The double homeobox transcription factor Duxbl is essential for the exit from totipotency in mice

Inaugural Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

doctor rerum naturalium (Dr. rer. nat.)

Angefertigt am Max-Planck-Institut für Herz- und Lungenforschung, Bad Nauheim

Vorgelegt beim Fachbereich 08 Biologie und Chemie der Justus-Liebig-Universität, Gießen

von

Felipe Francisco Lüttmann

2023

Arbeitstitel: Deciphering the role of the double homeobox transcription factor Duxbl in preimplantation stage development.

Erstgutachterin:

Prof. Dr. Sandra Hake

Institut für Genetik

Justus-Liebig-Universität Gießen

Heinrich-Buff-Ring 58-62 in 35392 Gießen

Zweitgutachterin:

Prof. Dr. Soni Pullamsetti

Abteilung für Entwicklung und Umbau der Lunge

Max-Planck-Planck-Institut für Herz- und Lungenforschung

Parkstraße 1 in 61231 Bad Nauheim

Center for Infection and Genomics of the Lung

Justus-Liebig-Universität Gießen

Aulweg 132 in 35392 Gießen

Declaration

“I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.”

-------------------------------- -------------------------------------

Place, Date Original Signature

Publications related to this dissertation

Parts of this dissertation are published as: Vega-Sendino, M., Lüttmann, F.F., Olbrich, T. et al. The homeobox transcription factor DUXBL controls exit from totipotency. Nat Genet 56, 697–709 (2024).

https://doi.org/10.1038/s41588-024-01692-z

Acknowledgements

Foremost, I would like to express my sincere gratitude to my Ph.D. supervisor Dr. Johnny Kim for the opportunity to conduct this fascinating research, for many thrilling and amusing scientific discussions, and for his constant support.

I would also like to thank Prof. Dr. Mario Looso and Prof. Dr. Dr. Thomas Braun for providing me the opportunity to start the journey of pursuing my Ph.D. in their laboratories, as well as for their feedback and advice.

I express my sincere appreciation to Prof. Dr. Sandra Hake, Prof. Dr. Soni Pullamsetti, Prof. Dr. Lienhard Schmitz and Prof. Dr. Nikola-Michael Prpic-Schäper for their time and willingness to take on the role of my examiners.

Moreover, I would like to thank members of the scientific service groups at the Max Planck Institute for Heart and Lung Research for their technical assistance and expertise. I am thankful to Dr. Carsten Kuenne for his help with analyzing raw RNA- and ChIP-seq data. I much appreciate the technically difficult microinjections performed by Susanne Kreutzer and Daniel Heil. Furthermore, I would also like to thank Dr. Stefan Günther for preparation of sequencing libraries and sequencing of RNA samples. Moreover, I would particularly like to thank the animal care takers for their work.

Many thanks also go to my colleague Dr. Jiasheng Zhong, who generated the *Duxbl*-KO mouse line.

My Ph.D. would not have been the staggering and joyful journey that it was without my dear friends and colleagues. I am very grateful for the amazing time we spent together, and I am looking forward to seeing you all again!

Last, but not least, my deepest gratitude goes to my family, who always supports me unwaveringly and unconditionally. I dedicate this dissertation to them.

Contents

Abstract

The preimplantation stage is the first phase in mammalian development and starts with fusion of egg and sperm. Thereby, fertilization gives rise to a totipotent single cell, called zygote, which holds the capacity to develop into an entire new organism. Within the zygote, the parental chromatin is reprogrammed to form a new embryonic genome. Activation of this genome results in the first transcriptional event in life, called zygotic genome activation (ZGA). In mice, this transcriptional burst occurs at the 2-cell stage and results in transient upregulation of thousands of genes. Immediately afterwards, at the 4-cell stage, those genes become silenced. Even though ZGA was discovered decades ago, its regulation remains poorly understood. Recently, the double homeobox transcription factor Dux was shown to drive a 2-cell-like conversion of pluripotent embryonic stem cells. In these cells, ectopic expression of Dux induces opening of ZGA-associated gene loci and consequent transcription of encoded genes. Thereby, Dux was discovered as a driver of ZGA. However, genetic depletion of Dux results in negligible impairment of ZGA and is compatible with development. Nevertheless, both activation and rapid decommissioning of the 2-cell-specific transcriptional profile are essential for life. However, how this is regulated has remained unknown. In this thesis, I investigated a potential role for another member of the DUX transcription factor family called Duxbl. Transcriptional profiling of developing murine embryos revealed that Duxbl is part of the ZGA-expression panel. In contrast to Dux, Duxbl does not interact with a histone acetyl transferase and does not induce expression of ZGA-associated genes when overexpressed in embryonic stem cells. Instead, Duxbl suppresses Dux-induced upregulation of 2-cell specific transcripts *in vitro*. Mechanistically, conservation of the DNA-binding domains of Duxbl and Dux is associated with occupation of Dux-binding sites by Duxbl within ZGA-promoters. Accordingly, preponed Duxbl expression in mouse embryos suppressed a large portion of ZGA-genes and resulted in a 2-cell stage developmental arrest. Inversely, inactivation of Duxbl led to upregulation of ZGA-associated transcripts and a developmental block at the 2- to 4-cell stage. This study identifies Duxbl as a master regulator in preimplantation development that is necessary for decommissioning the totipotency-associated ZGA-burst and thereby enabling the first cell fate decisions.

Zusammenfassung

Die Präimplantationsstufe ist die erste Phase der Säugetierentwicklung und beginnt mit der Verschmelzung von Ei- und Samenzelle. Dabei entsteht durch die Befruchtung eine totipotente einzelne Zelle, die Zygote, die die Fähigkeit besitzt, sich zu einem vollständigen neuen Organismus zu entwickeln. Innerhalb der Zygote wird das parentale Chromatin reprogrammiert, um ein neues embryonales Genom zu bilden. Die Aktivierung dieses Genoms führt zum ersten transkriptionellen Ereignis des Lebens, der sogenannten zygotischen Genomaktivierung (ZGA). Bei Mäusen findet dieser Transkriptionsschub im 2-Zell-Stadium statt und führt zu einer vorübergehenden Hochregulierung von Tausenden von Genen. Unmittelbar danach, im 4-Zell-Stadium, wird die Expression dieser Gene wieder abgeschaltet. Obwohl ZGA vor Jahrzehnten entdeckt wurde, ist das Verständnis von dessen Regulation noch immer unvollständig. Kürzlich wurde gezeigt, dass der Doppel-Homeobox-Transkriptionsfaktor Dux eine 2-Zell-ähnliche Konversion von pluripotenten embryonalen Stammzellen induziert. In diesen Zellen führt ektopische Dux-Expression zur Öffnung von ZGA-assoziierten Genloci und zur anschließenden Transkription der codierten Gene. Somit wurde Dux als Aktivator von ZGA entdeckt. Die genetische Deletion von Dux führt jedoch nur zu einer marginalen Beeinträchtigung von ZGA und ist mit normaler Entwicklung vereinbar. Dennoch sind beide die Aktivierung und die rasche Deaktivierung des 2-Zell-spezifischen Transkriptionsprofils lebensnotwendig. Wie dies reguliert wird, war jedoch bisher unbekannt. In dieser Arbeit untersuchte ich eine mögliche Rolle eines anderen Mitglieds der DUX-Transkriptionsfaktor-Familie, genannt Duxbl. Die Untersuchung von Transkriptionsprofilen von sich entwickelnden Mäuseembryonen ergab, dass Duxbl Teil des ZGA-Expressionsprogramms ist. Im Gegensatz zu Dux interagiert Duxbl nicht mit einer Histonacetyltransferase und induziert bei Überexpression in embryonalen Stammzellen keine Expression von ZGA-assoziierten Genen. Stattdessen unterdrückt Duxbl die durch Dux induzierte Hochregulierung von 2-Zell-spezifischen Transkripten *in vitro*. Mechanistisch gesehen ist die Konservierung der DNA-Bindedomänen von Duxbl und Dux mit der Besetzung von Dux-Bindungsstellen durch Duxbl innerhalb von ZGA-Promotoren assoziiert. Dementsprechend unterdrückte die vorzeitige Expression von Duxbl in Mausembryonen einen Großteil der ZGA-Gene und führte zu einem Entwicklungsarrest im 2-Zell-Stadium. Umgekehrt resultierte die Inaktivierung von

Duxbl in einer Hochregulierung von ZGA-assoziierten Transkripten und einem Entwicklungsstopp im 2- bis 4-Zell-Stadium. In dieser Studie wird Duxbl als Schlüsselfaktor in der Präimplantationsentwicklung identifiziert, der die Expression von mit Totipotenz-assoziierten ZGA-Genen abschaltet und dadurch die ersten Entscheidungen über das Zellschicksal ermöglicht.

1 Introduction

1.1 Preimplantation stage development

Mammalian life begins with the fusion of oocyte and sperm. These two unipotent gametes thereby give rise to the zygote, a cell that fulfills the strictest definition of totipotency: to hold the potential to develop into an entire new organism[1]. This journey begins at the site of fertilization, the ampulla, a pouch-like structure in the oviduct, which connects ovary and uterus[2-4] (Figure 1). As the embryo starts to divide and to develop within its maternally provided coat, called the zona pellucida, it migrates through the oviduct towards the uterus (Figure 1)[4,5]. When it reaches the uterus, the embryo hatches from its egg coat and implants into the maternal tissue (Figure 1)[4,5].

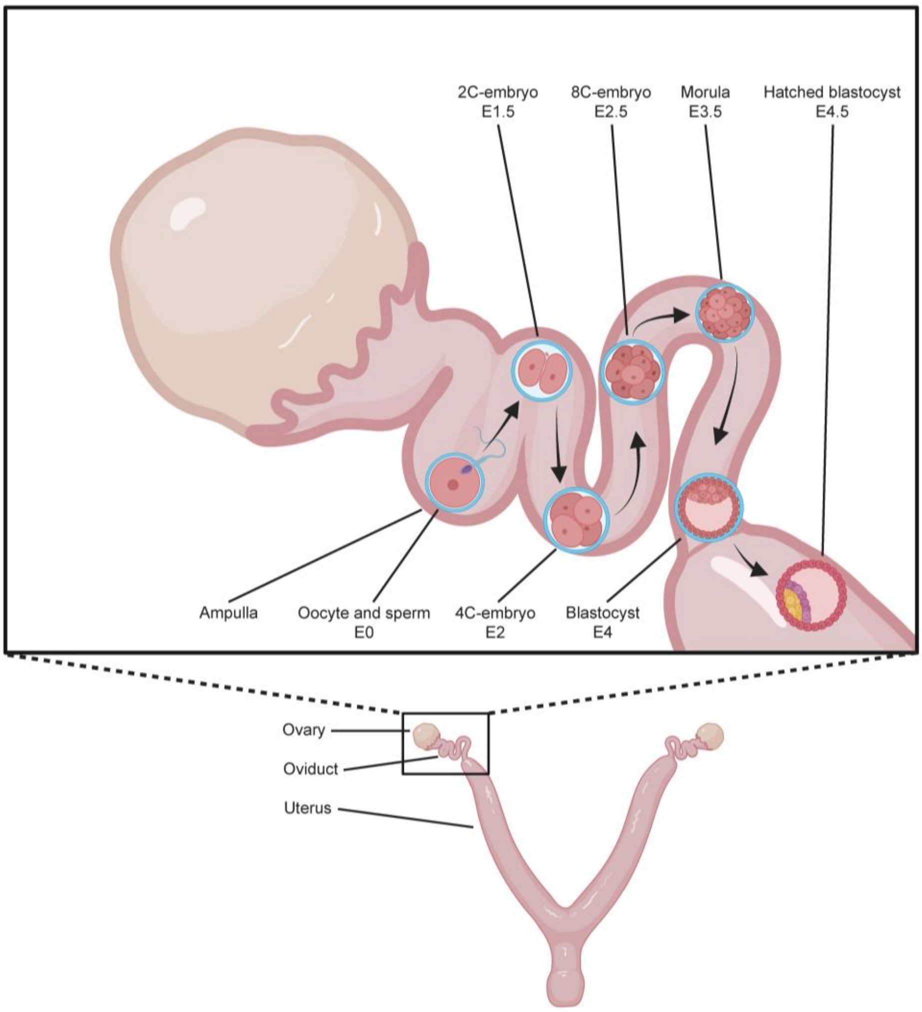

Figure 1: Preimplantation development in mice:

Schematic illustration depicting the migration of the developing murine embryo at different developmental stages during preimplantation development through the oviduct[2-6]. (Created with BioRender.com)

This first chapter of embryonic development, named preimplantation stage, is characterized by a sequence of tightly regulated and interlocked molecular events: the maternal-to-zygotic transition, activation of the embryonic genome, and the first cell lineage decisions[7,8]. These processes are coherent with and dependent on rapid changes at the genetic, epigenetic and transcriptional levels[9,10].

1.2 The maternal-to-zygotic transition (MZT)

Sexual reproduction is considered as one of the greatest accomplishments of evolution because it creates offspring with new genetic traits and potential benefits towards the environment[11]. This is achieved by recombining a new diploid genome from two haploid genetic sets[12]. This process is mainly executed by a maternally provided machinery that was previously stored in the oocyte[8]. Contribution to embryogenesis of most paternal components that are packaged into the sperm is rather confined[13]. About 40 % of the coding genome is stockpiled in the oocyte in form of mRNA transcripts[14], which are stabilized and translationally repressed by RNA-binding proteins during oocyte maturation[15-17]. Upon fertilization, maternal transcripts become translated by a polyadenylation-dependent mechanism[18-20]. For these reasons, the very first phase of embryonic development is referred to as the maternal-to-zygotic transition (MZT)[8,13]. During MZT, the parental chromatin undergoes massive epigenetic reprogramming, which is required to deploy the totipotent potential of the zygote[9,10]. This transformative event includes changes in the methylation status of the DNA, incorporation and modification of histones, as well as reshaping of the 3D-structure and accessibility of chromatin[21,22].

1.2.1 MZT - Emergence of a new genome by higher chromatin order formation

Eukaryotic chromatin is spatially organized and folded in the nucleus in a hierarchical manner (Figure 2)[23]. In order, so-called chromosome territories are further subdivided into two types of compartments, which can comprise several millions of bases[24,25]. Compartments A are defined as euchromatic regions where accessible, loosely packed chromatin is associated with transcriptional activity[26]. In contrast, compartments B are heterochromatic, generally more densely packed, and associated with a transcriptionally repressed state[26]. Both compartments A and B are further organized into topology associated domains (TADs)[27]. Within these domains, DNA loop formation enables physical contact between genetic elements that are, on a linear level, thousands of base pairs apart from each other[28,29]. Thereby, these interactions have been shown to be crucial for regulation of transcription[28]. Moreover, they can be detected by high-resolution chromosome conformation capture (Hi-C) and visualized by heat maps (Figure 2)[25]. As has been revealed by Hi-C, immature oocytes and sperm cells display highly compacted chromatin with clear compartmentalization and defined TADs[30-35]. However, fully matured oocytes prior to fertilization and sperm nuclei shortly after fertilization undergo chromatin decompaction[31,34]. Thus, in the zygote, maternal and paternal pronuclei display only weak TADs and low levels of compartmentalization (Figure 3)[32], indicating an erasure of the parental chromatin order. This enables the formation of a new embryonic genome structure as compartmentalization increases again and TADs become more defined while the embryo continues to develop (Figure 3)[34]. These changes in higher chromatin order are generally thought to be driven by dynamically changing phase separation-based interactions of DNA-associated proteins, such as histones[36].

