the Seahorse extracellular flux analyser.
- C** Glucose uptake rates in Zscan4⁺ (red) and 2-cell-like cells (green) were measured using a luciferase-based assay across four independent biological replicates and are represented relative to those of control ES cells (blue).
- D** Schematic representation of measured fluxes (left). In order to ascertain whether the increased glucose uptake observed leads to higher flux into the hexosamine biosynthesis pathway (HBP) or the pentose phosphate pathway (PPP), one enzyme of each pathway was disrupted through siRNA-mediated knockdown (right).
- E** Experimental design. ESC cultures were transfected with siRNAs targeting Gnpnat1 (HBP), G6pdx (PPP) or a negative control siRNA (Neg). After 48 h of culture, cells were FACS-sorted into a 96-well plate based on their fluorescent reporters and glucose uptake rates were measured using a luciferase-based assay.
- F** Glucose uptake rates upon knockdown of Gnpnat1 or a G6pdx were measured in ES, Zscan4⁺ and 2-cell-like cells. Measurements were quantified relative to the glucose uptake rate of ESCs transfected with a negative control siRNA. Shown are the mean \pm s.d. of the indicated number of independent cell cultures, performed across 2 or more independent biological replicates each.
- G** Glucose-6-phosphate dehydrogenase activity was measured in ES (blue), Zscan4⁺ (red) and 2-cell-like cell (green) lysates using a fluorometric assay. Measurements were obtained from three independent biological replicates, performed in three technical replicates each. n.s.—not significant; one-way ANOVA.

Data information: In panels (A–C and G), boxes indicate the range between the first and third quartile, the band depicts the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the measurements obtained in each individual technical replicate. Source data are available online for this figure.

Given the decreased respiration and overall lower mitochondrial activity in 2-cell-like cells, we next addressed whether ATP levels might be compromised in 2-cell-like cells. We FACS-sorted equal numbers of ES, Zscan4⁺ and 2-cell-like cells and measured ATP

levels using a luciferase-based assay. Unexpectedly, we did not detect significant changes in ATP levels in any of the three cell populations (Fig 2A). Our observation that 2-cell-like cells display similar ATP levels to ESCs suggests that the lower respiratory

activity observed in 2-cell-like cells is compensated by either decreased energy expenditure or increased glycolytic activity. To discern between the above possibilities, we next determined glycolytic activity in 2-cell-like cells by measuring the extracellular acidification rate and the glucose uptake rate of ESCs, *Zscan4*⁺ and 2-cell-like cells. Extracellular acidification is mainly the result of glycolytic activity and arises through the excretion of lactic acid, one of the major glycolytic end products, to the extracellular media. Under standard ESC culture medium conditions, *Zscan4*⁺ and 2-cell-like cells exhibited a lower extracellular acidification rate than ESCs, suggesting lower glycolytic output (Fig 2B). Surprisingly, however, glucose uptake rate measurements in all three populations indicated that *Zscan4*⁺ and 2-cell-like cells exhibited higher rates of glucose uptake than ESCs (Fig 2C), an observation at odds with the lower extracellular acidification rates measured in these two cell populations, which suggests an alternative, non-glycolytic fate for the consumed glucose. Because *Zscan4*⁺ and 2-cell-like cells derive primarily from naïve ESCs [16], we addressed whether the observed differences in glucose uptake between ESCs and 2-cell-like cells are related to their preferential origin from the naïve ESC state. Mouse ESCs exist in two metastable states—naïve and primed—, which differ in their metabolic state [10]. Naïve cells express high levels of the transcription factor *Rex1* (*Rex1*^{high} ESCs) [4,9] and rely on a mixture of glycolytic and aerobic metabolism. In contrast, primed pluripotent stem cells express low levels of *Rex1* (*Rex1*^{low}) and rely almost exclusively on glycolysis to satisfy their energetic demands [10]. Glucose uptake tends to be higher in ESCs grown in 2i—where ESCs are primarily in a naïve, *Rex1*^{high}—compared to serum/LIF conditions—in which ESCs cycle between *Rex1*^{high} and *Rex1*^{low} pluripotency states [31]. We FACS-sorted equal numbers of *Rex1*^{high} ESCs, *Rex1*^{low} ESCs and *Zscan4*⁺ cells and measured glucose uptake as before using a luciferase-based assay (Fig EV5A and B). We find that *Zscan4*⁺ cells exhibited higher glucose uptake than either primed or naïve cells, suggesting that the differences in glucose uptake between ESCs and early-embryonic-like cells are not related to their pluripotent state (Fig EV5C).

Our observations above, indicating lower mitochondrial respiration and lower lactate production in 2-cell-like cells, are at odds with their higher glucose uptake and suggest an alternative non-glycolytic fate for the bulk of the consumed glucose (Fig 2C). Therefore, we asked whether 2-cell-like cells divert their intracellular glucose towards other pathways such as the hexosamine biosynthetic pathway (HBP) or the pentose phosphate pathway (PPP). To address this, we downregulated each of these two pathways using RNAi—in such a way that flux through the corresponding pathway would be stalled upon knockdown of the targeted enzyme (Fig 2D and E)—and measured glucose uptake 48 h later. Because of the extended culture period required to achieve an efficient knockdown (Fig EV5D and E) and in order to maintain cellular viability, we were unable to include any glycolytic enzymes as a positive control in this assay. Downregulation of *Gnpat1*, which catalyses the transfer of an acetyl group from Ac-CoA to glucosamine-6-phosphate, did not reduce glucose uptake in either *Zscan4*⁺ or 2-cell-like cells, relative to the negative control siRNA (Figs 2F and EV5D). In contrast, RNAi for *G6pdx*, which catalyses the first and rate-limiting reaction of the oxidative branch of the pentose phosphate pathway (PPP), resulted in a decrease in glucose consumption of ESCs (−28%) and *Zscan4*⁺ (−65%), but barely affected the glucose

uptake of 2-cell-like cells (−13%), which remained mostly unchanged (Figs 2F and EV5E). Importantly, we did not detect any changes in G6PDX activity in lysates from the three cell types (Fig 2G). Thus, it would seem that increased glucose uptake supports higher flux into the pentose phosphate pathway in *Zscan4*⁺ cells. Altogether, our results indicate that early-embryonic-like cells exhibit decreased glycolytic and respiratory activity, altered mitochondrial morphology and increased glucose uptake, suggesting a shift of the metabolic programme during reprogramming to the 2-cell-like state.

Given the observed changes in metabolic activity described above, we hypothesized that the addition of specific metabolites may alter the number of early-embryonic-like cells present in mouse ESC cultures. Thus, we next addressed whether the number of *Zscan4*⁺ and 2-cell-like cells is affected upon addition of specific metabolites to the medium (Fig 3A). We incubated our double reporter cell line (*Zscan4*::mCherry, 2C::tbGFP) with varying concentrations of 20 selected metabolites for 48 h and quantified the number of *Zscan4*⁺ and 2-cell-like cells in each of these conditions using FACS (Fig 3B, Table 4 and Table EV1). Overall, we identified three metabolites that displayed a robust induction of 2-cell-like cells in a dose-dependent manner. These included sodium L-lactate, D-ribose and sodium acetate. Sodium acetate displayed the strongest effect in inducing the 2-cell-like cell population (Fig 3B), which reached up to 8% of the culture at the maximum dose applied (Fig 3C). Similarly, addition of D-ribose or sodium L-lactate also induced the 2-cell-like cell population in a dose-dependent manner. Addition of sodium acetate, sodium L-lactate or D-ribose also resulted in a clear induction of *Zscan4*⁺ cells—which reached up to ~60% of the total cell population at the highest sodium acetate dose applied—suggesting that these metabolites induce *bona fide* 2-cell-like cells (Fig 3D). Combining D-ribose with sodium acetate or sodium L-lactate resulted in an increased number of *Zscan4*⁺ and 2-cell-like cells compared to sodium acetate or sodium L-lactate alone (Fig 3E and F). However, addition of sodium acetate and sodium L-lactate together did not cause additive effects on the number of either *Zscan4*⁺ or 2-cell-like cells (Fig 3E and F). These results suggest that sodium acetate and sodium L-lactate may induce 2-cell-like cells through the same pathway.

Metabolite-induced 2-cell-like cells displayed the same molecular features of endogenous 2-cell-like cells, namely increased levels of ZSCAN4, loss of chromocentres and loss of OCT4 protein (Fig 4A). In addition, sodium acetate treatment induced a robust increase in *Zscan4* and *MERVL* transcripts (Fig 4B), although levels of L1 and IAP remained unchanged (Fig EV5F), consistent with the known transcriptional features of 2-cell-like cells [11,15]. The transcription factor DUX, which has been recently shown to bind to and regulate *MERVL* expression [12,22,32], was also upregulated upon acetate treatment (Fig 4B), and induction of 2-cell-like cells by acetate was significantly reduced upon Dux siRNA transfection (Fig 4C). In addition, acetate incubation led to a synergistic effect in the induction of 2-cell-like cells, when combined with siRNA for specific chromatin modifiers known to induce 2-cell-like cells (Fig 4D) [16]. This suggests that sodium acetate induces 2-cell-like cells through parallel pathways to those of the chromatin modifiers tested. The induction of 2-cell-like cells by acetate is in line with the known increase in levels of histone acetylation in 2-cell-like cells [11,15]

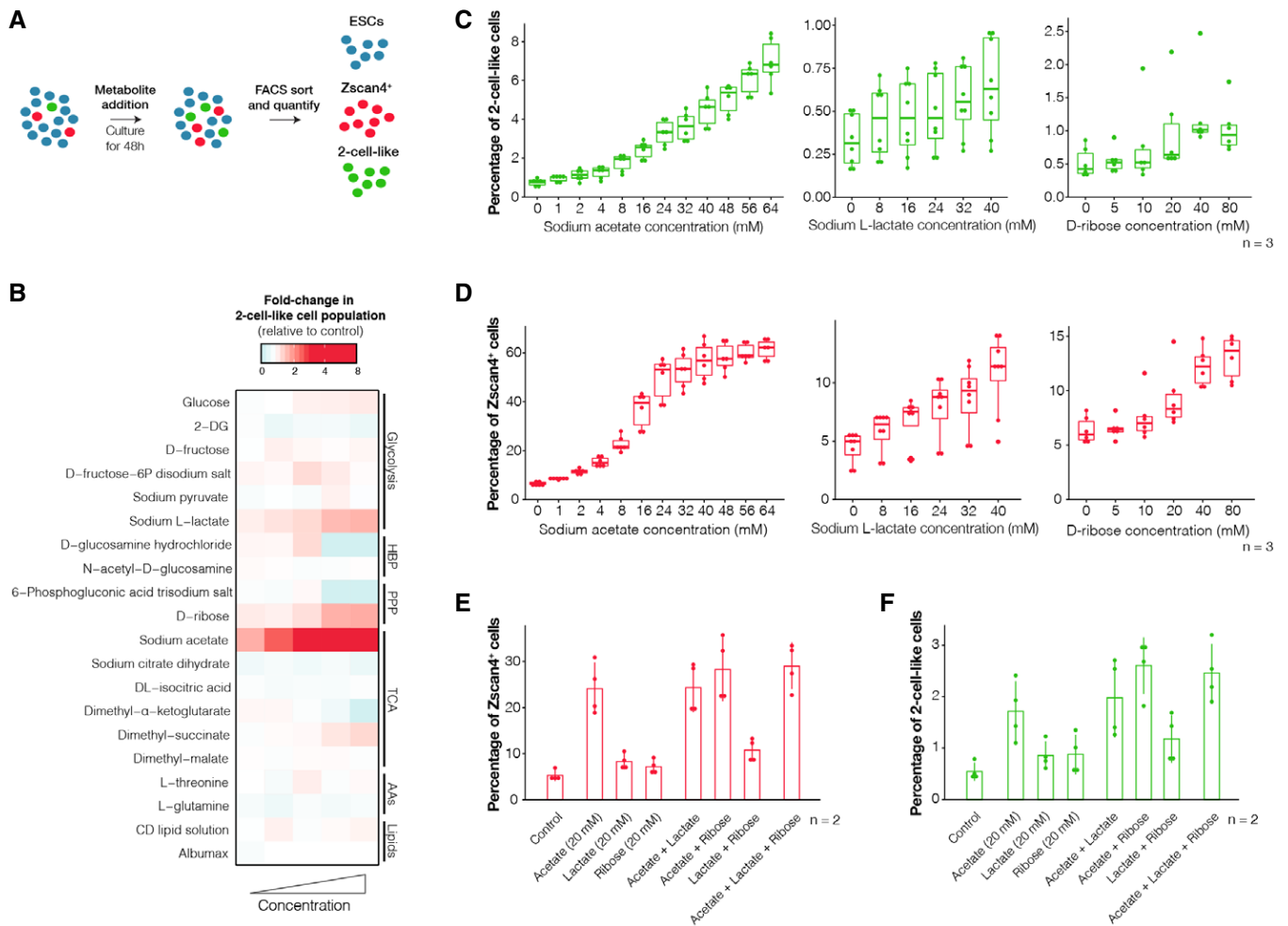


Figure 3. Induction of Zscan4⁺ and 2-cell-like cells by defined metabolites.

A Experimental design. ESC cultures were supplemented with increasing concentrations (see Table EV1) of 20 distinct metabolites for 48 h. The abundance of the Zscan4⁺ and 2-cell-like cells was quantified by FACS according to the reporters shown in Fig 1B.

B Heatmap showing the effect of the 20 metabolites tested on 2-cell-like cell levels.

C–F Percentage of Zscan4⁺ and 2-cell-like cells in cultures treated with increasing concentrations (C and D) or varying combinations (E and F) of sodium acetate, sodium L-lactate or D-ribose. Shown are the mean \pm s.d. of two independent cell cultures, performed across 3 (C and D) or 2 (E and F) biological replicates. Boxes indicate the range between the first and third quartile, the band depicts the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the measurements obtained in each of the 6 (C and D) or 4 (E and F) technical replicates.

Source data are available online for this figure.

and suggests that at least part of the effect observed upon acetate supplementation might be linked to increased levels of histone acetylation. Indeed, we find that acetate supplementation led to increased global levels of histone acetylation (Fig EV5G), consistent with previous reports [33]. We also addressed whether sodium acetate and L-lactate increase the 2-cell-like cell population by promoting either maintenance or induction using time-lapse microscopy with a Zscan4 reporter [16] after removal of Zscan4⁺ and 2-cell-like cells. Our results suggest that both sodium acetate and L-lactate induce rather than stabilize Zscan4⁺ cells (Fig 4E and F). Further studies are needed to determine the mechanism through which these metabolites induce 2-cell-like cells, which may encompass metabolic as well as epigenetic mechanisms. Lactate, for example, may act through glycolytic

metabolism, but potentially also through the inhibition of HDACs [34]. Thus, we conclude that 2-cell-like cells display a global metabolic shift compared to ES cells and that specific metabolites can induce the emergence of 2-cell-like cells in culture.

Overall, our work shows that early-embryonic-like cells (Zscan4⁺ and 2-cell-like cells) differ in their metabolic activity from ESCs. Similarly to the 2-cell-stage embryo [25,28,35–37], 2-cell-like cells seem to exhibit a “quiet” metabolism, characterized by low glycolytic and respiratory activity, as well as altered mitochondrial morphology and lower ROS production (Fig 5). However, our results also show that some differences between early-embryonic-like cells and early embryos exist, most notably in terms of their substrate uptake rates. Intriguingly, we observed that early-embryonic-like cells are characterized by higher glucose uptake

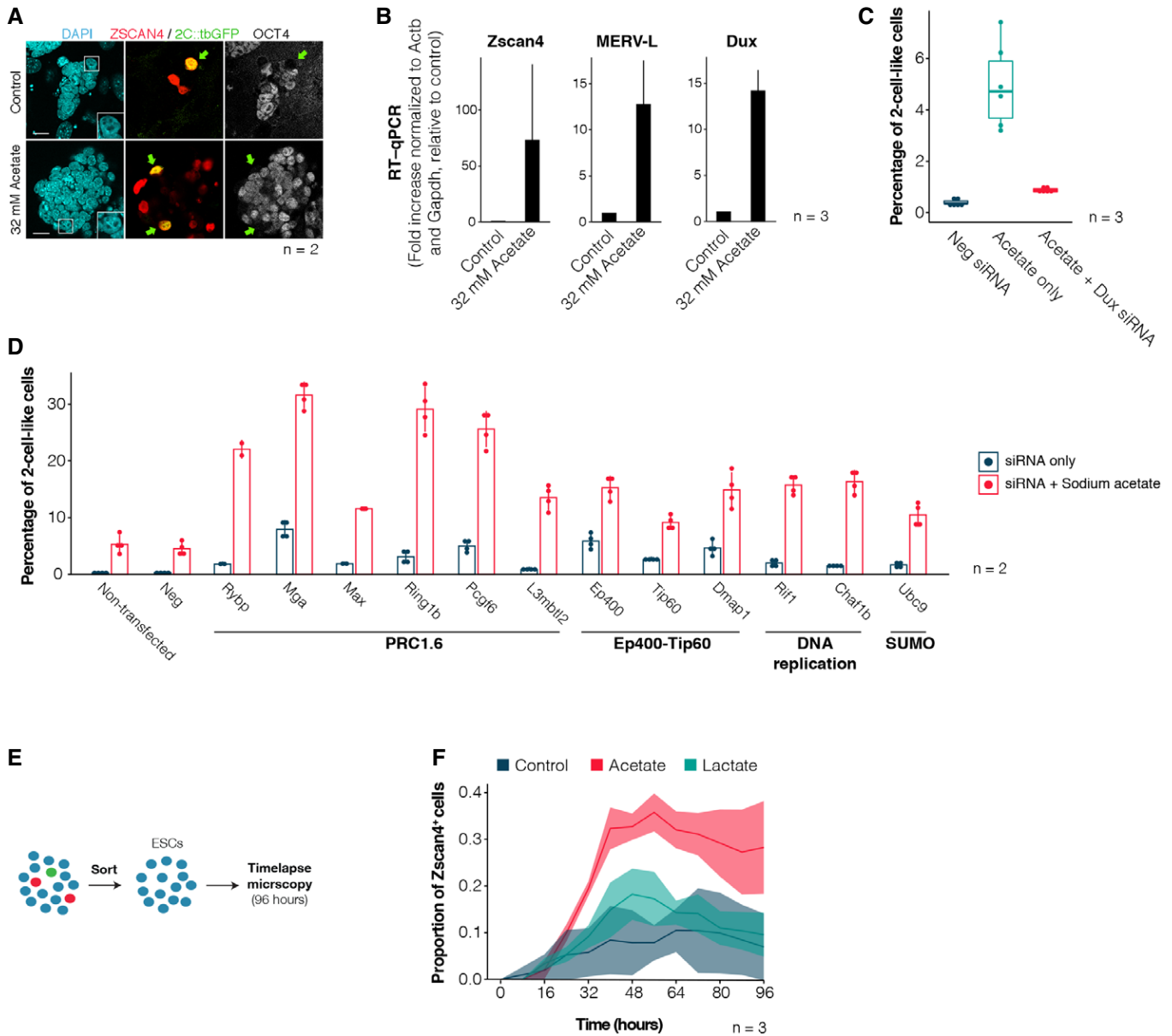


Figure 4. Sodium acetate induces Zscan4⁺ and 2-cell-like cells.

A Immunofluorescence staining for OCT4, ZSCAN4 and 2C::tbGFP in control and acetate-treated ESC cultures. Green arrows indicate 2-cell-like cells, and inlets highlight their DAPI structure. Scale bar, 20 μ m.

B RT-qPCR of the indicated genes in ESC cultures treated with sodium acetate for 24 h. Shown are the mean \pm s.d. of three independent cell cultures, performed in two technical replicates.

C Percentage of 2-cell-like cells obtained upon transfection of control or Dux-targeting siRNAs in control conditions or in combination with sodium acetate treatment. Measurements were obtained from two independent cell cultures, performed across three independent biological replicates. Boxes indicate the range between the first and third quartile, the band depicts the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the measurements obtained in each technical replicate.

D Percentage of 2-cell-like cells obtained upon transfection siRNAs targeting the indicated chromatin factors in control conditions or in combination with sodium acetate treatment. Shown are the mean \pm s.d. of four independent cell cultures, performed across two independent biological replicates.

E Experimental design. ESC cultures were FACS-sorted to remove Zscan4⁺ and 2-cell-like cells and plated in a glass bottom 96-well plate. Cells were then imaged for 96 h in the presence of sodium acetate (32 mM), sodium L-lactate (32 mM) or in control conditions.

F Proportion of Zscan4⁺ cells at various timepoints during the time-lapse experiment. Shown are the mean \pm s.d. of three independent experiments.

Source data are available online for this figure.

rates than ESCs and the inability to increase respiration using exogenous pyruvate. It is unclear to what extent such differences between early embryos and early-embryonic-like cells reflect differences in

culture conditions or a fundamental difference in metabolic requirements between a transient totipotent embryo and the self-renewing pluripotent state from which early-embryonic-like cells arise.

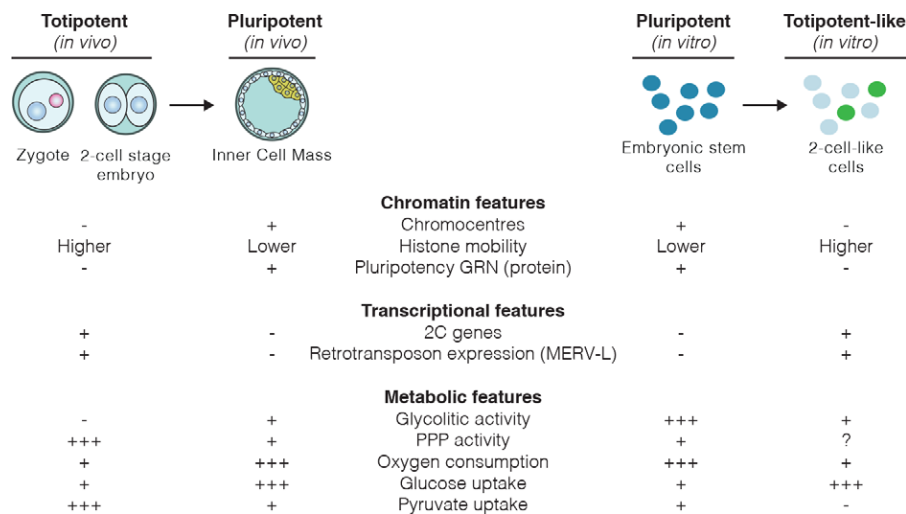


Figure 5. Summary of known molecular features of totipotent and totipotent-like cells.

Overview of the known molecular characteristics of totipotent and pluripotent cells *in vivo* and *in vitro*, now including metabolic features.

Morula and blastocyst-stage embryos rely on glucose to produce energy through a combination of glycolysis and oxidative phosphorylation, while cleavage embryos up to the 8-cell stage rely exclusively on monocarboxylates such as pyruvate and lactate [23]. These changes in substrate requirements have been proposed to reflect the embryo's need to provide sufficient supplies of specific metabolites—such as acetyl-CoA or α -ketoglutarate—that are required for the activation of the embryonic genome [38], while maintaining an overall low metabolic activity in order to restrict ROS production and oxidative damage [35]. Intriguingly, early-embryonic-like cells upregulate glucose uptake although without the consequent, expected increase in lactate production, suggesting alternative fates for the extra glucose consumed. PPP disruption led to a significant decrease in glucose consumption in Zscan4⁺ cells, but not in 2-cell-like cells, suggesting high glucose flux into this pathway in Zscan4⁺ cells. Other pathways might be active in 2-cell-like cells and may contribute to the high glucose consumption seen in 2-cell-like cells. These may include respiration-uncoupled pyruvate metabolism into acetyl-CoA or PPP activity through the non-oxidative branch. The former is supported by the fact that acetate strongly increases 2-cell-like cells in culture, which display higher levels of global histone acetylation compared to ESCs [11,15].

Materials and Methods

Cell culture

All cell lines used in this study, unless otherwise stated, were grown in media containing DMEM-Glutamax-I, 15% foetal calf serum, 2× LIF, 2-β-mercaptoethanol, non-essential amino acids, penicillin and streptomycin on feeders. Medium supplemented with 2i (3 μM CHIR99021 and 1 μM PD0325901, Miltenyi Biotec) was used for the establishment of stable cell lines and for their expansion and maintenance. After removal of 2i, cells were cultured for at least 5 days

in serum/LIF conditions on feeder cells before being used for experiments, unless otherwise stated. Lipofectamine RNAi MAX (Life Technologies) was used for siRNA transfection.

Reporter cell lines

Six different reporter ES cell lines derived from the E14 cell line were used in this study and are described in detail in Tables 1 and 2. 2C-reporter cell lines harbouring either a tdTomato, an EGFP or a turboGFP cassette were described previously [15]. A 2C::turboGFP reporter cell line with constitutive tdTomato expression driven by the CAG promoter was used for the electron microscopy experiments and is described in further detail in [16]. All measurements on the Zscan4⁺ and 2-cell-like populations were performed on a triple

Table 2. Reporter cell lines used in each experiment.

Experiment	Cell line
Seahorse extracellular flux assay	tbg4ZH
ATP content measurements	tbg4ZH
Glucose uptake measurements	tbg4ZH
Metabolite incubations	tbg4ZH and tbg4-12
ROS measurements	tbg4
Immunofluorescence experiments	tbg4
Electron microscopy	tbg4-12
RT-qPCR on control and acetate-treated cells	2C-EGFP
Glucose uptake measurements on naïve, primed and Zscan4 cells	Rex1-Zscan4
G6PDH activity assay	tbg4ZH
Histone acetylation western blot	tbg4
Autophagy measurements	2C-tdTomato
Time-lapse experiments	tbg4ZH
Chromatin factor knockdowns	2C-EGFP

reporter cell line carrying 2C::tbGFP, Zscan4c::mCherry and constitutive H2B-iRFP constructs. This cell line incorporates a Zscan4 reporter construct (kindly provided by M. Ko) and is described elsewhere [16]. The Rex1 reporter cell line used for the glucose uptake measurements on naïve and primed pluripotent ES cells was kindly provided by A. Smith [4], and the generation of the Rex1 reporter line with the Zscan4 reporter has been described previously [16].

Fluorescence-assisted cell sorting

Cells were washed with room temperature sterile PBS, trypsinized and resuspended in ice-cold sterile 0.5% BSA PBS solution. Sorting was performed using a BD BioSciences FACS Aria II or III. During sorting, cells were collected in culture medium and kept at 4°C during the sort. Analysis of FACS data was performed using the FlowJo software.

Measurements of cellular oxygen consumption rates and extracellular acidification rates

To measure the oxygen consumption and extracellular acidification rates, a Seahorse XFe96 Flux Analyser was used. Five $\times 10^4$ cells were FACS-sorted, centrifuged and seeded in XF96 plates 3–4 h prior to the assay in standard ESC medium as described above. Prior to the assay, culture medium was gently replaced with Seahorse assay medium containing glucose (25 mM), L-glutamine (2 mM) and sodium pyruvate (1 mM). A gentle wash using Seahorse medium was performed to minimize carry-over of the DMEM medium. Basal, maximal (FCCP, 250 nM) and non-mitochondrial respiration (rotenone/antimycin A, 500 nM) as well as extracellular acidification were determined over 3 min of measurement. Basal respiratory capacity results shown derive from three independent biological replicates. Maximal and non-mitochondrial respiratory capacity results shown derive from two independent biological replicates.

Measurements of cellular oxygen consumption rates upon lactate or pyruvate supplementation

To measure the oxygen consumption rate upon acute sodium lactate or sodium pyruvate supplementation, a Seahorse XFe96 Flux Analyser was used. Five $\times 10^4$ cells were FACS-sorted, centrifuged and seeded in XF96 plates 3–4 h prior to the assay in standard ES cell medium as described above. Before the assay, the ES culture medium was gently replaced with glucose-free Seahorse assay medium containing solely L-glutamine (2 mM). A gentle wash using Seahorse medium was performed to minimize carry-over of the DMEM medium. Basal, pyruvate or lactate-stimulated (20 mM), maximal (FCCP, 250 nM) and non-mitochondrial respiration (rotenone/antimycin A, 500 nM) was determined over 3 min of measurement. Respiratory rate measurements for each metabolite derive from three independent biological replicates.

ROS measurements

Cells were plated over gelatin-coated cell culture dishes and treated with CellROX Deep Red reagent (ThermoFisher Scientific) diluted in culture medium at a final concentration of 5 μ M for 30 min. For image acquisition, cells were washed three times with PBS, stained

with Hoechst 33342 diluted in culture medium and imaged using a 60 \times 1.4 NA Plan-Apochromat VC objective on a Nikon Ti-E equipped with a Yokogawa CSU-X1 spinning disc head and a Photometrics Evolve 512 EMCCD camera. For flow cytometry analysis, cells were washed three times in PBS, trypsinized, resuspended in 0.5% BSA PBS and analysed using a FACS Aria II.

ATP content measurements

ATP content was measured using the luciferase-based CellTiter-Glo assay (Promega) according to the manufacturer's instructions with a few modifications in order to couple it to FACS sorting. Briefly, 1,000 ES cells, Zscan4⁺ cells or 2-cell-like cells were FACS-sorted in biological triplicates into 100 μ l of sterile PBS deposited on the wells of a white 96-well plate. Following addition of 100 μ l of CellTiter-Glo reagent, plates were shaken for 2 min at room temperature and luminescence was allowed to stabilize for 10 min. Finally, luminescence was measured on an Orion II microplate luminometer (Berthold Titertek) with 1 s integration time. Readings from negative control wells where no cells were sorted were subtracted from all other measurements to account for background.

Glucose uptake measurements

Glucose uptake rates were measured using the luciferase-based Glucose Uptake-Glo assay (Promega) according to the manufacturer's instructions with a few modifications in order to couple it to FACS sorting. Briefly, 2,500 ES cells, Zscan4⁺ cells or 2-cell-like cells were FACS-sorted in biological triplicates into 20 μ l of sterile PBS deposited on the wells of a white 96-well plate (for a total volume of 35 μ l). Following addition of 15 μ l of 3.33 mM 2-DG diluted in PBS (for a final concentration of 1 mM), plates were shaken for 1 min and incubated at room temperature for 20 min. Following addition of stop, neutralization and luciferase solution according to manufacturer's instructions, luminescence was measured on an Orion II microplate luminometer (Berthold Titertek) luminometer with 1 s integration time. Readings from negative control wells to which 2,500 ES cells were sorted but no 2-DG was added were subtracted from all other measurements to account for background. In the case of the glucose uptake measurements performed on naïve, primed and Zscan4⁺ cells, experiments were carried as described above but sorting was performed on Rex1-high and Rex1-low populations instead.

Glucose uptake measurements after Gnpnat1 or G6pdx knockdown

Forty-eight hours before measurements, 0.375×10^6 ES cells were transfected with siRNAs targeting Gnpnat1, G6pdx or a non-targeting siRNA in biological triplicates using Lipofectamine RNAi MAX (ThermoFisher Scientific). Cell culture medium was replaced after 24 h. Following this time period, cells were washed with room temperature sterile PBS, trypsinized, resuspended in ice-cold sterile 0.5% BSA PBS solution and FACS-sorted as described above for the rest of the glucose uptake measurements. All three siRNA-transfected cultures were sorted into the same plate in biological duplicates and were assayed simultaneously. Knockdown efficiency was determined as described below.

Table 3. Oligonucleotide sequences used in this study.

Gene	Forward primer	Reverse primer	siRNA target sequence
Actb	CACGGTTGGCCTTAGGGTTCAG	GCTGTATCCCTCCATCGTG	
Gapdh	GCCTGCTTACCACCTTCTT	CATGGCTTCCGTGTTCTA	
Zscan4	GCTGTTGTTCAAAGCTTGATGACTTC	GAGATTATGGAGAGTCTGACTGATGAGTG	
MERVL	GAGGCTCCAACAGCATCTCTA	CTCTACCACTGGACCATATGAC	
IAP	AAGCAGCAATCACCCACTTTGG	CAATCATTAGATGCGGCTGCCAAG	
L1	GGACCAGAAAAGAAATTCCTCCCG	CTCTCTGGCTTTCATAGTCTCTGG	
Dux	AGGCCCTGCTATCAACTTTCA	CTCCTCTCCACTGCGATTCC	
Gnpnat1	CGTCCAGTGCGACTTTA	TGGGTGCAGCAACAAAAAT	GGCAAACUGUUUUUAUCAA
G6pdx	GATCATCAGCGATGTTATGC	CTCTGAGATACACTTCAACAC	GAGGAGUUUUUUGCCCGUA
Non-targeting siRNA #1			UAGCGACUAAACACAUCAA
Ubc9			AGAUCUAAGUCGCUCCGUA

Real-time RT-qPCR

Total RNA was extracted from ES cells using the ReliaPrep miniprep kit (Promega), and reverse transcription was performed with SuperScript II (ThermoFisher Scientific) with oligodT oligos. For the MERVL and Zscan4 RT-qPCR analysis, total RNA was extracted from ES cells with the RNeasy Plus mini kit (Qiagen) and treated with turbo DNase (ThermoFisher Scientific) to remove genomic DNA. Reverse transcription was performed with SuperScript II (ThermoFisher Scientific) with random hexamers. Real-time PCR was performed with Lightcycler 480 SYBR Green I Master Mix (Roche) on a LightCycler 96 Real-time PCR System (Roche). The relative expression level was normalized to *Gapdh* and *Actb* (for MERVL, Zscan4, Dux, IAP and L1), and to *Actb* only (for RT-qPCR analysis of siRNA efficiency). Primers used in this study are described in Table 3.

Metabolite incubations

For the metabolite incubation experiments, solid metabolites were diluted in PBS to generate a high concentration stock solution, and liquid metabolites were added directly to the culture media. Detailed references for all the metabolites used are listed in Table 4. Various volumes of these stock solutions were then added to ES cell cultures at the concentrations described in Table EV1. ES cells were grown in 24-well plates for 48 h over feeders, and the media were replaced daily, using metabolite-supplemented media. Following this culture period, cells were trypsinized and FACS-sorted for quantification.

Immunofluorescence

Cells were cultured over feeder-coated coverslips, fixed in 4% PFA for 10 min at room temperature and permeabilized with 0.2% Triton X-100 for another 10 min. A 3% BSA PBS blocking solution was used for blocking for 1 h. Primary antibodies were incubated overnight in blocking solution and were followed by three washes in PBS. Secondary antibodies were incubated for 1 h. Mounting was done in VECTASHIELD Hardset Mounting Medium (Vector Labs). Image acquisition was performed using a Leica SP8 confocal microscope.

Table 4. Metabolites used in this study.

Compound name	Reference
L-threonine	T8441-25G
Sodium citrate dihydrate	W302600-1KG
Dimethyl-(S)-(-)-malate	374318-5G
Dimethyl-succinate	W239607
Dimethyl- α -ketoglutarate	349631-5G
Monomethyl fumarate	A651419-10G
L-glutamine	G-3126-100G
N-acetyl-D-glucosamine	A3286-5G
D-fructose 6-phosphate disodium salt	F3627-500mg
Sodium L-lactate	L7022
DL-isocitric acid	I1252-1G
Sodium pyruvate	P2256-5G
Sodium acetate	S5636-250G
Nicotinamide mononucleotide	N0636-100G
Albumax I Lipid-Rich BSA	11020021
D-(-)-ribose	R7500-5G
6-phosphogluconic acid trisodium salt	P7877-100mg
D-(+)-glucosamine hydrochloride	G4875-25
D-(-)-fructose	F0127-100G
D-(+)-glucose	G8270-1KG
Chemically defined lipid concentrate	11905031
UK-5099	PZ0160-5mg
Sodium oxamate	O2751-5g
2-Deoxy-D-glucose	D8375-1g

Antibodies

Antibodies used in this work were the following: mouse turboGFP (OT12H8, Origene), rabbit Zscan4 (AB4340, EMD Millipore), panH4ac (3HH4 2C2, ThermoFisher), H4K5/K8/K16ac (in-house produced at IGBMC), H3 (ab1791, Abcam) and goat Oct4 (sc-8628, Santa Cruz).

Electron microscopy

Embryos at the zygote (~ 16 hpc, $n = 4$), 2-cell (~ 30 hpc, $n = 5$) and 8-cell (~ 54 hpc, $n = 4$) stages were collected after natural matings of B6CBAF1/J mice, fixed in 2% formaldehyde + 2.5% glutaraldehyde in 0.1 M cacodylate buffer for 2 h at 37°C, post-fixed 1 h at 4°C in 1% osmium tetroxide and *en bloc* stained with 1% uranyl acetate for 1 h at 4°C. Samples were then dehydrated in graded ethanol solutions (50, 70, 90, 100%) and then infiltrated with epoxy resin by a graded series of dilutions (30, 70, 100%). Due to the size of the embryos, they were flat embedded in a sandwich of Aclar (200 μm) in order to be observed using binoculars. Ultrathin sections (70 nm) were performed using an ultracut UCT ultramicrotome (Leica Microsystems, Vienna, Austria) and mounted on pioloform-coated slot grids to avoid crossing mesh in the nucleus. They were then stained for 20 min with uranyl acetate and 5 min with lead citrate and observed with a transmission electron microscope (CM12, Philips; FEI Electron Optics, Eindhoven, the Netherlands) operated at 80 kV. Images were acquired using an Orius 1000 CCD camera (Gatan, Pleasanton, CA). ESCs ($n = 49$ sections) and 2-cell-like cells ($n = 57$ sections) were sorted based on 2C::turboGFP fluorescence using a FACS Aria II and then cultured for 3 h at 37°C before fixing with 2% formaldehyde + 2.5% glutaraldehyde in 0.1 M cacodylate buffer for 2 h at 37°C and treated as described above for embryos. Following acquisition, images were corrected for illumination bias using an automatic method based on intensity gradients and a bivariate polynomial modelling as previously [14] and processed using unsharp masking in Fiji.

RNA-seq data analysis

RNA-seq data for 2-cell-like cells (2C::GFP⁺) and ESCs (2C::GFP⁻) were generated and described previously [15]. Heatmaps were generated using DESeq2-derived fold-changes between ES cells and 2-cell-like cells.

Knockdown of chromatin factors

Two days before transfection, cells were plated in gelatin-coated dishes. The 2i inhibitors were removed from the medium 1 day before transfection. Lipofectamine RNAi MAX (Life Technologies) was used for siRNA transfection according to the manufacturer's instructions. A total of 75,000 cells were reverse-transfected in 24-well-gelatin-coated plates using 30 nM siRNA final concentration (the siRNAs employed are listed in Table 5). We used Silenced Negative Control No.1 siRNA (Life Technologies) as a negative control for siRNA treatment. The effect of RNAi was examined 2 days after transfection. Sodium acetate was applied 24 h before measurements at a concentration of 32 mM.

Autophagy measurements

Measurement of autophagic activity was carried out using the CYTO-ID autophagy detection kit (ENZO Life Sciences) according to the manufacturer's instructions. Briefly, one 6-well-plate well of 2C::tdTomato ES cells was trypsinized, centrifuged and washed in PBS once. Following centrifugation, cells were resuspended in 250 μl assay buffer, and afterwards, 250 μl of staining solution (1 μl Cyto-ID dye per ml of assay buffer) was added. Cells were then

Table 5. siRNAs used in this study.