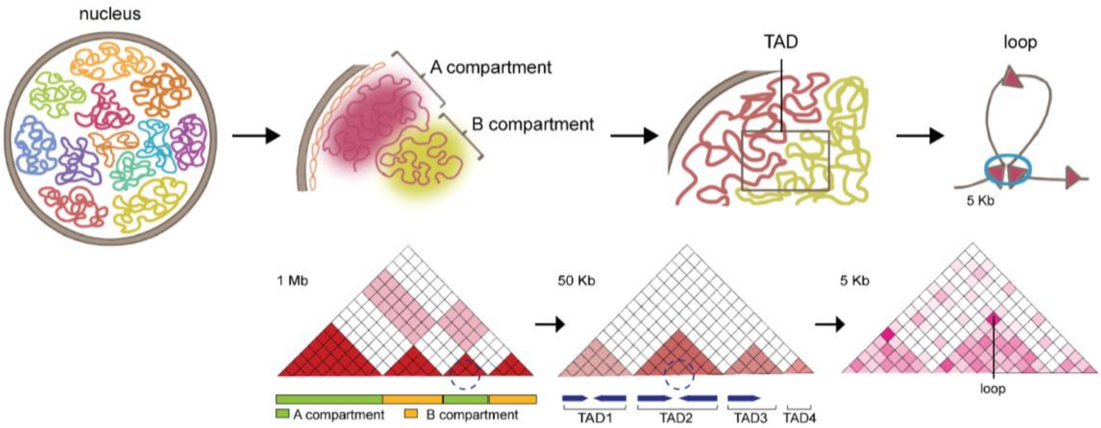

Figure 2: Hierarchical organization of higher chromatin order:

Schematic illustration showing organization of chromatin in chromosome territories, A/B compartments, topology associated domains (TADs) and loops. Generic heat maps visualizing interactions of DNA regions. Adapted from[37].

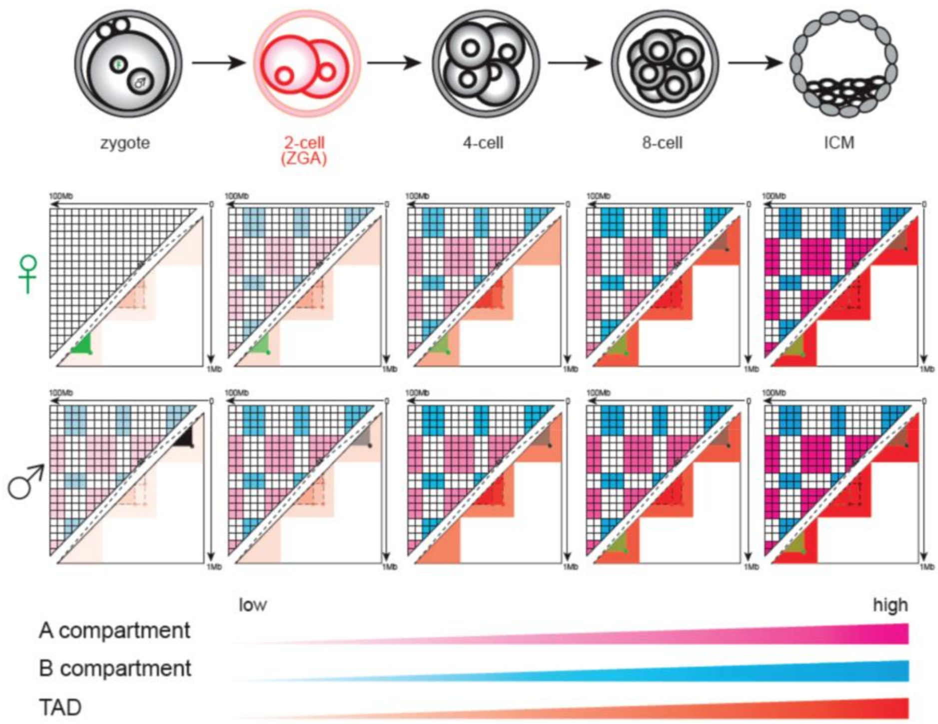

Figure 3: Higher chromatin order in murine preimplantation development:

Schematic illustrating increasing higher chromatin order throughout murine preimplantation development in form of compartments A and B as well as TADs in both maternal- and paternal-inherited chromatin. Adapted from[38].

1.2.2 MZT- Emergence of a new genome – histones and their modifications

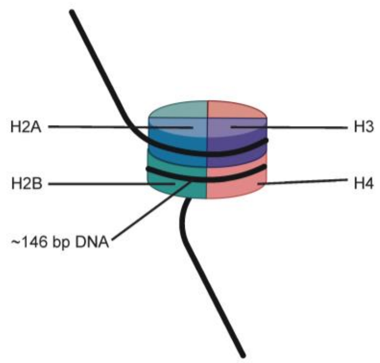

Figure 4: Histone octamer as the basal unit of chromatin organization:

Simplified schematic illustration of a histone octamer[39]. (Created with BioRender.com)

The histones H2A, H2B, H3 and H4 form octamers, which are wrapped by approximately 146 base pairs of DNA (Figure 4)[39]. The resulting nucleosomes provide the basic unit of chromatin organization in most cells[40]. Interestingly, during spermatogenesis, histones are replaced by protamines, enabling a high degree of chromatin compaction[41]. Upon fertilization, protamines are then exchanged again with maternally-derived histones (Figure 5)[42].

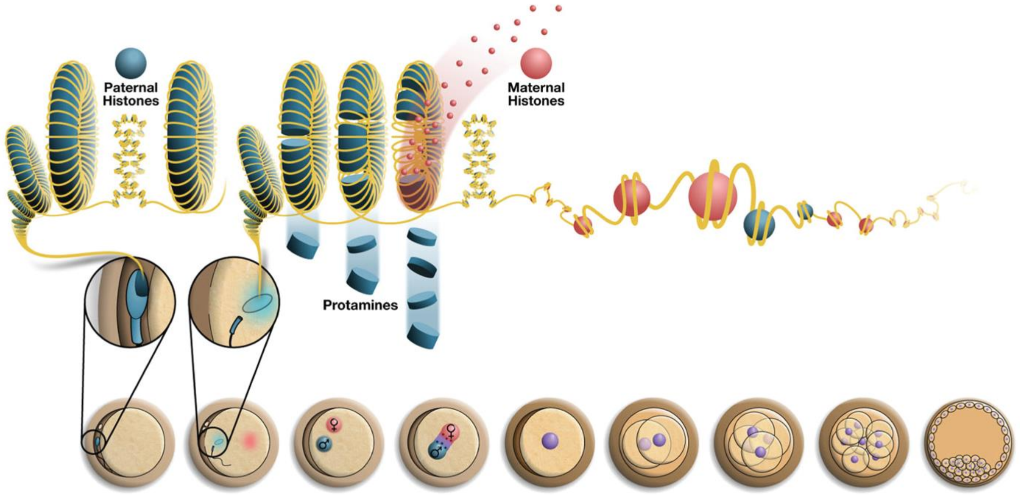

Figure 5: Protamine-histone exchange in paternal chromatin:

Schematic illustration of the changes in packaging of the paternal chromatin upon fertilization. Paternal protamines as the basic unit of chromatin packaging are exchanged by maternal histones in the zygote[42]. Adapted from[43].

The maternal histones display a particular composition, comprising an unusually high proportion of the histone variants H3.3, H2A.X and TH2A[44,45]. Since these variants bind DNA more loosely than other variants, it is generally accepted that they substantially contribute to decompaction of the chromatin[46-48]. Furthermore, the reshaping of the paternal genome by the protamine-histone exchange leads to an erasure of many histone modifications[42,49]. This is of particular importance since these posttranslational modifications (PTMs) critically influence the packing status and hence chromatin accessibility[50]. Thereby, an open, euchromatic state facilitates binding of transcription factors (TFs) to promoters and enhancers (Figure 6a)[51]. In contrast, densely packaged heterochromatin impedes accessibility of TFs to their targets (Figure 6b)[52,53]. Thereby, expression of the corresponding genes is regulated by nucleosomes and the affiliated PTMs[54]. Most of these PTMs are deposited at the N-terminally protruding tails of histones and include, among others, phosphorylation, methylation and acetylation[50]. These modifications alter the transcriptional status of associated chromatin, either directly or indirectly[55]. For example, acetylation of lysine residues neutralizes the positive charge of the amino acid, resulting in a less tight interaction of the histone core with the negatively charged DNA[56,57]. Different PTMs can also serve as recognition signals for the transcriptional machinery and chromatin remodeling complexes, which further alter the epigenetic landscape[55]. The recruitment of specific epigenetic modifiers, as well as the subsequent effect on gene expression, is highly dependent on different types of PTMs and on where they occur[58].

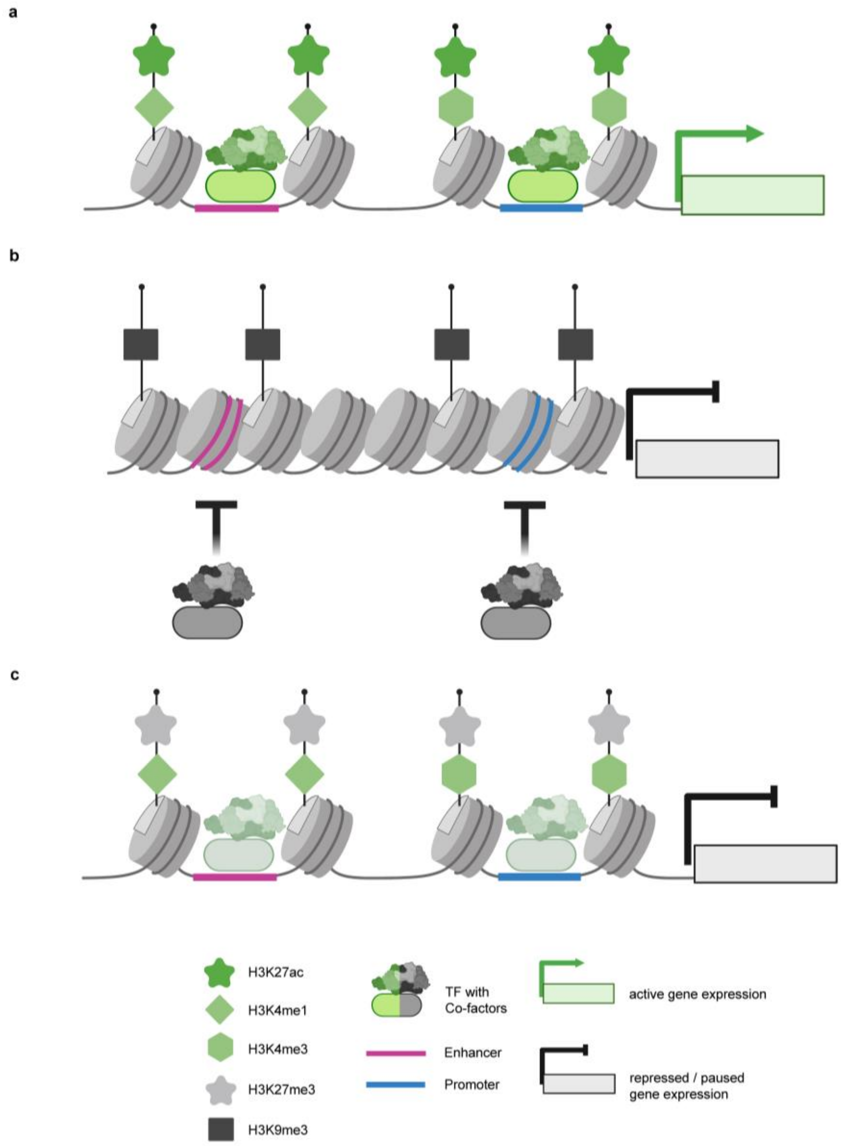

Figure 6: Posttranslational modifications (PTMs) of histones modulate packaging status of the chromatin and transcriptional activity:

Schematic illustration of a model for transcriptional regulation by PTMs of histones: **(a)** Euchromatic regions are associated with active histone PTMs, accessible promoters and enhancers as well as active transcription[51,59-64]. **(b)** Constitutive heterochromatin is marked by repressive PTMs and omits binding of TFs and transcription[52,65,66]. **(c)** Facultative heterochromatin contains bivalent domains, which exhibit active and repressive PTMs, while the transcriptional machinery is in a paused state[64,67-72]. (Created with BioRender.com)

Methylation of histone H3 at lysine 4 (H3K4me) is found at open chromatin and is associated with TF-occupancy (Figure 6a)[59]. At those binding sites, the number of methyl groups varies between different types of regulatory elements. In proximity to enhancers, nucleosomes are mainly marked by one methyl group (H3K4me1) (Figure 6a), while active promoters are flanked by histones displaying mostly three methyl groups (H3K4me3) (Figure 6a)[59]. H3K4me3 has been shown to enable recruitment of the preinitiation complex and transcriptional elongation, being strongly correlated with RNA synthesis[60,61]. Interestingly, transcriptional silencing in mature oocytes is affiliated with formation of broad distal H3K4me3 domains in the maternal chromatin[73-75], which can also be observed in the paternal pronucleus shortly after fertilization[74]. It is assumed that non-canonical H3K4me3 domains sequester the transcriptional machinery away from promoter sites in the zygotic genome[74].

Chromatin in the zygote also displays broad non-canonical domains for acetylation of lysine 27 at histone H3 (H3K27ac)[76]. This PTM gets deployed by the histone acetyl transferases (HATs) CREB-binding protein (CBP) and EP300, which are recruited by TFs to active enhancers (Figure 6a)[59,62,63,77]. At these sites, H3K27ac associates with ATP-dependent chromatin remodelers[64], which loosen DNA-histone interactions and thereby enable opening of the chromatin[78]. This euchromatic state enables recruitment of the transcription initiation complex and facilitates release of RNA polymerase II (RNA Pol II) from a transcriptionally inactive yet paused state[79]. Indeed, H3K27ac is highly correlated with transcriptional activation and distinguishes active enhancers from poised enhancers[64]. Poised enhancers are primed for transcription, displaying a similar degree of loosened chromatin as well as CBP/EP300 occupancy[64]. Importantly, poised enhancers do not enable transcription and, instead of acetylation, display tri-methylation of lysine 27 at histone H3 (H3K27me3) (Figure 6c)[64].

H3K27me3 deposition is mediated by the Polycomb Repressive Complex 2 and is associated with transcriptional silencing of developmental genes[69,80,81]. These genes often control transcriptional programs and are responsible for differentiation and lineage specification[69,82]. Therefore, both preservation and redirection of cell identity require tight control of transcription of developmental genes themselves. This highly depends on compaction of lineage-specific enhancers and promoters[69,80,81]. These regions are termed facultative heterochromatin, which can inhibit transcription[67] but remains accessible to TFs and RNA polymerases (Figure 6c)[68]. As a result, facultative heterochromatin plays a crucial role in development as a transcriptional switch and can display both active and repressive PTMs, such as H3K27me3 and H3K4me3 (Figure 6c)[69,70]. Interestingly, these bivalent domains are not observed during the initial cleavages of preimplantation development[75]. Instead, embryos display a negative correlation between H3K27me3 and H3K4me3[75]. While H3K27me3 is linked to transcriptional silencing in maturing oocytes, it is removed from promoter sites upon fertilization[83]. This indicates an eradication of the parental regulatory landscape during the reprogramming of the embryonic genome to totipotency[9].

Acquisition of totipotency further requires removal of an additional epigenetic barrier that imparts lineage identity, specifically tri-methylation of lysine 9 at histone H3 (H3K9me3)[65,84]. H3K9m3 is considered a repressive histone mark (Figure 6b), typically found at facultative heterochromatin and depleted at promoter sites shortly after fertilization[65,85]. Interestingly, H3K9me3 is also removed from long terminal repeats (LTRs)[85]. This phenomenon is particularly noteworthy during preimplantation development because LTRs are part of repetitive transposable elements (TEs)[86]. TEs, being self-replicating elements, pose a threat to genomic stability and are usually sequestered in constitutive heterochromatin[66,87]. In contrast to facultative heterochromatin, these domains are silenced across different lineages later in development[66]. Here, H3K9me3 typically omits TF binding and recruits the Heterochromatin protein 1 (HP1)[53,71,72], which in turn orchestrates a repressive machinery[88], including DNA methyltransferase 1 (DNMT1)[89].