Gene	Provider	Reference
Rybp	GE Healthcare	D-042769-01
Mga	GE Healthcare	D-045405-01
Max	GE Healthcare	D-047274-03
Ring1b	GE Healthcare	D-042180-01
Pcgf6	GE Healthcare	D-049359-01
L3mbtl2	GE Healthcare	D-065321-01
Ep400	GE Healthcare	D-058750-01
Tip60	GE Healthcare	D-057795-17
Dmap1	GE Healthcare	D-059463-02
Rif1	GE Healthcare	D-040028-01
Chaf1b/p60	Life Technologies	s99864

incubated for 30 min at 37°C, washed once in assay buffer and resuspended in 500 μl assay buffer. Measurements were performed on a FACS Aria III. Chloroquine- and/or rapamycin-treated cells were used as positive controls and exhibited a stronger fluorescence intensity, as expected. For these experiments, chloroquine and rapamycin were diluted in the culture media at 10 μM for 5–7 h and at 500 nM for 24 h, respectively.

G6pdh activity measurements

Measurement of glucose-6-phosphate dehydrogenase activity was carried out using the PicoProbe Glucose-6-Phosphate Dehydrogenase Activity fluorometric assay kit (BioVision) according to the manufacturer's instructions with a few modifications in order to couple it to FACS sorting. Briefly, 2,500 ES cells, Zscan4⁺ cells or 2-cell-like cells were FACS-sorted in biological triplicates into 35 μl of assay buffer deposited on the wells of a white 96-well plate (for a total volume of 50 μl). An additional three wells containing ESCs were FACS-sorted to serve as background wells. The plate was then placed on ice while positive controls, reaction mix and background mix were prepared. Lastly, 50 μl of the appropriate reaction or background mix was added to each well and measurements were performed immediately for 1 h using a CLARIOstar (BMG Labtech) fluorescent plate reader in kinetic mode at 37°C.

Time-lapse experiments

Prior to time-lapse analysis, 3000 ESCs (Zscan4⁻ and 2C::tbGFP⁻) were FACS-sorted into individual wells of a gelatin-coated glass bottom 96-well plate (ThermoFisher) containing 50 μl of ES cell media. Afterwards, ES cell media containing sodium acetate, sodium lactate or no added metabolites were added to each well, to a final concentration of 32 mM and a final volume of 150 μl . Cells were then allowed to attach for a couple hours. Image acquisition was carried out in four positions within each well with a 20 \times 0.75 NA Plan-Apochromat objective lens every 30 min for 96 h using a Nikon Ti-E system equipped with the Bruker Opterra II multipoint confocal system. Images were recorded on an EMCCD camera using emission filters for turboGFP (BP520/40), mCherry (570LP) and iRFP (655LP) mounted on a FLI filter wheel. Spontaneously arising Zscan4⁺ or

2-cell-like cells were manually identified using the ImageJ software and quantified relative to the total number of cells present in the field of view at each specific timepoint. Time-lapse experiments were carried out in three independent biological replicates.

JC-1 immunostaining

Animal experiments were carried out in compliance with local regulations (Government of Upper Bavaria). Embryos were collected from 5- to 8-week-old F1 (C57BL/6J × CBA/H) super-ovulated females crossed with F1 males. Superovulation was induced by intraperitoneal injection of pregnant mare serum gonadotropin (PMSG, Intervet, 5 IU) and human chorionic gonadotropin (hCG, Intervet, 7.5 IU) 46–48 h later. Embryos were collected at the following times after human chorionic gonadotrophin injection (phCG): 2-cell stage (41–43 h) and blastocyst stage (89–91 h). Embryos were randomly allocated to experimental groups, incubated for 30 min at 37°C with a 1 μM JC-1 solution (Abcam, ab113850) or with the dilution buffer as control and then imaged in dilution buffer. Confocal microscopy was performed on a 40× oil objective on a TCS SP8 inverted confocal microscope (Leica). We used an excitation wavelength of 475 nm and emission wavelengths of 530 ± 15 nm (monomer JC-1) and 590 ± 17.5 nm (aggregate JC-1). Z-sections were taken every 5 μm. Image analysis was performed using the software Fiji. For each image, the sum of slices z-projection was performed in order to obtain the fluorescence intensity from the whole embryo. The embryo was manually segmented, and the mean intensity per embryo was calculated for both channels. The mean value for the control embryos was subtracted from experimental values for each biological replicate. The aggregate to monomeric ratio for each embryo was then calculated.

Expanded View for this article is available online.

Acknowledgements

Work in the Torres-Padilla laboratory is funded by the Helmholtz Association, the German Research Council (CRC 1064) and H2020 Marie-Curie Actions ITN EpiSystem and ChromDesign. We thank C. Spiegelhalter for technical assistance with processing of electron microscopy samples and A. Burton for critical reading of the manuscript.

Author contributions

DR-T, GH, AE, MG, AI and XG performed and designed experiments. DR-T, GH and M-ET-P conceived the project. DR-T and M-ET-P wrote the manuscript. M-ET-P supervised the work.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Gascón S, Murenu E, Masserdotti G, Ortega F, Russo GL, Petrik D, Deshpande A, Heinrich C, Karow M, Robertson SP *et al* (2016) Identification and successful negotiation of a metabolic checkpoint in direct neuronal reprogramming. *Cell Stem Cell* 18: 396–409
- Liberti MV, Locasale JW (2016) The Warburg effect: how does it benefit cancer cells? *Trends Biochem Sci* 41: 211–218
- Festuccia N, Osorno R, Wilson V, Chambers I (2013) The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. *Curr Opin Genet Dev* 23: 504–511
- Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, Smith A (2011) Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol* 13: 838–845
- Kalmar T, Lim C, Hayward P, Muñoz-Descalzo S, Nichols J, Garcia-Ojalvo J, Martinez Arias A (2009) Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol* 7: e1000149
- Morgani SM, Canham MA, Nichols J, Sharov AA, Migueles RP, Ko MSH, Brickman JM (2013) Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep* 3: 1945–1957
- Nichols J, Smith A (2011) The origin and identity of embryonic stem cells. *Development* 138: 3–8
- Hackett JA, Dietmann S, Murakami K, Down TA, Leitch HG, Surani MA (2013) Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Reports* 1: 518–531
- Marks H, Kalkan T, Menafra R, Denisov S, Jones K, Hofemeister H, Nichols J, Kranz A, Francis Stewart A, Smith A *et al* (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149: 590–604
- Zhou W, Choi M, Margineantu D, Margaretha L, Hesson J, Cavanaugh C, Blau CA, Horwitz MS, Hockenbery D, Ware C *et al* (2012) HIF1 α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *EMBO J* 31: 2103–2116
- 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
- Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim JW, 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
- Eckersley-Maslin MA, Svensson V, Krueger C, Stubbs TM, Giehr P, Krueger F, Miragaia RJ, Kyriakopoulos C, Berrens RV, Milagre I *et al* (2016) MERVL/Zscan4 network activation results in transient genome-wide DNA demethylation of mESCs. *Cell Rep* 17: 179–192
- Bošković A, Eid A, Pontabry J, Ishiuchi T, Spiegelhalter C, Raghu Ram EVS, Meshorer E, Torres-Padilla ME (2014) Higher chromatin mobility supports totipotency and precedes pluripotency *in vivo*. *Genes Dev* 28: 1042–1047
- Ishiuchi T, Enriquez-Gasca R, Mizutani E, Bošković A, Ziegler-Birling C, Rodriguez-Terrones D, Wakayama T, Vaquerizas JM, Torres-Padilla ME (2015) Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat Struct Mol Biol* 22: 662–671
- Rodriguez-Terrones D, Gaume X, Ishiuchi T, Weiss A, Kopp A, Kruse K, Penning A, Vaquerizas JM, Brino L, Torres-Padilla ME (2018) A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat Genet* 50: 106–119
- Zalzman M, Falco G, Sharova LV, Nishiyama A, Thomas M, Lee S-L, Stagg CA, Hoang HG, Yang H-T, Indig FE *et al* (2010) Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* 464: 858–863
- Hirata T, Amano T, Nakatake Y, Amano M, Piao Y, Hoang HG, Ko MSH (2012) Zscan4 transiently reactivates early embryonic genes during the generation of induced pluripotent stem cells. *Sci Rep* 2: 208
- 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
20. Eckersley-Maslin M, Alda-Catalinas C, Blotenburg M, Kreibich E, Krueger C, Reik W (2019) Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev* 33: 194–208
 21. De Iaco A, Coudray A, Duc J, Trono D (2019) DPPA2 and DPPA4 are necessary to establish a 2C-like state in mouse embryonic stem cells. *EMBO Rep* 20: e47382
 22. 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
 23. Kaneko KJ (2016) Metabolism of preimplantation embryo development: a bystander or an active participant? *Curr Top Dev Biol* 120: 259–310
 24. Leese HJ, Barton AM (1984) Pyruvate and glucose uptake by mouse ova and preimplantation embryos. *Reproduction* 72: 9–13
 25. Houghton FD, Thompson JG, Kennedy CJ, Leese HJ (1996) Oxygen consumption and energy metabolism of the early mouse embryo. *Mol Reprod Dev* 44: 476–485
 26. Baharvand H, Matthaehi KI (2003) The ultrastructure of mouse embryonic stem cells. *Reprod Biomed Online* 7: 330–335
 27. Dumollard R, Duchon M, Carroll J (2007) The role of mitochondrial function in the oocyte and embryo. *Curr Top Dev Biol* 21–49 [https://doi.org/10.1016/s0070-2153\(06\)77002-8](https://doi.org/10.1016/s0070-2153(06)77002-8)
 28. Stern S, Biggers JD, Anderson E (1971) Mitochondria and early development of the mouse. *J Exp Zool* 176: 179–191
 29. Mishra P, Chan DC (2016) Metabolic regulation of mitochondrial dynamics. *J Cell Biol* 212: 379–387
 30. Murphy MP (2009) How mitochondria produce reactive oxygen species. *Biochem J* 417: 1–13
 31. Carey BW, Finley LWS, Cross JR, Allis CD, Thompson CB (2014) Intracellular α -ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* 518: 413–416
 32. Whiddon JL, Langford AT, Wong CJ, Zhong JW, Tapscott SJ (2017) Conservation and innovation in the DUX4-family gene network. *Nat Genet* 49: 935–940
 33. Moussaieff A, Rouleau M, Kitsberg D, Cohen M, Levy G, Barasch D, Nemirovski A, Shen-Orr S, Laevsky I, Amit M et al (2015) Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab* 21: 392–402
 34. Latham T, Mackay L, Sproul D, Karim M, Culley J, Harrison DJ, Hayward L, Langridge-Smith P, Gilbert N, Ramsahoye BH (2012) Lactate, a product of glycolytic metabolism, inhibits histone deacetylase activity and promotes changes in gene expression. *Nucleic Acids Res* 40: 4794–4803
 35. Baumann CG, Morris DG, Sreenan JM, Leese HJ (2007) The quiet embryo hypothesis: molecular characteristics favoring viability. *Mol Reprod Dev* 74: 1345–1353
 36. Nasr-Esfahani MH, Aitken JR, Johnson MH (1990) Hydrogen peroxide levels in mouse oocytes and early cleavage stage embryos developed *in vitro* or *in vivo*. *Development* 109: 501–507
 37. Brinster RL (1965) Studies on the development of mouse embryos *in vitro*. IV. Interaction of energy sources. *J Reprod Fertil* 10: 227–240
 38. Nagaraj R, Sharpley MS, Chi F, Braas D, Zhou Y, Kim R, Clark AT, Banerjee U (2017) Nuclear localization of mitochondrial TCA cycle enzymes as a critical step in mammalian zygotic genome activation. *Cell* 168: 210–223.e11



License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Expanded View Figures

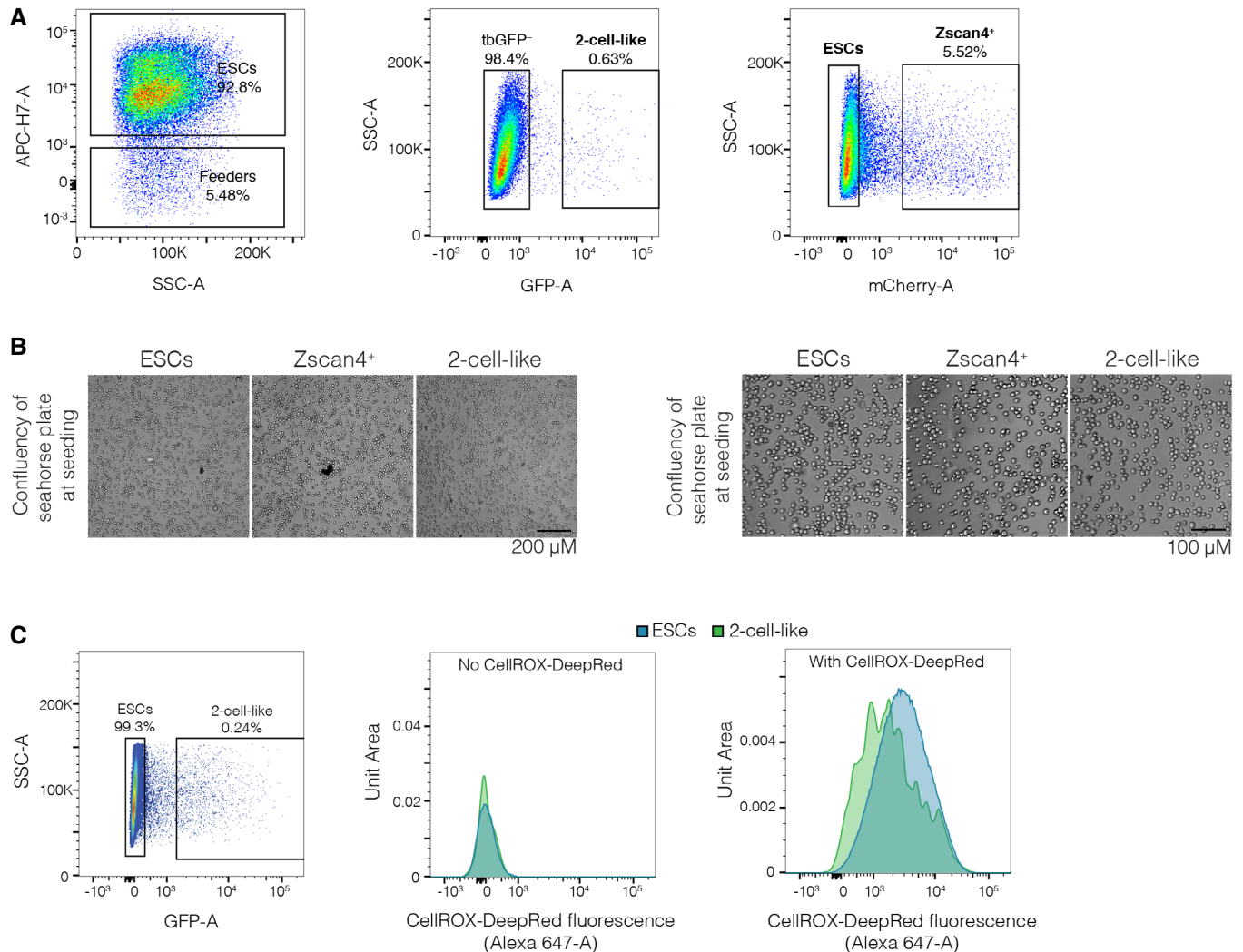


Figure EV1. Controls and set-up for Seahorse and ROS measurements.

- A Representative sorting gates used for the isolation of ES, Zscan4⁺ and 2-cell-like cells used throughout this study. Feeder cells were removed on the basis of their lack of far-red fluorescence, which is higher in ES cells because of the presence of an H2B-iRFP cassette. ESCs were defined as double negative for both the Zscan4 (Zscan4c::mCherry⁻) and the MERV-L reporters (2C::tbGFP⁻). Zscan4⁺ cells were defined as positive for the Zscan4 reporter but negative for the MERV-L reporter, and 2-cell-like cells were defined as positive for both reporters.
- B Brightfield microscopy images indicating the confluency of the three populations shortly after plating in the Seahorse extracellular flux analyser plates. Representative images, in two magnifications, for the three independent biological replicates presented in Fig 1C and D are shown.
- C Representative sorting gate for ES and 2-cell-like cells (left) used for the FACS-assisted ROS measurements (right). Fluorescence intensity distributions for ES cells (blue) and 2-cell-like cells (green) in control (centre) and CellROX-treated samples (right) are shown.

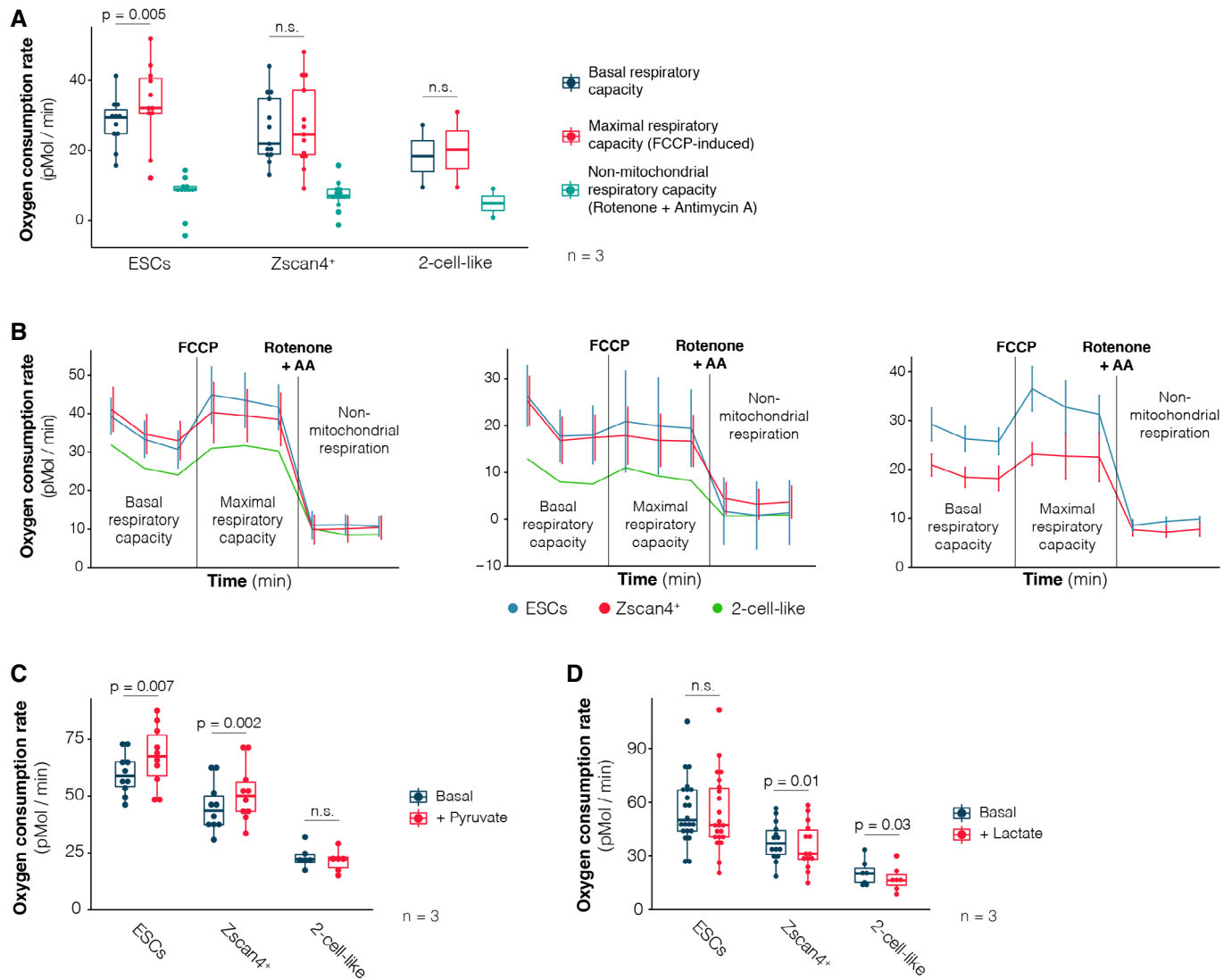


Figure EV2. Seahorse oxygen consumption rate measurements.

- A** Basal, maximal (FCCP-induced) and non-mitochondrial (rotenone + antimycin A-induced) oxygen consumption rate measurements of ES, Zscan4⁺ and 2-cell-like cells performed on the Seahorse extracellular flux analyser. Assay medium was formulated to recapitulate standard ES cell culture conditions and contained glucose, L-glutamine and pyruvate. Measurements were carried out in three independent biological replicates. Note that 2-cell-like cells could only be profiled in two of those replicates. *P*-value corresponds to a paired *t*-test.
- B** Additional individual replicates of the oxygen consumption rate measurements of ESCs (blue line), Zscan4⁺ (red line) and 2-cell-like cells (green line) measured on the Seahorse extracellular flux analyser, as in Fig 1C. Basal, maximal (FCCP-induced) and non-mitochondrial (rotenone-mediated) respiratory rates are indicated. Due to the low number of 2-cell-like cells in ESC cultures compared to ES and Zscan4⁺ cells, one technical replicate of the former was analysed per biological replicate, while three technical replicates were performed for the two other populations. Accordingly, mean \pm s.d. of technical replicates is shown for ES and Zscan4⁺ cells.
- C, D** Oxygen consumption rate measurements of ES, Zscan4⁺ and 2-cell-like cells performed on the Seahorse extracellular flux analyser. Basal measurements in glucose-free media and upon acute injection of sodium pyruvate (C, 20 mM) or sodium L-lactate (D, 20 mM) are shown. Note that L-glutamine—but not glucose or pyruvate—was initially present in the assay medium. Measurements from three independent biological replicates are shown. *P*-values correspond to paired *t*-tests.

Data information: Boxes indicate the range between the first and third quartile, the band depicts the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the measurements obtained in each technical replicate. Source data are available online for this figure.

Figure EV3. Electron micrographs of mitochondria in ES cells and 2-cell-like cells.

- A, B Representative electron micrographs of mitochondria from ES and 2-cell-like cells. Scale bars, 1 μm .
- C Autophagic flux measurements were carried out using Cyto-ID. Non-stained cells were employed as a negative control and show a distinct fluorescent profile. Chloroquine- and/or rapamycin-treated cells were used as a positive control and exhibited increased fluorescence intensity.
- D FACS-assisted measurement of Cyto-ID fluorescence intensity in ES (blue) and 2-cell-like cells (green).
- E Quantification of Cyto-ID fluorescence intensity in ES (blue) and 2-cell-like (green) cells. Boxes indicate the range between the first and third quartile, the band indicates the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the median fluorescence intensity measurements obtained in each biological replicate. *P*-values were calculated using a paired *t*-test.

Source data are available online for this figure.

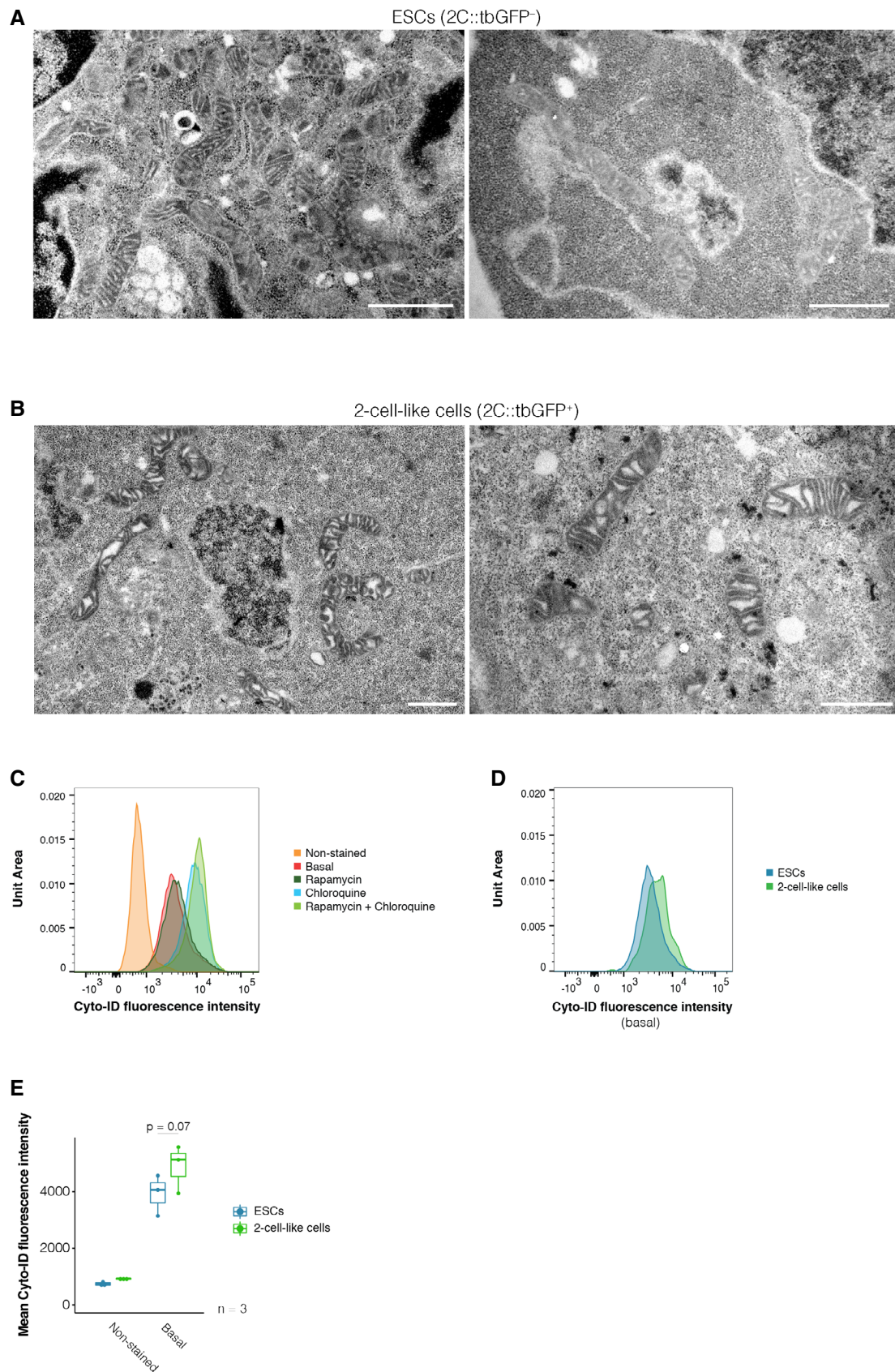


Figure EV3.

Figure EV4. Electron micrographs of mitochondria from pre-implantation embryos.

- A–C Representative electron micrographs of mitochondria from PN3 stage zygotes and 2-cell- and 8-cell-stage blastomeres. Scale bars, 1 μm .
- D Live-cell imaging of mouse 2-cell-stage embryos and blastocysts stained with the mitochondrial membrane potential probe JC-1.
- E Quantification of the ratio between aggregate and monomer emission in mouse 2-cell-stage embryos and blastocysts. Boxes indicate the range between the first and third quartile, the band specifies the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Measurements were obtained in three independent biological replicates, and each dot represents one individual embryo ($n = 34$, 2-cell stage; $n = 16$, blastocyst). n.s.—non-significant; t -test.

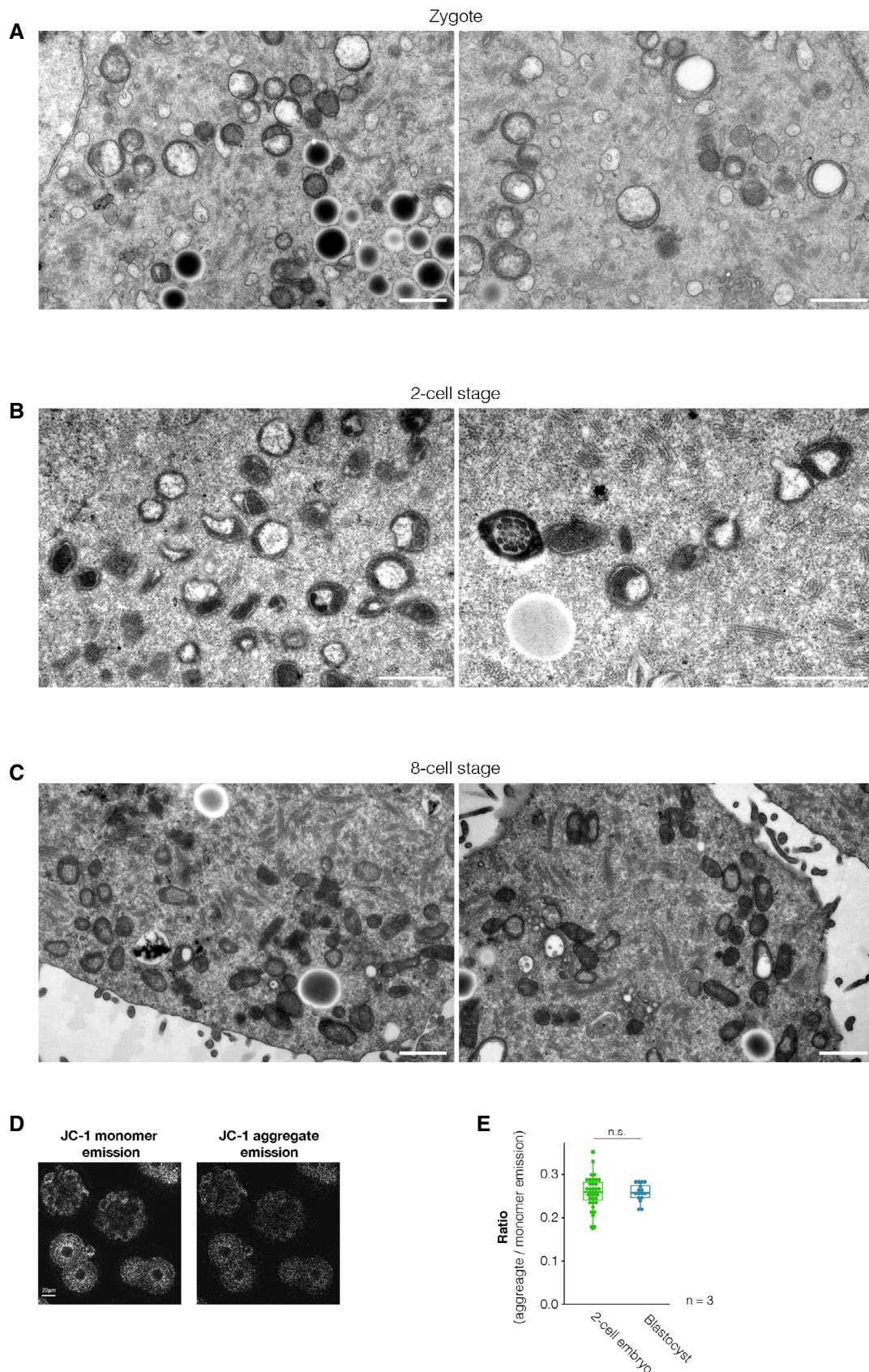


Figure EV4.

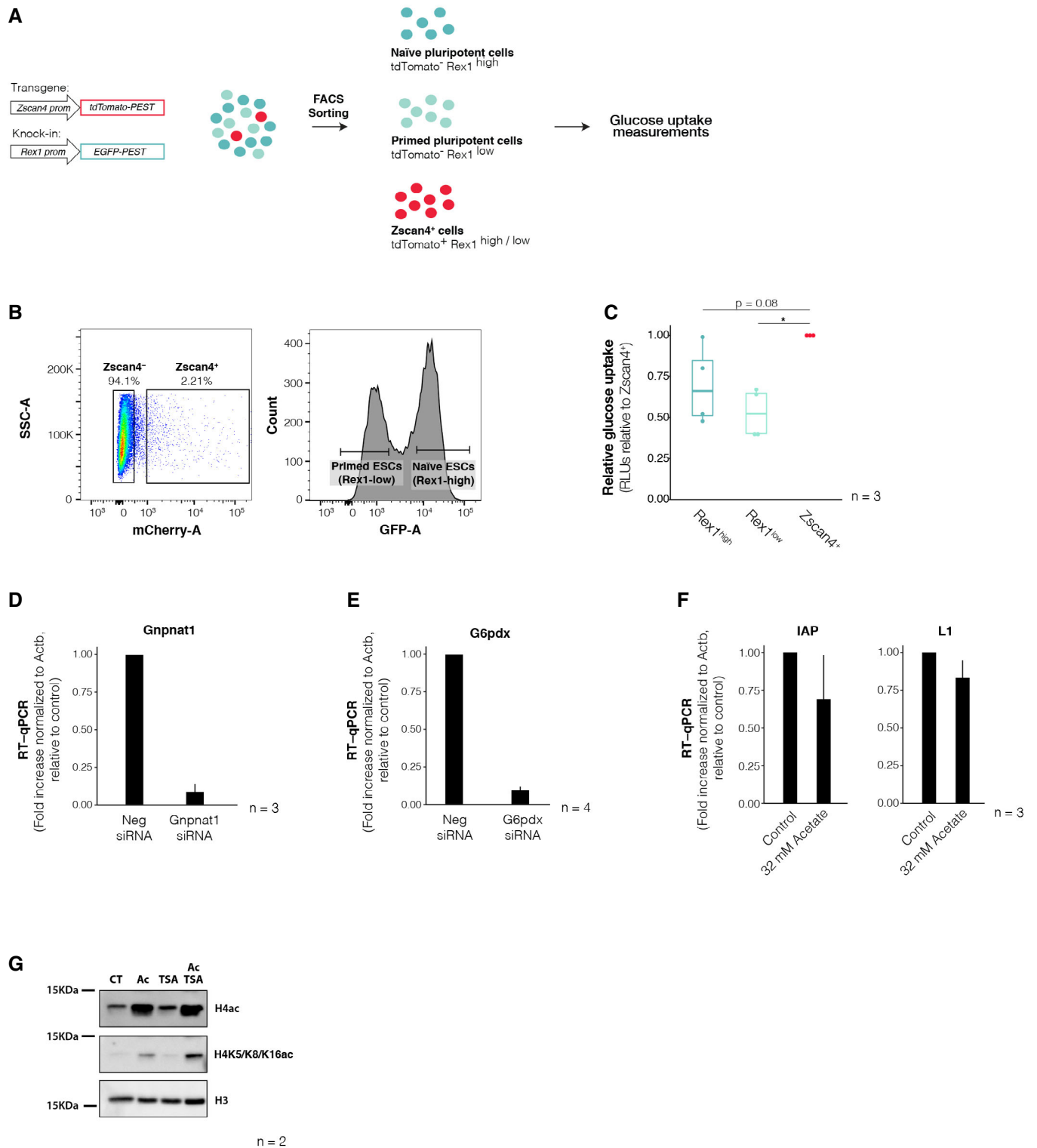


Figure EV5.

Figure EV5. Zscan4⁺ cells exhibit higher glucose uptake than both naïve and primed ES cells.

- A Experimental design. ES cells were cultured in serum/LIF conditions over feeders for at least 5 days in the absence of 2i, and subsequently FACS-sorted into naïve pluripotent, primed pluripotent or Zscan4⁺ cells. Glucose uptake rates were measured thereafter. Reporter constructs employed to identify all three distinct populations are represented on the left. An EGFP reporter driven by the Rex1 endogenous promoter was used to distinguish between Rex1-high (naïve pluripotent) and Rex1-low (primed pluripotent) cells, and a tdTomato cassette expressed downstream of an ectopic Zscan4c promoter was used to mark Zscan4⁺ cells.
- B Representative sorting gates used for the isolation of naïve pluripotent, primed pluripotent or Zscan4⁺ cells. Zscan4⁺ cells were defined as those positive for Zscan4c::tdTomato reporter (left), irrespective of their Rex1-EGFP fluorescence level. Naïve pluripotent and primed pluripotent stem cells were gated based on the bimodality of the Rex1-EGFP distribution (right).
- C Glucose uptake rates in Zscan4⁺ cells (red), naïve pluripotent stem cells (dark blue) and primed pluripotent stem cells (light blue) were measured using a luciferase-based assay across three independent biological replicates. Measurements are represented relative to the levels of Zscan4⁺ cells. Boxes indicate the range between the first and third quartile, the band specifies the median, and the whiskers span the range of the data while extending no further than 1.5 times the interquartile range. Individual dots indicate the measurements obtained in each technical replicate. * $P < 0.05$; one sample t -test.
- D, E RT-qPCR analysis of the indicated genes after transfection with the corresponding siRNAs. Shown are mean values \pm s.d. of two technical replicates from three independent biological replicates.
- F RT-qPCR of the indicated repeats upon 24 h of sodium acetate treatment. Shown are the mean \pm s.d. of three independent cell cultures, performed in two technical replicates.
- G Western blot for the indicated antibodies in lysates derived from control and acetate- and/or TSA-treated cultures.

Source data are available online for this figure.

**PART 2: ROLE OF LIQUID-LIQUID PHASE SEPARATION IN
HETEROCHROMATIN FORMATION DURING MOUSE PRE-
IMPLANTATION DEVELOPMENT**

**EXPRESSION AND PHASE SEPARATION POTENTIAL OF HETEROCHROMATIN
PROTEINS DURING EARLY MOUSE DEVELOPMENT**

Statement of contribution

I hereby state that my contribution to the publication:

Guthmann, M., Burton, A., Torres-Padilla, M.E. (2019). Expression and phase separation potential of heterochromatin proteins during early mouse development. EMBO reports. 20: e47952

consisted in the computational analysis of publicly available mass spectrometry data for heterochromatic proteins in different species, the analysis of the disorder content of those proteins, the in-depth analysis of phase separation potential of the “bona fide” heterochromatic proteins and the computational analysis of publicly available RNA sequencing mouse embryo data. Adam Burton helped with the analysis of the publicly available mass spectrometry mouse embryo data.

Manuel Guthmann

München, September 23rd, 2022


Confirmation of contribution

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

Prof. Dr. Maria-Elena Torres-Padilla

München, September 23rd, 2022

Expression and phase separation potential of heterochromatin proteins during early mouse development

Manuel Guthmann^{1,2}, Adam Burton^{1,2} & Maria-Elena Torres-Padilla^{1,2,*} 

Abstract

In most eukaryotes, constitutive heterochromatin is associated with H3K9me3 and HP1 α . The latter has been shown to play a role in heterochromatin formation through liquid–liquid phase separation. However, many other proteins are known to regulate and/or interact with constitutive heterochromatic regions in several species. We postulate that some of these heterochromatic proteins may play a role in the regulation of heterochromatin formation by liquid–liquid phase separation. Indeed, an analysis of the constitutive heterochromatin proteome shows that proteins associated with constitutive heterochromatin are significantly more disordered than a random set or a full nucleome set of proteins. Interestingly, their expression begins low and increases during preimplantation development. These observations suggest that the preimplantation embryo is a useful model to address the potential role for phase separation in heterochromatin formation, anticipating exciting research in the years to come.