1.2.3 MZT- Emergence of a new genome – DNA methylation

Methylation of cytosine (5mC) of cytosine-guanine dinucleotides (CpGs) is an epigenetic mark that is generally associated with transcriptional silencing, as it negatively correlates with H3K4me3, H3K4me1, the presence of RNA Pol II and CBP/EP300 occupancy[90]. Concordantly, genomes of the transcriptionally inactive germ cells display high levels of 5mC prior to fertilization[91]. Following fertilization, this epigenetic modification is removed by both active and passive mechanisms. The active removal of 5mC is mediated by Ten-eleven translocation (Tet) methylcytosine dioxygenases, while the passive removal of 5mC appears to occur in a DNA replication-dependent manner[92,93]. Interestingly, while removal of 5mC displays different dynamics in the two parental genomes, global 5mC levels remain consistently low on both parental alleles throughout preimplantation development[94-96].

1.3 Zygotic genome activation

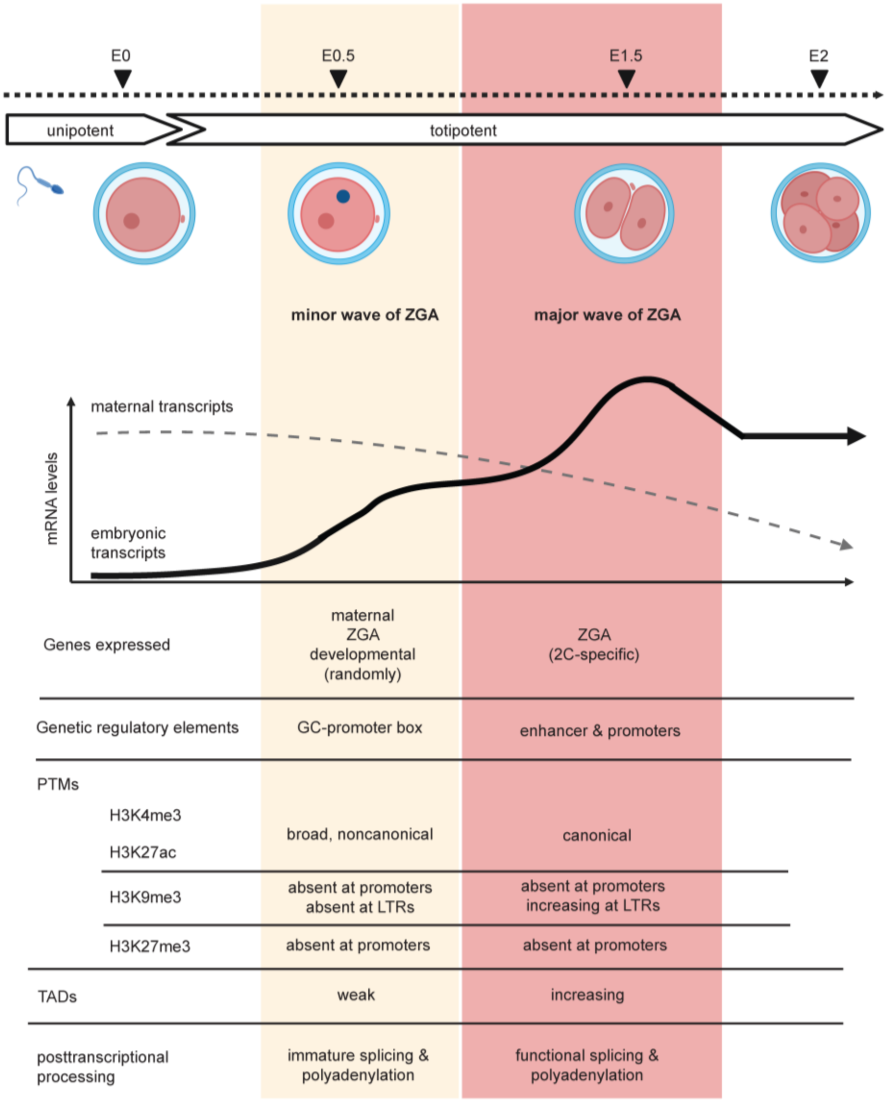

Figure 7: The minor and major waves of ZGA:

Illustration displaying the transcriptional dynamics and underlying epigenetic as well as posttranscriptional factors during the minor and major waves of zygotic genome activation[8,32,34,73,75,76,83,85,97-104]. (Created with BioRender.com)

1.3.1 Zygotic genome activation – the minor wave

During the epigenetic resetting of the parental genomes, early zygotes display no genomic occupancy of RNA Pol II[97]. The concomitant transcriptional silence of the embryo leaves the developmental program initially under maternal control[105,106]. As the maternal-to-zygotic transition progresses, maternal transcripts are gradually degraded[8]. Concomitantly, the epigenomes of the parental pronuclei undergo reprogramming, resulting in the emergence of a genome that is totipotent[21,22]. Following this epigenetic reshaping at the late 1-cell (1C) stage, this newly formed genome exhibits widespread active epigenetic marks, while repressive features are largely immature or absent[75,76,85,107]. Concurrently, the open, “primordial” chromatin landscape intersects with reengagement of RNA Pol II at GC-rich promoter sites[97]. During the S- to G2-phase of the first cell cycle[108], this results in genome-wide transcription of intra- and intergenic regions at low levels[98], known as the minor wave of zygotic genome activation (ZGA) (Figure 7). This promiscuous transcriptional event includes expression of LTR retrotransposons[109]. In mice, this comprises the murine endogenous retrovirus with leucine tRNA primer (MERVL)[110]. MERVL expression strongly correlates with the open chromatin state and is a unique entity of the totipotent stage of preimplantation development[111,112].

During this phase, the MERVL LTRs are believed to function as alternative promoters, driving the expression of genes associated with totipotency[112]. Remarkably, the expression of these genes does not require enhancers, as injected reporters containing only a GC-promoter box can still be successfully transcribed in late zygotes[99]. However, such reporter genes displayed a substantial delay in translation[113]. This observation has been corroborated by next-generation sequencing of endogenous transcripts from late 1C-stage embryos, revealing immature splicing and polyadenylation[98]. The resulting partition of transcription and translation is thought to protect the embryo against misexpression of developmentally harmful genes[114]. At the same time, this deferred posttranscriptional processing might favor the expression of short and intronless genes[114]. Yet, it remains enigmatic whether some of these genes play an essential role in development. Nevertheless, the minor wave of ZGA itself has been shown to be pivotal for progression beyond the 2-cell (2C) stage and for pre-configuration of RNA Pol II for the major wave of ZGA[97,115].

1.3.2 Zygotic genome activation – the major wave

The major wave of ZGA, immediately following minor ZGA, is a critical milestone during preimplantation stage that signifies the transition of developmental control from the maternal machinery to the embryonic genome (Figure 7)[8]. This momentous transcriptional burst occurring at the 2C-stage in mice involves the upregulation of a substantial number of genes, estimated to range from 4,000 to 10,000 genes[14,100-103]. Despite the sheer size of this transcriptional event, gene expression during the major wave of ZGA follows a more specific pattern compared to the minor wave of ZGA[116]. In fact, certain genes are temporarily expressed at high levels during major wave of ZGA and have been linked to specific functions. For instance, in mice, the zinc finger protein Zscan4, encoded by a cluster consisting of six-paralogs (*Zscan4a-f*), is exclusively expressed at high levels during the 2C-stage[117]. At this time, when transcriptional stress arising from ZGA threatens genomic integrity, Zscan4 has been shown to prevent DNA damage[118]. Other genes expressed during major ZGA are mainly associated with cell cycle, metabolism and transcription[103]. This directed expression of specific genes during major ZGA contrasts with the seemingly random gene expression during minor ZGA and is mirrored by incremental changes in the RNA Pol II occupancy profile[97]. At late 1C-stage, RNA Pol II binds equally to oocyte-specific genes, ZGA-specific genes and genes associated with later developmental stages[97]. In contrast, at early 2C-stage, an enrichment of RNA PolI binding at ZGA-genes is observed, which further increases at late 2C-stage[97]. A possible explanation for this more directed recruitment of the transcriptional machinery is a shift towards enhancer-dependent transcription after the second round of DNA replication[104]. Along this line, chromatin undergoes a switch to an intermediate state that is less accessible and permissive compared to minor ZGA[111]. Coincidentally, the histone variant H3.3 is replaced by the variants H3.1 and H3.2, potentially creating a more restrictive chromatin structure in a DNA replication-dependent manner[119]. Furthermore, ZGA also contributes to reshaping of the epigenetic landscape, as for example with lysine demethylases KDM5A and KDM5B being expressed at high levels during this stage[73]. These epigenetic modifiers decrease the width of the broad H3K4me3 domains[73], while histone deacetylases (HDACs) reduce non-canonical H3K27ac[76]. As a result, the broad active histone PTMs are remodeled to a canonical pattern[73,76]. Along this line, at the 4-cell- (4C) stage, deposition of H3K9me3 at LTRs, including those of *MERVL,* takes place and is accompanied by the transcriptional silencing of these TEs[85]. Indeed, the emergence of a more restrictive epigenetic landscape is thought to be imperative for lineage decision and toti- to pluripotent transition[9].

1.4 The transition to pluripotency and blastocyst formation

During the 4C- to 8-cell- (8C) stage, a substantial change takes place in the developmental capacity of the individual cells within an embryo that are called blastomeres[120,121]. Before this stage, each of the two blastomeres of a 2C-embryo still has the remarkable ability to generate an entire mouse when the sister blastomere is removed[122]. However, the scenario shifts at the 4- and 8C-stages, as single blastomeres are only capable of contributing to embryonic and extraembryonic lineages when they are transplanted into another embryo[123-125]. This implies that individual cells at these stages cannot completely support development on their own anymore[120,121].

The transition to pluripotency simultaneously marks the first cell fate commitment towards the trophectoderm, which encapsulates the embryo and later contributes to formation of the placenta upon implantation, and the inner cell mass (ICM)[7]. The ICM further differentiates into primitive endoderm, an additional epithelial-like layer of cells, and the pluripotent epiblast, which ultimately gives rise to all three germ layers of the embryo but not to extraembryonic tissue[7]. This segregation into distinct cellular identities is an incremental process that starts at the 8C-stage when the embryo undergoes compaction, resulting in an increase in cell-cell-contacts between the individual blastomeres and formation of the morula[126]. Consequently, as further cleavage occurs, cells become either encapsulated within the embryo or are facing the outside (Figure 8 a)[7]. This process establishes an apical-basolateral polarity for the individual cells[127]. This cell polarity is believed to drive the segregation into ICM and trophectoderm at 16-Cell- (16C) to 32-Cell- (32C) stage[128,129]. Concomitantly, cells of trophectoderm and ICM lose their potential to contribute to the other cell lineage at 32C- or 64-Cell- (64C) stage, respectively[130]. Additionally, at the 16C- to 32C-stage, an influx of water through the trophectoderm starts to occur, leading to cavitation[131]. This essentially results in formation of the blastocyst, which comprises the epiblast at one pole of the embryo, is delimited to the cavity by the primitive endoderm, and encapsulated as a whole by the trophectoderm (Figure 8b)[7]. Interestingly, the emergence of this structure coincides with reestablishment of DNA methylation[94,95]. Finally, the hatching of the blastocyst from the zona pellucida prior to implantation at embryonic day 4.5 marks the end of murine preimplantation development (Figure 8b)[7].

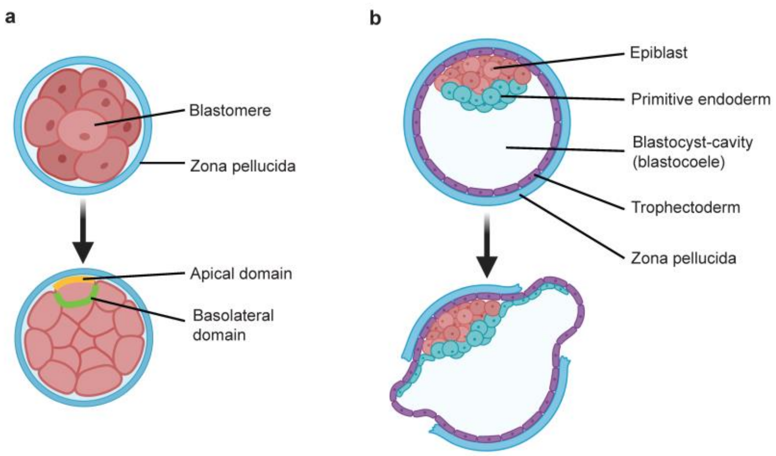

Figure 8: Morphological changes during the pluripotent phase of preimplantation development:

(a) Schematic illustration of morula compaction and establishment of an apical-basolateral polarity[7,126,127]. **(b)** Schematic illustration of an embryo at late blastocyst stage, hatching from the zona pellucida[7]. (Created with BioRender.com)

1.5 Transcriptional regulation of totipotency

The epigenetic and transcriptional processes underlying ZGA and totipotency remain enigmatic, despite extensive research over the past decades[132]. This challenge primarily arises from the technical difficulties associated with studying these processes during the early stages of preimplantation development, where the limited availability of material poses a significant obstacle. Even though sequencing applications for whole transcriptome profiling (RNA-seq) have become well-established even at single cell resolution in recent years[102,133], detection of DNA binding events for proteins still requires a much larger number of cells[134]. For example, only ten years ago, robust and reliable mapping of histone PTMs from standard chromatin immunoprecipitation coupled with high-throughput DNA sequencing (ChIP-seq) required at least 100,000 cells[134]. Even more recently, development of a customized microscale µChIP-seq protocol by Dahl and colleagues was shown to still require roughly 500 embryos to profile H3K4me3 occupancy[73]. Along this line, the small-scale Tn5-assisted chromatin cleavage with sequencing (Stacc–seq) technique recently developed by Liu and colleagues still demands at least 500 embryonic stem cells (ESCs) to profile RNA Pol II dynamics[97].

Transcription factors, which play a crucial role in regulating transcriptional programs in different developmental processes[135], typically require much larger cell numbers for ChIP-seq experiments due to their lower abundance compared to histone PTMs and RNA Pol II[136]. This poses a substantial challenge in understanding the complex regulation of gene expression by TFs specific to preimplantation, specifically during ZGA[21]. To address this, pluripotent ESCs derived from the ICM have often been used as a model to study TF occupancy in the context of preimplantation stage development. In this context, Double homeobox transcription factors (DUX) were identified as potential candidates involved in ZGA (Figure 9a)[137].

Hendrickson and colleagues detected temporary expression of murine Dux and human DUX4 at the onset of the major wave of ZGA in mice and humans, respectively[137]. They further investigated a potential involvement of these transcription factors by conducting ChIP-seq after Dux and DUX4 overexpression in ESCs[137]. This revealed that Dux and DUX4 bind to the promoter sites of major ZGA-specific genes, as well as to the LTRs of *MERVL* in mice and *human endogenous retrovirus with leucine tRNA primer (HERVL)* in humans, respectively (Figure 9b)[137]. Remarkably, this binding elicited strong expression of these TEs and many other ZGA-specific transcripts, which are normally absent in pluripotent ESCs (Figure 9b,c)[137].

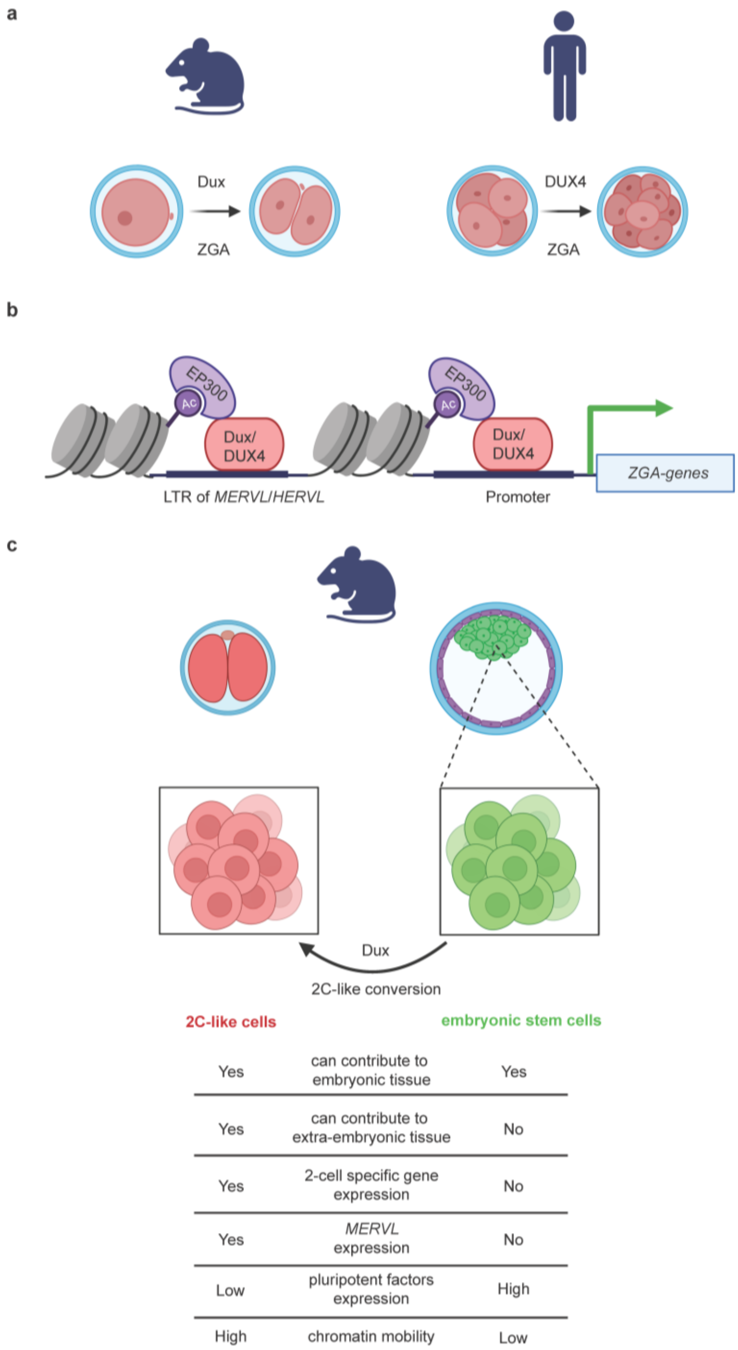

Figure 9: Dux and DUX4 as drivers of ZGA-gene expression in human and mice:

(a) Schematic illustrating the mechanism of Dux/DUX4-mediated expression of ZGA-genes[137-139]. **(b)** Dux and DUX4 are drivers of ZGA in mice and humans[137]. **(c)** Dux can drive a 2C-like conversion in murine embryonic stem cells[137]. Adapted from[140]. Ac = Acetylation of histone H3 at lysine 27. (Created with BioRender.com)

Moreover, these “2C-like” murine ESCs (mESCs) overexpressing Dux displayed downregulation of the pluripotency marker Pou5f1 (better known as Oct4) and exhibited an open chromatin landscape similar to 2C-embryos (Figure 9c)[137]. Although, it had been previously reported that a rare subpopulation of mESCs can exhibit 2C-associated features, this only occurs at very low ratios, and the specific transcriptional regulators responsible for eliciting this 2C-like state remained unknown[112]. Indeed, Dux and DUX4 were identified as the first activators of murine and human ZGA *in vitro*, as they were sufficient to drive a substantial portion of ESCs into a ZGA-like state[137]. Interestingly, subsequent *in vivo* studies using knockout (KO) mice depleted for the Dux encoding array surprisingly gave rise to normally developing, fertile offspring[141]. Moreover, *Dux*-KO-embryos displayed only minor alterations in ZGA[141]. Therefore, it was hypothesized that loss of Dux might be compensated by other TFs to ensure ZGA[141]. Such compensation is often established by functional redundancy of TFs within one family[142]. Indeed, the DUX family appears to comprise additional paralogs of Dux and DUX4 in both mice and humans (Figure 10)[143].

1.6 DUX transcription factors

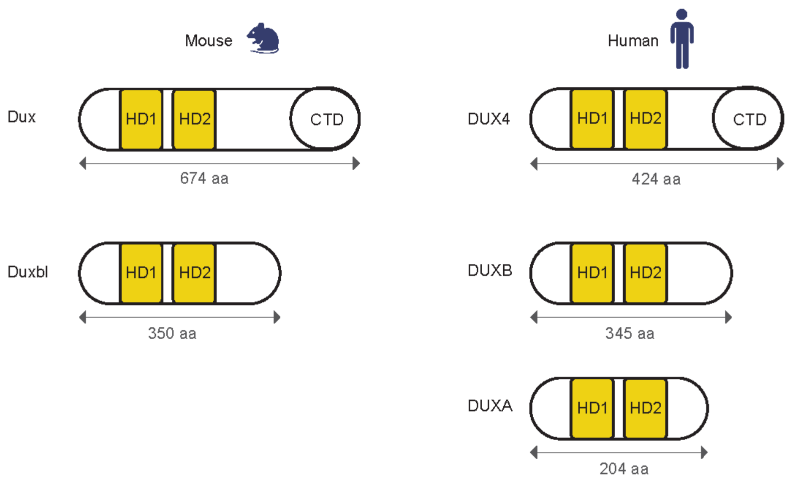

Figure 10: DUX-factors in mice and humans:

Schematic illustration of the murine and human DUX transcription factors[143]. HD = Homeobox domain. CTD = C-terminal domain. Information about protein length retrieved from UniProt[144]. (Created with BioRender.com)

Notably, DUX transcription factors were first discovered in the context of a human muscular dystrophy[145]. Remarkably, this disease is characterized by death of muscle fibers accompanied by expression of genes that are driven by DUX4 and normally only expressed during ZGA[146]. The ability of DUX4 to elicit both cell death and transcriptional activation has been linked to its C-terminal domain (CTD)[138]. Mechanistically, the CTD of DUX4 interacts with the histone acetyltransferases CBP and EP300 and thereby induces global acetylation of H3K27 (Figure 9a)[138]. While this transactivation ability is conserved between DUX4 and Dux, a C-terminal transactivation domain has not been identified in the other DUX family members, namely DUXA and DUXB in humans and Duxbl in mice (Figure 10)[147]. However, all DUX factors comprise two homeobox domains in their N-termini (Figure 10)[143]. These domains, which mediate interaction with DNA[148], exhibit substantial conservation across species in their predicted DNA-binding residues, thereby suggesting that DUXA, DUXB and Duxbl might also act as TFs[147].

Duxbl expression has been detected in brain, eyes and the reproduction organs[149]. Functionally, Duxbl has been linked to development of muscle and immune cells, in particular T-cells[150,151]. In the immune system, Duxbl has been shown to mediate the erasure of cells that fail proper rearrangement of the antigen-recognizing receptor by inducing cell death[150]. In addition, in muscle development, Duxbl has been shown to facilitate cell proliferation[151]. While these studies provided initial insights into the potential physiological roles of an alternative Dux factor during organogenesis and at adult stages, a role for Duxbl in preimplantation development remains elusive, despite its expression had been detected at 2C-stage[152]. Most interestingly, Duxbl expression has been observed in *Dux*-KO-embryos[152].

2 Aim of the project

This project aims to determine a potential role of the transcription factor Duxbl in preimplantation development, especially during ZGA. This idea is based on the finding that the Duxbl-paralog Dux has been identified as an activator of a ZGA-associated expression profile *in vitro*, but is dispensable for ZGA *in vivo*[137,141]. Taking the observation into account that Duxbl has been shown to be expressed in *Dux*-KO-embryos during ZGA[152], I hypothesized that Duxbl might functionally compensate for the loss of Dux.

To test this idea, I characterized the functional and molecular consequences of ectopic Duxbl overexpression *in vitro* and compared them to those of Dux. Moreover, I manipulated the expression dynamics of Duxbl *in vivo* to study its potential function in the regulation of ZGA and preimplantation development.

3 Materials

3.1 Cell culture

Table 1: List of cell lines.

Cell line	Description
293[HEK-293] ATCC® CRL-1573™	Human Embryonic Kidney Cells 293
ES-E14TG2a ATCC® CRL-1821™	blastocyst-derived, mouse embryonic stem cell line

Table 2: List of growth media.

Cell Type	Composition
Preimplantation embryos	EmbryoMax® Advanced KSOM, Sigma-Aldrich (MR-101-D)
ES-E14	Dulbecco's Modified Eagle Medium, Gibco (42430-025) Non-Essential Amino Acids Solution 1x, Gibco (11140050) Penicillin-Streptomycin-Glutamine 1x, Gibco (12090216) Fetal calf serum 15 %, Gibco (16141079) 2-mercaptoethanol 50 nM, Gibco (31350010) ESGRO® Recombinant Mouse LIF Protein 1000 units/ml , Merck (ESG1107) CHIR99021 3 nM, Sigma-Aldrich (SML1046) PD 0325901 1 nM, Sigma-Aldrich (PZ0162)
HEK293	Dulbecco's Modified Eagle Medium, Gibco (42430-025) Penicillin-Streptomycin-Glutamine 1x, Gibco (12090216) Fetal calf serum 10 %, Gibco (16141079)

3.2 Plasmids

Table 3: List of plasmids.

Name	Promoter	Expressed Gene
pRP-Dux	TRE	*Dux-3xHA*
pRP-Duxbl	TRE	*Duxbl-V5*
pVB-Duxbl	EIF1a	*Duxbl-V5*
pVB-Control	EIF1a	none
pPB-rtTA	CAG	*rtTA*
pLVX-NLS-mCherry	TRE	*NLS-mCherry*
pLVX-Duxbl-V5-2A-NLS-mCherry	TRE	*Duxbl-V5-2A-NLS-mCherry*

3.3 Chemicals

Table 4: List of chemicals, enzymes and commercial buffers.

Chemicals / enzymes / commercial buffers	**Manufacturer**
2 × Taq Plus Master Mix	Vazyme (P212)
7.5% Mini-PROTEAN® TGX™ Precast Protein Gel	BioRad (4561023)
Accutase cell detachment solution	Merck Milipore (SCR005)
BSA (bovine serum albumin), Fraction V	Merck (112018)
DAPI (4′,6-diamidino-2-phenylindole)	Invitrogen (D1306)
Dithiothreitol	Merck (11583786001)
DNA-Gel loading dye	Thermo Scientific (11541575)
Doxycycline hyclate	Sigma-Aldrich (D9891)
Dulbecco′s Phosphate Buffered Saline (DPBS)	Sigma-Aldrich (D8537)
Ethidium bromide	AppliChem (A1152,0100)
Gelatin from porcine skin	Sigma-Aldrich (G1890)
GeneRuler 1 kb Plus	Thermo Scientific (10101240)
HiMark prestained protein standard	Invitrogen (LC5699)
Hyaluronidase from bovine testes	Sigma-Aldrich (H4272)
Lipofectamine™ 3000	Invitrogen (L3000008)
MES (2-(N-morpholino)ethanesulfonic acid) running buffer	Invitrogen (NP0002)
Opti-MEM™	Gibco (31985062)
Phanta Max Super-Fidelity DNA Polymerase	Vazyme (P505)
Pierce™ 16 % Formaldehyd methanol-free	Thermo Scientific (28906)
Protein standard marker	Invitrogen (LC5800)
Proteinase inhibitor cocktail tablets	Roche (04693116001)
Proteinase K	Thermo Scientific (EO0491)
RNase A	Thermo Scientific (EN0531)
TAE (Tris-acetate EDTA) buffer	Roth (CL86.3)
Trypsin from porcine pancreas	Sigma (T6567)
Tyrode′s Solution, Acidic	Sigma-Aldrich (T1788)
SureBeads™ Protein A Magnetic Beads	BioRad (1614013)
Benzonase®	(Millipore, E1014)
Protein A beads	Diagenode (C03020002)
Laemmli Sample Buffer	BioRad (161-0747)
Light mineral oil	Sigma-Aldrich (ES-005-C)
Clarity Western ECL Substrate	BioRad (1705061)

3.4 Buffers

Table 5: List of buffers.

Buffer	Composition
RIPA	Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1 mM NaCl, 140 mM Triton X-100, 1 % Sodim deoxycholate, 0.1 %
RIPA+SDS	Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1 mM NaCl, 140 mM Triton X-100, 1 % SDS, 0.1 % Sodim deoxycholate, 0.1 %
RIPA+SDS(high-salt)	Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1 mM NaCl, 500 mM Triton X-100, 1 % SDS, 0.1 % Sodim deoxycholate, 0.1 %
LiCl	Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1 mM LiCl, 250 mM Triton X-100, 1 % NP-40, 0.5 % Sodim deoxycholate, 0.5 %
TE	Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1 mM
Elution	Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 5 mM NaCl, 300 mM SDS, 0.5 %
Blocking	DPBS fetal calf serum, 5% Triton X-100, 0.1 %
IF wash	DPBS Triton X-100, 0.1 %
Western blot wash	Tris buffered saline (TBS) Tween 20, 0.1 %

3.5 Antibodies

Table 6: List of antibodies.

Antibody	Manufacturer
anti-V5 (for ChIP-qPCR/IF)	Cell Signaling (D3H8Q)
anti-V5 (for ChIP-qPCR/Western blot)	Abcam (ab9116)
anti-rabbit-IgG Alexa Fluor 647 (for IF)	Invitrogen (A32795)
anti-p300 (for Western blot)	Invitrogen (RW128)
anti-Mouse HRP (for Western blot)	Pierce (1858413)
anti-Rabbit HRP (for Western blot)	Pierce (31460)
normal-rabbit-IgG control (for IP)	Cell Signaling (2729)

3.6 Primers

All primers were ordered from and synthesized by Sigma-Aldrich.