Keywords development; epigenetics; heterochromatin establishment; phase separation

Subject Categories Chromatin, Transcription & Genomics; Development

DOI 10.15252/embr.201947952 | Received 19 February 2019 | Revised 3 October 2019 | Accepted 16 October 2019 | Published online 7 November 2019

EMBO Reports (2019) 20: e47952

Introduction

In eukaryotes, around 145 basepairs of DNA are wrapped around octamers of the four canonical histones H2A, H2B, H3 and H4 to form the nucleosome. The nucleosome is the building block of the chromatin, which in addition includes other chromatin-associated proteins that bind nucleosomes and also the linker histone H1. Functionally, chromatin has been traditionally divided into two categories: hetero- and euchromatin [1], which were first recognised cytologically by Emil Heitz [2]. Heterochromatin appeared as regions of the nucleus that do not decondense after mitosis, which he considered to be a non-functional part of the genome. Nowadays,

the definition of heterochromatin has broadened to include features such as (i) histone modifications such as histone 3 lysine 9 trimethylation (H3K9me3), H3K27me3, DNA methylation and potentially also H3K56me3 [3,4]; (ii) a (mostly) transcriptionally silent state; (iii) a late replicating nature; (iv) an electron-dense and condensed state in electron microscopy [5], and more recently (v) a higher resistance to sonication [6]. Heterochromatin can be further broadly divided into constitutive heterochromatin—which is located at centromeric and telomeric regions, as well as at most repeat elements throughout most eukaryotic genomes—and facultative heterochromatin, which harbours the H3K27me3 mark and often localises to temporally or spatially regulated genes [5].

Over the last two decades, a rather unified model for constitutive heterochromatin establishment has emerged whereby the Suv39h1/h2 (Su(var)9-1) enzymes initiate a feedback cascade by catalysing H3K9me3, which in turn recruits heterochromatin protein 1 (HP1) proteins, primarily through their chromodomain [7–9]. Downstream recruitment of Suv420h1/h2 (Su(var)4-20) reinforces a heterochromatic loop by catalysing H4K20me3 [10], while as yet unknown enzymes deposit H3K64me3 [11]. Subsequent recruitment of Suv39h1/h2 by both HP1 and H3K9me3 enables spreading and amplification of the heterochromatin domain. In addition, RNA-mediated interactions of HP1 and the Su(var) enzymes themselves have also been implicated in maintaining constitutive heterochromatin in mouse, human and yeast [12–15]. However, relatively little is known about the mechanisms that direct heterochromatin formation *in vivo*, at the beginning of development.

It has recently been suggested that heterochromatin can form by phase separation through the local accumulation of HP1 α [16,17]. Phase-separated compartments appear as immiscible liquid droplets that emerge through multivalent, weak interactions between biological polymers, which can be either proteins or nucleic acids [18,19]. Multivalent interactions can be provided by intrinsically disordered domains (IDRs) or structured domains. Liquid droplets can undergo fission, coalesce into larger droplets and relax to their original spherical shape after shear stress [20,21]. Since the discovery that P granules form by liquid–liquid phase separation in the *Caenorhabditis elegans* germline around 10 years ago, many studies have shown that several membrane-less

¹ Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München, München, Germany

² Faculty of Biology, Ludwig-Maximilians Universität, München, Germany

*Corresponding author. Tel: +49 89 3187 3317; E-mail: torres-padilla@helmholtz-muenchen.de

organelles may in fact form through phase separation [22–26]. These include the nucleolus, which has physical properties of a phase-separated liquid-like droplet formed of several immiscible liquid sub-compartments [21,27], but also stress granules and paraspeckles [28,29] as well as cajal bodies [23]. More recently, some studies have also suggested a role for phase separation in transcription initiation, by facilitating the recruitment of the transcriptional machinery [30–35]. Similarly, liquid–liquid phase separation was suggested to play a role in facultative heterochromatin formation by enabling the assembly of the polycomb repressive complex 1 [36].

In the phase-separation-based model for constitutive heterochromatin formation [16,17,37], the binding of HP1 α to H3K9me3 would lead to a local increase in HP1 α concentration, which in turn would nucleate a phase-separated compartment that could then grow and fuse, enabling the formation of constitutive heterochromatin. The liquid–liquid phase separation biophysical properties would also explain the selective exclusion of certain proteins from these heterochromatin compartments. In such a model, exclusion from domains may be due to the inability to interact with phase-inclusive components, but it can also result from the emergent biophysical properties of the domain. However, a recent report shows that IDR-rich liquid condensates tend to exclude chromatin, which is at odds with the proposed growth and fusion of phase-separated heterochromatin compartments. In fact, when promoting droplet formation at heterochromatin using a synthetic “CasDrop” approach, condensates appear at the periphery of such regions [38]. Thus, these conceptual frameworks to understand the formation and physical properties of heterochromatic genomic regions are still in their early days, and have not yet incorporated all the additional proteins known to be present at constitutive heterochromatin, and which may therefore play a role in regulating heterochromatin establishment.

How and whether these mechanisms operate in the early mammalian embryo at the onset of epigenetic reprogramming are unknown. Even though heterochromatin has been extensively studied, little is known about its biophysical properties as well as the mechanisms that underlie heterochromatin formation, as opposed to maintenance, *in vivo*. Here, we have undertaken an analysis to investigate the properties of heterochromatin-associated proteins and their potential to phase separate as well as their expression pattern at the earliest developmental stages in the mouse embryo. Finally, we propose possible avenues for addressing phase separation as a potential mechanism for heterochromatin formation at the beginning of development.

Results and Discussion

Several mass spectrometry studies have been carried out in mammalian cells to better understand the pathways involved in constitutive heterochromatin maintenance and integrity. Most of them focused on the identification of proteins that bind H3K9me3 using peptides or modified nucleosomes pulldowns [39–41] or chromatin immunoprecipitation [42–44]. More recently, heterochromatin proteins have been identified by mass spectrometry of the sonication-resistant fraction of the chromatin [6]. Functionally, however, much of our knowledge on heterochromatin stems from

genetic screens in model organisms including *Schizosaccharomyces pombe*, *C. elegans* and *Drosophila melanogaster* [45–47]. In *Drosophila*, position-effect variegation analyses have identified proteins important for heterochromatin maintenance and/or spreading [48]. Likewise, genetic screens in *S. pombe* have uncovered genes involved in heterochromatin integrity using a pericentromeric insertion of the *ade6*⁺ reporter for example [49]. In *C. elegans*, many repressors have been identified in screens for defects in vulva development or nuclear peripheral localisation [46,50].

In an effort to identify the most relevant protein components of constitutive heterochromatin—and thereby potential proteins that may promote heterochromatin phase separation—we undertook a bioinformatic analysis, initially based on 7 mass spectrometry studies performed in mammalian cells [6,39–44]. We focused primarily on H3K9me3 as a proxy for constitutive heterochromatin, since it is its most prevalent mark across most, albeit not all, eukaryotes. We selected proteins as heterochromatic based on their ability to bind H3K9me3-modified peptides, H3K9me3-modified nucleosomes with and without methylated DNA, or to their presence in the sonication-resistant fraction of the chromatin. Our analysis of all these studies revealed 672 proteins identified as heterochromatic by at least one study, with 148 of these proteins being present in more than one study (Table EV1). To increase stringency in our selection, we then explored the conservation across evolution of the proteins identified by mass spectrometry. For this, we searched for the ortholog genes encoding the 672 proteins in *Danio rerio*, *S. pombe*, *D. melanogaster* and *C. elegans*. Our results show that 205 (31%) genes had orthologs in all the species that we investigated. In addition, 36 (24%) of the 148 genes coding for the proteins found in more than one mass spectrometry study had orthologs in all species (Table EV1). Among these, 36 genes are the well-characterised *Cbx1*, *Cbx3* and *Cbx5*, which encode the three mammalian HP1 isoforms known to bind H3K9me3 and to play a role in constitutive heterochromatin maintenance and/or spreading. We thus speculate that a thorough investigation of the remaining 33 genes will lead to the discovery of other proteins that may play a role in constitutive heterochromatin.

Because a biochemical identification does not necessarily imply that these proteins and their corresponding orthologs functionally regulate heterochromatin formation and/or maintenance, we mined our results against datasets derived from previous genetic screens. This was possible in three species (*S. pombe*, *D. melanogaster* and *C. elegans*) but not in *D. rerio*, as we were unable to find publicly available compilations of screens in this species [46,48,49]. Interestingly, we found very little overlap between the 672 proteins identified based on the biochemical studies performed with mammalian cell culture models, and the genetic screens across other model organisms. In fact, only *Cbx1*, *Cbx3* and *Cbx5* were common across all datasets and species. This raises interesting questions, as to whether non-“core” heterochromatin proteins in different species may be important to potentially specify additional heterochromatin features. Alternatively, redundancy could potentially prevent identification of proteins in *in vivo* screens. Due to the small number of hits obtained through the analysis of genetic screenings, we decided to perform our downstream analyses below on the common 148 proteins identified from the biochemical studies, which, for simplicity, will be referred to as heterochromatic proteins hereafter.

The physical properties of phase separation and heterochromatin

Membrane-less organelles are thought to form through the nucleation of protein and nucleic acid scaffolds, which will be enriched in the phase-separated compartment, compared with the surrounding solution [20]. A key parameter determining the composition of the droplet is the scaffold's concentration [51]. The scaffold proteins that mediate phase separation often contain IDRs, thought to be important for nucleating liquid droplets [29,52–55]. However, IDRs can be present in “nucleating” components as well as “recruited” components. Most attention in the field has been devoted to IDRs, but it is important to keep in mind that structured domains may also contribute to phase separation.

IDRs are structural features of protein domains, which are often found in linker regions between folded domains as well as in post-translational modification sites, lack a unique three-dimensional structure and tend to have low-complexity sequences [56,57]. IDRs are thought to drive liquid–liquid phase separation by forming multivalent interactions through their amino acid side chains [19]. We asked whether the heterochromatin proteins that we identified have a higher propensity to exhibit disorder properties or IDRs. To characterise the potential of the 148 proteins to contribute to heterochromatin phase separation, we generated disorder estimates for them using two prediction algorithms, PONDR-VLXT [58] and IUPRED [59]. IUPRED and PONDR take into account the context of individual amino acids to calculate disorder scores for each amino acid in a given protein context. The predicted scores are thus presented as percentage disorder, mean disorder and length of disordered segments. The results obtained with both predictors were not

always similar. However, the tendency was the same, and therefore, we averaged the results obtained with both algorithms. Heterochromatin proteins displayed a significantly higher disorder score, as compared to either a random group of total proteins or nuclear proteins of the same size (median = 0.47, compared with 0.31 and 0.37, respectively; Fig 1A). The median percentage length of disordered domains, measured as percentage of amino acids of the total protein length, was 44% (Fig 1A), which is similar to the percentages calculated for the proteome of several phase-separated membrane-less organelles and is higher than the value for organised structures such as the proteasome [60]. In addition, the percentage of the protein (length) containing disordered domains was also significantly higher compared with a random (22%) or the nuclear (30%) set of proteins, indicating that heterochromatin proteins are more disordered than a random set of proteins or compared with nuclear proteins in general. Interestingly, not only the percentage of amino acids within disorder domains but also the length of disorder domains was significantly higher in the heterochromatin group of proteins (Fig 1A). Of note, heterochromatin proteins tend to be longer, compared with both groups of proteins, but also when compared with a set of global chromatin proteins or of DNA-binding proteins (Fig 1B). The comparisons with the proteins constituting the nuclear protein groups clearly show that the subset of heterochromatin proteins displays features consistent with higher disorder scores.

We then asked whether this feature is exclusive to heterochromatin proteins or whether chromatin proteins in general and DNA-binding proteins possess IDRs as well. For this, we calculated disorder scores, overall percentage (in a.a.) disorder and length of

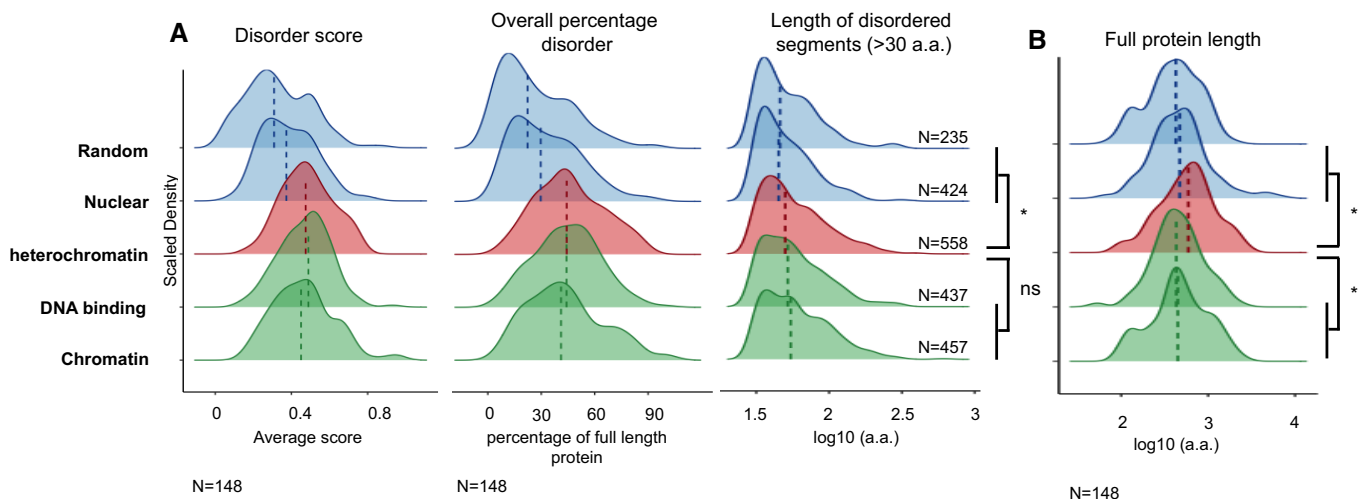


Figure 1. Analysis of the disorder content of the selected heterochromatin proteins.

A Analysis of three factors to measure disorder behaviour using both the PONDR-VLXT and IUPRED predictors. In the left panel, the disorder score per protein. In the centre panel, the percentage of predicted disorder per protein. In the right panel, the lengths of the predicted disordered regions for each protein set (length of disordered segments (> 30 a.a.)). For length of disordered regions, segments shorter than 30 amino acids were removed (based on Forma-Kay *et al* [56] and Ward *et al* [105]). The 148 heterochromatin proteins were compared with control protein sets of the same number generated from random sampling of chromatin, nuclear, DNA binding or total proteomes. The dotted lines correspond to the median value for the distributions shown. * $P \leq 0.05$ and ns > 0.05 by two-sided unpaired Wilcoxon rank-sum test.

B Length in amino acids of the proteins analysed in the indicated groups. The 148 heterochromatin proteins were compared with control protein sets of the same number generated from random sampling of chromatin, nuclear, DNA binding or total proteomes. The dotted lines correspond to the median value for the distributions shown. * $P \leq 0.05$ and ns > 0.05 by two-sided unpaired Wilcoxon rank-sum test.

disorder segments for these two additional groups of proteins. Interestingly, our analyses revealed that proteins with the potential to bind DNA and chromatin have a higher disorder score as calculated using IUPRED and PONDR-VLXT predictors, as well as higher overall percentage disorder score, compared with a random set of proteins, or to nuclear proteins (Fig 1A). We conclude that the specific part of the nucleome, which constitutes the chromatin and has the ability to bind DNA, has a higher potential to phase separate, based on IDR constitution.

To further assess the possible phase separation propensity of the 148 proteins, we used a different predictor for phase separation based on potential planar protein–protein contacts [61] (not shown). In fact, 38 of them were predicted to have a propensity to reversibly and dynamically self-associate. However, this predictor only takes the planar Pi–Pi interactions into consideration and further in-depth analysis of other interactions is typically required in order to better predict phase separation propensity. HP1 α , for example, which is known to phase separate, was not present in this list of proteins predicted to self-associate, advocating the use of several features in parallel when making predictions for phase separation potential.

Further to IDRs, interactions between amino acids with opposing charges as well as cation–pi interactions are likely to play a role in liquid droplet formation [54]. Molecular interactions between positively charged amino acids with nucleic acids also certainly play a role in the establishment of membrane-less organelles enriched in RNA and RNA-binding proteins [55,62]. In agreement with the importance of electrostatic interactions between macromolecules with different charges, phosphorylation and acetylation have been shown to perturb phase separation and dissolve membrane-less organelles [62–65]. Hydrophobic interactions have also been suggested to play an important role in phase separation [35,66]. Pi–Pi interactions between aromatic amino acids (Phe, Tyr, Trp and His) but also amino acids containing amide (Asn, Gln), carboxyl (Glu, Asp) or guanidinium (Arg) groups in their sidechain as well as amino acids with exposed backbone peptide bonds (Gly, Ser, Thr and Pro) are relevant for phase separation mediated by IDRs [61]. Tyrosines and arginines have, for example, been shown to play a predominant role in the liquid droplet formation by the FUS family proteins [67].

We thus undertook a more thorough analysis of all these features. For this, we aimed to generate a more restricted group of “bona fide” heterochromatin proteins, whose location in chromocentres and/or impact on heterochromatin functions have been validated by cell biological or genetic experiments. Specifically, we used a set of proteins identified as enriched at major satellites by PiCH in mouse embryonic stem cells [68]. From these, we selected those proteins, which are lost from the major satellites upon Suv39h1/h2 depletion, and which had been identified as suppressors of variegation (Su(var)) and modifiers of murine metastable epialleles (Mommies). This led to a list of seven proteins: CBX1 (HP1 β), CBX5 (HP1 α), ATRX, UHRF1, DNMT1, SUV420H2 and SUV39H2 (Table EV2). Excepting SUV420H2 and SUV39H2, the remaining five proteins exhibited disorder scores and overall percentage disorder values higher than the median values of the random set and nuclear proteomes (Table EV2).

We then expanded our analysis to other features indicative of a potential to phase separate, including IUPRED and FOLD disorder

scores, presence of predicted prion-like domains, propensity for Pi–Pi contacts, fraction of charged residues and net charge per residues across each protein as well as hydrophobicity (Figs 2A–C and EV1A–E). In addition, to provide a relevant comparison, we performed the same analysis with the transcription factor FUS (Fig 2A), which has been shown to phase separate both *in vitro* and *in vivo* [67,69]. This uncovered, for example, a clear prion-like domain (PLD) in ATRX as well as high IUPRED scores in ATRX, but also in CBX5 (Fig 2B and C), as previously reported [17]. Additionally, the N-terminal domain of SUV39H2, known to interact with RNA, exhibited also high IUPRED score (Fig EV1B). Interestingly, SUV39H2 is highly enriched in mouse zygotes [70], and therefore, the study of its role in heterochromatin formation, and potentially in phase separation, *in vivo*, should be an exciting research avenue. We find that the “bona fide” heterochromatin proteins contain various segments of high hydrophobicity and with a high fraction of charged residues (Figs 2A–C and EV1A–E), which could potentially favour phase separation. These features may be hard to interpret however, since they may not be sufficient *per se* to drive liquid–liquid phase separation, as recently shown for the FUS low-complexity domain [69]. Overall, these analyses suggest that the “bona fide” heterochromatin proteins that we selected have additional features linked to the potential to phase separate.

The above biophysical and biochemical characteristics are in general used as a proxy to assess if a given molecular—and in some instances cellular—process could be explained by phase separation. However, they are only an indicator. In fact, local concentration and post-translational modifications are key. For example, in HP1 α , phosphorylation is required for structural changes that promote phase separation [16]. While such additional features should be taken into account, overall, our analysis reveals that several proteins associated biochemically with constitutive heterochromatin present characteristics of proteins within membrane-less organelles and some of them are predicted to phase separate.

Establishment of heterochromatin *in vivo*

A significant rearrangement and reprogramming of constitutive heterochromatin occurs during germ cell and subsequently early embryonic development [71,72]. During preimplantation development, H3K9me3 is dramatically decreased and re-established on both parental genomes, albeit with different temporal dynamics [73–75], while H4K20me3 and H3K64me3, two modifications downstream of H3K9me3 [76], are both removed at the 2-cell stage and not re-established until post-implantation [11,77]. In addition, chromocentres only emerge from the late 2-cell stage onwards, while HP1 α , the primary heterochromatin protein suggested to be responsible for its phase separation [16,17], is not thought to be expressed during preimplantation development [78].

We suggest that in order to understand the role of phase separation in heterochromatin function, it will be particularly revealing to describe the dynamics of phase-separated heterochromatin during these periods of development, when heterochromatin is dynamic. In addition, a clearer temporal correlation could be made between the known markers of heterochromatin and the phase-separated heterochromatin state. For example, which, if any, histone modifications or protein readers typical of classical constitutive heterochromatin (such as H3K9me3, H4K20me3 and HP1 isoforms) or features such

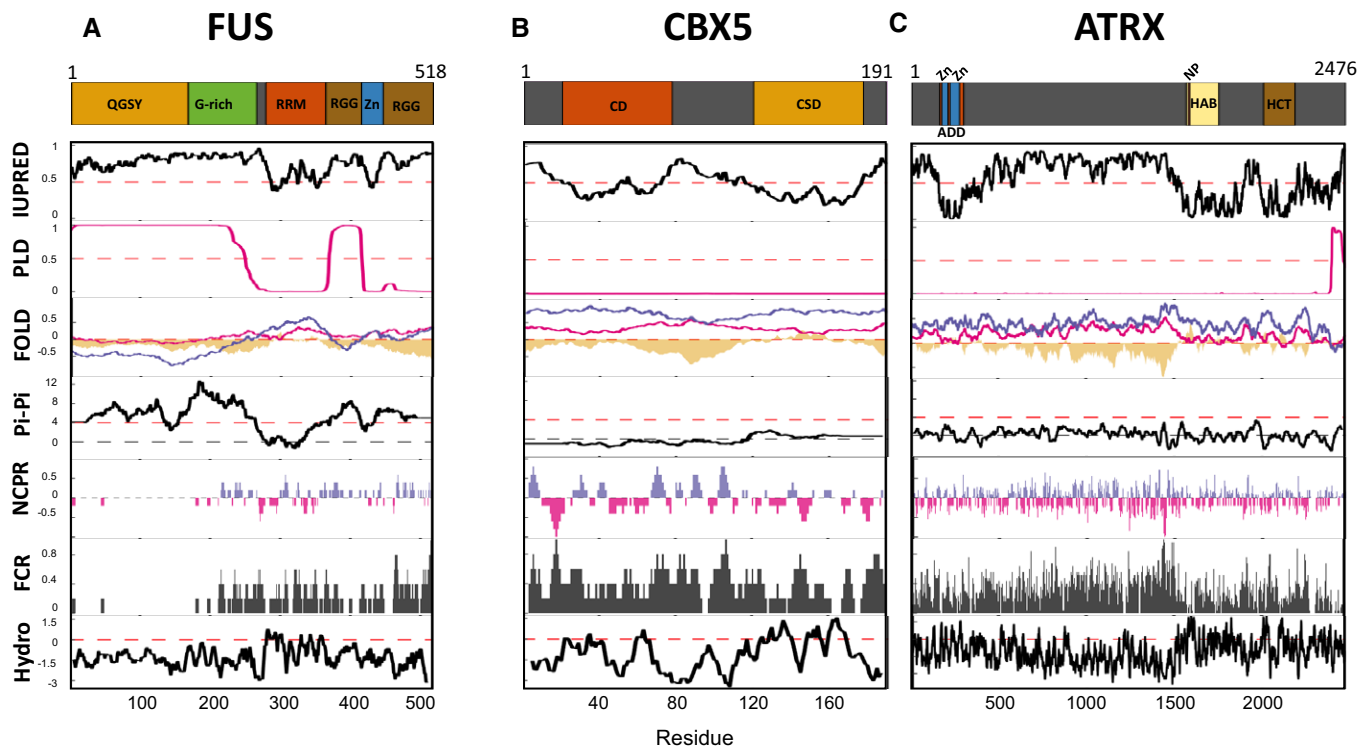


Figure 2. In-depth analysis of phase separation potential for FUS, CBX5 and ATRX.

The analysis of regions of protein primary sequence potentially contributing to liquid–liquid phase separation for FUS, CBX5 and ATRX (see also Fig EV1) was implemented following the same methodology as published in Alberti *et al* [101]. At the top, a schematic representation of the proteins is shown highlighting the different domains catalogued in UniProt. IUPRED; intrinsic disorder prediction using the IUPRED algorithm where values above 0.5 are considered disordered. PLD; prion-like domain prediction using the PLAAC algorithm where a value above 0.5 is considered a prion-like domain. FOLD; intrinsic disorder prediction with PLAAC (pink) or the PAPA (purple) algorithms and the fold index (yellow). PI-Pi; phase separation predictor based on propensity for Pi-Pi contacts where a region of a protein is predicted to phase separate when its mean value is above 4. NCPR; net charge per residue and FCR; fraction of charged residues (sliding window of 5 using the localCIDER version 0.1.14). Hydro; hydrophobicity (sliding window of 9 using the Kyte and Doolittle scale).

- A For FUS, the following domains or regions are depicted: QGSY, glutamine/glycine/serine/tyrosine-rich region (yellow); G-rich, glycine-rich region (green); RRM, RNA recognition motif domain (orange); RGG, arginine/glycine-rich region (brown); Zn, zinc finger domain (blue).
- B For CBX5, the chromo (CD in orange) and the chromo shadow (CSD in yellow) domains are shown.
- C For ATRX, the following domains or regions are depicted: ADD, ATRX-Dnmt3-Dnmt3L domain (orange); Zn, zinc finger domains (blue); HAB, helicase ATP binding (beige); NP; nucleotide (ATP) binding (red); HCT, helicase C-terminal (brown).

as chromocentres, temporally and spatially correlate with the appearance of a phase-separated heterochromatic state?

Can we predict phase transition occurrence during mouse preimplantation development? We reasoned that an analysis of the patterns of expression of heterochromatin proteins that we identified (Table EV1) during these stages of development, in combination with the knowledge of their predicted phase separation properties, can give a first forecast of the dynamics of phase-separated heterochromatin in mouse embryos. An analysis of publicly available RNAseq datasets [79] indicated a clear average upregulation of the genes encoding the 148 heterochromatin proteins at the 4-cell stage (Fig 3A). This suggests firstly that, for the most part, these genes do not exhibit the typical dynamics of maternally inherited transcripts, a fact not insignificant considering the large pool of such transcripts. Additionally, this trend was markedly different to the expression dynamics of the other groups of genes analysed, which included genes coding for chromatin proteins, in general, DNA-binding proteins, as well as the complete nucleome (Fig 3A). Thus, it is likely that constitutive heterochromatin is largely remodelled after fertilisation, fitting with the

known dynamics of heterochromatin markers by immunostaining and of H3K9me3 ChIPseq [74,80]. Interestingly, the timing of this increase also correlates with the reported increase in chromatin compaction between the 2-cell and 8-cell stages [81,82] and the establishment of chromocentres from the late 2-cell stage [83].

Analysis of mass spectrometry data [84] showed that the 108 (73%) of 148 heterochromatin proteins detected displayed a collective increase in protein levels towards the blastocyst stage (Fig 3B). In fact, this tendency is more consistent at the protein level than for the mRNA levels. The heterochromatin proteins displayed increasing expression over the preimplantation period, with a clear, sharp increase after the 8-cell stage (Fig 3B). Thus, constitutive heterochromatin may gradually mature during the early period of mammalian development. While we did not observe any correlation between the degree of predicted disorder and expression level (not shown), the clear increase in both mRNA (at the 4-cell stage) and protein (at the morula–blastocyst stage) suggests that the proteins more likely to promote heterochromatic phase separation are on average expressed at later timepoints during mouse preimplantation development.

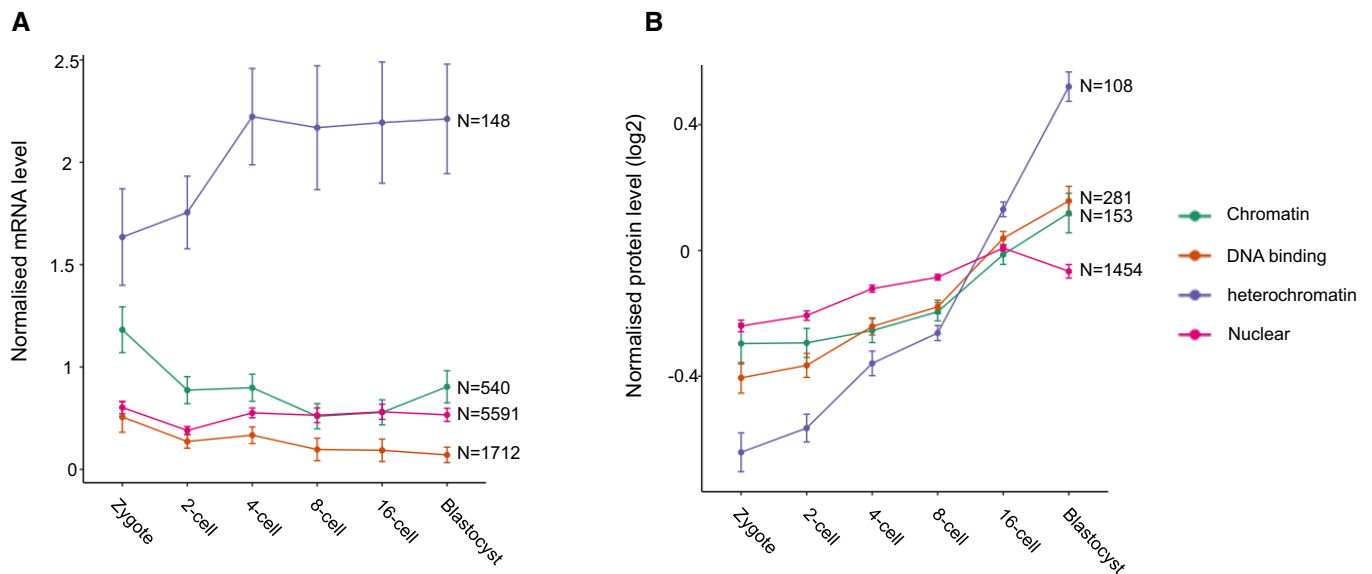


Figure 3. mRNA and protein levels of the selected heterochromatin and control datasets during mouse preimplantation embryonic development.

A Mean \pm SEM mRNA levels normalised to the sum of expression across detected genes during preimplantation development (data from Deng *et al*). The analysis was carried out for the 148 genes coding for the heterochromatin proteins as well as all the genes with “chromatin” (540), “DNA binding” (1,712) or “nuclear” (5,591) in their GO terms.

B Mean \pm SEM protein levels during preimplantation development by mass spectrometry, normalised to average expression of all detected proteins (data from Gao *et al*). The analysis was carried out for the 108 detected heterochromatin proteins as well as all the proteins with “chromatin” (153), “DNA binding” (281) or “nuclear” (1,454) in the GO terms of their corresponding genes.

Thus, heterochromatin, which is atypical in numerous other aspects in the period of development immediately after fertilisation [85], may also not phase separate at this stage. Potentially, phase separation of heterochromatin only occurs as it matures, after chromocentre formation at the late 2-cell stage, and chromatin compaction and silencing of repetitive elements at the 8-cell stage. It will be interesting to determine the point at which heterochromatin is able to initiate phase separation and its functional contribution to the embryo.

Current *in vivo* assays to address phase separation in heterochromatin establishment

To date, all methods to study phase separation *in vivo* are microscopy-based, primarily using differential interference contrast microscopy or fluorescence microscopy to visualise the sphericity, number and dynamics of condensates [20,21,54]. Indeed, the liquid state of a membrane-less organelle can be called by demonstrating their ability to fuse or fission [17,21,27,28,33,64]. Fluorescence recovery after photobleaching (FRAP) can also be used to determine whether proteins diffuse inside the phase-separated compartment as well as between the surrounding environment [18]. Some studies target part of the membrane-less organelle in order to assess internal diffusion of tagged proteins of interest [20,27]. In addition, bleaching the whole condensate assesses the diffusion of the protein of interest between the condensate and its environment [28,33,36,63,86]. Importantly, FRAP has been used to measure the mobile and immobile HP1 α fractions to uncover liquid-like properties of heterochromatin in the developing *Drosophila* embryo [17].

Imaging analyses are in general amenable to early mouse embryos, but phototoxicity is a major problem and must be taken

into consideration when used in live embryos. FRAP has previously been used to study dynamics of histone proteins during preimplantation development [81,87]. Therefore, implementation of FRAP and differential interference contrast microscopy in embryos could address whether specific proteins and/or compartments can fuse, as well as determine diffusion dynamics, which has been done for, e.g., transcription factors [88]. However, additional strategies requiring higher photon absorption, such as the number and brightness (N&B) [89] and raster image correlation spectroscopy (RICS) [90], will require major adaptation. Indeed, the N&B method was used in *Drosophila* to show that HP1 α exhibits coordinated movement at the heterochromatin boundary, while the RICS method showed that HP1 α diffusion was slower in heterochromatin. As both of these observations are predicted to occur at the boundary of a liquid condensate, it was concluded that HP1 α dynamics are consistent with the heterochromatin domains being in a liquid state [17].

The liquid state of condensates *in vivo* can also be assessed using 1,6-hexanediol, an aliphatic alcohol that disturbs weak hydrophobic interactions and thus liquid condensates [91]. However, this compound can be rather toxic for eukaryotic cells and is therefore typically used within short time windows [17,33]. Mutating amino acids necessary for phase separation of the protein of interest may be another strategy to manipulate liquid condensates *in vivo*, in order to probe function. This has been done, for example, by mutating the tyrosines to serines in the IDR of FUS, which disturbs phase separation of FUS [64,92]. Modifying relevant serines and threonines to glutamic acid, which mimics phosphorylation, is also another means of the disturbing phase separation [64,93]. Acetylation of intrinsically disordered regions has also been shown to regulate

phase separation [65] and mimicking acetylation may provide additional experimental strategies.

Finally, it is important to note that we have not considered a possible role for RNA interactions in this current work. Membraneless organelles are enriched in RNAs and RNA-binding proteins [60,94]. The role of RNA interactions in phase separation has been extensively characterised *in vitro*, as well as *in silico*, and less often *in vivo*. Ribosomal RNA transcription, for example, regulates nucleoli assembly [95]. In *C. elegans*, P granule formation has been suggested to be mediated by interactions between mRNA and the PGL-3 protein [96]. mRNA also controls the phase behaviour of RNA-binding proteins such as TDP43 and FUS, which will form liquid droplets or solid aggregates depending on mRNA availability [97]. Several RNA-binding proteins have the ability to phase separate, such as the heterogeneous nuclear ribonucleoproteins hnRNPA1 and hnRNPA2 [28,98]. In this context, it is important to note that major satellites are robustly transcribed in zygotes and 2-cell stage embryos [80,99,100]. This raises the interesting possibility that this RNA may be a good candidate as a scaffold for phase-separated domains *in vivo* in the mouse embryo.

The plethora of these studies, as well as the nature of the open questions to address how, when and under which conditions heterochromatin phase separates *in vivo*, promises exciting research in the years to come. From the technical viewpoint, it will be important to define the standards of the experimental approaches used to study phase separation *in vivo*, as recently proposed [101]. From the developmental perspective, it will be exciting to apply different methodologies to determine whether and when phase separation regulates establishment of heterochromatin.

Materials and Methods

Merging mass spectrometry datasets

Unless otherwise stated, all analyses were performed in R studio (version 1.2.1335) with the R version (R version 3.5.2 (2018-12-20)). The bioinformatic analysis was based on 7 mass spectrometry studies performed in mammalian cells [6,39–44]. Proteins predicted to be heterochromatic were selected based on their ability to bind H3K9me3, H3K9me3-modified nucleosomes with and without DNA methylation, or to their enrichment in the sonication resistance fragment of the chromatin. Due to the little overlap between the mass spectrometry studies, the proteins present in more than one mass spectrometry study were kept for the analysis. Note that the antibodies used in these studies have been thoroughly characterised, as follows: Bartke, Becker, Engelen, Ji and Soldi all used the same antibody (Abcam ab8898), which was reported to be highly specific to H3K9me3, with no binding to H3K9me2 or H3K9me1, with only a slight cross-reactivity to H3K27me3. The two other studies used H3K9me3 peptides as bait in pulldowns.

Identification of orthologs across model organisms

The orthologs in *D. rerio*, *D. melanogaster* and *C. elegans* were identified using the Ensembl project website with the Ensembl release 94 [102] and downloading a dataset with the orthologs in the different species of the mouse genes (GRCm38.p6). For the

S. pombe orthologs, a dataset containing the human orthologs of *S. pombe* orthologs was downloaded from the PomBase project website [103].

Disorder analysis

The control groups for the analysis of disorder content were selected by retrieving, from the Ensembl project website with the Ensembl release 96 [102], all the mouse genes (GRCm38.p6) or the ones which have chromatin, nuclear or DNA binding in their GO Term Names. All the genes also present in the heterochromatin dataset were later removed from these control groups. In order to compare the control and the heterochromatin groups, 148 genes were randomly sampled without replacement from each of the control datasets to obtain the final control groups. The fasta files from all the selected proteins were then downloaded from NCBI using the `efetch` function of the Entrez package build in Biopython [104]. To calculate the length in amino acids of the proteins in each group, the fasta files were imported in Rstudio with the `read.fasta` function of the `seqinr` package (version 3.4.5). For the disorder analysis, disorder estimates were generated for the proteins in the different groups using two prediction algorithms, PONDR-VLXT [58] and IUPred2 long disorder [59]. The predictors give a value between 0 and 1 for each amino acid where above 0.5 is predicted to lie within a disordered region. For each predictor, the average value (average disorder score) and the percentage of amino acids with a value over 0.5 (overall percentage disorder) were then calculated for each protein. The latter analysis was done on the average of the values obtained with the two predictors. The analysis of the length of the disorder fragments was done with the PONDR-VLXT. This was done by counting the number of predicted disorder fragments of different size in amino acids across the different proteins of the same group. For length of disordered regions, segments shorter than 30 amino acids were removed (based on Forman-kay *et al* and Ward *et al* [56,105]). To assess the statistical significance of the difference between the heterochromatic group and the different control groups, a two-sided unpaired Wilcoxon rank-sum test was performed in R with the `wilcox.test` function with default settings, as the data were found to be non-parametric. All the plots were done using `ggplot2`.

Analysis of bona fide heterochromatin proteins

The 7 bona fide heterochromatin proteins were selected based on their specific association to major satellite genomic regions as described by Saksouk *et al* [68]. Briefly, proteins enriched at major satellite genomic regions, and therefore constitutive heterochromatin, were identified by proteomics of isolated chromatin segments (PiCH) in mouse embryonic stem cells. The 7 bona fide heterochromatin proteins are depleted at major satellites when Suv39h1 and Suv39h2 are knocked out and have been identified as suppressors of variegation and modifiers of murine metastable epialleles.