Table 7: List of primers

Primer Name	Purpose	Sequence (5'->3')
F-Primer T7 NLS-mCherry	Addition of T7 promoter	TTTTTTAATACGACTCACTATA GGGAGGGCGAGGCCACCAT GCCAAAAAAGAAGAGA
R-Primer T7 NLS-mCherry	Addition of T7 promoter	TTACTTGTACAGCTCGTCCAT GCC
F-Primer T7 Duxbl-V5	Addition of T7 promoter	TTTTTTAATACGACTCACTATA GGGAGGGCGAGGCCACCAT GGAGCTGAGC
duxbl-5arm1-F	*Duxbl* CRISPR-Cas9 left targeting vector	CACCGTAAAGCAGACTGTCC AGTAT
duxbl-5arm1-R	*Duxbl* CRISPR-Cas9 left targeting vector	AAACATACTGGACAGTCTGC TTTA
Duxbl-3arm1-F	*Duxbl* CRISPR-Cas9 right targeting vector	CACCGTAGAGGTGCCTACAG TTCTT
Duxbl-3arm1-R	*Duxbl* CRISPR-Cas9 right targeting vector	AAACAAGAACTGTAGGCACC TCTA
Duxbl-Mut-F1	*Duxbl* deletion genotyping	TGTTGACCTTCTCCAGGTGAC

Duxbl-Mut-R1	*Duxbl* deletion genotyping	GCCTCCATCACACTCCCTATC
Duxbl-exon1-F1	*Duxbl* wild type genotyping	CATCTGACTCCCCACAGTCAC
Duxbl-exon1-R1	*Duxbl* wild type genotyping	GAGTGGCCCACCCATTTCTTA
ChIP_Pramel17_FWD	*Pramel17* ChIP-qPCR	ACCAGGAATAGCCACTACAACAAT
ChIP_Pramel17_REV	*Pramel17* ChIP-qPCR	CCACTTCAGACAGAGAGACCAAG
ChIP_Tcstv3_FWD	*Tcstv3* ChIP-qPCR	GAAATGGCAAGCAGGCAGAG
ChIP_Tcstv3_REV	*Tcstv3* ChIP-qPCR	CCAGGACTGCTGGTACACAC
ChIP_Tdpoz3_FWD	*Tdpoz3* ChIP-qPCR	GGGATGAGTAGCATTGGAGCC
ChIP_Tdpoz3_REV	*Tdpoz3* ChIP-qPCR	TTACACGTACAGTCTGGAGGAC
ChIP_Tdpoz8_FWD	*Tdpoz8* ChIP-qPCR	AAGCATGGCAGTTCCAAATGT
ChIP_Tdpoz8_REV	*Tdpoz8* ChIP-qPCR	CGTTAATTCACAGGATTCAGCTTTT
ChIP_Zfp352_FWD	*Zfp352* ChIP-qPCR	AGAGAAGACCAGCTACGACG
ChIP_Zfp352_REV	*Zfp352* ChIP-qPCR	ATCCAACAGTAGCTACCTGAAGATG
ChIP_Zscan4a_FWD	*Zscan4a* ChIP-qPCR	AAAAGGGGGCATCAAATACAAGC
ChIP_Zscan4a_REV	*Zscan4a* ChIP-qPCR	CATTGGAGATGCAAAACTTGGGTT
ChIP_Zscan4b_FWD	*Zscan4b* ChIP-qPCR	GGGTGGGGCAGATTACCATT
ChIP_Zscan4b_REV	*Zscan4b* ChIP-qPCR	ATAGGCAGTGCTGGGGACTA
ChIP_Zscan4c_FWD	*Zscan4c* ChIP-qPCR	CGGAAGCCAGGCGCTATAAA
ChIP_Zscan4c_REV	*Zscan4c* ChIP-qPCR	CCGGTGAACTTCAAGGGTGAT

3.7 Morpholinos

Table 8: List of morpholinos

Morpholino	Details
Control	PCO-StandardControl-300, Gene Tools
anti-*Duxbl*	Sequence: ATCCCAAGAGGCCCAAGAGAGCTGT, Gene Tools

3.8 Kits

Table 9: List of kits.

Kit	Manufacturer
truChIP Chromatin Shearing Kit with Formaldehyde	Covaris (520154)
Direct-zolTM RNA MiniPrep Plus Kit	Zymo Research (R2072)
mMESSAGE mMACHINE™ T7 ULTRA Transcription Kit	Invitrogen (AM1345)
MEGAclear™ Transkriptionsreinigungs-Kit	Invitrogen (AM1908)
REPLI-g Single Cell Kit	Quiagen (150343)
Trans-Blot Turbo RTA Mini 0.45 µm LF PVDF Transfer Kit	BioRad (1704274)
DC™ Protein Assay Kit II	BioRad (5000112)
QIAquick PCR Purification Kit	Quiagen (28104)

3.9 Equipment

Table 10: List of equipment.

Device	Manufacturer
QuantStudio™ 3	ThermoFisher
SP8 confocal microscope	Leica
Zeiss AXIO Observer inverted fluorescence microscope	Zeiss
Master cycler gradient PCR	Eppendorf
VersaDoc^{TM} 3000	BioRad
E220evolution Focused-ultrasonicator	Covaris
Gel Imager	Intas
Trans-Blot® Turbo™ Transfer System	BioRad

3.10 Software

Table 11: List of softwares.

Software	Provider / Developer
ZEN 3.2 (blue edition)	Zeiss
Fiji	Johannes Schindelin
GraphPad Prism 9	Dotmatics
LAS X	Leica
Rstudio	Posit PBC
Jalview 2	Andrew Waterhouse, Jim Procter
Integrative Genomics Viewer	Broad Institute
Illustrator	Adobe
QuantStudio™ Design & Analysis Software	ThermoFisher
bio RENDER	BioRender (some schematics were created with Biorender.com using a BioRender Premium version from the Max Planck Society)

4 Methods

4.1 Cell culture

HEK293 and ES-E14TG2a cells (see Table 1) were cultured at 37 °C and 5 % CO_2 in media as needed (see Table 2). ES-E14TG2a cells were grown on 0.1 % gelatin (see Table 1). Cultured cells were transfected with plasmids (see Table 3) using Lipofectamine 3000 (ThermoFisher, L3000015) with 1 µg of total plasmid per 1x10^6 cells according to the manufacturer's instructions.

4.2 ChIP-qPCR

ES-E14TG2a cells were grown and transfected in suspension in 100 mm dishes. 16 h after plasmid transfection, gene expression was induced by addition of Doxycycline (Sigma, D9891) to a final concentration of 100 ng/ml for 18 h. Next, cells were washed once with PBS and detached from cell culture dishes by treatment with Accutase (Milipore, SCR005) for 5 min at 37 °C. Cells were collected in fresh culture medium and chromatin-protein complexes were cross-linked by addition of methanol-free formaldehyde (Thermo Scientific, 28906) at a final concentration of 1 % for 10 min at RT. Cells were then washed three times in PBS and chromatin was isolated and sheared to 200-700 bp using the truChIP Chromatin Shearing Kit (Covaris, 520154) and a Covaris E220evolution focused-ultrasonicator following the manufacturer's recommendations. Subsequently, chromatin of 8 x

10^6 cells was incubated with either 7 µg of anti-V5 tag antibody (either Abcam, ab9116 or Cell Signaling D3H8Q) or 7 µg of IgG control antibody (Cell Signaling 2729) at 4 °C overnight. Antibody-bound chromatin was then coupled to protein A magnetic beads (BioRad, 1614013) for 3 h at 4 °C. Samples were then washed four times with RIPA+SDS-buffer (see Table 5), three times with RIPA+SDS(high-salt)-buffer (see Table 5), two times with LiCl-buffer (see Table 5) and finally two times with TE Buffer (see Table 5). RNA was eliminated by addition of RNAseA (Thermo Fisher Scientific, EN0531) to a final concentration of 0.1 mg/ml in Elution-buffer and incubated at 37 °C for 30 min. Reverse crosslinking was achieved by adding Proteinase K (ThermoFisher, EO0491) to a final concentration of 0.4 mg/ml and incubation at 55 °C for 1 h and then at 65 °C overnight. DNA was purified using the QIAquick PCR Purification Kit (Quiagen, 28104) and diluted 1:100 in H_20 prior to qPCRs. qPCRs were performed using the according primers (see Table 7) and SsoAdvanced Universal SYBR® Green Supermix (Bio Rad, #1725274) on a QuantStudio™ 3 (ThermoFisher Scientific) following the manufacturer's recommendations. QuantStudio™ Design & Analysis Software was used with default settings to obtain CT-values. CT-values were used to calculate the percentage of input values.

4.3 RNA-isolation from mESC

16 h after plasmid transfection, gene expression was induced by addition of Doxycycline (Sigma, D9891-10G) to a final concentration of 100 ng/ml for 18 h. Next, RNA was isolated with the Direct-zol™ RNA MiniPrep Plus Kit (Zymo Research, R2072) according to the manufacturer's instructions.

4.4 Co-immunoprecipitation (Co-IP) and western blotting

24 h after transfection, gene expression was induced by addition of Doxycycline (Sigma, D9891-10G) to a final concentration of 100 ng/ml. 24 h later, cells were lysed in RIPA buffer (see Table 5) with 1 unit/µl Benzonase® (Millipore, E1014), followed by incubation at 4 °C for 5 h. Subsequently, 1 mg of total protein lysate was incubated with 5 µg anti-V5-antibody (Abcam, ab9116) or normal rabbit IgG control antibody (Cell Signaling, 2729) rotating at 4 °C overnight. Samples were then incubated with protein A beads (Diagenode, C03020002) at 4 °C for 3 h, followed by washing four times with RIPA buffer. Beads were then incubated with Laemmli Sample Buffer (BioRad, 161-0747) and 50 mM dithiothreitol and subjected to gel

electrophoresis. After Western blotting, membranes were incubated overnight either with an anti-V5 tag antibody (Abcam, ab9116) or an anti-P300 antibody (Invitrogen, RW128). The following day, primary antibodies were washed off from membranes before the incubation with secondary antibodies (see Table 6), the washing and the development of membranes using Clarity Western ECL Substrate (BioRad, 1705061).

4.5 Mice

All animal experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals published by the US National Institutes of Health (NIH Publication No. 85-23, revised 1996) and according to the regulations issued by the Committee for Animal Rights Protection of the State of Hessen (Regierungspraesidium Darmstadt). Wildtype C57BL6/J mice were obtained from The Jackson Laboratory (Bar Harbor, ME).

Heterozygous *Duxbl*-knockout (*Duxbl*$^{\Delta/+}$) mice were generated by a former Ph.D. student of the Johnny Kim group, Jiasheng Zhong, by selecting two sgRNA sequences using CRISPick (https://portals.broadinstitute.org/gppx/crispick/public accessed: 11th of December 2023). Corresponding oligos (Sigma-Aldrich) were inserted into px458 (addgene #48138) according to protocol (https://media.addgene.org/cms/filer_public/3e/e1/3ee1ce9c-99f9-4074-9a28-109d34971471/zhang-lab-sam-cloning-protocol.pdf accessed: 11th of December 2023). The two constructed CRISPR-Cas9 vectors were then electroporated into SV129/C57BL6 E14 mouse embryonic stem cells and selected via fluorescence-activated cell sorting two days after transfection. Genotyping was performed via PCR using primers listed in Table 7. Two independent founder lines were generated by microinjection of heterozygous *Duxbl*$^{\Delta/+}$ embryonic stem cells isolated from two independent colonies into blastocysts.

4.6 Mouse preimplantation stage embryo isolation and culture

Three week-old C57BL6/J female mice were superovulated by intraperitoneal injection (i.p.) of pregnant mare serum gonadotropin (10 IU). 48 h later, human chorion gonadotropin (HCG) (10 IU) was delivered i.p., followed by mating with C57BL6/J males. Next morning, zygotes were surgically isolated from ampulla of the oviducts and brought into culture in EmbryoMax® Advanced KSOM Embryo Medium (Sigma-Aldrich, MR-101-D) at 37 °C and 5 % CO_2.

4.7 *In vitro* transcription of mRNAs for microinjection

Prior to *in vitro* transcription (IVT), the T7 promoter was added to the constructs by PCR with according primers (see Table 7) using a Phanta Max Super-Fidelity DNA Polymerase (Vazyme, P505). Capped and polyadenylated mRNAs were produced using the mMessage mMachine® T7 Ultra Kit (ThermoFisher, AM1345) and cleaned up utilizing the MEGAclear™ Kit (ThermoFisher, AM1345) according to the respective manufacturer's instructions.

4.8 Microinjection of preimplantation stage mouse embryos

Microinjection of mRNAs and morpholinos was carried out by Susanne Kreutzer and Daniel Heil from the transgenic facility at the Max Planck Institute for Heart and Lung Research using a FemtoJet® 4i (Eppendorf, 5252000013).

mRNAs were injected into 1C- and late 2C-embryos 24 h or 48 h post HCG, respectively. 15-25 embryos per condition were injected with either 200 ng/µl or 400 ng/µl of *Duxbl^{V5}-2A-^{NLS}mCherry* or 400 ng/µl of *^{NLS}mCherry*. 6 h post microinjection, expression of the ^{NLS}mCherry-reporter was tested by live cell imaging and nuclear localization of Duxbl^{V5} was assessed by immunostaining followed by confocal microscopy. Impairment of ZGA in embryos injected at late 1C-stage was addressed by RNA-seq of respective embryos 48 h post HCG.

Morpholinos were injected into 1C-embryos 24 h post HCG. 0.5 mM of either a control morpholino (see Table 8) or a custom translation-blocking morpholino targeting the 5'UTR of *Duxbl* (see Table 8) were injected into 25-30 embryos per condition. To address to what extent mRNA based overexpression or morpholino-mediated ablation of *Duxbl* interferes with developmental progression, embryos were cultured until embryonic day 4.5 and blastocyst formation was quantified.

4.9 Immunofluorescence

Late 1C-stage (30 h post HCG) or late 2C-stage (53 h post HCG) embryos were fixed for 10 min at room temperature in 4 % methanol-free formaldehyde (ThermoFisher, 28906), washed in DPBS with 0.1 M glycine, permeabilized for 10 min with 0.5 % Triton X-100 in DPBS and incubated for 1 h in blocking buffer (see Table 5). Samples were incubated with primary anti-V5 antibody (Cell Signaling, D3H8Q) (concentration: 1:100) in blocking buffer overnight at 4 °C. The following day, embryos were washed three times in wash buffer (see Table 5) and incubated with Alexa 647 donkey anti-rabbit (ThermoFisher, A32795) (concentration 1:400) for overnight at 4 °C. Embryos were washed three times, incubated for 5 min with DAPI and washed twice again. Embryos were placed in small droplets covered by light mineral oil (Sigma-Aldrich, ES-005-C) and images were taken at 40x magnification using a Leica SP8 confocal microscope.

4.10 Genotyping of Duxbl-KO embryos and pups

Genomic DNA from tail biopsies was isolated by Proteinase K treatment and subsequent 2-propanol precipitation. Embryos were collected at late 2C-stage (48 h post HCG), 4C-stage (56 h post HCG) and 8C-stage (69 h post HCG). Genomic DNA was pre-amplified using the Repli-g Single Cell Kit (Qiagen, 150345) according to the manufacturer's instructions. Genotyping was carried out using 2 × Taq Plus Master Mix (Vazyme, P212). To detect the deletion of *Duxbl*, PCR was carried out with Primers Duxbl-exon1-F1 and Duxbl-exon1-R1, detecting the intact wildtype *Duxbl*-Locus and Duxbl-Mut-F1 and Duxbl-Mut-R1, detecting the deletion of *Duxbl* (see Table 7). Resulting PCR-products were run on a 2 % agarose gel, stained with ethidium bromide.

4.11 Collection of mouse preimplantation stage embryos for RNA-seq

Untreated WT embryos were collected at early (32 h post HCG) and late (48 h post HCG) 2C-stage. mRNA injected WT embryos and embryos obtained from *Duxbl*$^{\Delta/+}$ intercrosses were harvested at late 2C-stage (48 h post HCG). Embryos were washed five times in DPBS and frozen in PCR-stripes on dry ice.

4.12 Library preparation and sequencing

Library preparation and sequencing of RNA-samples were carried out at the Deep Sequencing Platform by Dr. Stefan Günther at the Max Planck Institute for Heart and Lung Research in Bad Nauheim.

In brief, RNA from single embryos was prepared using the SMART-Seq Stranded Kit (Takara Bio, 634442) and ribosomal cDNA was depleted using AMPure beads (Beckman Coulter, A63880). RNA quality and yield were measured using RNA integrity number (RIN) algorithm using the Fragment Analyzer (Agilent) or LabChip Gx Touch 24 (Perkin Elmer). Samples with 7.0-10.0 RIN value were used for library preparation. 2 µg of total RNA was used as input for VAHTS Stranded mRNA-seq Library preparation. Sequencing was performed on a NextSeq500 instrument (Illumina) using v2 chemistry.

4.13 ChIP-seq analysis

Raw ChIP-seq data analysis was carried out by Dr. Carsten Kuenne from the bioinformatics core unit at the Max Planck Institute for Heart and Lung Research.

Reference data from Hendrickson et al.[137] (GSE85632) were downloaded from the Gene Expression Omnibus.

Trimmomatic version 0.39 was employed to trim reads after a quality drop below a mean of Q15 in a window of 5 nucleotides and keeping only filtered reads longer than 15 nucleotides[153]. Reads were aligned versus Ensembl mouse genome version mm39 (Ensembl release 104) with STAR 2.7.10a[154]. Aligned reads were filtered to remove duplicates with Picard 2.27.4 (Picard: A set of tools (in Java) for working with next-generation sequencing data in the BAM format, http://broadinstitute.github.io/picard; accessed: 24th of October 2023), spliced, multi-mapping, ribosomal, or mitochondrial reads. Peak calling was performed with Macs version 3.0.0a7 with FDR < 0.01[155]. Remaining peaks were unified to represent a common set of regions for all samples and counts were produced with bigWigAverageOverBed (UCSC Toolkit)[156]. The raw count matrix was normalized with DESeq2 version 1.36.0[157]. Peaks were annotated with the promoter (TSS +-5000 nt) of the nearest gene based on Ensembl release 104. Contrasts were created with DESeq2 based on the normalized union peak matrix with all size factors set to one. Peaks were classified as significantly differential at average count > 10 and -1 < log2FC > 1. Sequencing depth normalized data were visualized using the Integrative Genomics Viewer[158].

4.14 RNA-seq analysis

Raw RNA-seq data analysis was carried out by Dr. Carsten Kuenne from the bioinformatics core unit at the Max Planck Institute for Heart and Lung Research.

Reference data from Deng et al.[159] (GSE45719) for Zygote to 8C-stage embryos and Hendrickson et al.[137] (GSE85632) for mESCs were downloaded from the Gene Expression Omnibus and analysed with the same settings as own RNA-seq data from mESCs and embryos. In brief, Trimmomatic version 0.39[153] was employed to trim reads after a quality drop below a mean of Q20 in a window of 20 nucleotides, keeping only filtered reads longer than 15 nucleotides. Reads were aligned versus Ensembl mouse genome version mm39 (Ensembl release 104) with STAR 2.7.10a[154]. Aligned reads were filtered to remove duplicates with Picard 2.27.1 (Picard: A set of tools (in Java) for working with next generation sequencing data in the BAM format, http://broadinstitute.github.io/picard, accessed: 24th of October 2023), multi-mapping, ribosomal, or mitochondrial reads. Gene counts were established with featureCounts 2.0.3 by aggregating reads overlapping exons on the correct strand, excluding those overlapping multiple genes[160]. The raw count matrix was normalized with DESeq2 version 1.30.1[157]. Contrasts were created with DESeq2 based on the raw count matrix. For data from mouse embryos, genes were classified as significantly differentially expressed at average count >= 5, multiple testing adjusted p-value < 0.05 and -1 =< log2FC >= 1, for data from mESCs average count >= 5, p-value < 0.05 and -1 =< log2FC >= 1. The Ensemble annotation was enriched with UniProt data (Activities at the Universal Protein Resource (UniProt))[144]. DEGs per contrast were split into separate lists by up- or downregulation and submitted to gene set overrepresentation analyses with KOBAS[161]. The resulting bar plot shows gene sets with Benjamini-Hochberg corrected p-value < 0.2[162]. Sequencing depth normalized data were visualized using the Integrative Genomics Viewer[158].

4.15 Definition of Dux-target genes

To identify Dux-target genes, reference data from Hendrickson and colleagues[137] were downloaded from the Gene Expression Omnibus (GSE85632) and were reanalyzed (see above). As Dux-targets, genes were defined that were upregulated in the RNA-seq data, comparing Dux-overexpression to control (unsorted-plusdox/unsorted-nodox, base read count mean > 5, log2fc > 0.59, FDR <= 0.05) and genes that displayed Dux-binding in ChIP-seq data near their TSS (+- 5000 nt). Here, only Dux-binding events were considered that were not significantly different between the two replicates in the data (unsorted-plusdox-hachip-18h/unsorted-plusdox-hachip-12h, non-de peaks: (mean < 10, 1 < log2fc > -1, FDR <1).

4.16 Comparison of homeobox and C-terminal domains of Duxbl and Dux

For comparison of the homeobox domains of Dux and Duxbl, information about the amino acid sequences was retrieved from UniProt[144]. For comparison of the C-terminal domains of Dux and Duxbl, the C-terminal 59 amino acids were used according to the comparison of Dux and DUX4 by Eidahl and colleagues[147]. Illustrations of amino acid alignments were generated and percent identities were calculated using Jalview version 2.11.2.7[163].

4.17 Statistics

P-values and FDRs were either calculated with DESeq2 version 1.30.1[157] or GraphPad Prism 9 version 9.3.1.

No statistical methods were used to predetermine sample size.

5 Results

5.1 *Duxbl* expression is downstream of *Dux* during ZGA

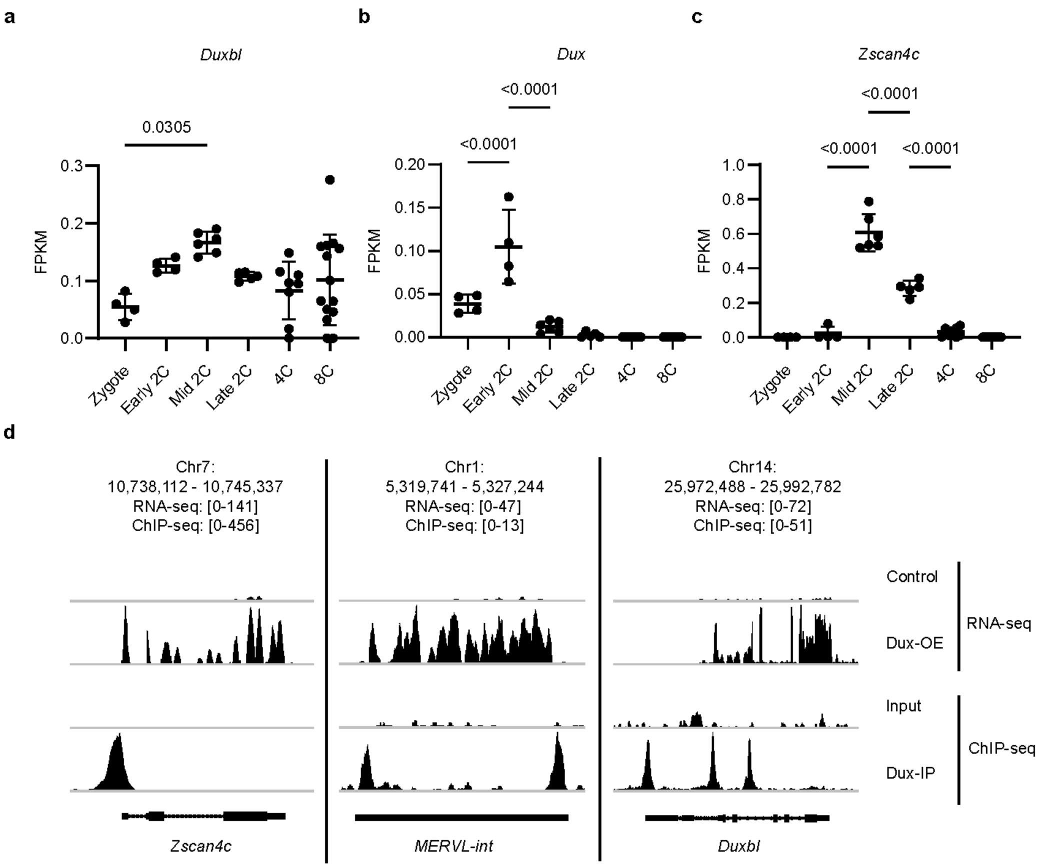

Figure 11: *Duxbl* is expressed during ZGA and downstream of *Dux*:

(a-c) Dot plots depicting relative gene expression values for *Duxbl* **(a)**, *Dux* **(b)** and *Zscan4c* **(c)** from 1C- to 8C-stage. Data represent mean ± standard deviation. One way ANOVA with Tukey's multiple comparison test was used to determine statistical significance. RNA-seq data from[159]. **(d)** Sequencing depth normalized genome browser tracks showing RNA- and ChIP-seq results after overexpression of Dux in mESCs at the loci of *Zscan4c*, *MERVL-int* and *Duxbl*. Data from[137]. OE = overexpression. IP = Immunoprecipitation.

ZGA is the first transcriptional event in life during which thousands of genes are transiently expressed with different temporal dynamics[14,100-103]. When examining expression dynamics of ZGA using published RNA-seq data[159], I found that the gene *Duxbl* starts to be expressed from the 1C-stage onwards, reaching its highest expression level at mid 2C-stage before decreasing at 4C-stage (Figure 11a). Thereby, *Duxbl* expression is deferred to that of *Dux*, which peaked at early 2C-stage (Figure 11b) and overlapped with the onset of the major wave of ZGA[164]. Moreover, *Duxbl* expression occurred concomitantly with *Zscan4c* expression at the peak of the major wave of ZGA[117] (Figure 11c). Since many ZGA-associated genes, including the *Zscan4* cluster, are known to be Dux targets[137], I hypothesized that Duxbl might also be regulated by its paralog. To test this idea, I analyzed publicly available RNA- and ChIP-seq data wherein Dux was ectopically overexpressed in mESCs[137]. Consistent with a previous report, Dux bound to the gene loci of *Zscan4c* and *MERVL*, which was concomitant with upregulation of their expression (Figure 11d)[137]. Most interestingly, Dux also bound to the genomic locus of *Duxbl* and this was associated with transcriptional activation of *Duxbl* (Figure 11d). These data, together with the observed expression dynamics of *Duxbl*, indicated that Duxbl acts downstream of its paralog Dux and suggested a potential role for Duxbl at mid to late 2C-stage.

5.2 Duxbl does not act as a transcriptional activator in mESCs

Since *Duxbl* has also been shown to be expressed during ZGA in *Dux*-KO embryos, which can develop normally[152], I sought to test whether Duxbl would be able to induce a 2C-like state in mESCs similar to its paralog Dux[137]. If true, Duxbl might be able to compensate for the loss of Dux during ZGA, ensuring proper initiation of the gene expression program[141]. To this end, I overexpressed Duxbl or Dux in mESCs and performed RNA-seq. This revealed that Duxbl elicited differential expression (2 >= FC =< -2, p-value <= 0.05, mean read count >= 5) of 28 genes, of which half were upregulated (Figure 12a). Interestingly, Duxbl overexpression did not induce transcription of 2C-specific ZGA-associated genes, as for example *Zfp352*, *Tcstv3*, *Tdpoz3*, *Tdpoz8* or genes from the *Zscan4*-cluster[137]. In contrast, Dux overexpression elicited upregulation of *Zfp352*, *Tcstv3*, *Tdpoz3*, *Tdpoz8* and the *Zscan4*-cluster, in line with previous reports (Figure 12b)[137]. In fact, Dux elicited differential expression of 173 genes, of which the vast majority (76 %) were upregulated (Figure 12b). This confirmed the previous finding that Dux is a potent transcriptional activator[137]. Comparison of gene expression levels of the Top 50 differentially expressed genes (DEGs) between Dux and Duxbl clearly demonstrated that Duxbl does not act as a transcriptional activator in mESCs in contrast to Dux (Figure 12c).

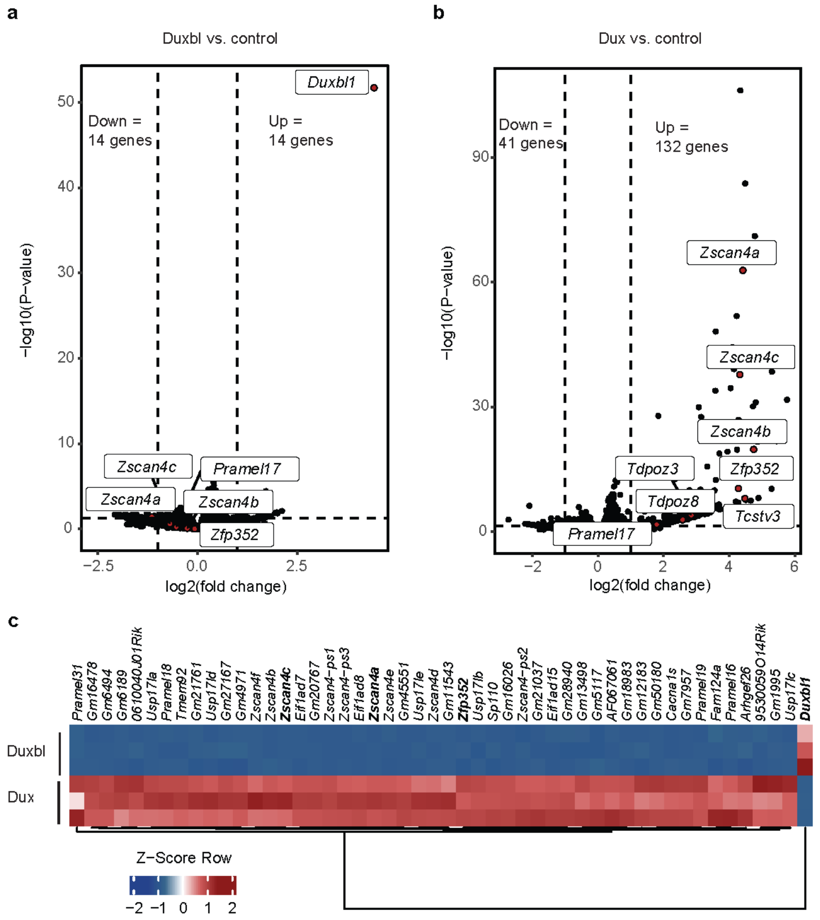

Figure 12: Duxbl does not induce a 2C-like transcriptome in mESCs:

(a, b) Volcano plots showing differential gene expression after overexpression of Duxbl **(a)** or Dux **(b)** in mESCs compared to control. **(c)** Heat map comparing expression levels of the top 50 DEGs in Dux- and Duxbl-overexpressing mESCs. (DEGs: 2 => FC <= -2; mean read count >= 5; *p*-value <= 0.05). n for each condition = 3.

The ability of Dux to induce transcriptional activation has been mechanistically linked to its interaction with the histone acetyltransferases EP300 and CBP[138,139]. This interaction is mediated via the C-terminal transactivation domain (C-TAD) of Dux, which comprises the C-terminal 59 amino acids of the protein[138,147]. In order to assess the extent to which evolutionary conservation between the two paralogs might have influenced the C-terminal regions of Dux and Duxbl, I conducted an alignment of the C-terminal 59 amino acids of Dux with the corresponding segment in Duxbl (Figure 13a). This analysis revealed a percent identity of 22.22 % (Figure 13b). To test whether this limited sequence overlap could be sufficient to support a conserved interaction between murine DUX factors and EP300, I conducted a co-immunoprecipitation (Co-IP) experiment for Duxbl and Dux following their overexpression in HEK293 cells. Despite the modest sequence similarity within their C-terminal domains, Western blot analysis revealed that Duxbl was not able to interact with EP300, unlike Dux (Figure 13c). These results indicated that Duxbl does not contain a C-TAD similar to that of Dux, which is required for transactivation of Dux-target genes[139].