The `drawProteins` (version 1.2.0) package was used to obtain the features of the 7 bona fide heterochromatin proteins from the UniProt Features API. The prediction of intrinsic disorder was done with the IUPred2 long-disorder algorithm [59]. The prion-like domains were predicted with the PLAAC algorithm using the website (<http://plaac.wi.mit.edu>) with the default settings [106]. The intrinsic disorder prediction with the PLAAC, the PAPA and the

fold index was obtained with the same website. To predict the phase separation property of the 7 bona fide heterochromatin proteins based on propensity for Pi-Pi contacts, the Pi-Pi predictor was used online on the Forman-Kay's laboratory website [61]. The net charge per residue and the fraction of charged residues were obtained using the localCIDER (version 0.1.14) [107] with a sliding window of 5. The hydrophobicity was calculated with the ExPASy website [108] with the Kyte and Doolittle scale [109] and a sliding window of 9. All the plots shown in Figs 2 and EV1 were done with ggplot2.

Analysis of gene expression in mouse preimplantation embryos

RNAseq dataset previously published [79] was analysed downloading the expression matrix provided in a GitHub repository ("jhsiao999/singleCellRNASeqMouseDengESC") which contains the data from National Center for Biotechnology Information Gene Expression Omnibus ("GSE45719"). The expression matrix was later normalised by library size by dividing the counts by the sum of expression across detected genes in each sample. Heterochromatin (148), chromatin (540), DNA binding (1,712) or nuclear (5,591) genes were extracted from the datasets based on GO terms, excepting for the "heterochromatin" dataset, which was selected in the current study as described above. The mean normalised mRNA levels and standard errors for each gene group and embryonic development stage were plotted using ggplot2.

Analysis of protein levels in mouse preimplantation embryos

The mass spectrometry study of preimplantation development by Gao *et al* [84] was analysed to investigate the expression pattern of the heterochromatin (106) and control groups. The control groups correspond to all the proteins with chromatin (153), DNA binding (281) or nuclear (1,454) in the GO terms of their corresponding genes. The protein levels were normalised to average expression of all detected proteins in each sample and transformed to a base 2 logarithmic scale. The normalised mean protein expression levels and standard errors for each protein group and embryonic development stage were plotted using ggplot2.

Expanded View for this article is available online.

Acknowledgements

The work in the Torres-Padilla laboratory is funded by the Helmholtz Association, the German Research Council (CRC 1064) and H202 Marie-Curie Actions ITN EpiSystem and ChromDesign.

Author contributions

MG, AB and M-ET-P conceived the work and wrote the manuscript. MG performed bioinformatic analyses. AB and M-ET-P supervised the work.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Woodcock CL, Ghosh RP (2010) Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol* 2: a000596
- Heitz E (1928) Das Heterochromatin der Moose. *Jahrb Wiss Bot* 69: 762–818
- Janssen A, Colmenares SU, Lee T, Karpen GH (2019) Timely double-strand break repair and pathway choice in pericentromeric heterochromatin depend on the histone demethylase dKDM4A. *Genes Dev* 33: 103–115
- Jack AP, Bussemer S, Hahn M, Punzeler S, Snyder M, Wells M, Csankovszki G, Solovei I, Schotta G, Hake SB (2013) H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization. *PLoS One* 8: e51765
- Jost KL, Bertulat B, Cardoso MC (2012) Heterochromatin and gene positioning: inside, outside, any side? *Chromosoma* 121: 555–563
- Becker JS, McCarthy RL, Sidoli S, Donahue G, Kaeding KE, He Z, Lin S, Garcia BA, Zaret KS (2017) Genomic and proteomic resolution of heterochromatin and its restriction of alternate fate genes. *Mol Cell* 68: 1023–1037 e1015
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD *et al* (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593–599
- Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410: 116–120
- Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410: 120–124
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 18: 1251–1262
- Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, Zeissler U, Lappe M, Schubeler D, Torres-Padilla ME, Schneider R (2009) H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol* 16: 777–781
- Velazquez Camacho O, Galan C, Swist-Rosowska K, Ching R, Gamalinda M, Karabiber F, De La Rosa-Velazquez I, Engist B, Koschorz B, Shukeir N *et al* (2017) Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife* 6: e25293
- Johnson WL, Yewdell WT, Bell JC, McNulty SM, Duda Z, O'Neill RJ, Sullivan BA, Straight AF (2017) RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* 6: e25299
- Muchardt C, Guilleme M, Seeler JS, Trouche D, Dejean A, Yaniv M (2002) Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep* 3: 975–981
- Maison C, Bailly D, Roche D, Montes de Oca R, Probst AV, Vassias I, Dingli F, Lombard B, Loew D, Quivy JP *et al* (2011) SUMOylation promotes *de novo* targeting of HP1alpha to pericentric heterochromatin. *Nat Genet* 43: 220–227
- Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ (2017) Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* 547: 236–240
- Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH (2017) Phase separation drives heterochromatin domain formation. *Nature* 547: 241–245

18. Hyman AA, Weber CA, Julicher F (2014) Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* 30: 39–58
19. Wheeler RJ, Hyman AA (2018) Controlling compartmentalization by non-membrane-bound organelles. *Philos Trans R Soc Lond B Biol Sci* 373: 20170193
20. Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, Gharakhani J, Julicher F, Hyman AA (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* 324: 1729–1732
21. Brangwynne CP, Mitchison TJ, Hyman AA (2011) Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc Natl Acad Sci USA* 108: 4334–4339
22. Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L et al (2018) Protein phase separation: a new phase in cell biology. *Trends Cell Biol* 28: 420–435
23. Courchaine EM, Lu A, Neugebauer KM (2016) Droplet organelles? *EMBO J* 35: 1603–1612
24. Shin Y, Brangwynne CP (2017) Liquid phase condensation in cell physiology and disease. *Science* 357: eaaf4382
25. Stanek D, Fox AH (2017) Nuclear bodies: news insights into structure and function. *Curr Opin Cell Biol* 46: 94–101
26. Banani SF, Lee HO, Hyman AA, Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 18: 285–298
27. Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, Brangwynne CP (2016) Coexisting liquid phases underlie nucleolar subcompartments. *Cell* 165: 1686–1697
28. Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, Mittag T, Taylor JP (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163: 123–133
29. Hennig S, Kong G, Mannen T, Sadowska A, Kobelke S, Blythe A, Knott CJ, Iyer KS, Ho D, Newcombe EA et al (2015) Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol* 210: 529–539
30. Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, Darzacq X, Zhou Q (2018) Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* 558: 318–323
31. Boehning M, Dugast-Darzacq C, Rankovic M, Hansen AS, Yu T, Marie-Nelly H, McSwiggen DT, Kocic G, Dailey GM, Cramer P et al (2018) RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* 25: 833–840
32. Cho W, Spille J, Hecht M, Lee C, Li C, Grube V, Cisse II (2018) Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 361: 412–415
33. Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC et al (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361: eaar3958
34. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA (2017) A phase separation model for transcriptional control. *Cell* 169: 13–23
35. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J et al (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149: 753–767
36. Tatabasian R, Kent S, Brown K, Yao T, Duc HN, Huynh TN, Zhen CY, Ma B, Wang H, Ren X (2018) Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J Biol Chem* 294: 1451–1463
37. Erdel F, Rippe K (2018) Formation of chromatin subcompartments by phase separation. *Biophys J* 114: 2262–2270
38. Shin Y, Chang YC, Lee DSW, Berry J, Sanders DW, Ronceray P, Wingreen NS, Haataja M, Brangwynne CP (2018) Liquid nuclear condensates mechanically sense and restructure the genome. *Cell* 175: 1481–1491 e1413
39. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143: 470–484
40. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG et al (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142: 967–980
41. Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M, Mann M (2013) A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol Cell* 49: 368–378
42. Engelen E, Brandsma JH, Moen MJ, Signorile L, Dekkers DH, Demmers J, Kockx CE, Ozgur Z, van Ijcken WF, van den Berg DL et al (2015) Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nat Commun* 6: 7155
43. Ji X, Dadon DB, Abraham BJ, Lee TI, Jaenisch R, Bradner JE, Young RA (2015) Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proc Natl Acad Sci USA* 112: 3841–3846
44. Soldi M, Bonaldi T (2013) The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components. *Mol Cell Proteomics* 12: 764–780
45. Allshire RC, Ekwall K (2015) Epigenetic regulation of chromatin states in *Schizosaccharomyces pombe*. *Cold Spring Harb Perspect Biol* 7: a018770
46. Ahringer J, Gasser SM (2018) Repressive chromatin in *Caenorhabditis elegans*: establishment, composition, and function. *Genetics* 208: 491–511
47. Mteirek R, Gueguen N, Jensen S, Brassat E, Vaury C (2014) Drosophila heterochromatin: structure and function. *Curr Opin Insect Sci* 1: 19–24
48. Elgin SC, Reuter G (2013) Position-effect variegation, heterochromatin formation, and gene silencing in *Drosophila*. *Cold Spring Harb Perspect Biol* 5: a017780
49. Bayne EH, Bijos DA, White SA, de Lima Alves F, Rappsilber J, Allshire RC (2014) A systematic genetic screen identifies new factors influencing centromeric heterochromatin integrity in fission yeast. *Genome Biol* 15: 481
50. Towbin BD, Gonzalez-Aguilera C, Sack R, Gaidatzis D, Kalck V, Meister P, Askjaer P, Gasser SM (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* 150: 934–947
51. Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, Rosen MK (2016) Compositional control of phase-separated cellular bodies. *Cell* 166: 651–663
52. Lin Y, Currie SL, Rosen MK (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem* 292: 19110–19120
53. Mitrea DM, Cika JA, Stanley CB, Nourse A, Onuchic PL, Banerjee PR, Phillips AH, Park CG, Deniz AA, Kriwacki RW (2018) Self-interaction of NPM1 modulates multiple mechanisms of liquid-liquid phase separation. *Nat Commun* 9: 842

54. Nott TJ, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowitz A, Craggs TD, Bazett-Jones DP, Pawson T, Forman-Kay JD *et al* (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57: 936–947
55. Pak CW, Kosno M, Holehouse AS, Padrick SB, Mittal A, Ali R, Yunus AA, Liu DR, Pappu RV, Rosen MK (2016) Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol Cell* 63: 72–85
56. Forman-Kay JD, Mittag T (2013) From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* 21: 1492–1499
57. Martin EW, Mittag T (2018) Relationship of sequence and phase separation in protein low-complexity regions. *Biochemistry* 57: 2478–2487
58. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK (2001) Sequence complexity of disordered protein. *Proteins* 42: 38–48
59. Dosztanyi Z (2018) Prediction of protein disorder based on IUPred. *Protein Sci* 27: 331–340
60. Sawyer IA, Sturgill D, Dundr M (2019) Membraneless nuclear organelles and the search for phases within phases. *Wiley Interdiscip Rev RNA* 10: e1514
61. Vernon RM, Chong PA, Tsang B, Kim TH, Bah A, Farber P, Lin H, Forman-Kay JD (2018) Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* 7: e31486
62. Aumiller WM Jr, Keating CD (2016) Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nat Chem* 8: 129–137
63. Rai AK, Chen JX, Selbach M, Pelkmans L (2018) Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* 559: 211–216
64. Bracha D, Walls MT, Wei MT, Zhu L, Kurian M, Avalos JL, Toettcher JE, Brangwynne CP (2018) Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* 175: 1467–1480 e1413
65. Saito M, Hess D, Eglinger J, Fritsch AW, Kreysing M, Weinert BT, Choudhary C, Matthias P (2019) Acetylation of intrinsically disordered regions regulates phase separation. *Nat Chem Biol* 15: 51–61
66. Yeo GC, Keeley FW, Weiss AS (2011) Coacervation of tropoelastin. *Adv Colloid Interface Sci* 167: 94–103
67. Wang J, Choi JM, Holehouse AS, Lee HO, Zhang X, Jahnel M, Maharana S, Lemaître R, Pozniakovskiy A, Drechsel D *et al* (2018) A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* 174: 688–699 e616
68. Saksouk N, Barth TK, Ziegler-Birling C, Olova N, Nowak A, Rey E, Mateos-Langerak J, Urbach S, Reik W, Torres-Padilla ME *et al* (2014) Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Mol Cell* 56: 580–594
69. Murthy AC, Dignon GL, Kan Y, Zerze GH, Parekh SH, Mittal J, Fawzi NL (2019) Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat Struct Mol Biol* 26: 637–648
70. Burton A, Muller J, Tu S, Padilla-Longoria P, Guccione E, Torres-Padilla ME (2013) Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep* 5: 687–701
71. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R, Surani MA (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 452: 877–881
72. Sasaki H, Matsui Y (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat Rev Genet* 9: 129–140
73. Arney KL, Bao S, Bannister AJ, Kouzarides T, Surani MA (2002) Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int J Dev Biol* 46: 317–320
74. 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
75. 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
76. Lange UC, Siebert S, Wossidlo M, Weiss T, Ziegler-Birling C, Walter J, Torres-Padilla ME, Daujat S, Schneider R (2013) Dissecting the role of H3K64me3 in mouse pericentromeric heterochromatin. *Nat Commun* 4: 2233
77. Kourmouli N, Jeppesen P, Mahadevaiah S, Burgoyne P, Wu R, Gilbert DM, Bongiorno S, Prantera G, Fanti L, Pimpinelli S *et al* (2004) Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J Cell Sci* 117: 2491–2501
78. Wongtawan T, Taylor JE, Lawson KA, Wilmot I, Pennings S (2011) Histone H4K20me3 and HP1alpha are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J Cell Sci* 124: 1878–1890
79. Deng Q, Ramskold D, Reinius B, Sandberg R (2014) Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* 343: 193–196
80. Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U, Kolb C, Otte AP, Koseki H, Orkin SH *et al* (2008) PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* 40: 411–420
81. Boskovic A, Eid A, Pontabry J, Ishiuchi T, Spiegelhalter C, Raghu Ram EV, Meshorer E, Torres-Padilla ME (2014) Higher chromatin mobility supports totipotency and precedes pluripotency *in vivo*. *Genes Dev* 28: 1042–1047
82. 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
83. Martin C, Beaujean N, Brochard V, Audouard C, Zink D, Debey P (2006) Genome restructuring in mouse embryos during reprogramming and early development. *Dev Biol* 292: 317–332
84. 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
85. Burton A, Torres-Padilla ME (2010) Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. *Brief Funct Genomics* 9: 444–454
86. Li P, Banjade S, Cheng HC, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, King DS, Banani SF *et al* (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483: 336–340
87. 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
88. 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
89. Digman MA, Dalal R, Horwitz AF, Gratton E (2008) Mapping the number of molecules and brightness in the laser scanning microscope. *Biophys J* 94: 2320–2332
90. Rossow MJ, Sasaki JM, Digman MA, Gratton E (2010) Raster image correlation spectroscopy in live cells. *Nat Protoc* 5: 1761–1774

91. Kroschwald S, Maharana S, Simon A (2017) Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments. *Matters* <https://doi.org/10.19185/matters.201702000010>
92. Wei MT, Elbaum-Garfinkle S, Holehouse AS, Chen CC, Feric M, Arnold CB, Priestley RD, Pappu RV, Brangwynne CP (2017) Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat Chem* 9: 1118–1125
93. Monahan Z, Ryan VH, Janke AM, Burke KA, Rhoads SN, Zerze GH, O’Meally R, Dignon GL, Conicella AE, Zheng W et al (2017) Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J* 36: 2951–2967
94. Fay MM, Anderson PJ (2018) The role of RNA in biological phase separations. *J Mol Biol* 430: 4685–4701
95. Berry J, Weber SC, Vaidya N, Haataja M, Brangwynne CP (2015) RNA transcription modulates phase transition-driven nuclear body assembly. *Proc Natl Acad Sci USA* 112: E5237–E5245
96. Saha S, Weber CA, Nusch M, Adame-Arana O, Hoege C, Hein MY, Osborne-Nishimura E, Mahamid J, Jahnel M, Jawerth L et al (2016) Polar positioning of phase-separated liquid compartments in cells regulated by an mRNA competition mechanism. *Cell* 166: 1572–1584 e1516
97. Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovskiy A, Poser I, Bickle M, Rizk S, Guillén-Boixet J, Franzmann TM et al (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* 360: 918–921
98. Xiang S, Kato M, Wu LC, Lin Y, Ding M, Zhang Y, Yu Y, McKnight SL (2015) The LC domain of hnRNP A2 adopts similar conformations in hydrogel polymers, liquid-like droplets, and nuclei. *Cell* 163: 829–839
99. Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME (2010) Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol* 12: 853–862
100. Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G (2010) A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* 19: 625–638
101. Alberti S, Gladfelter A, Mittag T (2019) Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* 176: 419–434
102. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Giron CG et al (2018) Ensembl 2018. *Nucleic Acids Res* 46: D754–D761
103. Lock A, Rutherford K, Harris MA, Hayles J, Oliver SG, Bahler J, Wood V (2018) PomBase 2018: user-driven reimplementations of the fission yeast database provides rapid and intuitive access to diverse, interconnected information. *Nucleic Acids Res* 47: D821–D827
104. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B et al (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25: 1422–1423
105. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 337: 635–645
106. Lancaster AK, Nutter-Upham A, Lindquist S, King OD (2014) PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* 30: 2501–2502
107. Holehouse AS, Das RK, Ahad JN, Richardson MO, Pappu RV (2017) CIDER: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys J* 112: 16–21
108. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31: 3784–3788
109. Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* 157: 105–132



License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Expanded View Figures

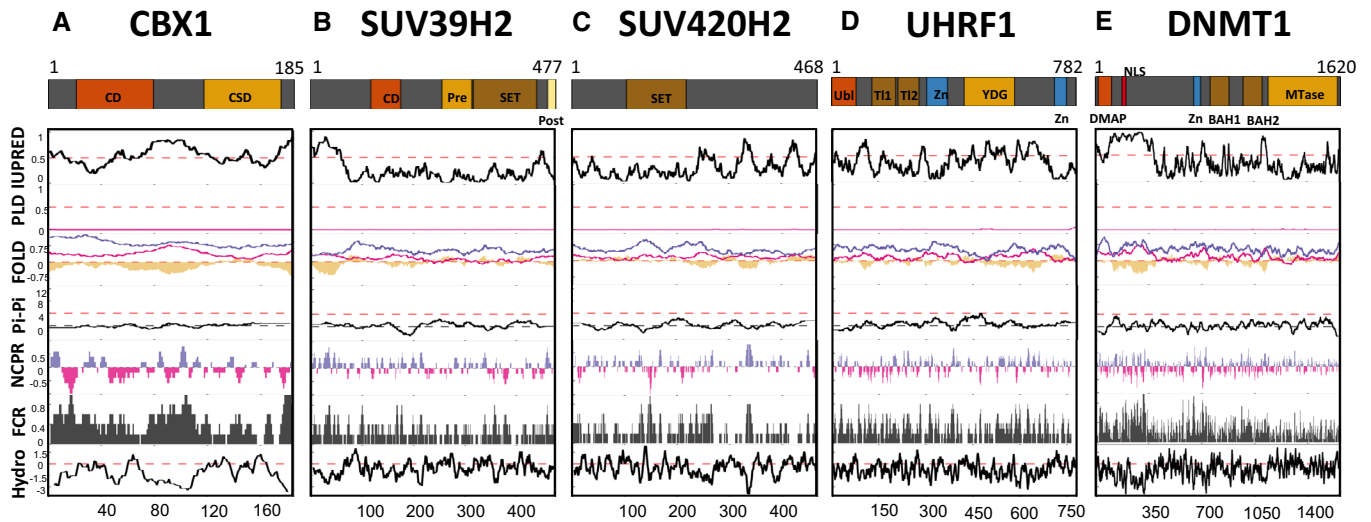


Figure EV1. In-depth analysis of phase separation potential for the bona fide heterochromatin proteins.

The analysis of regions of protein primary sequence potentially contributing to liquid–liquid phase separation for CBX1, SUV39H2, SUV420H2, UHRF1 and DNMT1 was implemented following the same methodology as in Fig 2.

- A For CBX1, the chromo (CD in orange) and the chromo shadow (CSD in yellow) domains are shown.
- B For SUV39H2, the following domains or regions are depicted: CD, chromodomain (orange); Pre, Pre-SET domain (yellow); SET, SET domain (brown); Post, Post-SET domain (beige).
- C For SUV420H2, the SET domain (brown).
- D For UHRF1, the following domains or regions are depicted: Ubl, ubiquitin-like domain (orange); TI1 and TI2, Tudor-like 1 and 2 regions (brown); Zn, zinc finger domains (blue); YDG, YDG domain (yellow).
- E For DNMT1, the following domains or regions are depicted: DMAP, DMAP-interaction domain (orange); NLS, nuclear localisation signal (red); Zn, zinc finger domain (blue); BAH1 and BAH2, bromo-adjacent homology 1 and 2 domains (brown); Mtase, SAM-dependent Mtase C5 type (yellow).

	Gene name	Ensembl id	Mammals	Orthologs	Disorder score	overall percentage disorder		Gene name	Ensembl id	Mammals	Orthologs	Disorder score	overall percentage disorder	
1	cbx3	ENSMUSG00000029836	5	5	0.49	44.53		75	rbm14	ENSMUSG00000006456	2	4	0.55	48.95
2	cbx5	ENSMUSG00000009575	5	5	0.5	45.55		76	rbmx	ENSMUSG00000031134	2	4	0.71	75.07
3	cbx1	ENSMUSG00000018666	4	5	0.54	54.86		77	syncrip	ENSMUSG000000032423	2	4	0.46	39.97
4	orc2	ENSMUSG00000026037	4	5	0.34	29.25		78	tarbcp	ENSMUSG00000041459	2	4	0.39	35.75
5	hnrnpk	ENSMUSG00000021546	3	5	0.51	47.95		79	tra2a	ENSMUSG00000029817	2	4	0.7	75.8
6	incenp	ENSMUSG00000024660	3	5	0.66	75.4		80	tra2b	ENSMUSG00000022858	2	4	0.67	68.75
7	orc1	ENSMUSG00000028587	3	5	0.38	33.39		81	trip12	ENSMUSG00000026219	2	4	0.41	37.78
8	rcc1	ENSMUSG00000028896	3	5	0.36	26.96		82	cby1	ENSMUSG00000059288	4	3	0.47	48.73
9	chd1	ENSMUSG00000023852	2	5	0.49	49.97		83	cdca8	ENSMUSG00000028873	3	3	0.53	48.96
10	ddx17	ENSMUSG00000055065	2	5	0.44	37.15		84	dek	ENSMUSG00000021377	3	3	0.62	65.13
11	ddx47	ENSMUSG00000030204	2	5	0.31	25.6		85	hnrnpa2b1	ENSMUSG00000004980	3	3	0.47	47.87
12	ddx5	ENSMUSG00000020719	2	5	0.37	27.36		86	hnrnpab	ENSMUSG00000020358	3	3	0.56	48.77
13	dnajc9	ENSMUSG00000021811	2	5	0.4	32.24		87	hnrnpd	ENSMUSG00000000568	3	3	0.46	44.65
14	h2afv	ENSMUSG00000041126	2	5	0.46	39.06		88	mga	ENSMUSG00000033943	3	3	0.54	55.46
15	heatr1	ENSMUSG00000050244	2	5	0.17	8.4		89	rcc2	ENSMUSG00000040945	3	3	0.35	27.69
16	hist1h2bc	ENSMUSG00000018102	2	5	0.5	40.48		90	suz12	ENSMUSG000000017548	3	3	0.45	44.4
17	hist1h2bm	ENSMUSG000000114279	2	5	0.51	40.48		91	dnmt1	ENSMUSG00000004099	2	3	0.39	33.36
18	hmg2b	ENSMUSG00000054717	2	5	0.62	68.1		92	ncf	ENSMUSG00000026234	2	3	0.61	62.73
19	hnrnpjm	ENSMUSG00000059208	3	5	0.49	47.12		93	cenpf	ENSMUSG00000018509	2	3	0.45	42.46
20	hnrnpu1	ENSMUSG00000040725	2	5	0.56	57.22		94	fus	ENSMUSG00000030795	2	3	0.73	82.63
21	mcm2	ENSMUSG00000002870	2	5	0.39	31.75		95	h1fo	ENSMUSG00000096210	2	3	0.73	84.02
22	mcm3	ENSMUSG00000041859	2	5	0.39	31.1		96	hdgf	ENSMUSG00000004897	2	3	0.64	66.03
23	mcm4	ENSMUSG00000022673	2	5	0.38	30.05		97	ifit3	ENSMUSG00000032178	2	3	0.56	61.47
24	mcm5	ENSMUSG00000005410	2	5	0.34	23.4		98	khdrbs1	ENSMUSG00000028790	2	3	0.6	63.88
25	rbbp7	ENSMUSG00000031353	2	5	0.29	18.47		99	nol11	ENSMUSG00000018433	2	3	0.21	12.93
26	rbm39	ENSMUSG00000027620	2	5	0.46	39.9		100	phip	ENSMUSG00000032253	2	3	0.42	39.84
27	rpl4	ENSMUSG00000002399	2	5	0.48	43.32		101	rbp4	ENSMUSG000000057236	2	3	0.32	20.71
28	smc1a	ENSMUSG00000041133	2	5	0.45	44.24		102	saib	ENSMUSG00000071054	2	3	0.73	85.86
29	smc3	ENSMUSG00000024974	2	5	0.42	38.74		103	saib2	ENSMUSG00000042625	2	3	0.67	77.7
30	srr1	ENSMUSG00000037364	2	5	0.59	55.83		104	stfm	ENSMUSG00000032212	2	3	0.45	42.49
31	tkt	ENSMUSG00000021957	2	5	0.25	12.04		105	smarca5	ENSMUSG00000031715	2	3	0.38	31.49
32	top1	ENSMUSG00000070544	2	5	0.49	42.18		106	snrpa	ENSMUSG00000061479	2	3	0.5	45.64
33	top2a	ENSMUSG00000020914	2	5	0.37	31.12		107	taf11	ENSMUSG00000020680	2	3	0.62	72.55
34	top2b	ENSMUSG00000017485	2	5	0.35	29.40		108	usp48	ENSMUSG00000043411	2	3	0.29	17.63
35	u2af2	ENSMUSG00000030435	2	5	0.45	42.31		109	ylpm1	ENSMUSG00000021244	2	3	0.68	71.18
36	wdr43	ENSMUSG00000041057	2	5	0.36	24.59		110	lrwd1	ENSMUSG000000029703	5	2	0.3	22.99
37	lmbd1	ENSMUSG00000024590	4	4	0.54	55.36		111	adnp	ENSMUSG000000051149	4	2	0.45	43.86
38	nono	ENSMUSG000000031311	4	4	0.6	64.69		112	smcch1	ENSMUSG00000024054	4	2	0.28	13.93
39	eed	ENSMUSG00000030619	3	4	0.27	19.16		113	uhrf1	ENSMUSG00000001228	4	2	0.39	34.91
40	rif1	ENSMUSG00000036202	3	4	0.45	47.17		114	ahdc1	ENSMUSG00000037692	3	2	0.65	65.87
41	ape1	ENSMUSG00000035960	3	4	0.34	27.29		115	h1fx	ENSMUSG00000044927	3	2	0.7	76.06
42	atrx	ENSMUSG000000031229	3	4	0.56	58.7		116	hplbp3	ENSMUSG00000028759	3	2	0.62	61.28
43	chd4	ENSMUSG00000063870	3	4	0.46	43.26		117	prdm10	ENSMUSG00000042496	3	2	0.51	54.61
44	h2afy	ENSMUSG00000015937	3	4	0.37	31.05		118	baz1b	ENSMUSG00000002748	2	2	0.48	44.83
45	hnrnp1	ENSMUSG00000015165	3	4	0.44	37.71		119	hells	ENSMUSG00000025001	2	2	0.33	27.95
46	hnrnp2	ENSMUSG00000066037	3	4	0.47	42.64		120	hist1h1a	ENSMUSG000000049539	2	2	0.71	80.75
47	lmba	ENSMUSG00000028063	3	4	0.58	58.87		121	hist1h1b	ENSMUSG000000058773	2	2	0.76	88.34
48	lmbd2	ENSMUSG00000062075	3	4	0.58	59.82		122	hist1h1e	ENSMUSG000000051627	2	2	0.75	81.74
49	parp1	ENSMUSG00000026496	3	4	0.34	21.47		123	hist1h4a	ENSMUSG00000060093	2	2	0.51	40.78
50	pssp1	ENSMUSG00000028484	3	4	0.69	74.24		124	hist1h4c	ENSMUSG00000060678	2	2	0.51	40.78
51	ptbp1	ENSMUSG00000006498	3	4	0.35	27.80		125	hnrnpf	ENSMUSG00000042079	2	2	0.36	24.58
52	sfq	ENSMUSG00000028820	3	4	0.72	74.54		126	hnrnp1	ENSMUSG00000007850	2	2	0.38	25.95
53	zfr	ENSMUSG00000022201	3	4	0.49	47.95		127	hnrnp2	ENSMUSG00000004527	2	2	0.38	26.5
54	hdac2	ENSMUSG00000019777	2	4	0.31	26.23		128	hnrnp3	ENSMUSG00000020069	2	2	0.43	30.35
55	setdb1	ENSMUSG00000015697	2	4	0.45	43.76		129	matr3	ENSMUSG000000037236	2	2	0.54	56.56
56	adar	ENSMUSG00000027951	2	4	0.41	34.72		130	mecp2	ENSMUSG000000031393	2	2	0.75	85.02
57	ahctf1	ENSMUSG00000026491	2	4	0.44	44.45		131	naca	ENSMUSG000000061315	2	2	0.61	57.67
58	anp32a	ENSMUSG000000032249	2	4	0.56	48.79		132	ncoa5	ENSMUSG000000039804	2	2	0.66	69.17
59	ataf2	ENSMUSG00000022360	2	4	0.37	28.32		133	rpl17	ENSMUSG00000062328	2	2	0.46	42.12
60	aurkb	ENSMUSG00000020897	2	4	0.32	21.74		134	set	ENSMUSG00000054766	2	2	0.57	59.86
61	cpsf6	ENSMUSG00000055531	2	4	0.68	71.32		135	thrap3	ENSMUSG00000043962	2	2	0.72	83.02
62	elavl1	ENSMUSG00000040028	2	4	0.28	15.03		136	tmpo	ENSMUSG00000019961	2	2	0.54	57.94
63	fubp1	ENSMUSG00000028034	2	4	0.62	67.9		137	tpx2	ENSMUSG00000027469	2	2	0.52	55.03
64	hnrnpa0	ENSMUSG00000007836	2	4	0.45	41.64		138	trim28	ENSMUSG00000005566	2	2	0.47	43.65
65	hnrnpb	ENSMUSG00000060373	2	4	0.58	60.86		139	u2af1	ENSMUSG000000061613	2	2	0.4	28.87
66	hnrnpu	ENSMUSG000000039630	2	4	0.49	48.94		140	wiz	ENSMUSG00000024050	2	2	0.52	53.18
67	hnrnpu2	ENSMUSG0000001659	2	4	0.53	48.86		141	zfp326	ENSMUSG00000029290	2	2	0.52	49.14
68	khsp	ENSMUSG000000007670	2	4	0.64	67.31		142	zfp638	ENSMUSG000000030016	2	2	0.59	64.95
69	luc712	ENSMUSG00000029823	2	4	0.62	62.63		143	zfp280c	ENSMUSG00000036916	3	1	0.43	40.36
70	luc713	ENSMUSG00000020863	2	4	0.64	67.13		144	zfp462	ENSMUSG00000066205	3	1	0.42	39.2
71	mcm7	ENSMUSG00000029730	2	4	0.31	17.11		145	bnmf1	ENSMUSG00000024844	2	1	0.27	19.1
72	nudt21	ENSMUSG00000031754	2	4	0.23	11.89		146	ifl204	ENSMUSG000000073489	2	1	0.45	38.98
73	raty	ENSMUSG00000027593	2	4	0.52	48.23		147	orc4	ENSMUSG00000026761	2	1	0.17	8.43
74	rbm12b	ENSMUSG00000008924	2	4	0.33	23.27		148	zfp518b	ENSMUSG000000046572	2	1	0.47	45.4

Table EV1. Conserved heterochromatin components.

List of the 148 heterochromatic proteins identified in more than one mass spectrometry study in mammalian cells [6,39-44] (Gene information). The number of studies in which the protein was found is shown under the Mammals column. The number of orthologs found in *D. renio*, *S.pombe*, *D. melanogaster* and *C. elegans* can be seen in the Orthologs column. This analysis was done using the Ensembl project website with the Ensembl release 94 [102] for the orthologs in *D. renio*, *D. melanogaster* and *C. elegans*, and with the PomBase project website [103] for *S.pombe*. The two Disorder estimates show the disorder score and overall percentage disorder obtained with the PONDR-VLXT and IUPRED predictors.

	Protein length	disorder score	Overall percentage disorder
Heterochromatin (median)	615.5	0.47	44.32
<i>CBX5</i>	<i>191</i>	<i>0.5</i>	<i>45.55</i>
<i>CBX1</i>	<i>185</i>	<i>0.54</i>	<i>54.86</i>
<i>ATRX</i>	<i>2476</i>	<i>0.56</i>	<i>58.70</i>
<i>DNMT1</i>	1620	0.39	33.36
<i>UHRF1</i>	782	0.39	34.91
<i>SUV39H2</i>	477	0.28	20.86
<i>SUV420H2</i>	468	0.32	22.86

Table EV2: Disorder analysis of bona fide heterochromatic proteins

Disorder analysis for the 148 heterochromatin proteins (median) and the 7 'bona fide' heterochromatic proteins. The first column shows the protein length in amino acids. The second and third column show the disorder score and overall percentage disorder calculated using the PONDR-VLXT and IUPRED predictors similarly to Figure 1.

**A PHASE TRANSITION ACCOMPANIES HETEROCHROMATIN FORMATION IN
MOUSE EMBRYOS**

A phase transition accompanies heterochromatin formation in mouse embryos

Manuel Guthmann¹, Chen Qian², Irene Gialdini², Tsunetoshi Nakatani¹, Andreas Ettinger¹, Tamas Schauer¹, Igor Kukhtevich³, Robert Schneider³, Don C. Lamb², Adam Burton¹ & Maria-Elena Torres-Padilla^{1,4*}

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

²*Department of Chemistry and Center for NanoScience (CeNS), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany*

³*Institute of Functional Epigenetics (IFE), Helmholtz Zentrum München, D-85764 Neuherberg, Germany*

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

ABSTRACT

The majority of our genome is composed of repeated DNA sequences, which assemble into heterochromatin, a highly compacted structure that constrains their mutational potential. How heterochromatin forms during development and how its structure is maintained is not fully understood. Here, we show that mouse heterochromatin phase separates after fertilization, during the earliest stages of mammalian embryogenesis. Using high resolution, quantitative imaging and molecular biology approaches, we show that pericentromeric heterochromatin displays liquid-like properties at the 2-cell stage, but it transitions into a more solid-like or gel-like state at the 4-cell stage, when chromocenters mature and heterochromatin becomes silent. Disrupting the condensates results in altered transcript levels of pericentromeric heterochromatin, suggesting a functional role for phase separation in heterochromatin function. Thus, our work shows that mouse heterochromatin forms membrane-less compartments with biophysical properties that change during development and provides new insights into the self-organization of chromatin domains during mammalian embryogenesis.

MAIN TEXT

Mammalian development starts upon fertilization of an oocyte by the sperm. Following fertilization, the two parental genomes undergo a major process of epigenetic reprogramming, which involves the establishment of chromatin domains (Burton and Torres-Padilla, 2010). How chromatin forms at the beginning of development remains a fundamental question in biology. In particular, how heterochromatin acquires its silencing signatures is poorly

understood. Several molecular pathways contribute to the establishment of heterochromatin but the potential contribution of its biophysical properties to this process remains completely unknown.

In mammals, heterochromatin formation at pericentromeric regions of the chromosomes is accompanied by a spatial reorganization into multichromosomal domains, the chromocenters. The major satellite repeats that constitute the pericentromeric heterochromatin in mice undergo a dramatic remodelling in their shape and nuclear positioning during the 2-cell stage, changing from a ring-like configuration around the nucleoli precursors into spherical chromocenters (Casanova et al., 2013; Jachowicz et al., 2013; Probst et al., 2010) (*Fig. 1a*). This organization persists subsequently throughout development and differentiation. Pericentromeric chromatin progressively acquires heterochromatin marks such as H3K9me3 and HP1 β during development (Puschendorf et al., 2008). Initially, major satellites are actively transcribed in zygotes and 2-cell stage embryos (Probst et al., 2010), and this transcriptional activity is thought to promote heterochromatin formation (Casanova et al., 2013; Santenard et al., 2010) and is essential for embryo development (Probst et al., 2010). Functionally, silencing of pericentromeric heterochromatin occurs concomitantly with chromocenter formation and requires changes in nuclear organization (Casanova et al., 2013; Jachowicz et al., 2013; Probst et al., 2010).

In *Drosophila* embryos, heterochromatin displays condensation behaviour via liquid-liquid phase separation (LLPS) (Strom et al., 2017). Whether mammalian heterochromatin displays similar properties has been debated (Erdel et al., 2020; Huo et al., 2020; Larson et al., 2017; Strickfaden et al., 2020). Differences in the cellular systems used may explain some of the apparent discrepancies observed, but also, studying the biophysical properties of chromatin in a cellular environment remains technically challenging. Notwithstanding, the dynamic properties of mammalian embryonic heterochromatin *in vivo* have not been studied. Whether state transitions of heterochromatin occur in the mammalian embryo and whether such transitions are required for heterochromatin formation and transcriptional silencing remains unknown.

Considering the significant changes in the shape of heterochromatin leading to the formation of chromocenters (*Fig. 1a*), we reasoned that large scale changes in their biophysical properties may take place after fertilization. Indeed, these morphological changes result in the formation of spherical domains (*Fig. 1a*) (Probst et al., 2010), a hallmark of LLPS *in vitro* (Alberti et al., 2019; Brangwynne et al., 2011). We thus established imaging conditions with high spatial and temporal resolution to investigate and quantify the biophysical properties of

pericentromeric chromatin after fertilization and verified that these do not impair development (*Fig. 1b* and *Extended Data Fig. 1a-b*).

We focused on five key features to determine the biophysical state of pericentromeric domains: i) fusion and fission events; changes in ii) sphericity and iii) volume over time; iv) the diffusion coefficients of its components inside and outside the domain, and v) the presence of a boundary constraining the diffusion of its components. To investigate the biophysical properties of chromatin in embryos, we chose to image pericentromeric heterochromatin itself, as opposed to associated chromatin proteins (Novo et al., 2022). To image pericentromeric heterochromatin in embryos, we used the TALE-based Genome Visualization (TGV) system that we developed previously, which uses fluorescent TALEs to specifically visualise the major satellites (Miyanari et al., 2013). We visualised nuclei by co-injection of mRNA for histone H2B-tdiRFP and imaged heterochromatin for ~24 hours in more than 90 embryos (*Fig. 1b* and *Extended Data Fig. 1c*). Over this time period, we observed the formation of bright, condensed pericentromeric foci, reflecting chromocenter formation (*Extended Data Fig. 1c*) (Probst et al., 2010; Puschendorf et al., 2008). We then performed quantitative image analysis in 3D by implementing a segmentation pipeline (*Extended Data Fig. 1d* and *Methods*). We observed both fusion and fission of pericentromeric domains in 2-cell stage embryos (*Fig. 1c* top and bottom panels, respectively) and an increase in their sphericity over time (*Fig. 1d*). Of note, the pericentromeric domains do not always recover a spherical shape immediately after fusion (*Fig. 1c*). This is in contrast to *in vitro* fusion events of liquid-like droplets and may reflect the underlying nature of the nucleosome polymer, which can greatly distort condensate shape *in vivo*. Interestingly, during the 2-cell stage, pericentromeric domains initially increased in volume, but then started to decrease in size prior to a sharp increase in sphericity (*Fig. 1e* and *1f*). The changes in volume were highly reproducible across embryos even though the absolute volume of the pericentromeric domain is highly heterogeneous both within and between nuclei (*Extended Data Fig. 2*). The changes we observed are not due to our long-term imaging, as we reproduced these results using DNA FISH for major satellites on freshly collected embryos (*Fig. 1g, h* and *Extended Data Fig. 3a*). Together, this suggests that events of fusion and fission underlie the changes in shape undergone by pericentromeric chromatin at the 2-cell stage. Indeed, volume changes were consistently related to changes in sphericity, with most pericentromeric domains shifting from larger and less spherical to smaller domains of higher sphericity towards the end of the 2-cell stage (*Fig. 1i*), indicating partitioning of heterochromatin domains into smaller, more rounded clusters. However, we did not observe the dynamic changes in volume and sphericity of pericentromeric chromatin at the 4-cell stage when chromocenters are already formed (*Extended Data Fig. 3b-d*), indicating that such

changes specifically accompany the period of heterochromatin maturation into chromocenters.

To address whether the progressive changes in sphericity that we observed at the 2-cell stage are a general feature of other nuclear bodies in early embryos, we imaged the embryonic nucleoli (referred to as nucleolar-like bodies or NLBs)(Burns et al., 2003; Inoue and Aoki, 2010) (*Fig. 1j,k and Extended Data Fig.4a*). Nucleoli are known to undergo LLPS in other developmental settings and correspondingly, display highly spherical shapes(Brangwynne et al., 2011; Lafontaine et al., 2021). As expected, NLBs exhibited fusion and fission events, preferential internal rearrangement and high sphericity, but their sphericity remained relatively constant throughout the 2-cell stage (*Fig. 1k and Extended Data Fig. 4b-e*). Importantly, the values of sphericity for the major satellite domains reached similar levels as those of the NLBs at the late 2-cell stage (*Extended Data Fig. 4e*). Thus, pericentromeric chromatin undergoes fission and fusion and displays specific changes in shape in early mouse embryos, which are consistent with condensate properties.

If embryonic heterochromatin is in a phase-separated state, we would expect distinct diffusion of its internal components within the condensate compared to the surrounding nucleoplasm(Frank and Rippe, 2020; Strom et al., 2017). To address this, we adapted RICS (Raster Image Correlation Spectroscopy)(Hendrix et al., 2016) to mouse embryos to measure diffusion coefficients inside and outside pericentromeric heterochromatin. We performed RICS for H3.1, since H3.1 is preferentially incorporated at pericentromeric chromatin in 2-cell stage embryos (*Fig. 2a*)(Santenard et al., 2010). To distinguish pericentromeric chromatin in live embryos, we co-expressed TALE-MajSat-mClover to delimit the heterochromatin domains (*Fig. 2b*). We found that the H3.1 diffuses at a significantly different rate within the pericentromeric domain compared to the nucleoplasm (*Fig. 2b*), suggesting that embryonic heterochromatin is phase-separated. In agreement, the concentration of H3.1 was significantly higher inside the heterochromatin domain (*Extended Data Fig. 5a*). Next, we reasoned that if heterochromatin is in a liquid-like state, as opposed to a gel or solid-like state, a boundary that partitions heterochromatin and the nucleoplasm should be present, resulting in different recovery kinetics of its internal components compared to the surrounding nucleoplasm. To test this, we performed FRAP (Fluorescence Recovery After Photobleaching) inside and outside heterochromatin at the 2-cell stage. We measured H3.1 dynamics at both, NLB-associated and chromocenter domains as a mixture of both are present at this timepoint (*Fig. 2c*). H3.1 displayed high mobility in the nucleoplasm of 2-cell stage embryos (*Fig. 2d, e and Supplementary Table 1*), consistent with previous work (Boskovic et al., 2014; Ooga et al., 2016). In addition, H3.1 was also highly mobile inside the pericentromeric domains (*Fig. 2d, e*

and *Supplementary Table 1*), in line with our RICS data. While the H3.1 mobile fraction did not differ significantly between the nucleoplasm and the pericentromeric domains (*Fig. 2e* and *Supplementary Table 1*), the fast recovery kinetics at heterochromatin were significantly slower in the pericentromeric domain (*Fig. 2f*, *Extended Data Fig. 5b* and *Supplementary Table 1*). Since the fast recovery kinetics measures a fraction of H3.1 that can exchange between two compartments¹⁸ (Kimura and Cook, 2001), we conclude that H3.1 exchange has slower kinetics inside the pericentromeric condensates at the 2-cell stage, suggesting the presence of a boundary between the pericentromeric domain and the nucleoplasm. Importantly, we confirmed these results by performing FRAP with another histone, H2B (*Fig. 2g-i*, *Extended Data Fig. 5c* and *Supplementary Table 1*), thus indicating that histone recovery kinetics differ between the heterochromatin and the surrounding nucleoplasm. Remarkably, at the 4-cell stage, when chromocenters are fully mature (Probst et al., 2010; Puschendorf et al., 2008), both the mobility and recovery kinetics of H3.1 are similar between heterochromatin domains and the surrounding nucleoplasm (*Extended Data Fig. 5d-f* and *Supplementary Table 1*). RICS at the 4-cell stage indicated that the diffusion of H3.1 inside the heterochromatin, while still differing statistically compared to the nucleoplasm (*Fig. 2j*), becomes more similar between the two compartments (*Fig. 2k*; *dashed line*). Thus, the differences in internal mobility between the heterochromatin domain and the nucleoplasm, in this case measured via the diffusion of H3.1, become progressively reduced from the 2- to the 4-cell stage. In contrast, the relative enrichment of H3.1 inside heterochromatin is even higher at the 4-cell stage, compared to the 2-cell stage (*Extended Data Fig. 5g*). Thus, all the above support the interpretation that embryonic heterochromatin shows condensate properties that change during development. Initially heterochromatin displays liquid-like features during the process of chromocenter formation, including increasing sphericity, fusion/fission, and distinct recovery kinetics and diffusion of its internal components. However, the similar recovery kinetics inside and outside the heterochromatin domains, when chromocenters are fully mature suggest a transition into a more solid-like or gel-like state (Frank and Rippe, 2020). Overall, we conclude that pericentromeric chromatin forms membrane-less compartments with evolving biophysical properties during this time of development.

Liquid-liquid phase separation of condensates can be mediated by several type of interactions, among which hydrophobic interactions have been the most extensively studied (Alberti et al., 2019; Simon et al., 2017). Thus, we next probed whether weak hydrophobic interactions are necessary for chromocenter integrity and their biophysical properties by exposing embryos to 1,6-Hexanediol (Elbaum-Garfinkle, 2019). Because NLBs displayed LLPS features in mouse embryos, including fusion, high sphericity, and preferential internal rearrangement (*Fig. 1i* and *Extended Data Fig. 4b-e*), we first identified conditions in which 1,6-Hexanediol alters the

number of NLBs (*Extended Data Fig. 6a*). To address whether 1,6-Hexanediol disrupts the pericentromeric compartment, we performed DNA FISH for major satellites, which revealed a severe alteration in the nuclear localization of pericentromeric heterochromatin after 1,6-Hexanediol treatment, which led to a clear dispersion of the heterochromatin domains to the periphery of the nuclei in all of the embryos analysed (*Fig. 3a and Extended Data Fig. 6b*). Live imaging of pericentromeric domains recapitulated this observation (*Extended Data Fig. 6c,d*). This behaviour was similar to that observed for the NLBs (*Fig. 3b*). Prolonged treatment with 1,6-Hexanediol led to a further dispersion of the TGV fluorescent signal (*Extended Data Fig. 6c*), in line with our DNA-FISH data indicating that pericentromeric heterochromatin became dispersed in response to 1,6-Hexanediol treatment (*Fig. 3a*). These data suggest that liquid-liquid demixing of pericentromeric heterochromatin at the 2-cell stage is enabled through weak hydrophobic interactions. We note that 1,6-Hexanediol may have side effects, including effects on nuclear area, which we did observe upon incubation of zygotes at several 1,6-Hexanediol concentrations (*Extended Data Fig. 6a*). However, such non-specific side effects would be expected in both 2- and 4-cell stage embryos. Importantly, 1,6-Hexanediol treatment at the 4-cell stage did not affect major satellite localization (*Fig. 3c and Extended Data Fig. 6b*), suggesting that mature, fully formed chromocenters do not depend upon weak hydrophobic interactions. These observations, together with our findings showing similar recovery of internal components between the pericentromeric domains and the surrounding nucleoplasm and the the lack of detectable fusion and fission events at the 4-cell stage point towards a transition of heterochromatin from a liquid-like state to a more gel- or solid-like state during this early developmental period.

To test this hypothesis further, we sought to manipulate the material properties of pericentromeric heterochromatin by tethering the low complexity domain (LC) of human FUS, recently shown to induce liquid-like properties *in vivo* (*Bose et al., 2022*), to pericentromeric heterochromatin. We reasoned that if pericentric heterochromatin is in a liquid-like state, the targeted recruitment of the FUS domain would not lead to any alterations in heterochromatin. In contrast, if heterochromatin is in a solid or gel-like state, the recruitment of FUS LC, which has been shown to induce a liquid state, would affect heterochromatin clustering. For this, we fused the FUS LC to the TALE-MajSat construct and microinjected mRNA in mid to late 2-cell stage embryos to enable FUS LC expression at the 4-cell stage. DNA-FISH analyses revealed that heterochromatin became dispersed in embryos in which FUS LC was targeted to MajSat, compared to the LLPS-deficient form of FUS LC, which we used as negative control (*Fig. 3d*). Thus, we conclude that pericentromeric heterochromatin at the 4-cell stage is sensitive to manipulation of material properties towards a liquid state, thus suggesting that the 4-cell stage heterochromatin is not in a liquid-like state.

To examine the functional consequences of perturbing heterochromatin condensate integrity upon 1,6-Hexanediol treatment, we measured the transcript levels of major satellite repeats as a proxy for heterochromatin silencing. Pericentromeric heterochromatin is actively transcribed in zygotes and 2-cell stage embryos (Probst et al., 2010) and this transcriptional activity is thought to promote heterochromatin formation and chromocenter formation (Casanova et al., 2013; Santenard et al., 2010). Treatment with 1,6-Hexanediol led to a significant reduction of major satellite transcripts as detected by RT-qPCR (*Fig. 3e*), but not of control, β -actin and rDNA transcripts in these conditions (*Extended Data Fig. 7a*), suggesting defective heterochromatin integrity. We confirmed these results by using RNA-FISH (*Fig. 3f*). Importantly, 1,6-Hexanediol treatment did not elicit major transcriptional changes as determined by RNA-seq of control and 1,6-Hexanediol treated embryos (*Extended Data Fig. 7b-d and Supplementary Table 2*). We also addressed whether RNA from major satellite repeats themselves contribute to H3.1 dynamics within the pericentromeric domain. However, FRAP analysis indicated that depletion of major satellite RNA does not significantly perturb heterochromatin dynamics at this stage of development (*Extended Data Fig. 8 and Supplementary Table 1*). This suggests that, while transcriptional output of pericentromeric heterochromatin may be influenced upon disruption of weak hydrophobic interactions, the transcripts themselves do not participate in the dynamics of the internal components of the condensate.

Next, to investigate the molecular basis of our observations, we screened for proteins expressed in embryos that have a potential to phase separate and a functional role in heterochromatin (Guthmann et al., 2019; McDowell et al., 1999). Of note HP1 α is not expressed prior to the blastocyst stage (Wongtawan et al., 2011). Our previous work identified ATRX as a protein with a potential to phase separate (Guthmann et al., 2019). ATRX is required for heterochromatin integrity in oocytes and lack of ATRX in female gametes leads to chromosomal instability in pre-implantation embryos (Baumann et al., 2010; De La Fuente et al., 2015; De La Fuente et al., 2004; Liu et al., 2020). ATRX contains a large intrinsically disordered region (IDR) and a C-terminal prion-like domain (PLD) (Guthmann et al., 2019) which, are commonly found in proteins that promote or are enriched in phase separated compartments (Franzmann and Alberti, 2019; Hennig et al., 2015; Lin et al., 2017; Nott et al., 2015; Pak et al., 2016; Simon et al., 2017; Wang et al., 2018b). Immunostaining for ATRX revealed enrichment in foci at the pericentromeric chromatin around the NLBs in zygotes (*Fig. 4a*), in agreement with previous work (De La Fuente et al., 2015; Liu et al., 2020). At the 2-cell stage ATRX displayed a marked pericentromeric heterochromatin enrichment, exclusively at chromocenters but not around NLBs (*Fig. 4a*).

To address whether ATRX is involved in the biophysical and morphological changes of the pericentromeric heterochromatin during this time of development, we performed acute protein depletion of ATRX using Trim-away (Clift et al., 2018), which led to efficient ATRX degradation in 2-cell stage embryos (*Extended Data Fig. 9a*). Depletion of ATRX did not affect the early changes in the volume of heterochromatin domains (*Fig. 4b and Extended Data Fig. 9b*). However, depletion of ATRX completely prevented the partitioning into smaller rounded heterochromatin domains observed in control embryos at the late 2-cell stage (*Fig. 4b and Extended Data Fig. 9c*). These data suggest that the initial heterochromatin phase transition is independent of ATRX function, but that ATRX is required for later maturation into chromocenters. In agreement, ATRX depletion did not affect the fast recovery kinetics of H3.1 inside the heterochromatin domains or in the nucleoplasm (*Fig. 4c*), suggesting that ATRX is not required for the integrity of the pericentromeric condensate. The role of ATRX in heterochromatin dynamics is unlikely to be related to its function as a H3.3 chaperone, since *de novo* incorporated H3.3 at this stage is not enriched within the pericentromeric domains (*Extended Data Fig. 9d*). Phase separated condensates rely on scaffold proteins for the integrity of the condensate, and on client proteins, which interact with the components of the condensate through low valency interactions (Banani et al., 2016). Because our data indicate that ATRX is not required for the condensate integrity, we thus tested whether ATRX acts as a client protein in the process of heterochromatin formation in the early embryo. Indeed, we found the specific localization of ATRX at chromocenters at the 2-cell stage was highly sensitive to 1,6-Hexanediol treatment (*Fig. 4d*). Overall, our data suggest that ATRX recruitment to heterochromatin as well as the initial integrity of chromocentres (*Fig. 3a, 4d*) depends on weak hydrophobic interactions, and that ATRX is subsequently required for the morphological and biophysical changes that lead to chromocenter formation.

The lack of partitioning into smaller heterochromatin domains in the absence of ATRX that we observed is in line with the known defects in aneuploidy resulting from maternal depletion of ATRX (Baumann et al., 2010; De La Fuente et al., 2015). While all our observations are compatible with embryonic heterochromatin displaying condensate features, it is also likely that ATRX may contribute additional roles through its ability to bind and remodel chromatin (Eustermann et al., 2011; Iwase et al., 2011; Xue et al., 2003) by biochemical, rather than biophysical means. For example, maternal depletion of ATRX in the oocyte results in increased major satellite transcript accumulation, indicating a role for heterochromatin formation and/or maintenance in the germline (De La Fuente et al., 2015). However, we note that the contribution of both *in vivo* is plausible and in line with the complexity of cellular developmental systems.

The fusion events that we observed in mouse embryos occur significantly slower than most other liquid-like condensates studied so far (Brangwynne et al., 2009; Brangwynne et al., 2011). However, most other studies have studied fusion kinetics only *in vitro*, and thus the influence of the cellular environment on the fusion kinetics of condensates remain to be studied in depth. Indeed, our observations relate to a naturally occurring process *in vivo*, in the absence of the induction of shear stress. In addition, the inherent properties of the chromatin template may also influence the properties of heterochromatic condensates inside the cell. Our results indicate that interfering with compartmentalisation of heterochromatin compromises its transcription, which is in turn essential for full developmental competence (Probst et al., 2010).

It is noteworthy that the properties of condensates *in vivo*, including protein dynamics, may differ significantly less than its surrounding environment as has been documented for phase separation systems studied *in vitro*¹³. This is expected considering the constraints imposed by their cellular environment and biochemical function, in this case the chromatin itself. Recent work has pointed out the importance of undertaking several approaches to investigate material properties and phase separation *in vivo* (Alberti et al., 2019; Frank and Rippe, 2020). This is because the interpretation of individual parameters is not conclusive on its own. For example, the interpretation of FRAP data in terms of absolute values of recovery kinetics is dependent on multiple parameters but a side-by-side comparison of recovery kinetics between the condensates and the surrounding space can be used as a relative measure to differentiate the properties between the two regions (Frank and Rippe, 2020). Likewise, 1,6-Hexanediol, while widely used in the field, remains a largely non-specific chemical. In keeping with the complexity of studying chromatin properties *in vivo*, our interpretations are based on the combination of orthogonal approaches and not on a single criterion.

Our work suggests a model whereby dynamic changes in biophysical properties underlie heterochromatin formation at the beginning of mammalian development (Fig. 5). While heterochromatin displays properties of a liquid-like state initially during development, heterochromatin maturation into chromocenters is accompanied by a phase transition into a more solid- or gel-like state. These data imply that the phase separation features of heterochromatin are cell-type variable and could potentially be linked to cellular plasticity, in this case in a developmental setting. In the future, it will be important to investigate the properties of proteins known to drive phase separation in other systems, and to determine their potential contribution to the changes in biophysical properties of embryonic heterochromatin that we document here.

Overall, our data provide novel insights into self-organization of chromatin domains at the beginning of mammalian development and indicates that the physical state of heterochromatin may change depending on the biological cellular context.

MATERIAL AND METHODS

Plasmid construction and mRNA production

The TALEMajSat-mClover (Addgene 47878), H2B-tdiRFP (Addgene 47884) and mGFP (Addgene 139402) plasmids were previously characterised in our lab (Borsos et al., 2019; Miyanari et al., 2013) and are available in Addgene. For the mRuby-NPM2 and mEOS3.2-NPM2 plasmids, we replaced eGFP for either mRuby or mEOS3.2, which was amplified from a plasmid obtained from the Burtscher lab (Zhang et al., 2012) in a eGFP-NPM2 construct kindly provided by F. Aoki (Inoue and Aoki, 2010). The mCherry-Trim21 plasmid was obtained from Addgene (105522, (Clift et al., 2018)) and the pRN3P.Trim21 plasmid was generated by removing the mCherry sequence. The H3.1 plasmid was constructed from the pRN3P.H3.1.GFP plasmid (Santenard et al., 2010) by replacing the GFP sequence by mRFP. The SNAP-H3.1 and SNAP-H3.3 plasmid contains an N-terminal SNAP tag sequence. H2B-mRFP was previously cloned in our laboratory (Miyanari et al., 2013), which we inserted into the pRN3P vector for mRNA synthesis. For the TALEMajSat-FUS-LC and TALEMajSat-FUS-MUT, we replaced the mClover by the wild-type or a phosphomimetic mutant version (FUS E12) of the low complexity (LC) domain of human FUS (gift from Anne Ephrussi, EMBL) (Bose et al., 2022). mRNAs were transcribed *in vitro* using either a T3 or a T7 mMACHINE mMACHINE kit (Ambion) or mMACHINE ultra kit (Ambion).

Embryo collection, microinjection and culture

Animal experiments were carried out in compliance with local regulations (Government of Upper Bavaria). For the ATRX immunostainings, CD1 female mice (4–8 weeks old) were mated with CD1 male mice (3–6 months old), and zygotes and 2-cell stage embryos were collected at 17h and 40h post-coitum, respectively. For all other experiments, embryos were collected from 5–6-week-old F1 (C57BL/6J×CBA/H) females mated with F1 males (3–6 months old) after hormonal induction with pregnant mare serum gonadotropin (Intervet, 5 IU) and human chorionic gonadotropin (hCG; Intervet, 7.5 IU) 46–48-h later. Embryos were collected at the following time points after hCG injection: zygotes were collected between 18h and 21h post-hCG except for Extended Data Fig. 6a, in which zygotes were collected at 25h

post-hCG. Throughout all the experiments, zygotes were microinjected with 1–2pl of the indicated mRNA or antibody. Embryos were cultured in K-modified simplex optimized medium (KSOM) microdrops under paraffin oil at 37 °C, 5% CO₂. For the analysis of H3.1 localization in Fig. 3a, *in vitro*-transcribed mRNA of SNAP tagged H3.1 (50ng/μL) was injected into zygotes at 17-18h post-hCG. Injected embryos were cultured until 47.5h post-hCG and treated with 5μM SNAP-Cell Oregon Green (NEB, S9104S) for 30 min in KSOM. Embryos were subject to triton pre-extraction to detect incorporated SNAP tagged H3.1(ref. (Hajkova et al., 2010)). For the analysis of H3.3 localisation in Extended Data Fig.9d *in vitro*-transcribed mRNA of SNAP tagged H3.3 (10ng/μL) was injected into zygotes at 28h post-hCG. Depletion of major satellite RNA was performed as described before(Casanova et al., 2013). The sequences for the MajSat and the control ASO are 5'mC*mU*mG*mU*mU*T*T*C*T*T*G*C*C*A*T*mA*mU*mU*mC*mC-3 and 5'-mU*mC*mA*mC*mC*T* T*C*A* C*C*C* T*C*T*mC*mC*mA*mC*mU-3', respectively(Ideue et al., 2014). Briefly, 10μM of MajSat or control ASO were injected into zygotes at 17-18hphCG. For the FUS experiments, one of the two blastomeres of a 2-cell embryo was injected with 200ng/μL of mGFP and 300ng/μL of either TALEMajSat-FUS-LC or TALEMajSat-FUS-MUT to enable FUS LC expression at the 4-cell stage. mGFP was used as a marker for injection and only blastomeres with mGFP expression were analysed.

Immunostaining and confocal microscopy

Embryos were fixed as previously described (Torres-Padilla et al., 2006). Briefly, the zona pellucida was removed with Acid Tyrode solution, followed by two washes in PBS and fixation for 20 min in 4% paraformaldehyde, 0.04% Triton, 0.3% Tween-20, 0.2% sucrose at 37 °C. Embryos were then washed with PBS and permeabilized with 0.5% Triton X-100 for 20min. After permeabilization, embryos were washed three times in PBST (0.1% Tween 20 in PBS), quenched in 2.6mg/ml freshly prepared ammonium chloride, washed three times in PBST, blocked for 3–4h at 4 °C in blocking solution (3% BSA in PBST) and incubated with primary antibodies in blocking solution. Antibodies used were as follows: anti-ATRX (Santa Cruz sc55584, 1:250 dilution), anti-H3K9me3 (Active Motif, 39286, 1:250 dilution) and anti-SNAP (BioLabs, p9310S, 1:1000 dilution). After overnight incubation at 4 °C, embryos were washed three times in PBST, blocked and incubated for 3h at room temperature in blocking solution containing secondary antibodies labelled with Alexa-488 or Alexa-555 fluorophores (Invitrogen A32731 or A32727, 1:500 dilution). After washing, embryos were mounted in Vectashield (Vector Laboratories) containing DAPI. For the immunostainings after Triton pre-extraction, embryos were first permeabilized with pre-extraction buffer (50mM NaCl, 3mM MgCl₂, 300mM sucrose, 25mM HEPES, 0.5% Triton X-100, adjust pH to 7.4) for 10 min on

ice and washed 3 times in washing buffer (50mM NaCl, 3mM MgCl₂, 300mM sucrose, 25mM HEPES, adjust pH to 7.4) before fixing in 4% PFA at RT. Embryos were then processed in the same way as a normal immunostaining after the permeabilization step. Confocal microscopy was performed using a plan-Apo 63x NA 1.4 oil immersion objective on a TCS SP8 inverted confocal microscope (Leica). Z-sections were taken every 0.5–1µm. Image analysis was performed using FIJI (Schindelin et al., 2012). For all experiments, acquisition parameters were set to obtain fluorescence intensity signals in the linear range of the hybrid detectors. These detectors have negligible detector noise and linearly amplify incoming photons into photoelectrons, thereby allowing the counting of measured photons, provided the detector is not saturated. Hence, given identical acquisition settings, the recovered fluorescence signal accurately reflects the level of antigen present in the sample.

Live imaging of Major Satellite (MajSat) regions and Nucleolar-Like Bodies (NLBs)

For live imaging, embryos were microinjected with 1-2pl of the following combinations of mRNAs: 1) 600 ng/µl of TALEMajSat-mClover mRNA and 200 ng/µl of H2B-tDiRFP mRNA for MajSat visualization; 2) 600 ng/µl of TALEMajSat-mClover mRNA, 200 ng/µl of H2B-tDiRFP mRNA and 100 ng/µl of mRuby-NPM2 mRNA for imaging NLBs. Embryos were cultured in microdrops of KSOM in the incubator until the start of the live imaging, when they were transferred to an imaging chamber containing micrometric wells with 1mL of in KSOM covered with 1mL of paraffin oil prior the beginning of the start of the live imaging experiment. We used either a custom-made chamber (for most imaging experiments) or the Embryo Immobilisation Chip (Dolomite Microfluidics). Each of our custom-made chamber consists of two functional elements: a layer with wells for single embryo isolation and imaging; a wall around the wells, which allows media keeping and exchange. The layer with wells was made as a sandwich of a coverslip and a layer of Ostemer 322 (Mercene Labs AB, Sweden) with holes, which was bonded on top of the coverslip. The holes in Ostemer 322 were formed using a polydimethylsiloxane (PDMS) master mold, which was produced using standard soft lithography. Briefly, a master mold for replication in PDMS was fabricated from SU-8 photoresist (MicroChem, USA) spin-coated on a Si wafer. The SU-8 master mold was then used to produce the PDMS master mold from Sylgard 184 (Dow Corning, USA). In total, there are 361 wells with 200 µm diameter and 110 µm height in each chamber. The wall was 3D printed from High Temp Resin (Formlabs, USA) using 3D printer Form 2 (Formlabs, USA). The outer diameter of the wall is 35 mm, the inner area is 16x16 mm, and the height is 10 mm. The layer with wells and the wall were centered to have the wells within the inner area of the wall and then bonded to each other using Ostemer 322. Each chamber was extensively washed in distilled water and under sonication to remove residual non-cross-linked materials. Live imaging experiments were performed using a plan-Apo 100x NA 1.35 silicon oil

immersion objective on a dragonfly 304 spinning disc system attached to a Nikon Ti-2 microscope. Images were acquired on a Andor iXon 888 live EMCCD camera. Temperature was maintained at 37 °C with a microscope enclosure incubator (OloLab, Italy) and 5% CO₂ was supplied by a gas mixer into a stage-top chamber with magnetic holder for 35 mm dishes (Okolab, Italy). For imaging of MajSat (mClover) and NPM2 (mRuby) in early and late 2-cell stage embryos, live imaging was started between 40 and 42 h or 48h post-hCG respectively. For the 4-cell stage imaging live imaging was started between 54 and 55h post-hCG. Note that live imaging procedure may result in some developmental delay and therefore the post-hCG timepoints should be considered relative. To generate the 2-cell and 4-cell stage datasets, Z-sections were taken every 3µm and for live imaging of NPM2, Z-sections were taken every 8.33µm. For all the live imaging experiments images were taken every 20 minutes.

Image Analysis and Processing

The MajSat channel (mClover) was first denoised using the Rudin-Osher-Fatemi algorithm (Getreuer, 2012; Leonid I. Rudin, 1992) with a denoising weight of 10. The NPM2 or MajSat channel was then segmented using Imaris software (Bitplane). For segmentation a smoothing parameter of 0.4µm was applied and a size threshold of 0.674µm³ for the MajSat channel and of 10µm³ for the NPM2 channel. The values of sphericity and volume of the different segmented objects were selected from Imaris. For volume analyses, we mostly present data as Z-score in order to capture all the heterogeneity of the data because the absolute volume values across embryos vary. A representative sample of randomly selected individual nuclei illustrating absolute volume values are shown in Extended Data Fig. 2. All statistical analysis was done using Rstudio.

Fluorescence Recovery After Photobleaching (FRAP) and Photoconversion

For H3.1 or H2B FRAP at MajSat regions, zygotes were microinjected with 600ng/µl of TALEMajSat-mClover and 300ng/µl of H3.1-mRFP or 300ng/µl of H2B-mRFP mRNAs respectively. FRAP for H3.1 and for H2B was performed on the RFP channel. For FRAP and photoconversion of NPM2 at the NLBs, zygotes were microinjected with 100ng/µl of mRuby-NPM2 or 100ng/µl of mEOS3.2-NPM2 mRNAs, respectively. FRAP and photoconversion was performed on a Nikon Ti-e microscope equipped with a Bruker Opterra 2 confocal system using a plan-Apo 100x NA 1.4 oil immersion objective. Images were acquired on a Photometrics EVO EMCCD camera. Temperature was maintained with a microscope enclosure incubator adjusted to 37°C at the sample (In Vivo Scientific). Embryos were placed in drops of KSOM medium on a glass-bottom dish covered with paraffin oil. A circular region of interest of 2.28µm² was selected either in the nucleoplasm or in the region comprising the MajSat domains as determined by the TALEMajSat-mClover signal, or inside the NLBs as

determined using the NPM2 signal, and was subjected to FRAP or photoconversion. Acquisition was started with images taken every 0.1s for 4s during which bleach was performed. Subsequently, images were acquired every 1s for 10s and then every 5s for 50s.

FRAP curve fitting and statistical analysis

The FRAP raw data were processed with FIJI (ImageJ). All analyses were done on background-subtracted values, which were normalized by the average H3.1 or total NPM2 fluorescence intensity in the nucleus or NLB respectively. The curves obtained were normalized using the full-scale normalization method so that the first post-bleach frame was set to 0. Normalized curves were then subjected to curve fitting. Experimentally obtained and normalized recovery curves were fit using Rstudio. A two-phase exponential association equation, $Y = Y_{max1} \times [1 - e(-K1 \times X)] + Y_{max2} \times [1 - e(-K2 \times X)]$, was used to obtain mobile fractions and reaction rates of H3.1, as this has been previously described to be appropriate for nuclear proteins (Boskovic et al., 2014; Phair and Misteli, 2000). A single exponential association equation $Y = Y_{max} \times [1 - e(-K \times X)]$ was used to obtain the mobile fraction and reaction rate of NPM2. Unpaired *wilcoxon*-tests were used for statistical comparisons. Y_{max2} values were used for mobile fraction estimation of H3.1, as they reflect the steady-state protein pool, unless otherwise stated.

DNA and RNA FISH

DNA FISH was performed as previously described (Miyinari and Torres-Padilla, 2012), using a protocol that preserves the nuclear 3D structure of embryos. Briefly, the zona pellucida was removed with Acid Tyrode solution, followed by two washes in PBS and fixation in 4% paraformaldehyde, 0.05% Triton-X at RT for 15 min. Embryos were then permeabilised in 0.5% Triton-X 100, 0.02% RNaseA for 1h at RT and treated with HCl solution (0.1N HCl, 0.7 Triton-X 100 and 1 mg/ml PVP in water) for 3min at RT. Embryos were washed in prehybridization buffer (50% formamide, 2x SSC, 10% Dextran, 1 mg/ml PVP, 0.05% TritonX, 0.5 mg/ml BSA) and incubated at 55 °C for 3 h. Embryos were then incubated 10 min at 80 °C and transferred into drops of 0.2 µl hybridization buffer containing the MajSat probe (Jachowicz et al., 2013) at 5 ng/µl, which was previously denatured at 80 °C for 10 min. After overnight hybridization at 42 °C, embryos were washed once in 2x SSC at room temperature followed by washing three times 10 min in 0.1x SSC at 42 °C and transferred in drops of Vectashield containing DAPI. Super resolution STED microscopy was performed using a plan-Apo 100x NA 1.4 oil immersion objective on a TCS SP8 inverted confocal microscope (Leica). STED was performed with a 775 nm pulsed laser. Z-sections were taken every 0.1µm. For RNA FISH, the zona pellucida was removed with Acid Tyrode solution, followed by two washes in PBS and fixation in 4% paraformaldehyde, 0.05% Triton-X at RT for 15 min. Embryos were

then permeabilised in 0.5% Triton-X 100 for 1h at RT. Embryos were washed into prehybridization buffer (50% formamide, 2x SSC, 10% Dextran, 1 mg/ml PVP, 0.05% TritonX, 0.5 mg/ml BSA) and incubated at 55 °C for 3 h before being transferred into drops of 0.2 µl hybridization buffer containing the MajSat probe at 5 ng/µl, which was previously denatured at 80 °C for 10 min under oil. After overnight hybridization at 42 °C, embryos were washed once in 2x SSC at room temperature followed by washing three times 10 min in 0.1x SSC at 42 °C and mounted in Vectashield (Vector Laboratories) containing DAPI. Confocal microscopy was performed using a plan-Apo 63x NA 1.4 oil immersion objective on a TCS SP8 inverted confocal microscope (Leica). Z-sections were taken every 0.5–1µm. DNA and RNA FISH signals were analysed using Imaris software (Bitplane). The nucleus was first segmented based on the DAPI channel with a smoothing parameter of 0.6. The MajSat signal was then segmented with a smoothing parameter of 0.2 and a size threshold of 0.674µm³. The values of sphericity, volume and intensity of the different segmented objects were obtained from Imaris. Statistical analyses were performed using Rstudio.

Raster Image Correlation Spectroscopy (RICS)

Zygotes were microinjected, at 17-18 or at 28h post-HCG for the 2-cell and 4-cell experiments respectively, with 600ngµl of TALEMajSat-mClover and 20ngµl of SNAP-H3.1 mRNA. Embryos were cultured in microdrops of KSOM until the start of the SNAP tagging, at 44 and 54h post-hCG for the 2-cell and 4-cell experiments, respectively, at which time embryos were first incubated with a tagging medium (100nM of SNAP-tagcell647 in KSOM) for 30 minutes followed by 2 washes and a 30 min incubation in KSOM without dye. Embryos were then washed 3 times in M2 medium (Sigma) and then transferred to an imaging chamber containing micrometric wells with 1mL of M2 covered with 1mL of paraffin oil. RICS measurements were performed at 37 °C on a home-built confocal laser scanning microscope (CLSM), as described elsewhere (Hendrix et al., 2015). Pulsed diode lasers of wavelengths 470 nm and 635 nm (LDH-P-C-470 and LDH-P-C-635M, PicoQuant) were used for the pulsed-interleaved excitation (Muller et al., 2005) of TALEMajSat-mClover and H3.1-SNAP-647-SiR at a laser power of 10 µW before the objective. A 100x oil immersion objective (Apo-TIRF 100x Oil/NA 1.49, Nikon) was used for all measurements. The descanned fluorescence emission was separated from the excitation pathway with a quad-line 405/488/561/635 beamsplitter (Semrock). The fluorescence emission of TALEMajSat-mClover and H3.1-SNAP-647-SiR was then separated with a 565DCXR beamsplitter (AHF) and spectrally filtered using a 420/40 nm bandpass (Chroma) and a 635 nm longpass (AHF) emission filters respectively before being detected with avalanche photodiode detectors (Count® Blue and Count® Laser Components). The alignment of the system was routinely checked by measuring an aqueous mixture of ATTO488-COOH and ATTO655-COOH with fluorescence correlation

spectroscopy. Per embryo, 300 frames of 12x12 μm , divided into 300 pixels per lines, were acquired with a frame time of $\tau_f = 1$ s, line time, $\tau_l = 3.33$ ms, pixel dwell time $\tau_p = 11.11$ μs , and pixel size $\delta r = 40$ nm. Image calculation from the raw photon data stream and subsequent analyses were performed with PIE Analysis in MATLAB (PAM)(Schrimpf et al., 2018). The RICS measurements were corrected for slow fluctuations and cellular movement by applying a moving average correction of 3 frames prior to image correlation (Hendrix et al., 2015). To analyse H3.1-SNAP-647-SiR fluctuations in different cellular regions, we used arbitrary region RICS (ARICS)(Hendrix et al., 2016) . Arbitrary ROIs were selected based on the TALEMajSat-mClover signal, to distinguish between inside and outside the chromocenters. The spatial autocorrelation functions (SACFs) were fitted with a one-component free diffusion model assuming a 3D Gaussian focal shape:

$$G(\xi, \psi) = \frac{\gamma}{N} \left(1 + \frac{4D|\xi\tau_p + \psi\tau_l|}{\omega_r^2} \right)^{-1} \left(1 + \frac{4D|\xi\tau_p + \psi\tau_l|}{\omega_z^2} \right)^{-\frac{1}{2}} \exp\left(-\frac{\delta r^2(\xi^2 + \psi^2)}{\omega_r^2 + 4D|\xi\tau_p + \psi\tau_l|} \right)$$

in which ξ and ψ indicate the spatial lags in pixels along the fast and slow scanning axes, respectively, and ω_r and ω_z represent the lateral and axial focus sizes, respectively. The shape factor γ is $2^{-3/2}$ for a 3D Gaussian. D represents the diffusion coefficient and N the average number of fluorescent particles in the observation volume. The correlation at zero lag time was omitted from analysis due to the contribution of uncorrelated shot noise. The concentration (C) of H3.1 was calculated according to the equation:

$$C = \frac{1}{G_0 N_A V_{eff}}$$

with G_0 being the correlation amplitude and equal to γ/N , N_A the Avogadro number and V_{eff} the effective detection volume (Rigler et al., 1993), equal to:

$$V_{eff} = \gamma \pi^{\frac{3}{2}} \omega_r^2 \omega_z$$

1,6-Hexanediol Treatment

To determine the effective concentration of 1,6-Hexanediol (Alfa Aesar, A12439) zygotes were incubated for 1h with 0.01% to 3.5% of 1,6-Hexanediol in KSOM prior to fixing and staining with DAPI. The number of NLBs and the pronuclear area was calculated in 3D using Imaris (Bitplane). For the RT-qPCR, the RNA-seq, the DNA and RNA FISH experiments, embryos were incubated 1 hour in 3.5% 1,6-Hexanediol in KSOM prior to collection. For the experiments in which embryos were analysed under live imaging, embryos were put in the imaging chamber containing micrometric wells with 1mL of 3.5% 1,6-Hexanediol in KSOM covered in oil and imaged directly. For FRAP, embryos were incubated for 10 minutes in 3.5%

1,6-Hexanediol in KSOM and then photobleached. For all the experiments at the 2-cell and 4-cell stage, embryos were incubated in KSOM until 48h and 54h post-hCG, respectively, prior to 1,6-Hexanediol treatment.

RNA sequencing and analysis.