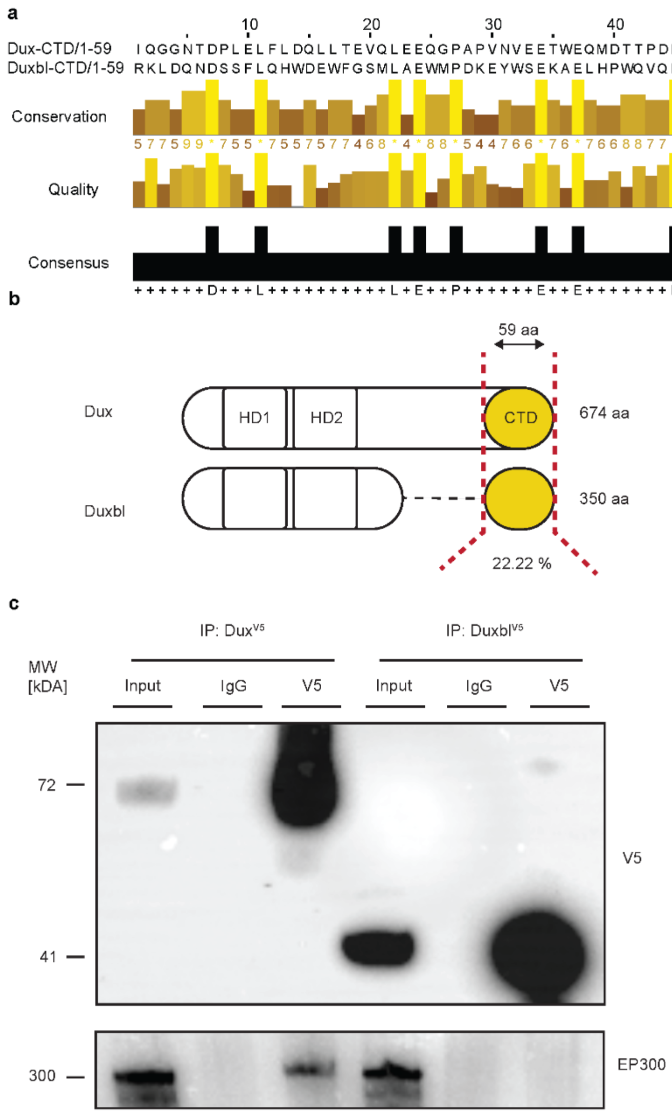

Figure 13: Duxbl does not interact with EP300:

(a) Amino acid sequence alignment of the CTDs (last 59 amino acids) of Dux and Duxbl. **(b)** Schematic depicting the percent identity of the CTDs of Dux and Duxbl. **(c)** Western blot results after Co-IP of V5 epitope-tagged Dux or Duxbl, overexpressed in HEK293 cells. IP = immunoprecipitation, MW = molecular weight, HD = homeobox domain, CTD = C-terminal domain

Taken together, these data showed that Duxbl is not able to elicit expression of ZGA-genes and does not act as a transcriptional activator in mESCs, potentially due to lack of a transactivation domain. Along this line, the inability of Duxbl to mediate

H3K27 acetylation via EP300 and to activate genes from the ZGA panel strongly indicated that Duxbl is not able to reprogram mESCs to a 2C-like state, in contrast to Dux[137].

5.3 Duxbl antagonizes Dux-induced transcription in a 2C-like system

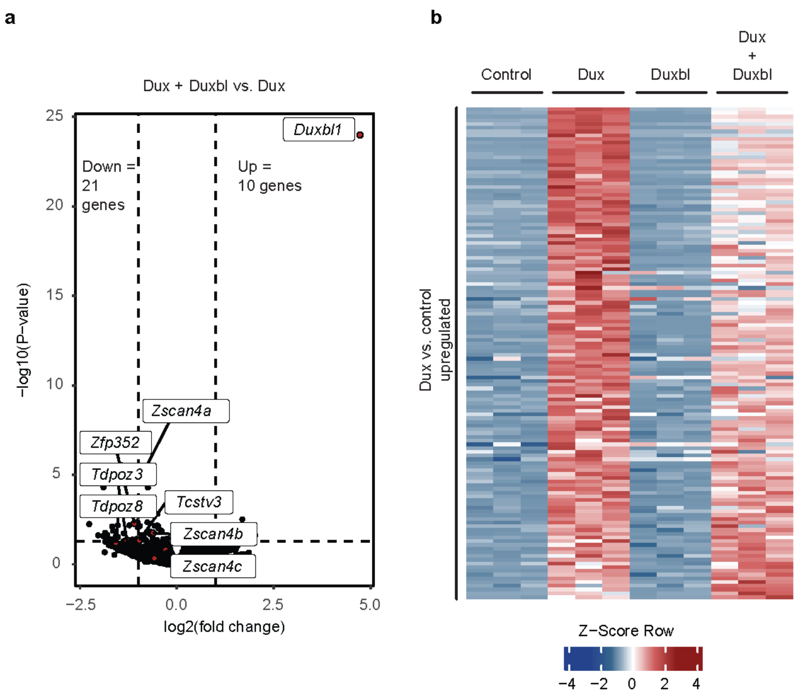

Figure 14: Duxbl suppresses Dux-mediated induction of ZGA-genes in 2C-like mESCs:

(a) Volcano plot showing differentially expressed genes upon co-overexpression of Dux and Duxbl in comparison to overexpression of Dux alone. **(b)** Heat map depicting expression levels of Dux-induced genes in indicated samples. (DEGs: 2 >= FC <= -2; mean read count >= 5; *p*-value <= 0.05). n for each condition = 3.

Since 2C-like cells display a clearly distinct gene expression panel and epigenetic landscape compared to pluripotent mESCs[137,165], I speculated that certain co-factors of Duxbl might not be present or that genomic targets might not be accessible to Duxbl in mESCs. To assess the transcriptional effect of Duxbl in a 2C-like system, I co-overexpressed both paralogs Dux and Duxbl in mESCs and performed RNA-seq. In this approach, ectopic Dux shifts the gene expression profile and chromatin accessibility towards a state similar to that of the 2C-embryo[137]. I conducted a comparative analysis to investigate the potential synergistic effects of co-overexpressing Duxbl and Dux as opposed to overexpressing Dux alone. Remarkably, Duxbl expression resulted in significant downregulation of ZGA-associated Dux-targets, such as *Zscan4a*, *Zfp352* and *Tcstv3* (Figure 12b, Figure

14a). Moreover, a large portion of Dux-induced genes displayed substantially lower expression levels when Dux was co-overexpressed with Duxbl (Figure 14b). These data indicate that Duxbl can function as a transcriptional repressor towards Dux-meditated gene expression in 2C-like cells.

5.4 Duxbl and Dux share conserved DNA-binding domains and genomic ZGA-targets

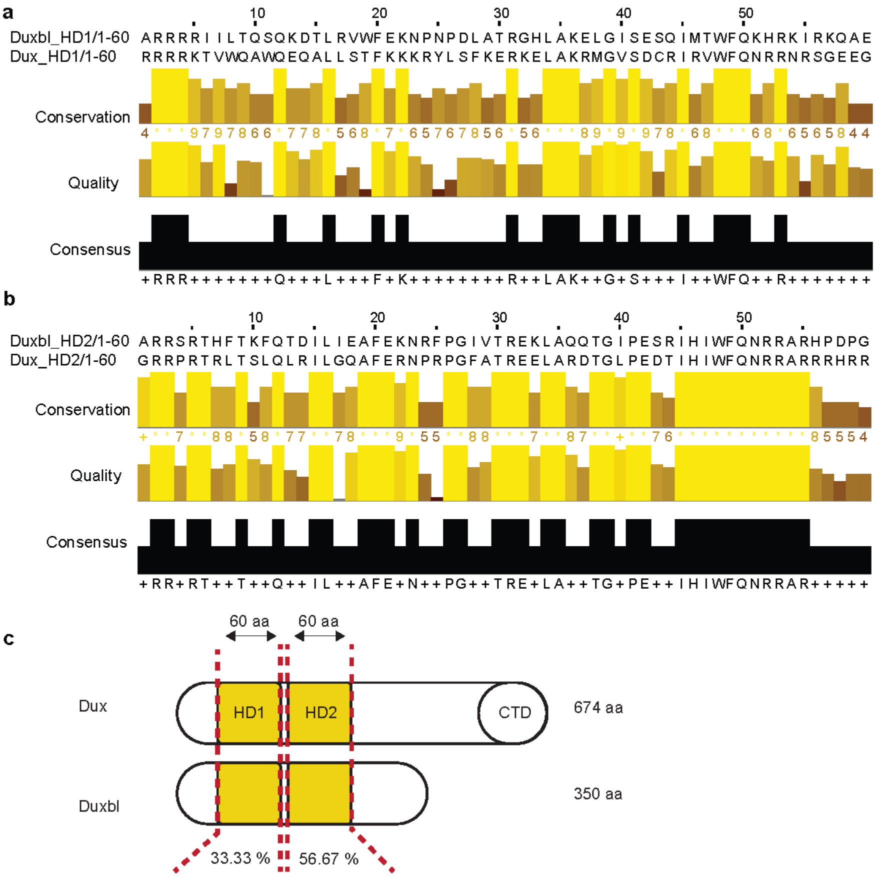

Figure 15: The DNA-binding domains of Dux and Duxbl are conserved:
(a,b) Amino acid sequence alignment of the first **(a)** and second **(b)** homeobox domains of Dux and Duxbl. **(c)** Schematic illustrating Dux and Duxbl with the percent identity of the DNA-binding domains. HD = homeobox domain, CTD = C-terminal domain

The ability of Dux to induce a 2C-like state in mESCs involves its C-TAD-mediated interaction with CBP/EP300 and direct binding of Dux to genomic loci of many ZGA-target genes[137-139]. Within the DUX-family, the DNA-binding ability is mediated by the two N-terminal homeobox domains[148]. Interestingly, while the C-TAD seems to be an exclusive entity of human DUX4 and murine Dux within the DUX-family, all DUX family members display a substantial amount of sequence similarity within their DNA-binding domains[147]. To determine the degree of similarity between the homeobox domains of Dux and Duxbl, the primary structures of the 60 amino acid-long DNA-binding domains were aligned to each other (Figure 15a,b). This analysis revealed that the first homeobox domains of Dux and Duxbl share 33.33 % sequence identity (Figure 15a,c), while the second homeobox domains exhibit a much larger overlap, sharing a sequence identity of 56.67 % (Figure 15b,c). Remarkably, human DUX4 and murine Dux show a similar degree of sequence homology (33 % and 56 %) in their DNA-binding domains and are both capable of eliciting expression of common binding-targets such as *Pramef12* or *Zscan4* in human myoblasts[147]. Based on these findings and taking into consideration that Duxbl seems to lack a C-TAD, I speculated that Duxbl might act as a direct transcriptional repressor on potentially shared ZGA-associated target genes by binding to the same loci as Dux. To test this idea, I conducted Chromatin immunoprecipitation coupled to quantitative polymerase chain reactions (ChIP-qPCRs) against a V5-epitope tagged Duxbl that was co-overexpressed with Dux in mESCs. Primers directed to genomic loci of Dux-induced ZGA-genes were based on publicly available Dux ChIP-seq data (Figure 16a)[137]. Strikingly, I observed strong enrichment at all selected ZGA-associated gene loci using two different anti-V5 antibodies compared to a control IgG antibody (Figure 16b). These data clearly demonstrated that Duxbl binds to the same ZGA-genes at the same sites as Dux within a 2C-like system. Importantly, this binding concurs with downregulation of those shared target genes when Dux and Duxbl are co-overexpressed in mESCs (Figure 14a, Figure 16c).

In summary, these data showed that Duxbl counteracts Dux-mediated transcription in a 2C-like *in vitro* system, potentially by binding to the same genetic regulatory elements of ZGA-associated genes as Dux. Taking into account that Duxbl is expressed downstream of the ZGA-activator Dux during preimplantation development (Figure 11), I reasoned that Duxbl might act as an antagonist towards the major wave of ZGA in a negative feedback loop.

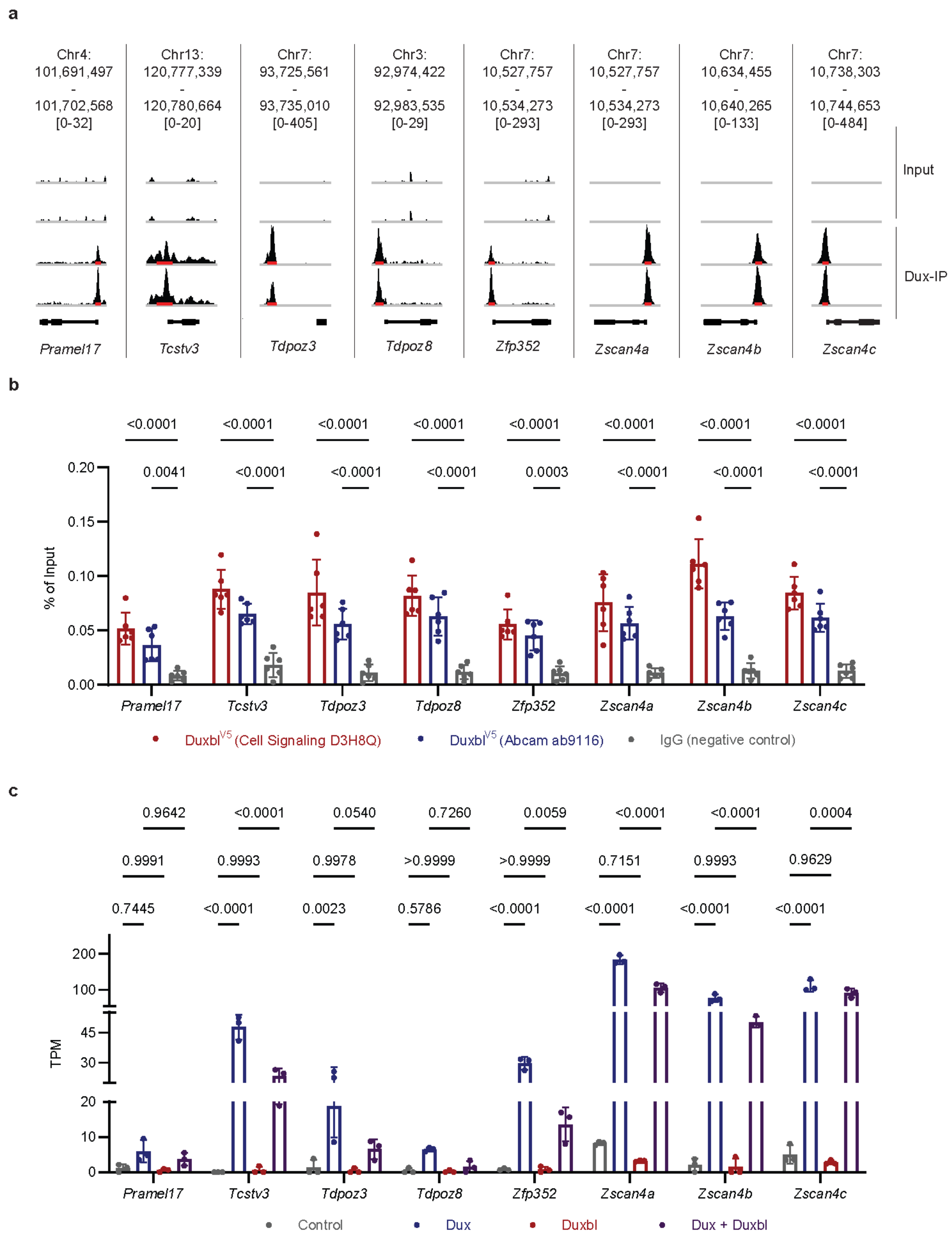

Figure 16: Duxbl binds regulatory sites and suppresses expression of ZGA-associated Dux-targets in 2C-like mESCs:

(a) Sequencing depth normalized genome browser tracks depicting ChIP-seq results after Dux-overexpression in mESCs. Red areas mark sequences used for ChIP-qPCR primer design in **(b)**. Data from[137]. **(b)** Bar graphs showing ChIP-qPCR results for V5 epitope-tagged Duxbl after co-overexpression with Dux in mESCs. n for each condition = 6. **(c)** Bar graphs comparing gene expression levels in mESCs after overexpression of Dux, Duxbl or Dux and Duxbl together compared to control. n for each condition = 3. **(b,c)** Data represent mean ± standard deviation. Two way ANOVA with Tukey's multiple comparison test was used to determine statistical significance.