Library preparation and sequencing. Embryos were cultured until the indicated timepoints post-hCG at which point a randomly selected pool of embryos were transferred to a drop of KSOM containing 3.5 % 1,6 Hexanediol 1h before RNA extraction. Embryos were obtained from three independent experiments. Individual embryos were washed with PBS, placed in tubes with 1x lysis buffer (Takara ST0948) containing ERCC spike-ins (External RNA Controls Consortium; Thermo) and flash frozen in liquid nitrogen. RNA-seq library preparation was carried out with the SMART-seq2 protocol (Picelli et al., 2013), and subjected to paired-end sequencing on the NovaSeq 6000 (Illumina) platform.

Data processing. Paired-end sequencing reads were mapped to the reference genome (version GRCh38, primary assembly fasta) using STAR (version 2.7.6a) with the parameter `quantMode GeneCounts` and the `gtf` annotation (version GRCh38.101), and ERCC92 (ThermoFisher). Raw read counts files (`ReadsPerGene.out.tab`) generated by STAR were used for downstream analysis.

Differential expression analysis. Downstream analysis and data visualization was performed in R (version 4.1.2). Samples were filtered for at least 500,000 genic reads and less than 20 percent ERCC as well as mitochondrial reads. Differential expression analysis was carried out by using DESeq2 (version 1.34.0) and custom helper functions (source: <https://github.com/tschauer/HelpersforDESeq2>). Genes with at least one read present in at least 25% of all samples were considered in the analysis. Experimental batch was included as covariate in the model. Significant differentially expressed genes were defined by an adjusted p-value cutoff of 0.05. Results were summarized on MA-plots showing log₂ fold change in treatment vs. control in relation to log₁₀ mean counts. Expression levels (CPM – counts per million) of endogenous control genes (*Actb* and *Gapdh*) were visualized as dot plots. All analysis steps were carried out in a reproducible pipeline using snakemake (version snakemake-minimal 5.2.4) and is available upon request.

Trim away for ATRX

For FRAP, zygotes were first microinjected with 1 mg/ml of either ATRX or IgG antibodies with 0.33% of Cascade blue (Invitrogen, D1976) or rhodamine B dextran (Invitrogen, D1841), to monitor the success of injection, at 18h post-HCG. After 4 h, zygotes were microinjected a second time with the following mRNAs: 130 ng/ml Trim21 with 600ng μ l of TALEMajSat-mClover and 300ng μ l of H3.1-mRFP. Embryos were placed in a 37 °C and 5% CO₂ incubator

until the beginning of the FRAP experiment at the 2-cell stage. For live imaging measurements, the same procedure was followed, except that the second microinjection was done with the following mRNAs: 200 ng/ μ l mCherry-Trim21 with 600ng/ μ l of TALEMajSat-mClover and 200 ng/ μ l of H2B-tidRFP. For the RNA FISH and qPCR experiments, the second microinjection included 200 ng/ml mCherry-Trim21 and 200 ng/ml mRNA for mGFP as a positive control for microinjection.

Quantitative RT-qPCR

Single embryo RTqPCR was performed as previously described (Burton et al., 2013). Control, 1,6-Hexanediol treated or ATRX and IgG Trim-away embryos were washed in PBS and snap frozen in 5 μ l of 2x reaction buffer at -80°C (CellsDirect One-Step qRT-PCR kit, 11753100, ThermoFisher) at 49 h post-hCG. TaqMan Gene Expression assays (20 \times Applied Biosystems), were pooled to a final concentration of 0.2x for each of the 3 assays. To each of the single-cell samples in 2x reaction buffer, 2.5 μ l of 0.2x assay pool, 0.5 μ l RT/Taq enzyme (CellsDirect One-Step qRT-PCR kit, 11753100, ThermoFisher) and 2.3 μ l of water were added. Cell lysis and sequence-specific reverse transcription were performed at 50 °C for 20 min. The reverse transcriptase was inactivated by heating to 95 °C for 2 min. The resulting cDNA was diluted tenfold before analysis with Universal PCR Master Mix and 1x TaqMan Gene Expression assays (Applied Biosystems). qPCR analysis was carried out with an initial denaturing at 95 °C for 5 min followed by 45 cycles of denaturing at 95 °C for 15 s and then annealing and amplification at 60 °C for 1 min. All raw Ct values were normalised to Gapdh for all conditions. The delta-delta Ct and then fold change were calculated compared to the mean of the control for the 1,6-Hexanediol experiment and to the mean of the IgG Trim-away for the ATRX Trim-away experiments. Box plots of the resulting dataset were generated using the ggplot2 R package.

Statistical Analysis

Statistical significance based on the two-sided Mann-Whitney U-test was performed for all the analysis except for the RICS experiment where a two-sided Wilcoxon signed-rank test was performed, and in the experiments in which we determined appropriate 1,6-Hexanediol concentrations (Extended Data Fig.6a) where a Kruskal-Wallis and Conover-Iman test with bonferonni correction were performed. All data derive from at least 3 independent biological experiments, excepting Extended Data Fig. 9d, in which 2 independent replicates were performed. All boxplots show the median and interquartile range; whiskers span the range of the data, while extending no further than 1.5 \times the interquartile range. N corresponds to the number of independent experiments and n to the total number of embryos analyzed. The actual *P* values up to the third decimal are indicated throughout the figures.

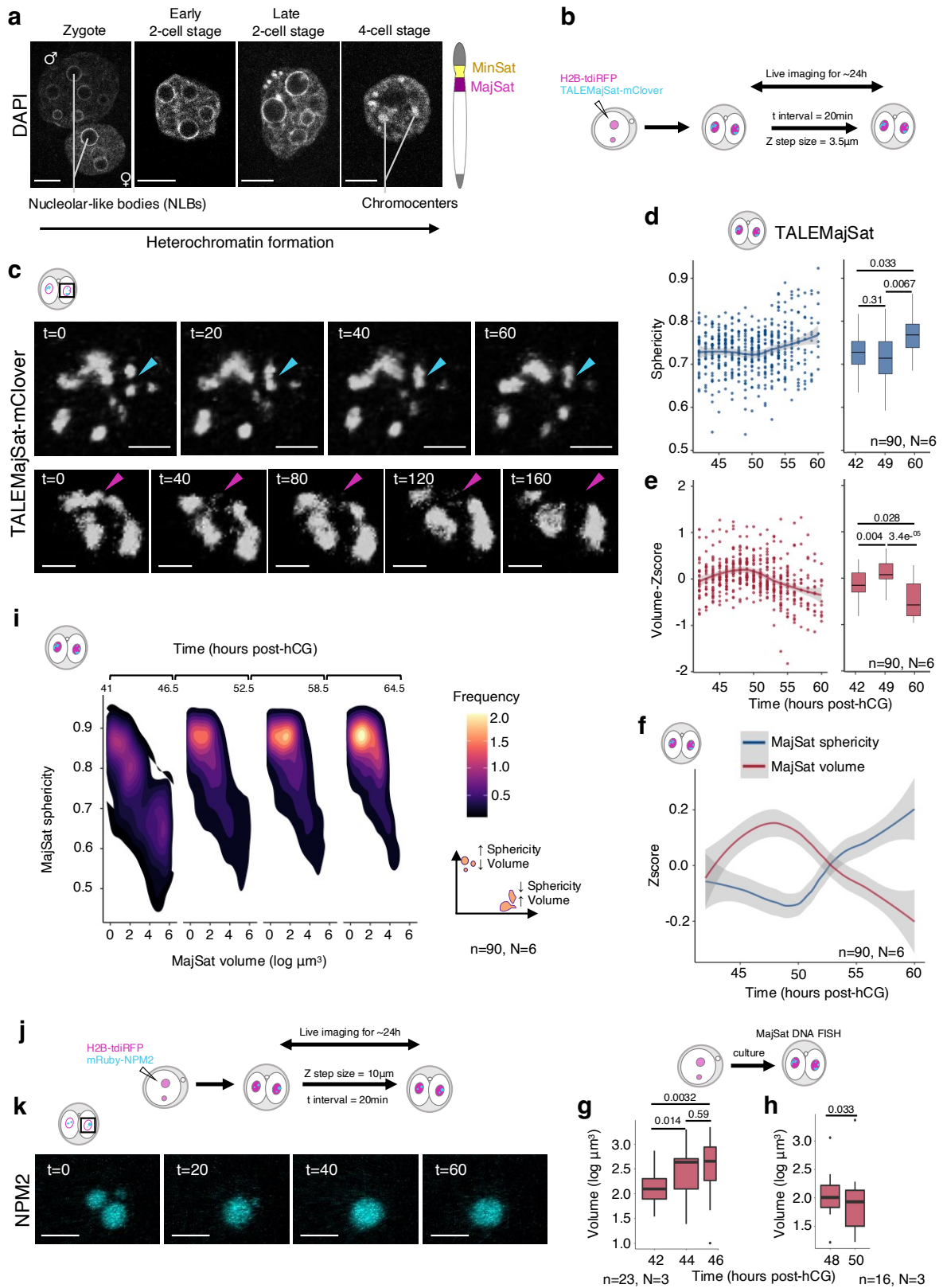


Fig. 1. Live imaging of pericentromeric heterochromatin in mouse embryos.

(a) DAPI stained nuclei of the corresponding mouse preimplantation embryo stage to visualise the nucleolar-like bodies and chromocenters. Right: schematic representation of the centromeric regions of a typical mouse chromosome. Scale bars, 10 μ m.

(b) Experimental design for visualising pericentromeric heterochromatin in live embryos.

(c) Still images of a single Z-plane of one nucleus from live imaging of a 2-cell stage embryo in the major satellite channel at the corresponding time in minutes. Arrows point to events of fusion (top row, cyan) and fission (bottom row, magenta). The black square in the 2-cell schematic depicts the imaged area. Scale bars, 5 μ m.

(d) Mean sphericity of major satellite regions in one nucleus represented with a dotplot and a locally estimated scatterplot smoothing (LOESS) curve (left) or a boxplot with data from the indicated post-hCG timepoints (right).

(e) Z-scores of the volume of pericentromeric domains per nucleus during the 2-cell stage plotted as in Fig. 1d.

(f) LOESS curves for the z-scores of sphericity and volume of major satellites per nucleus throughout the 2-cell stage.

(g) Volume of individual pericentromeric heterochromatin regions in freshly collected early 2-cell stage embryos at the corresponding timepoints, quantified using super-resolution microscopy of major satellite DNA FISH with stimulated emission depletion microscopy (STED).

(h) Volume of individual pericentromeric heterochromatin regions as in (g), but in embryos collected at the late 2-cell stage at the corresponding timepoints and quantified using superresolution microscopy of major satellite DNA FISH with stimulated emission depletion microscopy (STED).

(i) Density plot showing the sphericity and volume of individual pericentromeric domains across 4 different time windows covering the 2-cell stage. The small schematic shows the expected position of small spherical versus large irregular objects.

(j) Experimental design for live imaging of nucleolar-like bodies at the 2-cell stage.

(k) Still images of a single Z-plane from live imaging of one 2-cell stage nucleus in the NPM2 channel at the corresponding time in minutes. The black square in the 2-cell schematic depicts the imaged area. Scale bars, 10 μ m.

All statistical analysis was performed using the two-sided Mann–Whitney U-test.

Fig. 2. Pericentromeric heterochromatin phase separates in mouse embryos.

(a) Representative confocal section of a single nucleus of a late 2-cell stage embryo after immunostaining for SNAP-H3.1 and H3K9me3 after Triton X-100 pre-extraction to detect chromatin bound H3.1. The arrow points towards a pericentromeric domain. Scale bar, 10 μ m.

(b) Experimental design of Raster Image Correlation Spectroscopy (RICS) of histone H3.1 at pericentromeric heterochromatin. Representative confocal microscopy images of TALE-MajSat and SNAP-H3.1 used for RICS analysis. Diffusion coefficients for H3.1-mRFP mean \pm SEM values are 5.5 \pm 0.24 μ m²/s in the nucleoplasm versus 4.4 \pm 0.23 μ m²/s at pericentromeric domain. The boxplots show the median and the IQR and whiskers depict the smallest and largest values within 1.5 \times IQR. The black square in the 2-cell schematic depicts the imaged area. Scale bar, 5 μ m.

(c) Experimental design for Fluorescent Recovery After Photobleaching (FRAP) analysis of histone H3.1 at the 2-cell stage. Schematics and still images of FRAP performed in the nucleoplasm or inside pericentromeric heterochromatin, either when around the NLBs or at chromocenters. The black square in the 2-cell schematic depicts the imaged area. Bleach regions are represented by a red circle. Scale bars, 8 μ m.

(d) Recovery curves of H3.1-mRFP in the specified regions. The individual points are the mean \pm SEM. The curve is a double exponential fit of all individual points for each region.

(e) Mean \pm SEM of the estimated H3.1-mRFP mobile fraction from fitting a double exponential to each individual FRAP experiment.

(f) Mean \pm SEM of estimated H3.1-mRFP fast recovery kinetics from fitting a double exponential to each individual FRAP experiment.

(g) Recovery curves of H2B-mRFP in the specified regions. The individual points are the mean \pm SEM. The curve is a double exponential fit of all individual points for each region.

(h) Mean \pm SEM of the estimated H2B-mRFP mobile fraction from fitting a double exponential to each individual FRAP experiment.

(i) Mean \pm SEM of estimated H2B-mRFP fast recovery kinetics from fitting a double exponential to each individual FRAP experiment.

(j) Diffusion coefficients for H3.1-mRFP in the nucleoplasm (Nuc.) and at major satellites (MajSat) at the 4-cell stage determined using RICS where each dot represents one nucleus. Mean \pm SEM values are 4.22 \pm 0.23 μ m²/s in the nucleoplasm versus 3.8 \pm 0.16 μ m²/s at the pericentromeric domain. The boxplots show the median and the IQR and whiskers depict the smallest and largest values within 1.5 \times IQR.

(k) Relative diffusion of SNAP-H3.1 at the MajSat region over the nucleoplasm at the 2- and 4-cell stage.

All statistical analysis was performed using the two-sided Mann–Whitney U-test.

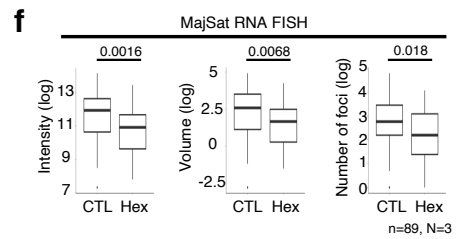
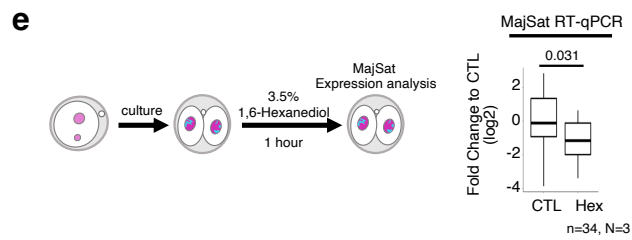
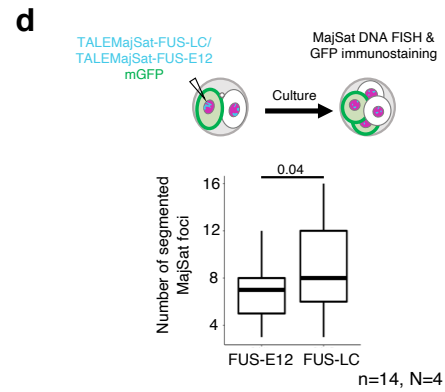
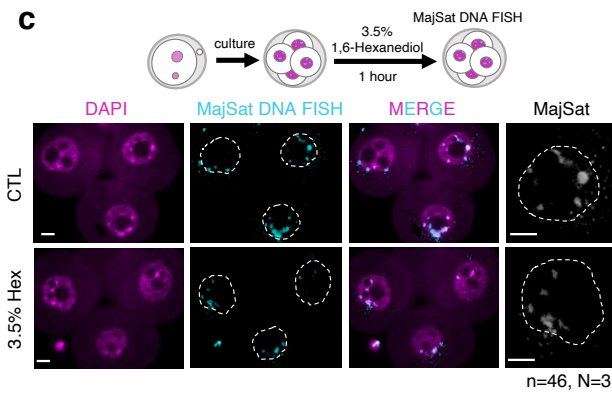
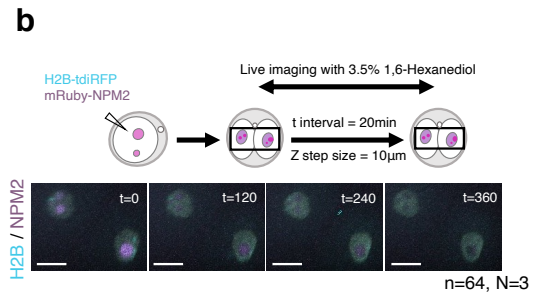
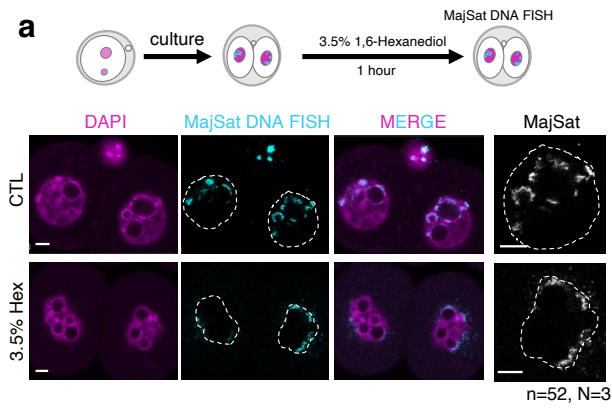


Fig. 3. Biophysical properties of pericentromeric heterochromatin at the 2-cell stage are dependent on weak hydrophobic interactions

(a) Experimental design and representative images of major satellite DNA FISH after 1,6-Hexanediol treatment at the 2-cell stage. The panels show the DAPI and the major satellite channels, a merge of both channels and an inset with a higher magnification of one of the two nuclei in the major satellite channel. White dashed lines demarcate the nuclear membrane. Scale bars, 5 μ m. See Extended Data Figure. 6b for additional representative images.

(b) Experimental design and still images from live imaging of nucleolar-like bodies (NLB) with 1,6-Hexanediol. Images show the merge of the NPM2 and H2B channels at the corresponding timepoints in minutes. The black square in the 2-cell schematic depicts the imaged area. Scale bars, 15 μ m.

(c) Experimental design and representative images of major satellite DNA FISH after 1,6-Hexanediol treatment at the 4-cell stage. White dashed lines demarcate the nuclear membrane. Scale bars, 10 μ m.

(d) Experimental design and results of the number of segmented MajSat foci detected by MajSat DNA FISH after TALEMajSat-FUS-LC or TALEMajSat-FUS-MUT expression following mRNA microinjection. mGFP was used as an injection marker to identify the injected blastomere expressing the TALE FUS fusion.

(e) RT-qPCR analysis of major satellite transcripts after 1,6-Hexanediol treatment. Data are normalized against *Gapdh* mRNA and shown as log₂ fold change to controls. Mean Ct values for *Gapdh* are 28.79 and 28.45 for the control and 1,6-Hexanediol-treated embryos, respectively.

(f) Experimental design and results of quantification of major satellite transcripts by RNA FISH after 1,6-Hexanediol treatment. Total intensity, volume and number of FISH foci per nucleus are shown. The boxplots show the median and the IQR and whiskers depict the smallest and largest values within 1.5 \times IQR.

All statistical analysis was performed using the two-sided Mann–Whitney U-test.

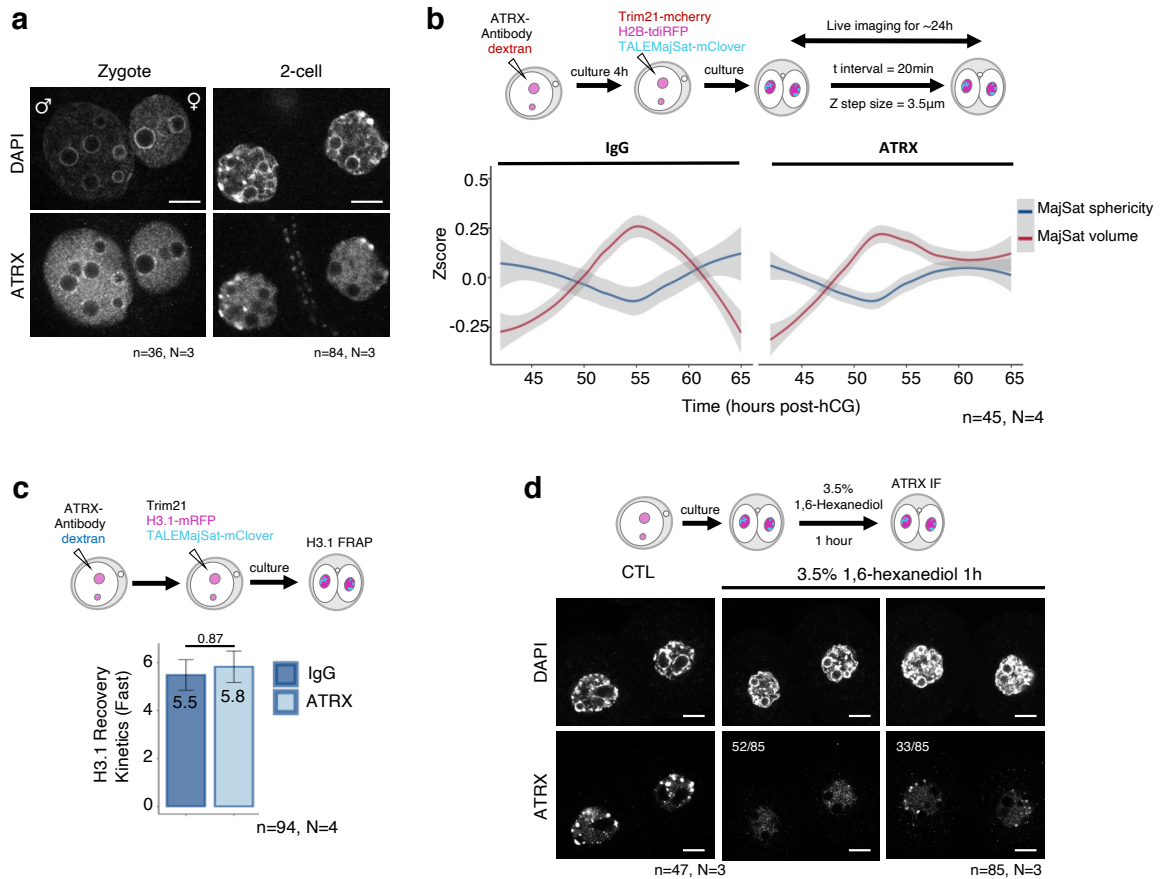


Fig. 4. Weak hydrophobic interactions-dependent ATRX recruitment to pericentromeric chromatin mediates chromocenter maturation.

(a) Representative confocal sections of ATRX immunostaining in zygotes and late 2-cell stage embryos. The DAPI channel is also shown.

(b) Experimental design for the analysis of pericentromeric heterochromatin in live embryos after ATRX depletion by Trim-Away. LOESS curves for the z-score of sphericity and volume of pericentromeric domains per nucleus across the 2-cell stage for control IgG and ATRX depleted conditions.

(c) Experimental design and quantification of fast recovery kinetics of H3.1-mRFP in the specified regions after ATRX Trim-Away, measured by FRAP in late 2-cell stage embryos.

(d) Experimental design and representative confocal sections of ATRX immunostaining after 1,6-Hexanediol treatment. The panels show the results of untreated embryos and 2 types of phenotypes observed after 1,6-Hexanediol treatment and their frequency.

Scale bars throughout: 10µm.

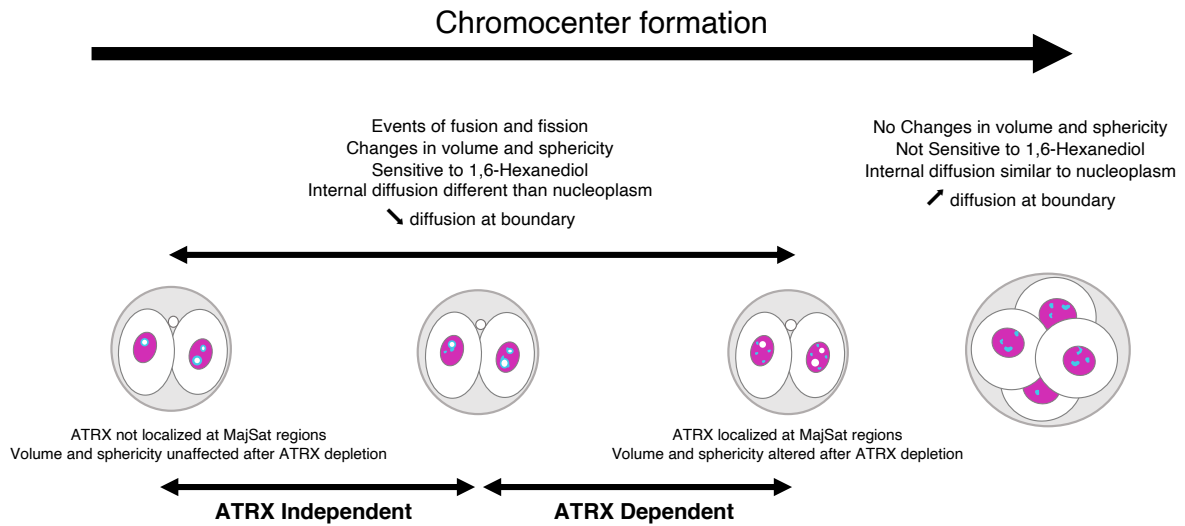
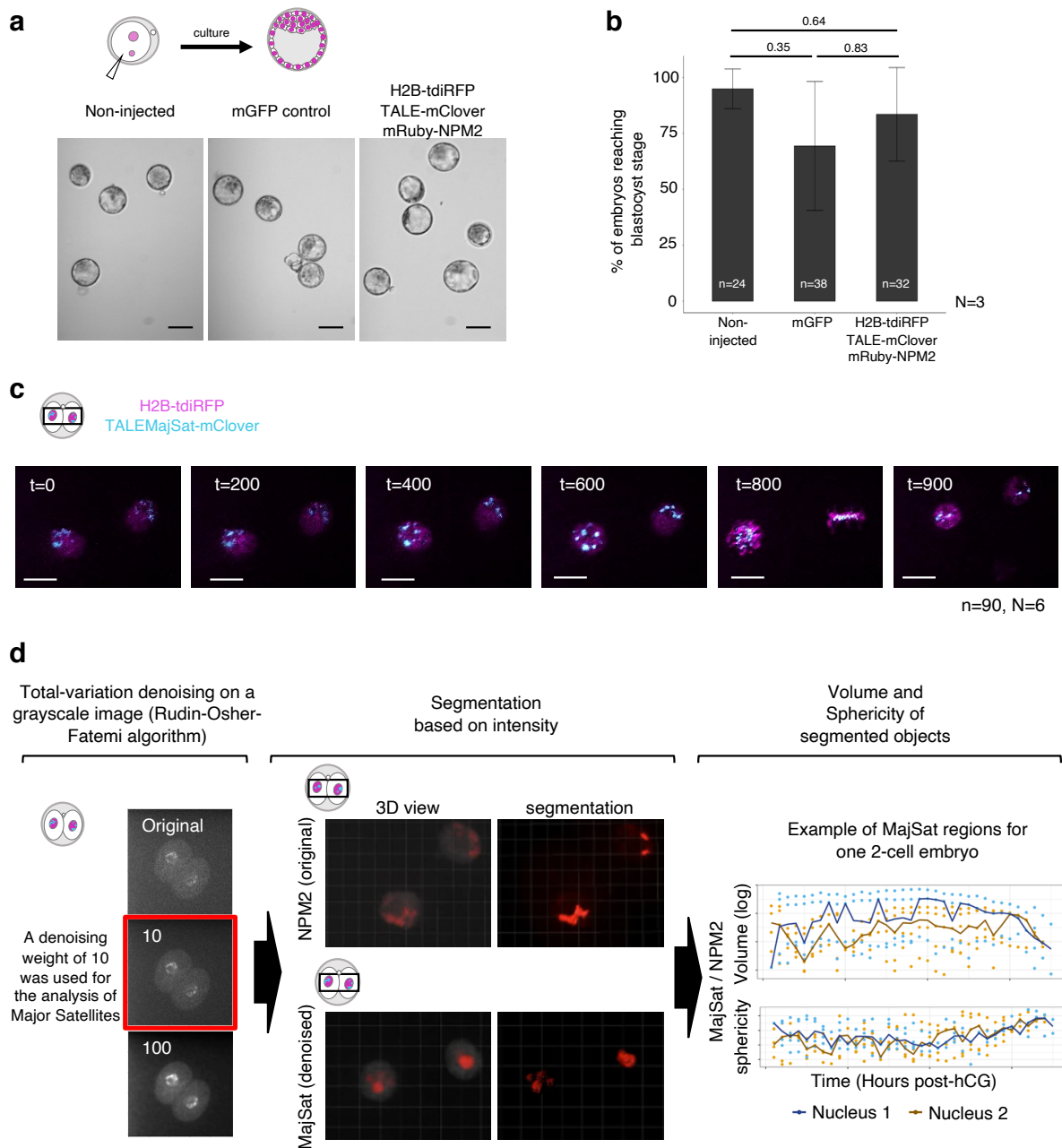


Fig. 5. Model summarizing our findings on the phase state transition of embryonic heterochromatin and the role of ATRX in this process. Pericentromeric heterochromatin transitions from a liquid-like to a more solid- or gel-like state during the process of chromocenter formation. ATRX is required for chromocenter formation but not for the initial heterochromatin phase transition.



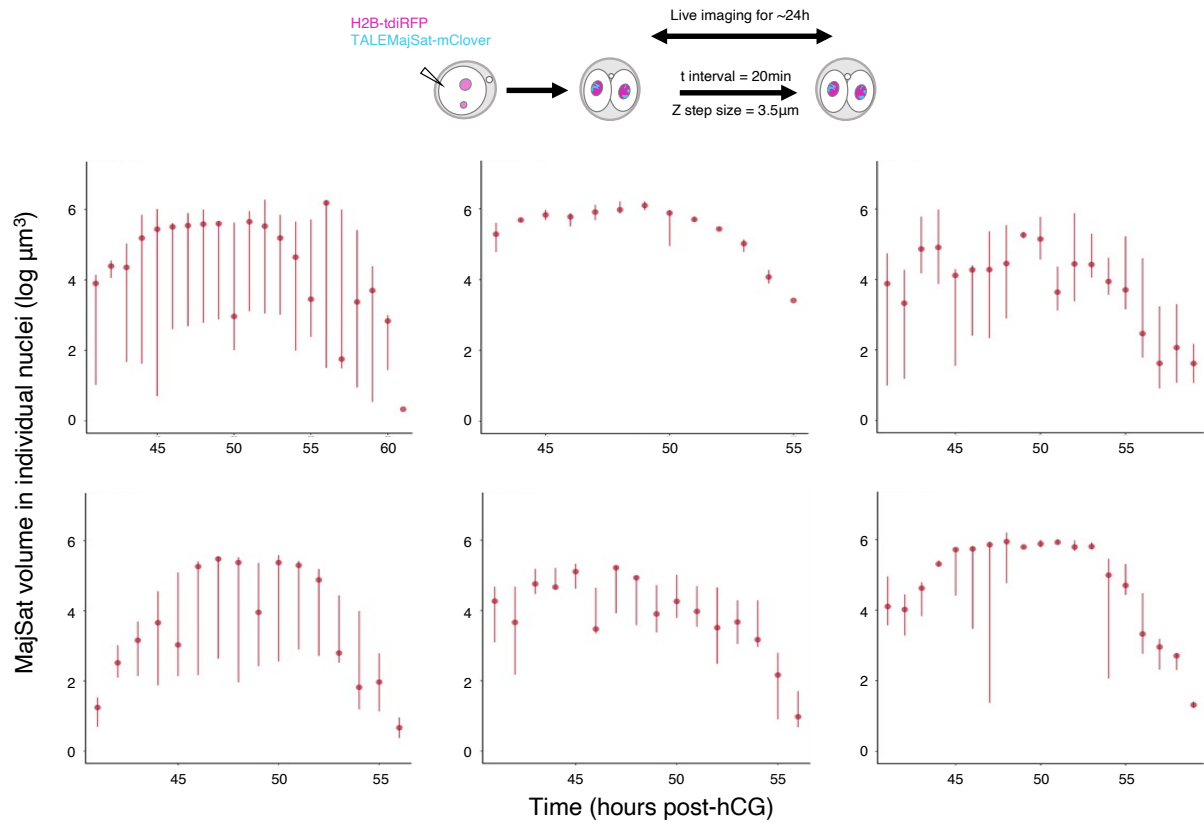
Extended Data Fig.1. Live imaging protocol and pipeline to study the dynamics and biophysical properties of pericentromeric heterochromatin in mouse embryos.

(a) Experimental design and representative bright-field images of embryos 3 days after microinjection with *H2B-tdiRFP*, *TALEMajSat-mClover* and *mRuby-NPM2* mRNA. Scale bars, 100 μ m.

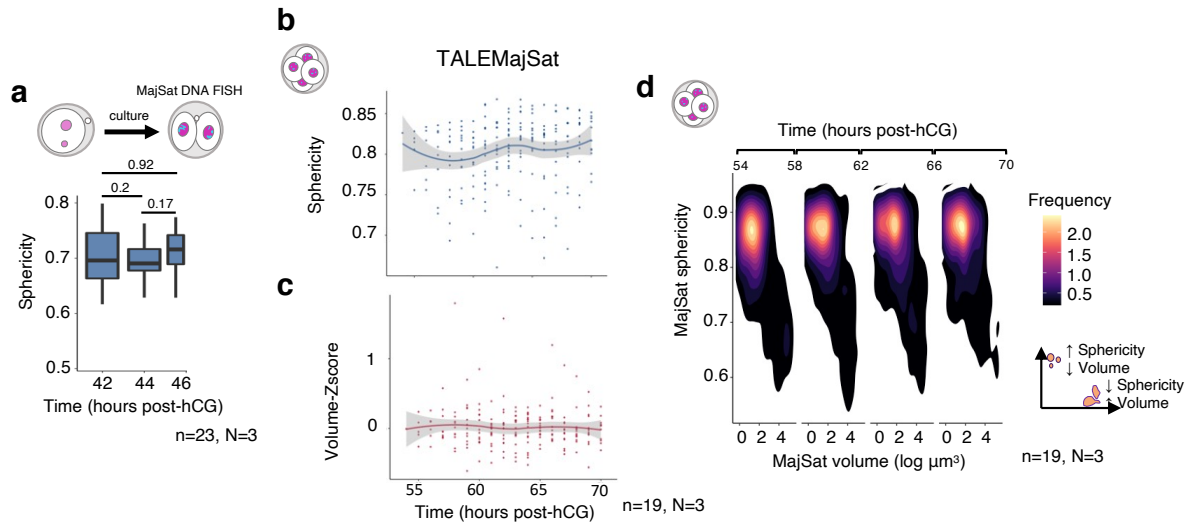
(b) Proportion of embryos that reached the blastocyst stage after 3 days of culture.

(c) Representative projections from 3D-reconstructions of Z-stacks acquired during live imaging of a 2-cell stage embryo in the major satellite and H2B channels at the corresponding time in minutes. The black square in the 2-cell schematic depicts the imaged area. Scale bars, 15 μ m.

(d) Diagram of the image analysis pipeline developed to calculate the volume and sphericity of 3D-reconstructed pericentromeric heterochromatin domains and nucleolar-like bodies (NLBs) from live imaging. The major satellite channel is denoised with the Rudin-Osher-Fatemi algorithm and subsequently segmentation and calculation of volume and sphericity for major satellites and NPM2 was performed using Imaris software. The black square in the 2-cell schematic depicts the imaged area. All statistical analysis was performed using the two-sided Mann–Whitney U-test.



Extended Data Fig.2. Examples of volume changes of major satellite for individual nuclei of 2-cell stage embryos. Median and 25/75 quantiles of the volume of major satellite segmented region (y-axis) at each timepoint based (x-axis) on live imaging with the TGV system. Each graph contains data for a randomly selected nucleus from the dataset compiled in Fig. 1e.



Extended Data Fig.3. Changes in volume and sphericity of heterochromatin domains are specific to the 2-cell stage.

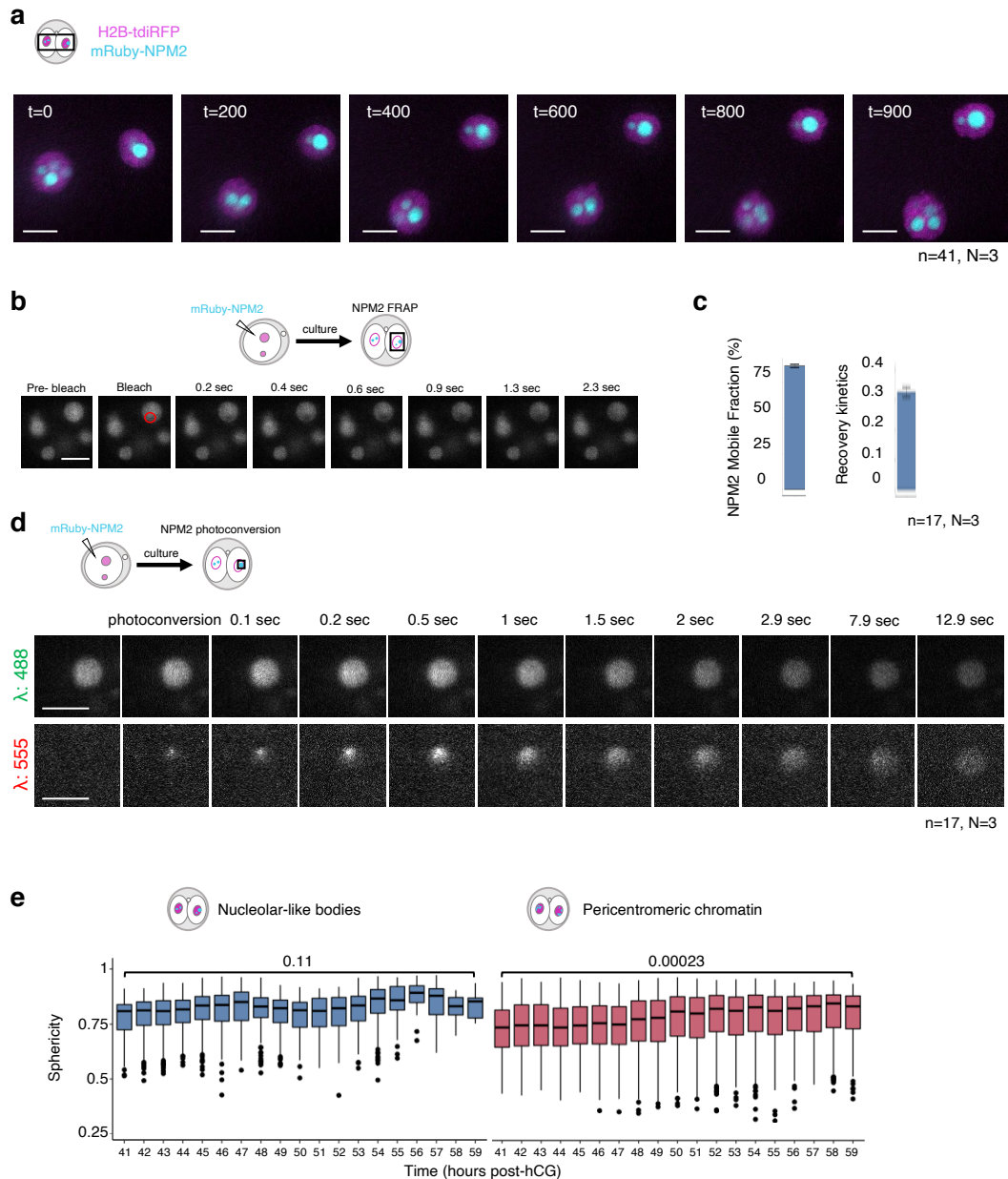
(a) Sphericity of individual pericentromeric heterochromatin regions in freshly collected early 2-cell stage embryos at the corresponding timepoints, quantified using superresolution microscopy of major satellite DNA FISH with stimulated emission depletion microscopy (STED).

(b) Mean sphericity of pericentromeric heterochromatin domains per nucleus from live imaging measurements in 4-cell stage embryos plotted as in Fig. 1d.

(c) Volume of pericentromeric heterochromatin per nucleus, calculated from live imaging during the 4-cell stage plotted as in Fig. 1e.

(d) Density plot showing the sphericity and volume of individual pericentromeric domains across 4 different time windows during the 4-cell stage. The small schematic shows the expected position of small spherical versus large irregular objects.

All statistical analysis was performed using the two-sided Mann–Whitney U-test.



Extended Data Fig.4. The nucleolar-like bodies have liquid-like properties at the 2-cell stage.

(a) Representative projections of 3D-reconstructions of Z-stacks acquired from live imaging of a 2-cell stage embryo in the NPM2 and H2B channels at the corresponding time in minutes. The black square in the 2-cell schematic depicts the imaged area. Scale bars, 15 μ m.

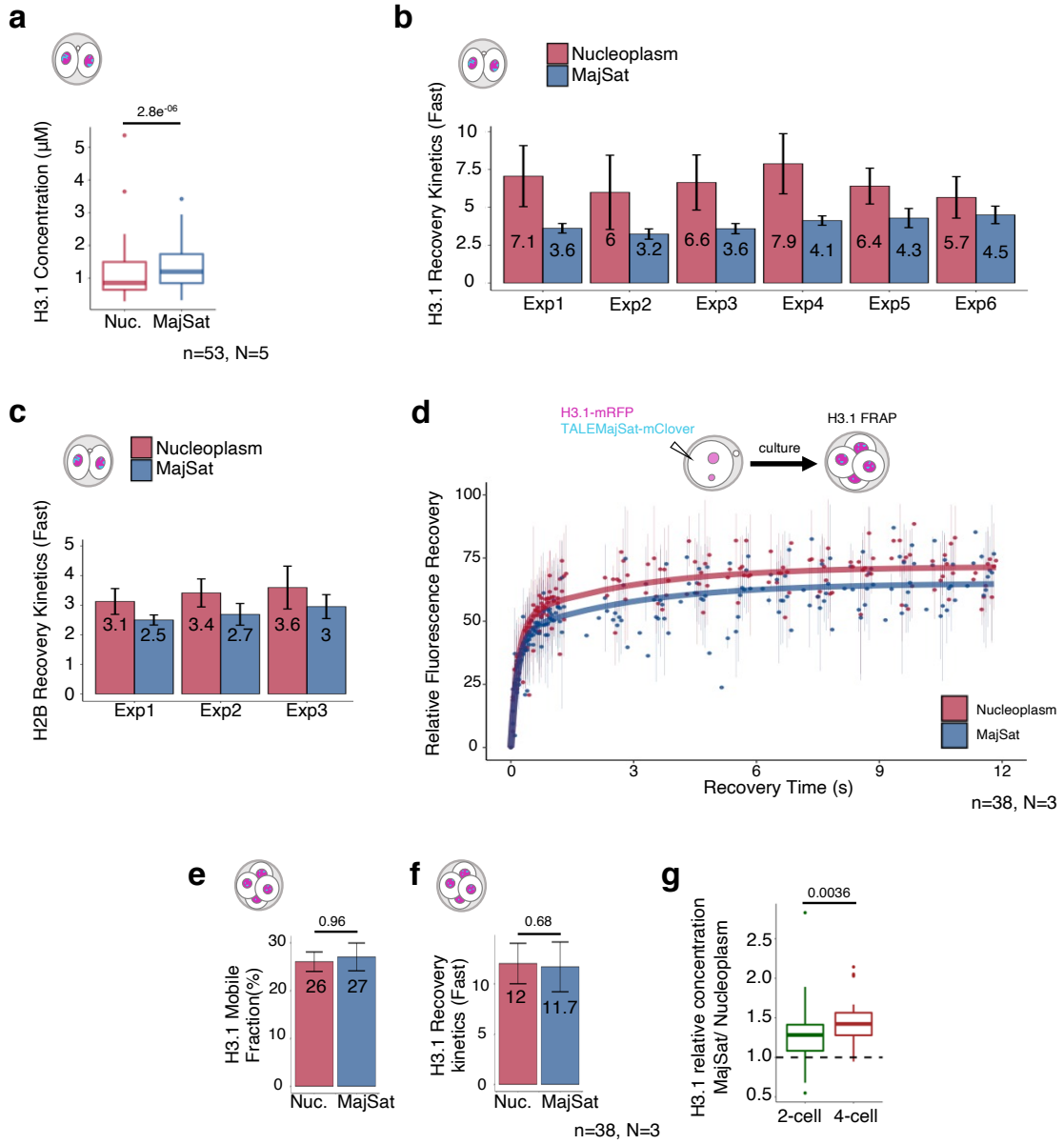
(b) Experimental design and still images from FRAP for NPM2 in the nucleolar-like bodies (NLBs) at the 2-cell stage. Bleach region is represented by a red circle. The black square in the 2-cell schematic depicts the imaged area. Scale bar, 8 μ m.

(c) Mean \pm SEM of NPM2 mobile fraction and recovery kinetics at the 2-cell stage from fitting a single exponential to each individual FRAP experiment.

(d) Experimental design and still images of photoconversion experiments for NPM2 in nucleolar-like bodies at the 2-cell stage. The black square in the 2-cell schematic depicts the imaged area and the signal distribution at the 555 wavelength indicates internal rearrangement of NPM2 within the NLBs. Scale bars, 8 μ m.

(e) Sphericity of nucleolar-like bodies (NLBs) and pericentromeric heterochromatin obtained from live imaging in 2-cell stage.

All statistical analysis was performed using the two-sided Mann–Whitney U-test.



Extended Data Fig.5. The liquid-like properties of pericentromeric heterochromatin are specific to the 2-cell stage.

(a) SNAP-H3.1 concentration at nucleoplasm (Nuc.) and at pericentromeric heterochromatin (MajSat) at the 2-cell stage as measured by RICS. The boxplots show the median and the IQR and whiskers depict the smallest and largest values within $1.5 \times \text{IQR}$. Mean \pm SEM values are $1.15 \times 10^{-6} \pm 1.2 \times 10^{-7}$ in the nucleoplasm and $1.34 \times 10^{-6} \pm 9.85 \times 10^{-8}$ at pericentromeric domain.

(b) Mean \pm SEM of estimated H3.1 fast recovery kinetics in the specified regions at the 2cell stage for each biological replicate used for Fig. 2f.

(c) Mean \pm SEM of estimated H2B fast recovery kinetics in the specified regions at the 2cell stage for each biological replicate used for Fig.2i.

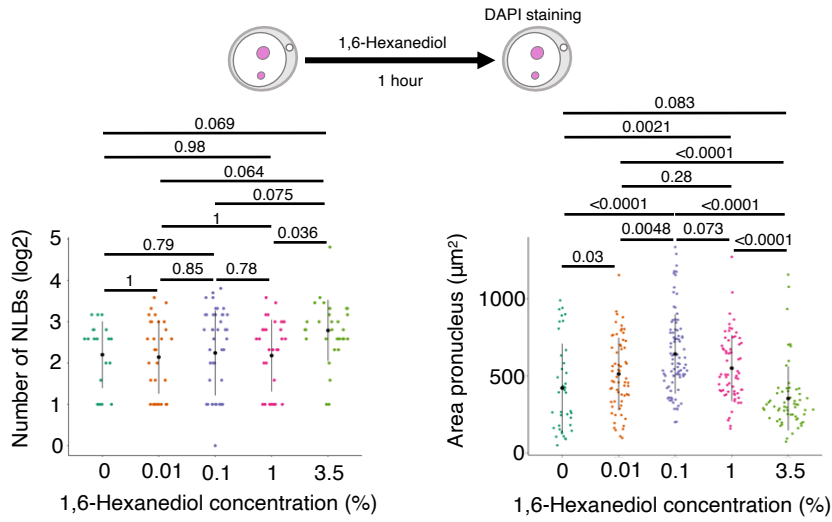
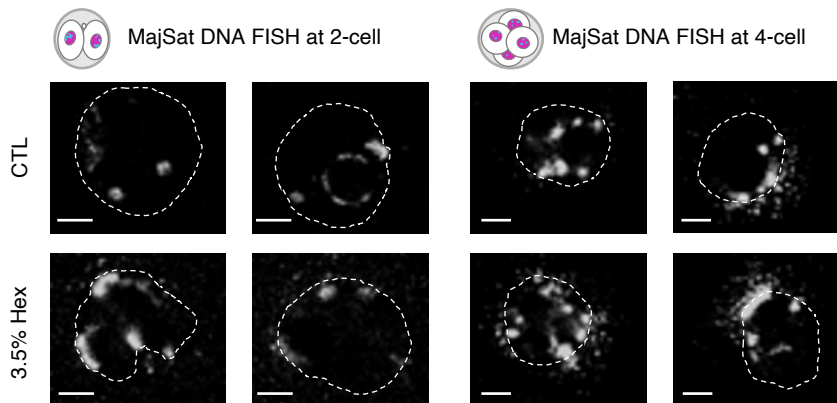
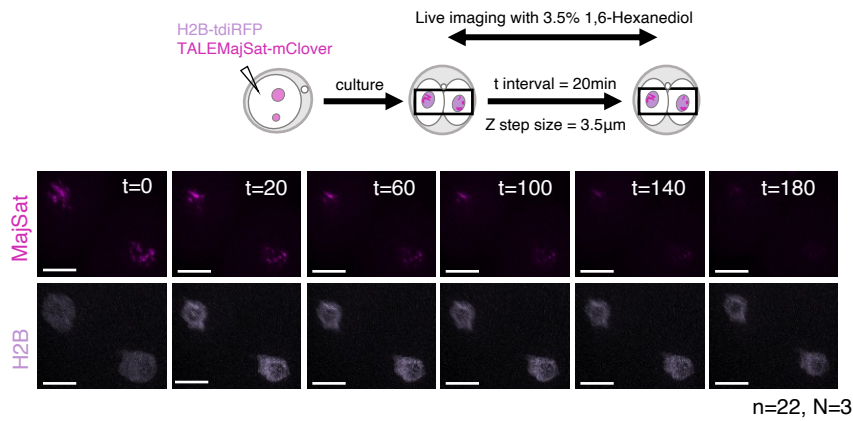
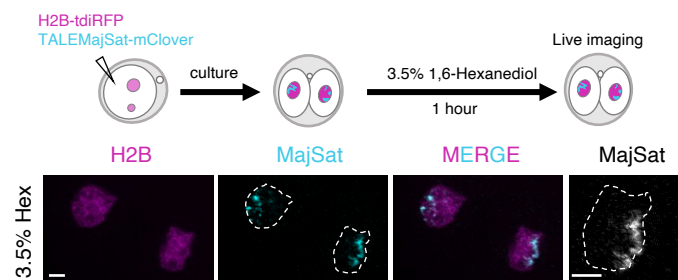
(d) Experimental design and FRAP recovery curves of H3.1-mRFP at the 4-cell stage in the specified regions. The individual points are the mean \pm SEM. The curve is a double exponential fit of all individual points for each region.

(e) Mean \pm SEM of estimated H3.1 mobile fractions from fitting a double exponential to each individual FRAP experiment at the 4-cell stage.

(f) Mean \pm SEM of estimated H3.1 fast recovery kinetics from fitting a double exponential to each individual FRAP experiment at the 4-cell stage.

(g) SNAP-H3.1 enrichment at the pericentromeric heterochromatin relative to the nucleoplasm at the 2- and the 4-cell stage. The enrichment was calculated as a ratio of H3.1 concentration at each stage inside and outside the pericentromeric domain.

All statistical analyses were performed using the two-sided Mann–Whitney U-test

a**b****c****d**

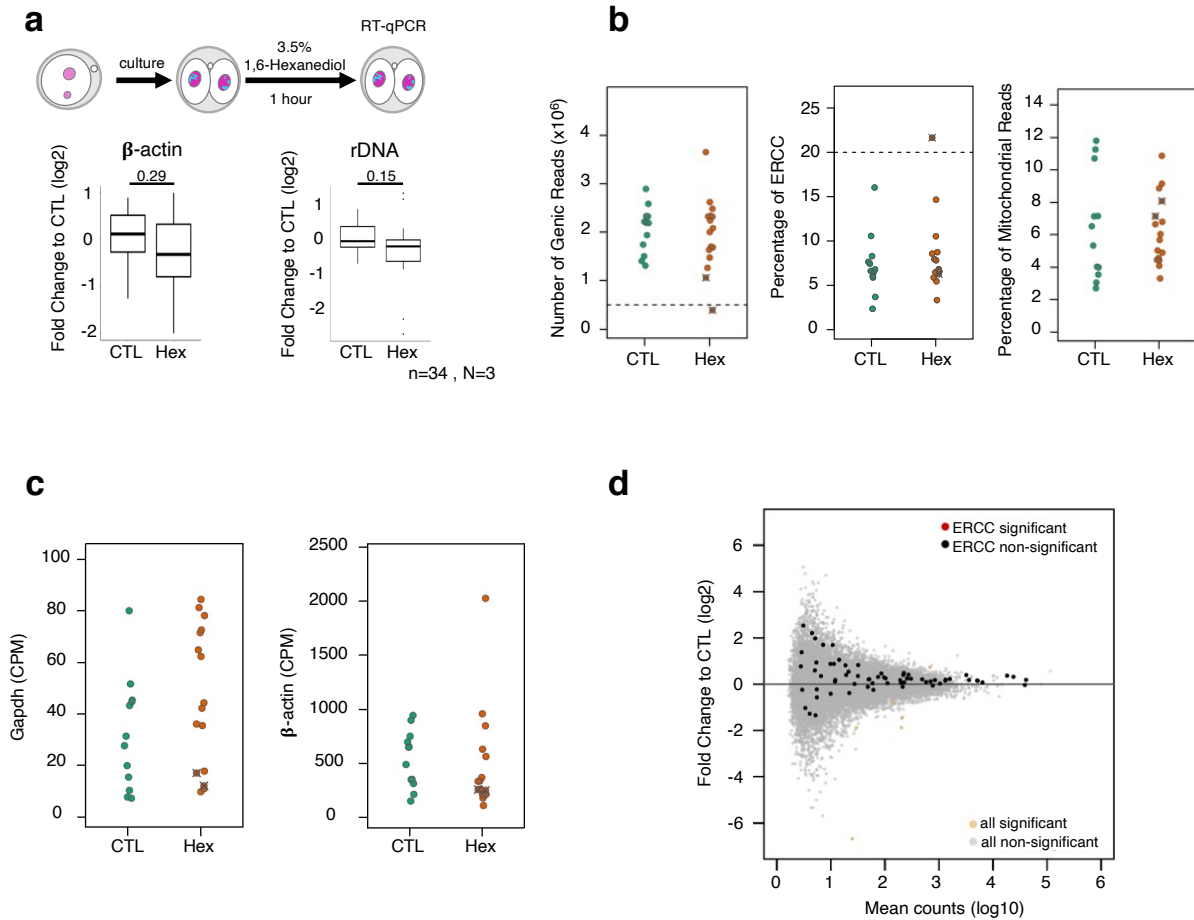
Extended Data Fig.6. Establishing conditions for 1,6-Hexanediol treatment in mouse embryos.

(a) Experimental design to determine the effective concentration of 1,6-Hexanediol in mouse embryos. Embryos were treated at the indicated concentrations of 1,6-Hexanediol, fixed and the number of nucleolar-like bodies was determined using confocal imaging of DAPI-stained embryos. The left SINA plot shows the number of nucleolar-like bodies per embryo under different 1,6-Hexanediol concentrations. The right SINA plot shows the area of the pronuclei, measured in the Z-section of their maximal diameter, after incubation with the indicated 1,6-Hexanediol concentrations. Statistical analysis was performed using a Kruskal-Wallis test followed by a Conover-Iman test with bonferonni correction.

(b) Examples of MajSat DNA FISH after 1,6-Hexanediol treatment in 2- and 4-cell embryos. Higher magnification images showing additional examples of control and 1,6-Hexanediol -treated embryos as in Fig. 3a and c. White dashed lines demarcate the nuclear membrane. Scale bars, 5 μ m.

(c) Experimental design and still images from live imaging of pericentromeric heterochromatin during 1,6-Hexanediol treatment. Images show the major satellite and H2B channels at the corresponding timepoints in minutes. The black square in the 2-cell schematic depicts the imaged area. Scale bars, 15 μ m.

(d) Still images from live imaging of pericentromeric heterochromatin with 1,6-Hexanediol 60 minutes after the start of the treatment. The panels show the H2B and the major satellite channels, a merge of both channels and images of major satellites in one nuclei at a higher magnification of one of the two nuclei in the major satellite channel. White dashed lines demarcate the nuclear membrane. Scale bars, 5 μ m.



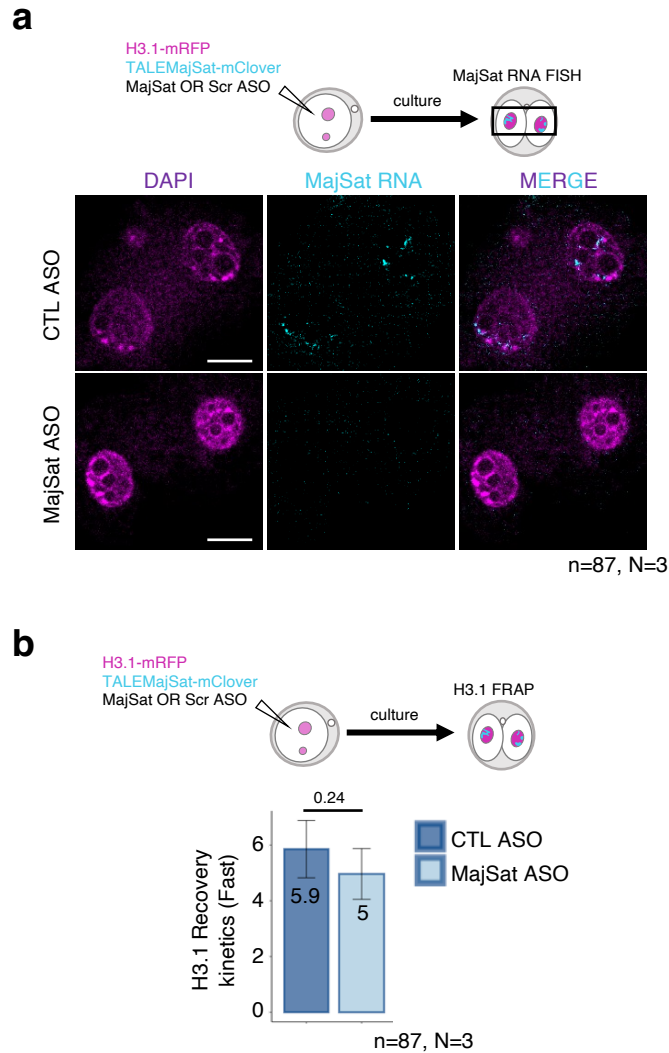
Extended Data Fig. 7. 1,6-Hexanediol does not affect global gene expression at the 2-cell stage.

(a) RT-qPCR analysis of β -actin and *rDNA* transcripts after 1,6-Hexanediol treatment. Data are normalized against *Gapdh* mRNA and shown as log₂ fold change to controls.

(b) Sequencing statistics and quality control for single embryo RNA-seq datasets in control and 1,6-Hexanediol-treated embryos. Two embryos (cross) were removed for downstream analysis after quality control filtering based on genic reads number and percentage of ERCC (dashed lines, see Methods).

(c) Counts per million (CPM) for *Gapdh* and β -actin in 1,6-Hexanediol-treated and control single embryo Smart-seq2 libraries.

(d) MA plot depicting differential gene expression as log₂ fold change between 1,6-Hexanediol-treated and control 2-cell stage embryos in relation to log₁₀ mean normalized counts. ERCC spike-ins are indicated in black (non-significant). Significantly differentially expressed genes are indicated in orange, non-significant changes in grey.

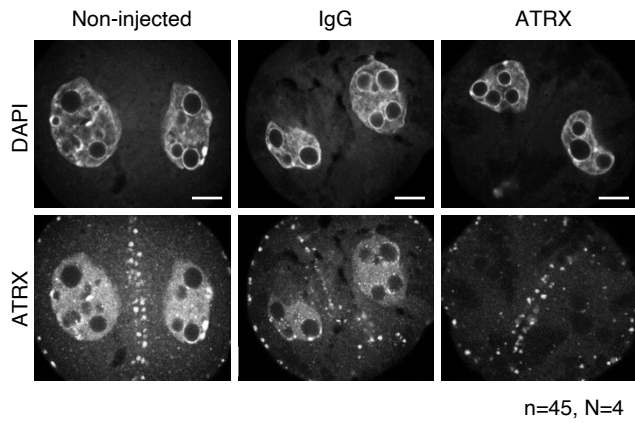
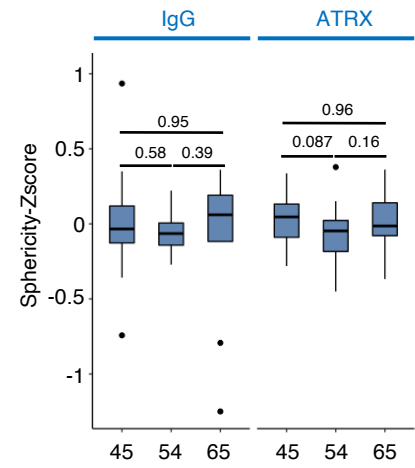
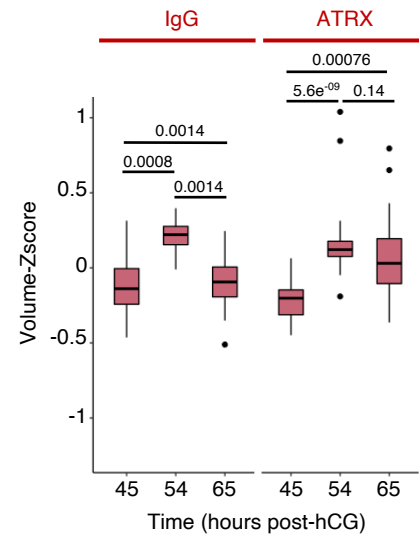
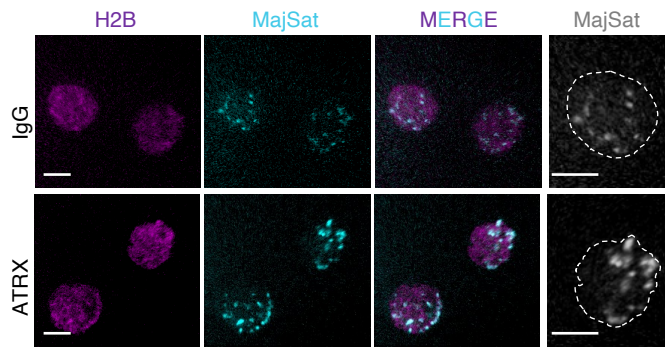
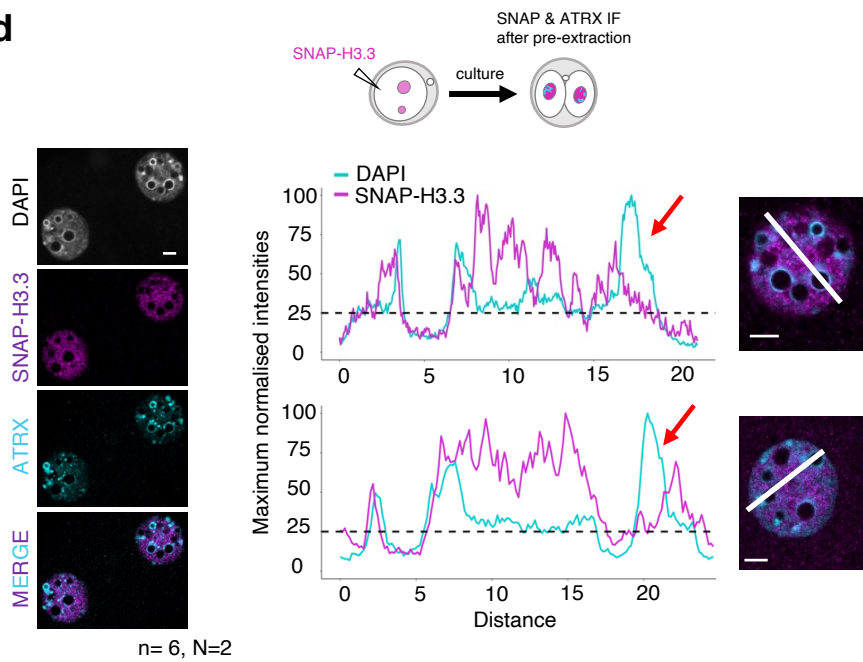


Extended Data Fig.8. Major satellite knock-down does not perturb heterochromatin dynamics in 2-cell stage embryos.

(a) Experimental design for the analysis of H3.1 kinetics at pericentromeric heterochromatin in live embryos after major satellite RNA knock down and validation by RNA FISH. RNA FISH for major satellites at 46 hours post-HCG of embryos microinjected with either a control or a major satellite antisense oligo (ASO). The panels show the DAPI, the major satellite RNA-FISH signal and a merge of both channels. Scale bars, 15µm.

(b) Experimental design and quantification of fast recovery kinetics (Mean±SEM) of H3.1-mRFP in the major satellite regions after major satellite RNA knock-down, as measured by FRAP in late 2-cell stage embryos.

All statistical analyses were performed using the two-sided Mann–Whitney U-test.

a**b****c****d**

Extended Data Fig.9. ATRX is efficiently degraded by Trim-Away at the 2-cell stage.

(a) Representative single confocal sections of an embryo immunostained for ATRX after Trim-Away with anti-ATRX antibody or IgG control. Scale bars, 10 μ m.

(b) Z-scores of the volume and sphericity of pericentromeric domains per nucleus after ATRX Trim-away represented as boxplots at the corresponding timepoints.

(c) Still images from live imaging of pericentromeric heterochromatin of 2-cell stage embryos at 65 hours post-HCG for control IgG and ATRX depleted conditions. The panels show the H2B and the major satellite channels, a merge of both channels and an inset with a higher magnification of one of the two nuclei in the major satellite channel. White dashed lines demarcate the nuclear membrane. Scale bars, 10 μ m.

(d) Experimental design to detect *de novo* incorporated H3.3. Representative images of a 2-cell stage embryo after immunostaining with a SNAP or an ATRX antibody (left). The panels on the right show representative line profiles for DAPI and SNAP-H3.3 along the line in the corresponding nuclei, shown on the right. The red arrow indicates the position of the chromocentres. Scale bar is 5 μ m.

All statistical analyses were performed using the two-sided Mann–Whitney U-test.

DISCUSSION

PART 1: METABOLIC STATE OF TOTIPOTENT CELLS

2-cell-like cells recapitulate several transcriptional and chromatin features of the 2-cell embryo but little is known about the metabolic state of these cells. The study included in this thesis showed for the first time that 2CLCs have a “quiet” metabolic state represented by a lower oxygen consumption and glycolytic activity reminiscent of the totipotent stages of mouse embryos (Rodriguez-Terrones et al., 2020). The lower oxygen consumption measured in 2CLCs indicated a lower respiratory rate which was also confirmed by lower levels of reactive oxygen species produced by these cells. These observations prompted us to study the differences in mitochondrial morphology and membrane potential between totipotent and pluripotent cells. 2CLCs, similar to zygotes and 2-cell embryos, exhibited mitochondria with an electron poor matrix and irregularly shaped cristae associated with lower oxygen consumption. Instead, blastocyst and ESCs in culture displayed mitochondria with an opaque matrix and more developed cristae. On the other hand, the membrane potential of mitochondria between the 2-cell and the blastocyst did not differ significantly. Therefore, contrary to their morphological differences, mitochondria display similar functional potentials in embryos during the stages supporting totipotency and pluripotency. This suggests that further analysis is required to fully understand the metabolic similarities and differences between totipotent and pluripotent cells in vitro and in vivo. Nonetheless, this study contributes greatly to a better understanding of the metabolic properties of totipotent cells and their potential importance in the early mammalian embryo. It also shows the importance of understanding the metabolic properties of distinct cell types as this is likely to enable a greater control over the artificial reprogramming and modulation of the potency of cells in culture, for cell-based therapeutic applications for example.

PART 2: ROLE OF LIQUID-LIQUID PHASE SEPARATION IN HETEROCHROMATIN FORMATION DURING MOUSE PRE-IMPLANTATION DEVELOPMENT

In recent years, a lot of effort has been put into studying the biophysical properties of constitutive heterochromatin and in identifying the major players that could nucleate its condensation. The main focus of my PhD has been to tackle this complicated question in the context of de novo heterochromatin formation in the mouse pre-implantation embryo. We reasoned that, based on changes in expression of the core heterochromatic proteins and the dynamic nature of heterochromatin at these early stages of development, the process of

chromocenter formation would be an ideal model to shed light on these complex processes. First, using *in silico* methods, we identified a list of potential scaffolds for heterochromatin condensation and showed that their expression correlated on average with chromocenter formation. Second, using *in vivo* microscopy methods, we demonstrated that the biophysical properties of heterochromatin can change over time during preimplantation development. Our work not only confirms that the process of chromocenter formation is a good model for identifying novel players in the regulation of the biophysical properties of heterochromatin in the living embryo, but also gives experimental tools for this new exciting line of research.

EXPRESSION AND PHASE SEPARATION POTENTIAL OF HETEROCHROMATIN PROTEINS DURING EARLY MOUSE DEVELOPMENT

This project started with a simple question: what are the core constitutive heterochromatic proteins and can we predict their propensity to form liquid condensates. In order to tackle this question, we started by collating all the publicly available mass spectrometry studies focusing on constitutive heterochromatin (Bartke et al., 2010; Becker et al., 2017; Eberl et al., 2013; Engelen et al., 2015; Ji et al., 2015; Soldi and Bonaldi, 2013; Vermeulen et al., 2010). From this combined analysis we identified core heterochromatin proteins, which we then continued to further study for their propensity to phase separate. As described in more detail in the introduction, most of the proteins that mediate phase separation contain IDRs, thought to be important for nucleating liquid droplets by forming multivalent weak interactions with other proteins or nucleic acids. We thus decided to generate disorder estimates for our core heterochromatic proteins and compare them with a random group of total or nuclear proteins. The results obtained with two different algorithms (Dosztanyi, 2018; Romero et al., 2001) showed that the core heterochromatic proteins have a significantly higher disorder score than both of the control groups (Fig 4). Moreover, the median percentage of disorder per protein for our heterochromatic proteins is similar to that which was calculated for the proteome of several membrane-less organelles (Sawyer et al., 2019) suggesting that indeed heterochromatic proteins have a higher potential to phase separate. This important result indicates that it is likely that more than one of the core heterochromatic proteins plays a role in heterochromatin's biophysical properties.

Since our goal was to study heterochromatin LLPS in the mouse pre-implantation embryo, we decided to study the expression pattern of the core heterochromatic proteins during these stages of development using publicly available RNA-seq (Deng et al., 2014) and mass

spectrometry (Gao et al., 2017) datasets. Our results show that, on average, heterochromatin proteins increase in expression both at the RNA and protein level during pre-implantation development. Furthermore, the increase in RNA levels happened before the 4-cell stage and thus concomitantly to the process of chromocenter formation (Fig 4). This sharp increase in zygotes and 2-cell embryos also suggest that most of these proteins are not maternally inherited, facilitating the study of their function in vivo via knock-down by siRNA. This in silico analysis thus demonstrates the relevance of the mouse embryo as a model to study heterochromatin LLPS.

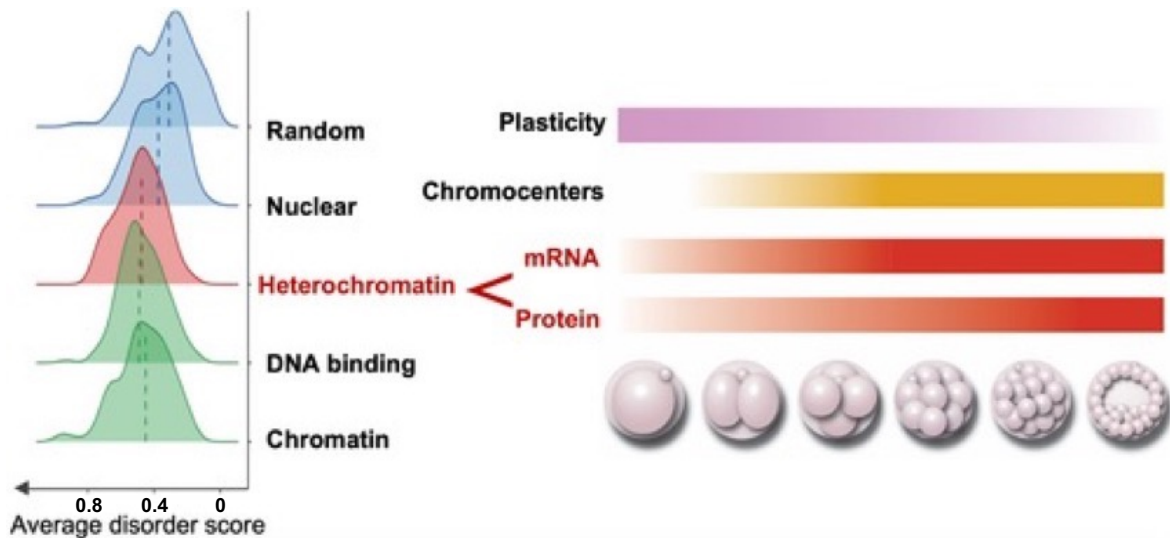


Figure 4. The 148 selected heterochromatin proteins have a significantly higher average disorder score compared to a random or nuclear set of proteins while having similar values compared to DNA binding or chromatin associated proteins (left). The heterochromatin proteins increase their expression both at the RNA and protein level during the process of chromocenter formation in mouse pre-implantation development (left).

After showing that heterochromatin proteins have a high average disorder score, we decided to continue our analysis to see if they have other features predictive for LLPS. Indeed, several types of weak interactions, like electrostatic, hydrophobic, cation and Pi-Pi, have been shown to drive LLPS. We thus measured, with the use of several algorithms, these additional features, namely the presence of a PLD (Lancaster et al., 2014), the propensity of Pi-Pi contacts (Vernon et al., 2018), the proportion of charged amino acids (Holehouse et al., 2017) and hydrophobicity (Gasteiger et al., 2003). In order to perform this manual analysis, we first had to reduce the number of heterochromatin proteins that we would study. We decided to focus on chromocenter-associated proteins, since these structures represent a major feature of constitutive heterochromatin in mouse cells and we aimed to study the formation of these structures in the developing mouse embryo. Thus, we cross-compared our core heterochromatin proteins with a dataset of proteins found to interact with major satellites originally identified using proteomics of isolated chromatin segments method (PiCH) in mouse

ESCs (Saksouk et al., 2014). We identified 7 proteins with 5 of those showing a high disorder score, namely HP1 α , HP1 β , ATRX, UHRF1 and DNMT1 that we compared to FUS, a well characterised protein shown to phase separate in vitro and in vivo (Murthy et al., 2019). Our analysis showed that the selected proteins had large segments with high hydrophobicity and a high fraction of charged residues that could facilitate LLPS. Moreover, we identified a PLD at the C-terminal of ATRX, prompting us to study the role of this protein in the biophysical properties of constitutive heterochromatin in mouse pre-implantation embryos. More recently several algorithms have been developed to predict the phase separation potential of a protein of interest based on the features described above and thus a more systematic analysis could now be performed (Chu et al., 2022; Hatos et al., 2022; Orlando et al., 2019; van Mierlo et al., 2021). However, at the time most of these predictors were not available, and our study was the first analysis of LLPS potential of heterochromatic proteins and constitutes a valuable dataset for the community.

A PHASE TRANSITION ACCOMPANIES HETEROCHROMATIN FORMATION IN MOUSE EMBRYOS

As detailed before, we decided to focus on chromocenter formation as a model to study biophysical properties of de novo establishment of constitutive heterochromatin. Indeed, in zygotes and early 2-cell embryos, the MajSat repeats are located around the NLBs and start forming the chromocenters at the late 2-cell stage. We first defined 5 hallmarks of LLPS that we could study, in the context of forming chromocenters, in vivo (Fig 5); (i) high sphericity, (ii) ability to fuse and fission, (iii) distinct diffusion and concentration of its internal components compared to the nucleoplasm, (iv) reduction of diffusion at the boundary of the condensate, and (v) sensitivity to 1,6-hexanediol. We then established a toolbox of microscopy experiments to test these hallmarks comprising of live imaging, FRAP and RICS.

Our live imaging results at the 2-cell stage show that MajSat domains are able to undergo fusion and fission, while their morphology changes, specifically during the process of chromocenter formation. Moreover, we show that the diffusion and concentration of H3.1, an internal component of the MajSat region, is significantly different inside the compartment compared to the surrounding nucleoplasm. In addition, results obtained with FRAP suggest that the diffusion of H3.1 is reduced at the boundary of the MajSat regions. Finally, treatment of 2-cell embryos with 1,6-hexanediol highly perturbs MajSat localisation and morphology. Interestingly, all these results were specific to the process of chromocenter formation at the 2-cell stage compared to the 4-cell stage embryo when chromocenters are fully mature.

Combining the data from all in vivo experiments, we came to the conclusion that MajSat regions transition from a liquid-liquid like state to a more solid-like or gel-like state during chromocenter formation (Fig 5). This work shows for the first time that the biophysical properties of heterochromatin can vary depending on the developmental state of the cells. This could explain the seemingly contradicting results present in the literature currently (Erdel et al., 2020; Larson et al., 2017; Strom et al., 2017). Indeed, our results show that the phase separation features of heterochromatin could be cell-type specific depending on the chromatin environment and set of heterochromatic proteins expressed. The experimental toolbox we developed during my PhD will also be of great help for the community as it will enable other labs to assess the phase separation potential of their proteins or RNA of interest in an in vivo setting (Fig 5).

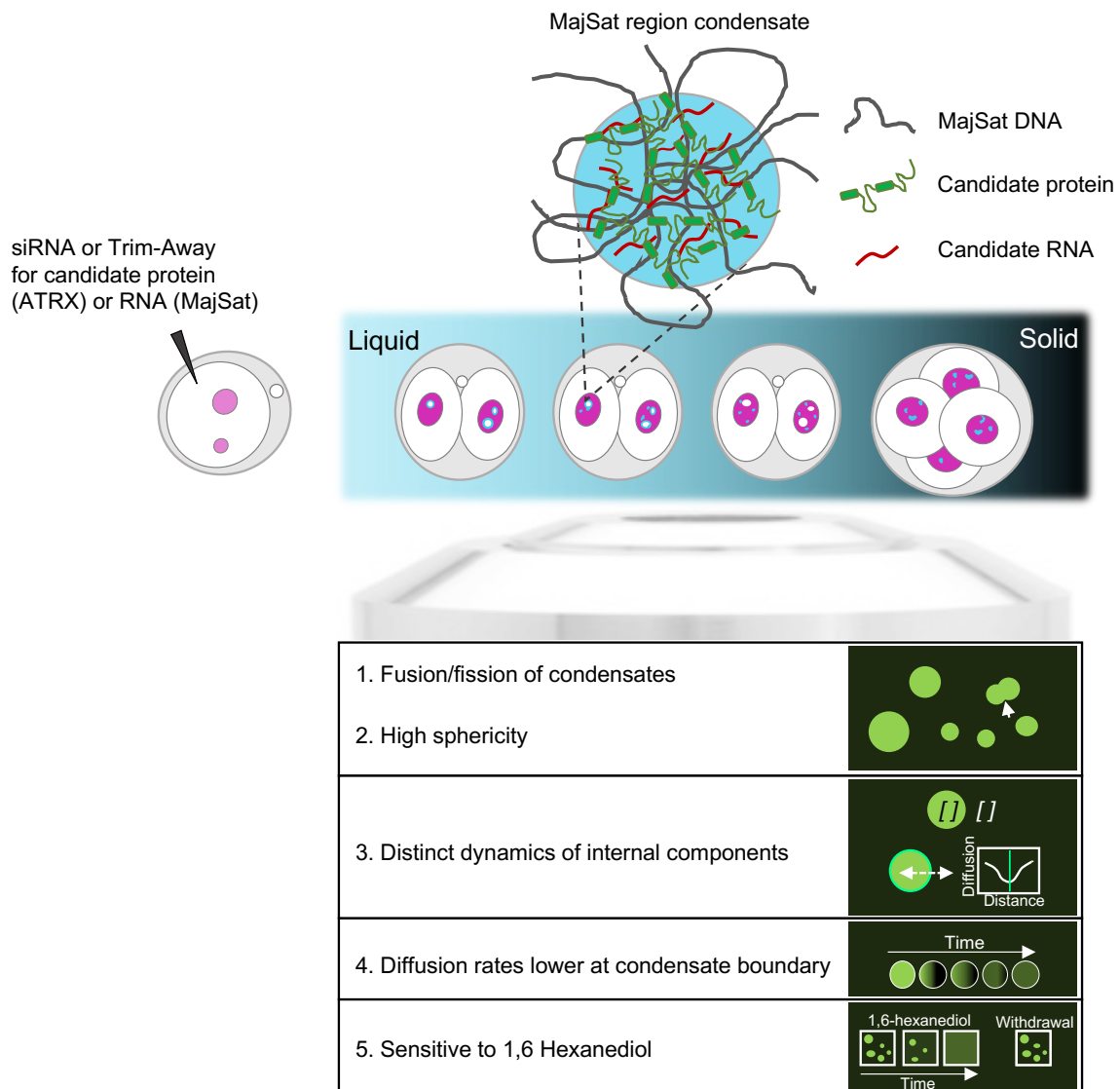


Figure 5. Model showing how chromocenter formation in the embryo can be used to study the phase separation potential of a protein or RNA of interest by means of siRNAs or Trim-away. The phase transition of embryonic heterochromatin happening in wild-type conditions is shown in the blue box. The 5 hallmarks studied in vivo during this study is shown at the bottom.

The dynamic changes in biophysical properties of MajSat regions during chromocenter formation makes it a perfect model for the functional study of these properties. For example, MajSats are expressed when they have liquid-like properties in 2-cell embryos, while their downregulation corresponds to the transition to a solid state, when chromocenters mature. We thus tested if disrupting MajSat condensates with 1,6-hexanediol treatment had an impact on their expression. Our results show indeed a significant reduction of MajSat expression when embryos are treated with 1,6-hexanediol. This suggests a functional role of the weak hydrophobic interactions and condensation of MajSat domains on their expression. Additionally, we developed a system to perturb the solid state of MajSat at the 4-cell stage by tethering the IDR of FUS, which was recently reported to induce liquid-like properties (Bose et al., 2022), to chromocenters. We show that this tethering induces the dispersion of MajSat regions, which enables the functional study of the transition of chromocenters to a solid-like state at the 2-cell stage. It would be interesting to test whether this manipulation also affects the expression of MajSat, for example.

Finally, the process of chromocenter formation can be used to study phase the separation potential of a protein or RNA of interest and thus the mechanism of chromocenter formation (Fig 5). Indeed, with the experimental tool box we developed, it becomes feasible to knock-down or over-express any potential scaffold and study how it impacts the biophysical properties of MajSats in vivo. During my PhD we focused on two candidates in particular, ATRX, which we had identified as a protein with potential to undergo LLPS in our in silico analysis, and the MajSat RNA, which had been shown to play a role in chromocenter formation (Casanova et al., 2013). Our results show that ATRX is a client of the MajSat condensate as 1,6-hexanediol perturbs its recruitment to MajSat while depleting it does not affect H3.1 dynamics at MajSats. On the other hand, knocking down MajSats does not impact H3.1 dynamics in the condensate, suggesting that MajSat RNA does not contribute to the biophysical properties of forming chromocenters at this stage of development. As detailed in the introduction, many other proteins have been shown to induce condensation of constitutive heterochromatin in other systems. During my PhD, I proposed several methods for the study of these known proteins, as well as the additional proteins identified in our in silico work, by presenting tools to measure the biophysical properties of constitutive heterochromatin in the mouse pre-implantation embryo.

Overall, the work shown in this thesis emphasises the functional relevance of the biophysical properties of pericentromeric chromatin in de novo heterochromatin formation as well as in proper pre-implantation development. Indeed, perturbing the liquid-state of pericentromeric chromatin affects MajSat expression and ATRX localisation at chromocenter which have both

been linked to proper chromocenter formation and pre-implantation development (Baumann et al., 2010; Casanova et al., 2013; Probst et al., 2010).

OUTLOOK

The totipotent embryo is a unique model system as it differs from most cells in terms of plasticity, gene expression, metabolism and nuclear organisation. During my PhD, I characterised the latter two, using a wide spectrum of microscopy techniques, expanding our knowledge on the features of totipotency. These results are, not only important to better understand how normal development at these early stages occurs, but also to develop better models to study totipotency in vitro. Most importantly, my work shows that the very dynamic nature of the totipotent embryo makes it an invaluable model system to functionally characterise the role of biophysical processes in the regulation of chromatin architecture and of cellular metabolism in reprogramming and cellular plasticity.

REFERENCES

- Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419-434.
- Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285-298.
- Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional Control of Phase-Separated Cellular Bodies. *Cell* **166**, 651-663.
- Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143**, 470-484.
- Baumann, C., Viveiros, M.M., and De La Fuente, R. (2010). Loss of maternal ATRX results in centromere instability and aneuploidy in the mammalian oocyte and pre-implantation embryo. *PLoS Genet* **6**, e1001137.
- Becker, J.S., McCarthy, R.L., Sidoli, S., Donahue, G., Kaeding, K.E., He, Z., Lin, S., Garcia, B.A., and Zaret, K.S. (2017). Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. *Mol Cell* **68**, 1023-1037 e1015.
- Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., *et al.* (2018). Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **28**, 420-435.
- 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.
- Borsos, M., and 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.
- Bose, M., Lampe, M., Mahamid, J., and Ephrussi, A. (2022). Liquid-to-solid phase transition of oskar ribonucleoprotein granules is essential for their function in *Drosophila* embryonic development. *Cell* **185**, 1308-1324 e1323.
- Boskovic, A., Eid, A., Pontabry, J., Ishiuchi, T., Spiegelhalter, C., Raghu Ram, E.V., Meshorer, E., and Torres-Padilla, M.E. (2014). Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. *Genes Dev* **28**, 1042-1047.
- Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoegge, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732.
- Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc Natl Acad Sci U S A* **108**, 4334-4339.
- Brinster, R.L. (1965). Studies on the Development of Mouse Embryos in Vitro. II. The Effect of Energy Source. *J Exp Zool* **158**, 59-68.
- Burns, K.H., Viveiros, M.M., Ren, Y., Wang, P., DeMayo, F.J., Frail, D.E., Eppig, J.J., and Matzuk, M.M. (2003). Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science* **300**, 633-636.
- 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.

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.

Burton, A., and Torres-Padilla, M.E. (2010). Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. *Brief Funct Genomics* 9, 444-454.

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, 723-734.

Casanova, M., Pasternak, M., El Marjou, F., Le Baccon, P., Probst, A.V., and Almouzni, G. (2013). Heterochromatin reorganization during early mouse development requires a single-stranded noncoding transcript. *Cell Rep* 4, 1156-1167.

Cho, W.K., Spille, J.H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 361, 412-415.

Chu, X., Sun, T., Li, Q., Xu, Y., Zhang, Z., Lai, L., and Pei, J. (2022). Prediction of liquid-liquid phase separating proteins using machine learning. *BMC Bioinformatics* 23, 72.

Clift, D., So, C., McEwan, W.A., James, L.C., and Schuh, M. (2018). Acute and rapid degradation of endogenous proteins by Trim-Away. *Nat Protoc* 13, 2149-2175.

Condic, M.L. (2014). Totipotency: what it is and what it is not. *Stem Cells Dev* 23, 796-812.

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

De La Fuente, R., Baumann, C., and Viveiros, M.M. (2015). ATRX contributes to epigenetic asymmetry and silencing of major satellite transcripts in the maternal genome of the mouse embryo. *Development* 142, 1806-1817.

De La Fuente, R., Viveiros, M.M., Wigglesworth, K., and Eppig, J.J. (2004). ATRX, a member of the SNF2 family of helicase/ATPases, is required for chromosome alignment and meiotic spindle organization in metaphase II stage mouse oocytes. *Dev Biol* 272, 1-14.

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.

Dosztanyi, Z. (2018). Prediction of protein disorder based on IUPred. *Protein Sci* 27, 331-340.

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.

Eberl, H.C., Spruijt, C.G., Kelstrup, C.D., Vermeulen, M., and Mann, M. (2013). A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol Cell* 49, 368-378.

Eeftens, J.M., Kapoor, M., Michieletto, D., and Brangwynne, C.P. (2021). Polycomb condensates can promote epigenetic marks but are not required for sustained chromatin compaction. *Nat Commun* 12, 5888.

Elbaum-Garfinkle, S. (2019). Matter over mind: Liquid phase separation and neurodegeneration. *J Biol Chem* 294, 7160-7168.

Engelen, E., Brandsma, J.H., Moen, M.J., Signorile, L., Dekkers, D.H., Demmers, J., Kockx, C.E., Ozgur, Z., van, I.W.F., van den Berg, D.L., *et al.* (2015). Proteins that bind regulatory regions

identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nat Commun* **6**, 7155.

Erdel, F., Rademacher, A., Vlijm, R., Tunnermann, J., Frank, L., Weinmann, R., Schweigert, E., Yserentant, K., Hummert, J., Bauer, C., *et al.* (2020). Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Mol Cell* **78**, 236-249 e237.

Eustermann, S., Yang, J.C., Law, M.J., Amos, R., Chapman, L.M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R.J., Rhodes, D., *et al.* (2011). Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* **18**, 777-782.

Fadloun, A., Le Gras, S., Jost, B., Ziegler-Birling, C., Takahashi, H., Gorab, E., Carninci, P., and 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.

Flyamer, I.M., Gassler, J., Imakaev, M., Brandao, 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.

Frank, L., and Rippe, K. (2020). Repetitive RNAs as Regulators of Chromatin-Associated Subcompartment Formation by Phase Separation. *J Mol Biol* **432**, 4270-4286.

Franzmann, T.M., and Alberti, S. (2019). Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J Biol Chem* **294**, 7128-7136.

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.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784-3788.

Genet, M., and Torres-Padilla, M.E. (2020). The molecular and cellular features of 2-cell-like cells: a reference guide. *Development* **147**.

Getreuer, P. (2012). Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman. *Image Processing On Line* **2**, 74-95.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **179**, 470-484 e421.

Guelen, L., Pagie, L., Brassat, 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.

Guthmann, M., Burton, A., and Torres-Padilla, M.E. (2019). Expression and phase separation potential of heterochromatin proteins during early mouse development. *EMBO Rep* **20**, e47952.

Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78-82.

Hatos, A., Tosatto, S.C.E., Vendruscolo, M., and Fuxreiter, M. (2022). FuzDrop on AlphaFold: visualizing the sequence-dependent propensity of liquid-liquid phase separation and aggregation of proteins. *Nucleic Acids Res*.

Hendrickson, P.G., Dorais, J.A., Grow, E.J., Whiddon, J.L., Lim, J.W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R., Wilcox, A.L., *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.

Hendrix, J., Baumgartel, V., Schrimpf, W., Ivanchenko, S., Digman, M.A., Gratton, E., Krausslich, H.G., Muller, B., and Lamb, D.C. (2015). Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. *J Cell Biol* **210**, 629-646.

Hendrix, J., Dekens, T., Schrimpf, W., and Lamb, D.C. (2016). Arbitrary-Region Raster Image Correlation Spectroscopy. *Biophys J* **111**, 1785-1796.

Hennig, S., Kong, G., Mannen, T., Sadowska, A., Kobelke, S., Blythe, A., Knott, G.J., Iyer, K.S., Ho, D., Newcombe, E.A., *et al.* (2015). Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol* **210**, 529-539.

Henninger, J.E., Oksuz, O., Shrinivas, K., Sagi, I., LeRoy, G., Zheng, M.M., Andrews, J.O., Zamudio, A.V., Lazaris, C., Hannett, N.M., *et al.* (2021). RNA-Mediated Feedback Control of Transcriptional Condensates. *Cell* **184**, 207-225 e224.

Hermant, C., and Torres-Padilla, M.E. (2021). TFs for TEs: the transcription factor repertoire of mammalian transposable elements. *Genes Dev* **35**, 22-39.

Hewitson, L.C., and Leese, H.J. (1993). Energy metabolism of the trophectoderm and inner cell mass of the mouse blastocyst. *J Exp Zool* **267**, 337-343.

Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O., and Pappu, R.V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* **112**, 16-21.

Houghton, F.D., Thompson, J.G., Kennedy, C.J., and Leese, H.J. (1996). Oxygen consumption and energy metabolism of the early mouse embryo. *Mol Reprod Dev* **44**, 476-485.

Huo, X., Ji, L., Zhang, Y., Lv, P., Cao, X., Wang, Q., Yan, Z., Dong, S., Du, D., Zhang, F., *et al.* (2020). The Nuclear Matrix Protein SAFB Cooperates with Major Satellite RNAs to Stabilize Heterochromatin Architecture Partially through Phase Separation. *Mol Cell* **77**, 368-383 e367.

Ideue, T., Cho, Y., Nishimura, K., and Tani, T. (2014). Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells* **19**, 528-538.

Inoue, A., and Aoki, F. (2010). Role of the nucleoplasmin 2 C-terminal domain in the formation of nucleolus-like bodies in mouse oocytes. *FASEB J* **24**, 485-494.

Ishiuchi, T., Enriquez-Gasca, R., Mizutani, E., Boskovic, A., Ziegler-Birling, C., Rodriguez-Terrones, D., Wakayama, T., Vaquerizas, J.M., and Torres-Padilla, M.E. (2015). Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat Struct Mol Biol* **22**, 662-671.

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P.W., Cochrane, J.C., Allis, C.D., Picketts, D.J., Patel, D.J., Li, H., *et al.* (2011). ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* **18**, 769-776.

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.

Ji, X., Dadon, D.B., Abraham, B.J., Lee, T.I., Jaenisch, R., Bradner, J.E., and Young, R.A. (2015). Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proc Natl Acad Sci U S A* **112**, 3841-3846.

Kaneko, K.J. (2016). Metabolism of Preimplantation Embryo Development: A Bystander or an Active Participant? *Curr Top Dev Biol* **120**, 259-310.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., *et al.* (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* *149*, 753-767.

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.

Keenen, M.M., Brown, D., Brennan, L.D., Renger, R., Khoo, H., Carlson, C.R., Huang, B., Grill, S.W., Narlikar, G.J., and Redding, S. (2021). HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. *Elife* *10*.

Keizer, V.I.P., Grosse-Holz, S., Woringer, M., Zambon, L., Aizel, K., Bongaerts, M., Delille, F., Kolar-Znika, L., Scolari, V.F., Hoffmann, S., *et al.* (2022). Live-cell micromanipulation of a genomic locus reveals interphase chromatin mechanics. *Science* *377*, 489-495.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *J Cell Biol* *153*, 1341-1353.

Kroschwald, S., Maharana, S., and Alberti, S. (2017). Hexanediol: a chemical probe to investigate the material properties of membraneless compartments. *Matters*, 1-7.

Lafontaine, D.L.J., Riback, J.A., Bascetin, R., and Brangwynne, C.P. (2021). The nucleolus as a multiphase liquid condensate. *Nat Rev Mol Cell Biol* *22*, 165-182.

Lancaster, A.K., Nutter-Upham, A., Lindquist, S., and King, O.D. (2014). PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* *30*, 2501-2502.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* *547*, 236-240.

Leonid I. Rudin, S.O., Emad Fatemi (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* *60*, 259-268.

Lin, Y., Currie, S.L., and Rosen, M.K. (2017). Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem* *292*, 19110-19120.

Liu, Z., Tardat, M., Gill, M.E., Royo, H., Thierry, R., Ozonov, E.A., and Peters, A.H. (2020). SUMOylated PRC1 controls histone H3.3 deposition and genome integrity of embryonic heterochromatin. *EMBO J* *39*, e103697.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* *487*, 57-63.

Martin, E.W., Holehouse, A.S., Peran, I., Farag, M., Incicco, J.J., Bremer, A., Grace, C.R., Soranno, A., Pappu, R.V., and Mittag, T. (2020). Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* *367*, 694-699.

Martin, K.L., and Leese, H.J. (1995). Role of glucose in mouse preimplantation embryo development. *Mol Reprod Dev* *40*, 436-443.

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

McDowell, T.L., Gibbons, R.J., Sutherland, H., O'Rourke, D.M., Bickmore, W.A., Pombo, A., Turley, H., Gatter, K., Picketts, D.J., Buckle, V.J., *et al.* (1999). Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc Natl Acad Sci U S A* *96*, 13983-13988.

McSwiggen, D.T., Mir, M., Darzacq, X., and Tjian, R. (2019). Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev* 33, 1619-1634.

Mintz, B. (1964). Synthetic Processes and Early Development in the Mammalian Egg. *J Exp Zool* 157, 85-100.

Miyanari, Y., and Torres-Padilla, M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. *Nature* 483, 470-473.

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.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163, 123-133.

Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. *Hum Mol Genet* 14 *Spec No 1*, R47-58.

Muller, B.K., Zaychikov, E., Brauchle, C., and Lamb, D.C. (2005). Pulsed interleaved excitation. *Biophys J* 89, 3508-3522.

Murthy, A.C., Dignon, G.L., Kan, Y., Zerze, G.H., Parekh, S.H., Mittal, J., and Fawzi, N.L. (2019). Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat Struct Mol Biol* 26, 637-648.

Muzzopappa, F., Hertzog, M., and Erdel, F. (2021). DNA length tunes the fluidity of DNA-based condensates. *Biophys J* 120, 1288-1300.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., *et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57, 936-947.

Novo, C.L., Wong, E.V., Hockings, C., Poudel, C., Sheekey, E., Wiese, M., Okkenhaug, H., Boulton, S.J., Basu, S., Walker, S., *et al.* (2022). Satellite repeat transcripts modulate heterochromatin condensates and safeguard chromosome stability in mouse embryonic stem cells. *Nat Commun* 13, 3525.

Ooga, M., Fulka, H., Hashimoto, S., Suzuki, M.G., and Aoki, F. (2016). Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics* 11, 85-94.

Orlando, G., Raimondi, D., Tabaro, F., Codice, F., Moreau, Y., and Vranken, W.F. (2019). Computational identification of prion-like RNA-binding proteins that form liquid phase-separated condensates. *Bioinformatics* 35, 4617-4623.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* 10, 475-478.

Pak, C.W., Kosno, M., Holehouse, A.S., Padrick, S.B., Mittal, A., Ali, R., Yunus, A.A., Liu, D.R., Pappu, R.V., and Rosen, M.K. (2016). Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol Cell* 63, 72-85.

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., and Knowles, B.B. (2004). Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 7, 597-606.

Phair, R.D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature* 404, 604-609.

Picelli, S., Bjorklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* *10*, 1096-1098.

Plys, A.J., Davis, C.P., Kim, J., Rizki, G., Keenen, M.M., Marr, S.K., and Kingston, R.E. (2019). Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev* *33*, 799-813.

Probst, A.V., Okamoto, I., Casanova, M., El Marjou, F., Le Baccon, P., and 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.

Probst, A.V., Santos, F., Reik, W., Almouzni, G., and Dean, W. (2007). Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* *116*, 403-415.

Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., Kolb, C., Otte, A.P., Koseki, H., Orkin, S.H., *et al.* (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* *40*, 411-420.

Qin, W., Stengl, A., Ugur, E., Leidescher, S., Ryan, J., Cardoso, M.C., and Leonhardt, H. (2021). HP1beta carries an acidic linker domain and requires H3K9me3 for phase separation. *Nucleus* *12*, 44-57.

Rigler, R., Mets, U., Widengren, J., and Kask, P. (1993). Fluorescence Correlation Spectroscopy with High Count Rate and Low-Background - Analysis of Translational Diffusion. *Eur Biophys J Biophys* *22*, 169-175.

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.

Rodriguez-Terrones, D., Hartleben, G., Gaume, X., Eid, A., Guthmann, M., Iturbide, A., and Torres-Padilla, M.E. (2020). A distinct metabolic state arises during the emergence of 2-cell-like cells. *EMBO Rep* *21*, e48354.

Rodriguez-Terrones, D., and Torres-Padilla, M.E. (2018). Nimble and Ready to Mingle: Transposon Outbursts of Early Development. *Trends Genet* *34*, 806-820.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. *Proteins* *42*, 38-48.

Sabari, B.R., Dall'Agnes, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., *et al.* (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science* *361*.

Saksouk, N., Barth, T.K., Ziegler-Birling, C., Olova, N., Nowak, A., Rey, E., Mateos-Langerak, J., Urbach, S., Reik, W., Torres-Padilla, M.E., *et al.* (2014). Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Mol Cell* *56*, 580-594.

Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A.J., and Torres-Padilla, M.E. (2010). Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol* *12*, 853-862.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* *241*, 172-182.

Sawyer, I.A., Sturgill, D., and Dundr, M. (2019). Membraneless nuclear organelles and the search for phases within phases. *Wiley Interdiscip Rev RNA* *10*, e1514.

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.

Schrimpf, W., Barth, A., Hendrix, J., and Lamb, D.C. (2018). PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys J* *114*, 1518-1528.

Shin, Y., Berry, J., Pannucci, N., Haataja, M.P., Toettcher, J.E., and Brangwynne, C.P. (2017). Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* *168*, 159-171 e114.

Simon, J.R., Carroll, N.J., Rubinstein, M., Chilkoti, A., and Lopez, G.P. (2017). Programming molecular self-assembly of intrinsically disordered proteins containing sequences of low complexity. *Nat Chem* *9*, 509-515.

Soldi, M., and Bonaldi, T. (2013). The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components. *Mol Cell Proteomics* *12*, 764-780.

St John, J.C., Facucho-Oliveira, J., Jiang, Y., Kelly, R., and Salah, R. (2010). Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum Reprod Update* *16*, 488-509.

Strickfaden, H., Tolsma, T.O., Sharma, A., Underhill, D.A., Hansen, J.C., and Hendzel, M.J. (2020). Condensed Chromatin Behaves like a Solid on the Mesoscale In Vitro and in Living Cells. *Cell* *183*, 1772-1784 e1713.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. *Nature* *547*, 241-245.

Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. *Nature* *184*, 1286-1287.

Tatavosian, R., Kent, S., Brown, K., Yao, T., Duc, H.N., Huynh, T.N., Zhen, C.Y., Ma, B., Wang, H., and Ren, X. (2019). Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J Biol Chem* *294*, 1451-1463.

Teslaa, T., and Teitell, M.A. (2015). Pluripotent stem cell energy metabolism: an update. *EMBO J* *34*, 138-153.

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

Trimarchi, J.R., Liu, L., Porterfield, D.M., Smith, P.J., and Keefe, D.L. (2000). Oxidative phosphorylation-dependent and -independent oxygen consumption by individual preimplantation mouse embryos. *Biol Reprod* *62*, 1866-1874.

van der Heijden, G.W., Dieker, J.W., Derijck, A.A., Muller, S., Berden, J.H., Braat, D.D., van der Vlag, J., and de Boer, P. (2005). Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* *122*, 1008-1022.

van Mierlo, G., Jansen, J.R.G., Wang, J., Poser, I., van Heeringen, S.J., and Vermeulen, M. (2021). Predicting protein condensate formation using machine learning. *Cell Rep* *34*, 108705.

Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G., *et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* *142*, 967-980.

Vernon, R.M., Chong, P.A., Tsang, B., Kim, T.H., Bah, A., Farber, P., Lin, H., and Forman-Kay, J.D. (2018). Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* 7.

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

Wang, J., Choi, J.M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., *et al.* (2018b). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* 174, 688-699 e616.

Wang, L., Gao, Y., Zheng, X., Liu, C., Dong, S., Li, R., Zhang, G., Wei, Y., Qu, H., Li, Y., *et al.* (2019). Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. *Mol Cell* 76, 646-659 e646.

Wheeler, R.J., and Hyman, A.A. (2018). Controlling compartmentalization by non-membrane-bound organelles. *Philos Trans R Soc Lond B Biol Sci* 373.

Wongtawan, T., Taylor, J.E., Lawson, K.A., Wilmut, I., and Pennings, S. (2011). Histone H4K20me3 and HP1alpha are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J Cell Sci* 124, 1878-1890.

Woodland, H.R., and Graham, C.F. (1969). RNA synthesis during early development of the mouse. *Nature* 221, 327-332.

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.

Xue, Y., Gibbons, R., Yan, Z., Yang, D., McDowell, T.L., Sechi, S., Qin, J., Zhou, S., Higgs, D., and Wang, W. (2003). The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc Natl Acad Sci U S A* 100, 10635-10640.

Zhang, M., Chang, H., Zhang, Y., Yu, J., Wu, L., Ji, W., Chen, J., Liu, B., Lu, J., Liu, Y., *et al.* (2012). Rational design of true monomeric and bright photoactivatable fluorescent proteins. *Nat Methods* 9, 727-729.

ACKNOWLEDGMENTS

I am grateful to the reviewers for taking the time to read my thesis. I hope you were at least 10% as excited as I was with this work! I would also like to thank my TAC committee, Thomas Cremer, Nadine Vastenhouw and Sebastian Bultmann for the great advice and constructive criticism I had during my PhD.

I would like to thank Prof. Dr. Torres-Padilla for her great supervision during my PhD. Jokes aside, you have been incredible with me during these 5 years Maria-Elena! Thank you for accepting me in your lab, thank you for believing in me, thank you for teaching me so much, thank you for the patience you had with me over the years and thank you for always having your door open whenever I needed advice, scientific or not! I feel like the METP lab was the perfect place for me to do my PhD and I have grown up so much scientifically and personally during my stay in the IES.

I started with Maria-Elena but supervising me was actually a 2-person job! Thank you, Adam, for teaching me everything I know and always being there whenever I had a stupid question. You could have told me earlier how to properly say “foci” but apart from that you were a great supervisor! You made me feel comfortable from the start in the lab and I’m very happy to call you my friend instead of supervisor now.

During my PhD I had the chance to collaborate with a lot of great people I would also like to acknowledge. I’m grateful to the Lamb lab for the amazing collaboration and especially to Chen and Irene for being so patient with me when they were trying to explain to me how the RICS worked. Thank you also to Robert and Igor for making the imaging chamber that I used extensively and that I’m sure people will continue using in the lab.

I am also so thankful that I was part of the IRTG! I am grateful to Elizabeth for being always enthusiastic and taking care of me over all these years. The chromatin dynamics community has been incredibly insightful for me and I feel privileged to have done my PhD surrounded by such brilliant people.

I would also like to acknowledge all the professors I have met during my studies and who helped me so much over the years. A special thanks to all the Tyteca lab for making my first real research experience so great and giving me the enthusiasm to continue in this path.

Donatienne and H el ene you were amazing supervisors and I learned so much from both of you.

It is time now to acknowledge the greatest lab in all the world! I never thought these 5 years in Munich would change me so much and it is in big part thanks to all of you. Marga, we haven't had the chance to interact a lot but I wish you all the best for the future and thank you for all the help in this last part of my PhD. Tamas, thank you for correcting me when my statistical background was going bananas and for being there to help when my coding skills desperately needed advice. Andreas, I promise I will never look into a laser! Thank you for all the help and knowledge you gave me during my PhD and thank you for being the best german version of the blues brother! Amelie, thank you for all the coffee breaks and the stories about your rabbits! It is a shame we had to share you with the IFE but your good mood and joyfulness were great to have around. Ken sensei, we joined and left the lab almost at the same time, thank you for paving the way in the METP lab. Tsune sensei, it was great having you next to me in the lab for my first years. I learned so much from you over the years. I hope you will keep my picture for a very long time and that one day you will name a snake "Manuel". Yung-Li, I hope you will use the flask I gave you in a conference and wish you all the best for the future. I have no doubt you will do great in your future career and I hope we will keep in touch! Hiromi, you showed me the greatest places to take a nap in the building and just for that I will be eternally grateful. I wish you all the best for the future and hopefully see you soon. Fede, Marlies and Antoine, the three postdoc musketeers, I am grateful for all the help during our time together in the institute. Fede your constant good mood and great karaoke talent made my last years in the lab so much more enjoyable. Marlies, ON THE OTHER HAND, you were amazing! Thank you for helping correcting this thesis until the very last minute and being patient with my terrible writing. I learned so much from you and hope you will have a spot for me in your future lab. Antoine, or as my phone calls you, Antoine Curie, who knew we would end up in the same lab after the epigenetics course? Thank you for all the thesis corrections, wine, tarot, raclettes, movie recommendations, tennis and badminton. And never forget I am better than you at table tennis! Marion, I'm glad we became better friends over the years. I'm so grateful you were there with me when I changed desks and I will never get tired of our long discussions about figures, being for my paper or yours. I wish you and Jerome all the best for the future and I hope to see you again for a tarot or a raclette night!

Mrinmoy, you were a source of inspiration throughout my PhD! Thanking you for your great cooking skills would do you justice. I loved and learned so much from every moment with you in the lab! You were one of the only persons who came to see me almost daily "on the other side" and we had a great relationship from day one! I hope I was at least half as supportive

and helpful as you were for me. I am sure we will always be friends and you will always be welcomed in the Guthmann institute!

Clara, where to begin? You were the best smoking, skiing, partying and beer buddy in the whole world! Thank you first of all for the kitkats of course, but also for everything else. You were my biggest support and friend during my PhD! Our Frida meetings were one of the things that made my stay in Munich so enjoyable and I hope I was as supportive with you as you were for me. Tu resteras toujours dans mon cœur et je sais que ce n'est que le début d'une amitié à vie. À plus dans le bus!

The METP lab has changed a lot since I started and I also want to acknowledge all the people that I have met on the way. Juana, I think you were the first person after Maria-Elena who I have talked to in the lab. Thank you for all your help of course, but also for always caring about me and asking how my life was. Emily, you left way too early but your ever-present smile and good humour was great to have around at my beginnings. Ane, eskerrik asko for everything you did for me in the lab. It was great to have you around and you were a great source of inspiration for me. Luis, we got to spend a lot of time together and it was amazing to see you grow from that shy bachelor student you were to the great PhD student you are now. See you soon in Paris for a lot of new adventures. Diego and Mate, my two role models when I started my PhD. I think I worked as hard because I wanted to be as good as you when I started. Mate you were a great desk buddy when I started and you taught me so much at the beginning! Diego, the model PhD student, I am glad we got to spend all this time in and outside the lab! I will always remember the times you were leaving the lab at 4am and I was arriving. Camille, queen of the lab, merci pour tout! I hope I was not too much of a pain on a daily basis. You didn't manage to find me a wife but were helpful on everything else! Tash, I still have your postcard next to my bed. As you said, you were the best ski, table tennis, techno, micro-injector and german class friend. I hope the little zygote I was when you joined the lab is making you proud! Last but not least I am grateful to the students I had the opportunity to supervise during my thesis. I learned a lot from supervising you both. I hope you learned something from me too and that you had a good time in the lab! Special thanks to Fiona who was an amazing master student and, I'm sure, a better PhD student now at Yale.

Of course, I also have to thank the rest of the IES institute. First of all, thank you to Stephan, Antonio and Eva for their scientific mentoring and help during my PhD. Then all the other people from these teams like Gabriele, Mayra, Henning, Anna, Matthias, Elisabeth, Meghana, Thomas, Ana and many others over the years! I would then like to acknowledge Dr. Fiorentino, a good friend and source of inspiration for me. I hope you will come for my PhD defence and

if not, I am sure we will see each other soon enough. Last but not least, Elmiracle has followed me during most of my PhD, through good and bad times, through never ending parties and long hours in the lab. You in Berlin and me in Paris, it looks like somehow, we are going forward and I hope to party with you very soon.

This long PhD journey would not have been possible without my amazing roommates being there by my side. Aurore your good mood and energy are contagious and it was great to have you in the flat for the last months of my stay in Munich. Verena, it's a shame you had to leave Munich so early but I am super happy I got to meet you and become your friend. Paul, you little b, mulțumesc for everything! Thank you for being a good roommate and friend. Except that painful night we had to share a bed in Bozen, it was great to have you around in the flat.

Pelin, who would have thought we would become such good friends when we first interviewed you for the flat. You were one of my biggest supports in Munich! We have lived so much stuff together over this short amount of time that I feel we have been friends forever. From skiing to partying in Berlin and watching the weirdest movies ever on MUBI you made my last years in Munich so memorable. You better come and visit me in Paris soon and teşekkürler!

Salma, you are one the kindest persons I know. You probably don't realise all you did for me in Munich over these last years. I wouldn't have survived the tough times of my PhD without you by my side. I would tell you to come and visit me soon in Paris but I'm sure you have already bought some train tickets. On ne se perdra jamais de vue où que nous vivions. Gros bisous !

I would also like to thank all my friends (yes, I have more than one!) who made my stay in Munich an amazing time. So thank you Ekin, Franzi, Sophie, Rima, Boris and Houyem and many others over the years.

¡Muchas gracias a Mati y a todo el grupo de experiencia impro! Las clases del jueves siempre fueron uno de los momentos preferidos de la semana en el que podía sacar toda la frustración cada vez que mis experimentos fallaban. La primera cosa que haré en París es buscarme otro grupo de impro... ¡Aunque ya sé que nunca será tan bueno como experiencia impro!

À tous mes amis de Bruxelles, merci de ne m'avoir jamais oublié après toutes ces années à Munich. Aux «our house in the middle», merci d'être les meilleurs amis du monde ! Vous ne vous rendez pas compte à quel point cette semaine de vacances annuelle en septembre était une bouffée d'oxygène pour moi pendant mon doctorat. Un merci spécial à Virginia, Inigo,

Yassine, Arthur, Matthias et Charles qui sont venus me rendre visite et boire quelques bières à Munich.

Llega ahora el momento de agradecer a toda mi familia por todo su apoyo durante todos estos años lejos de casa. Ioana, muchísimas mulțumesc por todo tu apoyo y consejos estos últimos años. Iulia, todavía no te enteras de mucho pero seguro que algún día serás una gran bióloga. Elvira, eres una parte primordial de mi vida y educación. Sin ti no sería lo que soy y siempre tendrás una parte importante en mi vida. Lo mismo va para el Prof. Dr. Van Eeuwijk. Sin ti nunca hubiera hecho ese máster en estadística que me abrió tantas puertas. Gracias por todos tus preciosos consejos durante mi carrera. Muchísimas gracias a todos mis tíos, tías, primos y prima que siempre me apoyaron durante todos estos años. Michel y Marie, Juan-Pablo y Marielle, Marisa y Rafa y Cheché y Mariola os quiero mucho y espero veros a todos pronto. Y a los parisinos, nos vemos dentro de muy poco. Este doctorado se lo dedico a todos mis abuelos y abuelas. Sin todo su amor y apoyo durante toda mi vida no hubiera llegado adonde estoy ahora. Mari-Luz y Raym, siempre fueron un ejemplo de humanidad, bondad y trabajo para mí. Me hubiera encantado que estuvieran aquí conmigo para festejar este doctorado y sé que hubieran estado muy orgullosos de mí. Lola, lo mismo va para ti. ¡Espero que estés orgullosa de mí estés donde estés y nunca te agradeceré lo suficiente por todo el amor y apoyo! Y para acabar, Pepe, tu continuo amor e interés en mi trabajo fue un gran apoyo y motivación durante toda mi carrera. Saber que seguías lo que hacía y que te interesaba mi investigación siempre contó mucho para mí.

Antonius el enano, con los años cada vez eres menos tonto. Cuando empecé la carrera tenías 8 años y lo único que hacías era intentar pegarme. ¡Cuánto has crecido en estos últimos años hasta empezar la universidad! Yo sigo convencido de que estudias biología porque soy tu ejemplo en la vida. Estoy seguro de que te va a ir genial hagas lo que hagas. ¡Ahora que somos los dos mayores de edad y que voy a estar más cerca de Bélgica, tenemos que quedar más a menudo para tomar una cerveza!

Querida madre, grazas por tu apoyo incondicional, hasta cuando escribía “avecque” en los dictados en la escuela. Aunque haya hecho una carrera completamente diferente, siempre fue una gran inspiración tu trabajo, gran sabiduría y todos los libros que te tragas todas las noches. Siempre me sentí querido y apoyado en todo que quise hacer y este doctorado lo debo en gran parte a ti. ¡Y grazas por ser la mejor mamá del mundo, aunque con el peor carácter de toda Galicia!

Querido padre, sé la suerte que tengo de que me hayas apoyado y dejado hacer lo que quisiera. Siempre seguiste con el máximo interés todo lo que hice dándome valiosos consejos, pero no obligándome nunca a seguirlos. No todo siempre me fue bien en mi carrera, pero siempre me sentí apoyado, hasta cuando hacía boludeces. Como mamá, siempre fuiste una gran inspiración, por la gran pasión y trabajo con los que llevas todos tus proyectos para Haití, los refugiados, etc. ¡Muchas gracias monsieur Claude por ser el mejor y el más tête en l'air padre del mundo!

I didn't expect these acknowledgments to be this long but that shows how lucky I am to have such an amazing family and friends. Thank you all and see you very soon!