5.5 Duxbl as a potential repressor of the major wave of ZGA

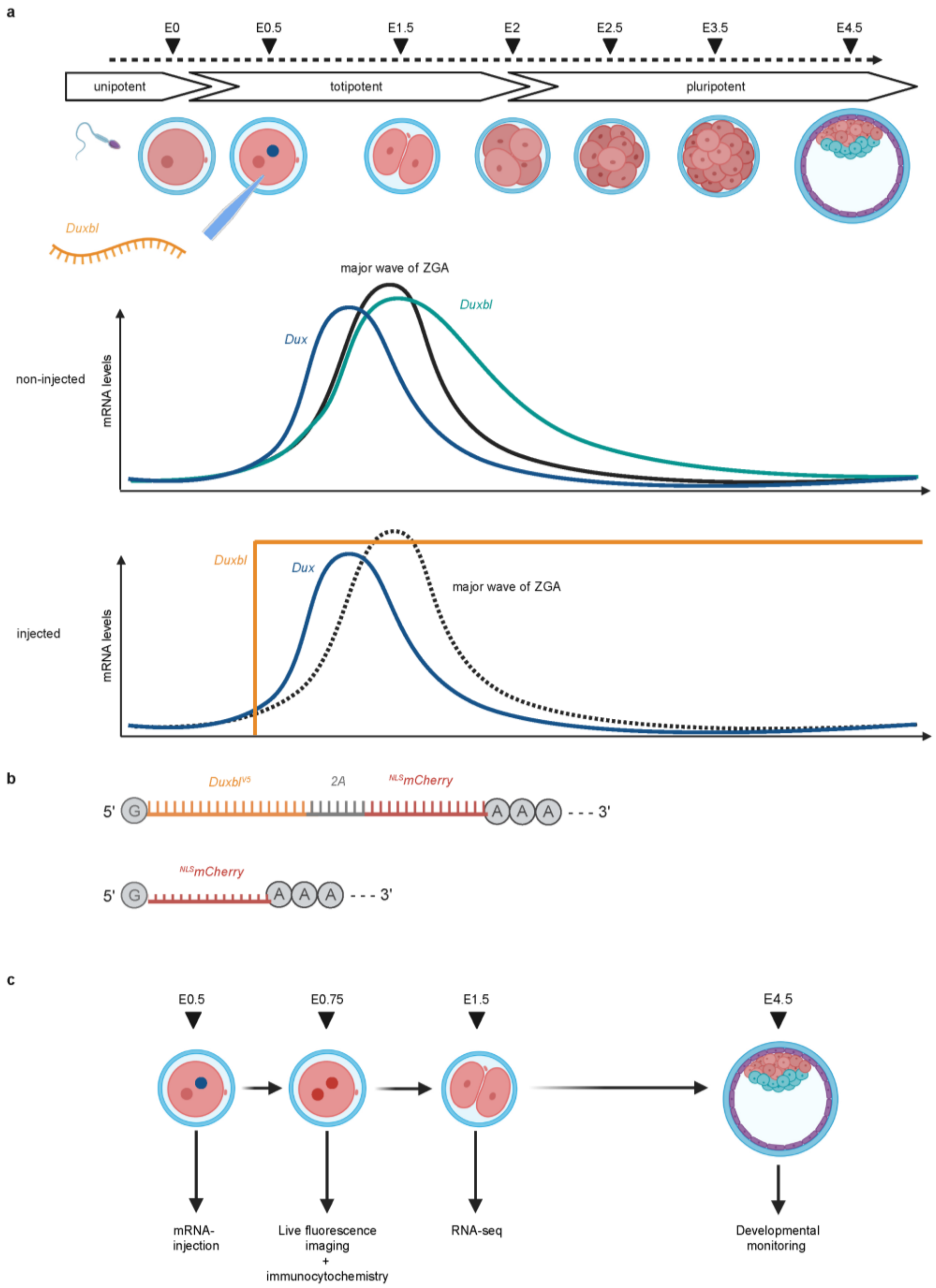

Figure 17: Precocious expression of Duxbl via mRNA-injection:
(a) Schematic showing gene expression dynamics of *Dux*, *Duxbl* and the major wave of ZGA in non-injected controls and embryos injected at E0.5 with *Duxbl*-encoding mRNA. **(b)** Schematics of mRNA constructs. **(c)** Schematic of experimental workflow. (Created with BioRender.com)

In their pioneering experiments in 1974, Warner and Versteegh showed that ZGA is an essential process during early mammalian development by inhibiting transcription in 2C-stage mouse embryos through treatment with alpha-amanitin[166]. This treatment abolished ZGA and resulted in a developmental arrest at the 2C-stage[166]. This prompted me to assume that if Duxbl would act as a repressor towards the major wave of ZGA, shifting Duxbl expression to the 1C-stage, prior to the initiation of Dux-mediated transcription and prior to the major wave of ZGA[164], might result in a similar phenotype (Figure 17a). To test this idea, I designed an mRNA-construct encoding a V5-epitope tagged Duxbl (Duxbl^{V5}) connected by a self-cleaving 2A peptide in tandem to a nuclear localization signal tagged mCherry (^{NLS}mCherry) (Figure 17b). In this setup, expression of the fluorescence reporter reflects expression of Duxbl, while the functional consequences of the transcription factor can be investigated independently of the reporter after translation of both[167].

5.6 Preponed expression of Duxbl leads to developmental arrest at 2C-stage

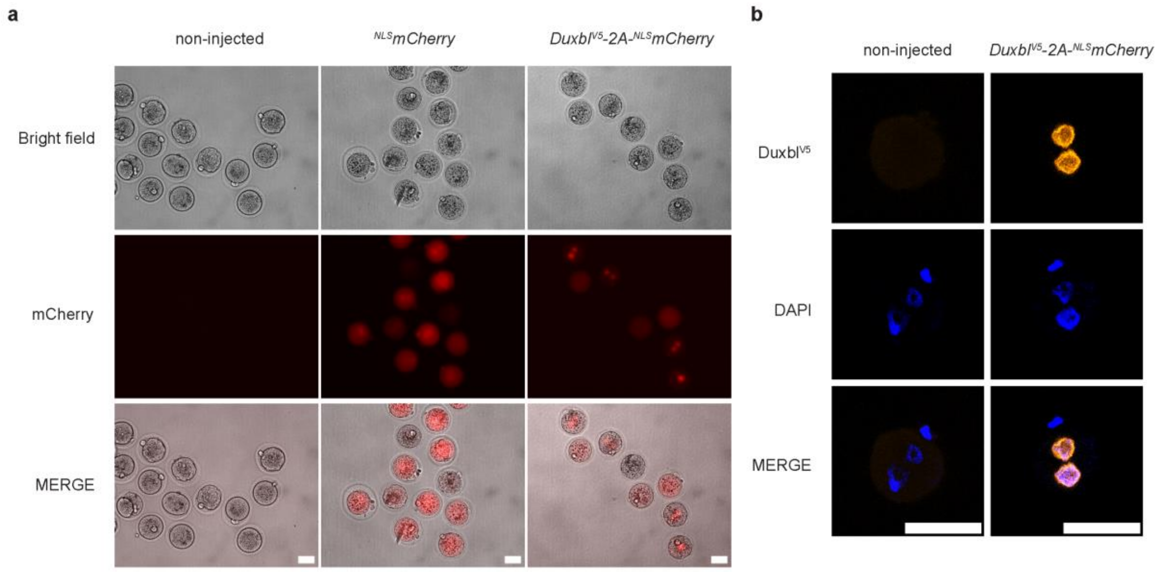

Figure 18: Validation of precocious Duxbl^{V5} expression:
Embryos were injected either with *^{NLS}mCherry*- or *^{NLS}mCherry-2A-Duxbl^{V5}*-encoding mRNAs at E0.5 (24 h post HCG). Images were taken six hours after injection. **(a)** Representative live images of embryos. **(b)** Confocal images of embryos after immunocytochemistry with anti-V5 epitope antibody. DAPI was used for nuclear staining. Scale bars = 100 µm.

To express Duxbl before the major wave of ZGA, *Duxbl^{V5}-2A-^{NLS}mCherry*-mRNA or *^{NLS}mCherry*-mRNA as a control were microinjected into zygotes at embryonic day 0.5 (E0.5) (Figure 17c). Six hours after injection, expression of the fluorescence reporter was confirmed by live cell imaging (Figure 17c, Figure 18a) and nuclear localization of Duxbl^{V5} could be observed by confocal microscopy (Figure 17c, Figure 18b). These findings demonstrated that mRNA-injection at the 1C-stage faithfully enables preponed Duxbl protein expression prior to the onset of the major wave of ZGA[164].

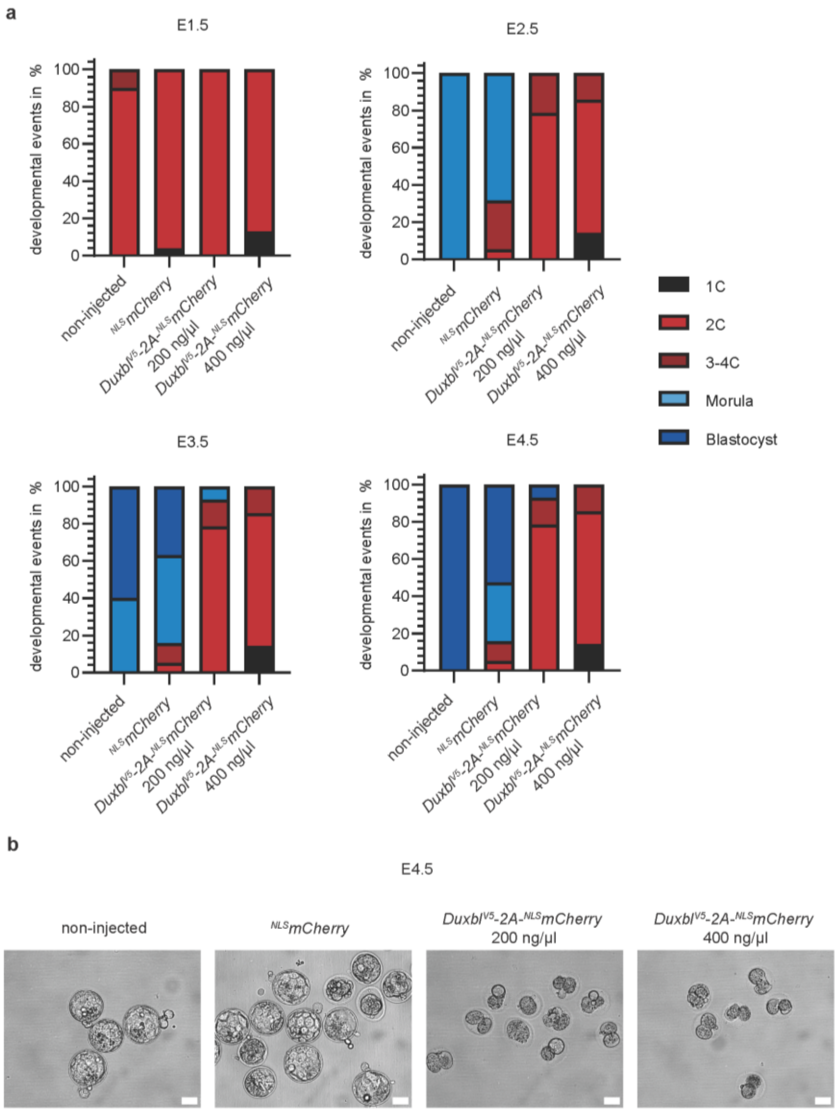

Figure 19: Precocious expression of Duxbl results in developmental arrest at 2C-stage:
(a) Staggered bar graphs depicting developmental ratios in indicated groups of embryos at E1.5 (48 h post HCG), E2.5 (72 h post HCG), E3.5 (96 h post HCG) and E4.5 (120 h post HCG). **(b)** Representative live images of microinjected or control embryos at E4.5. Scale bars = 100 µm. non-injected control: n = 5, ^{NLS}*mCherry*: n = 19, ^{NLS}*mCherry-2A-Duxbl*^{V5} 200 ng/µl: n = 14, ^{NLS}*mCherry-2A-Duxbl*^{V5} 400 ng/µl: n = 21.

To assess to which extent this would interfere with preimplantation development, microinjected embryos and non-injected controls were cultured for developmental monitoring until embryonic day 4.5 (E4.5) (Figure 17c). At embryonic day 1.5 (E1.5), all embryos, including those injected with two different doses of Duxbl^{V5}-encoding mRNA, developed normally to the 2C-stage (Figure 19a). However, at embryonic day 2.5 (E2.5), the majority of Duxbl^{V5}-overexpressing embryos arrested at the 2C-stage, while control embryos progressed normally to the morula stage and later, at embryonic day 4.5 (E4.5), to blastocyst stage (Figure 19a,b). Interestingly, a potentially dosage-dependent effect was observed with the injection of a higher dosage of Duxbl^{V5}-encoding mRNA, resulting in 100 % of the embryos undergoing developmental arrest (Figure 19a). These data showed that preponed expression of Duxbl at the 1C-stage phenocopies alpha-amanitin-induced inhibition of ZGA[166].

5.7 Duxbl is sufficient to suppress the major wave of ZGA

To gain insight into the molecular mechanism by which preponed Duxbl expression elicits developmental arrest, I collected embryos that were injected with Duxbl-encoding mRNA at the late 1C-stage for RNA-seq at late 2C-stage, when the major wave of ZGA reaches its peak[164]. Remarkably, differential gene expression analysis (2 >= FC <= -2, FDR <= 0.05, mean read counts >= 5) revealed that the majority of differentially expressed genes (DEGs) were downregulated (2437 out of 3708) (Figure 20a). This clearly supported the idea that Duxbl acts as a transcriptional repressor rather than an activator. Importantly, more than 50 % (1254 out of 2437) of the downregulated genes belonged to the major wave of ZGA (early to mid 2C, FC > 2, FDR <= 0.05, mean read count >= 5, data[159]) (Figure 20a). Indeed, more than one third (34.1 %) of all genes belonging to the major wave of ZGA were suppressed by preponed expression of Duxbl (Figure 20b). Accordingly, at late 2C-stage, expression levels of these downregulated ZGA-genes in embryos with preponed expression of Duxbl were more similar to the ones in control embryos at early 2C-stage, prior to the major wave of ZGA (Figure 20c)[164]. Strikingly, preponed expression of Duxbl *in vivo* resulted in downregulation of the ZGA-associated Dux-targets that were bound and suppressed by Duxbl in mESCs *in vitro* (Figure 16, Figure 20d). Moreover, average expression of all Dux-targets (data[137], RNA-seq: Dux vs. Control, mean read count > 5, log2fc > 0.59, FDR <= 0.05, ChIP-seq: peak +- 5000 nt to TSS) at 2C-stage was significantly downregulated upon precocious expression of Duxbl (Figure 20e). These data clearly demonstrated that Duxbl alone is sufficient to suppress a substantial portion of genes belonging to the major wave of ZGA. Taken together, this strongly indicated a role for Duxbl as a transcriptional antagonist of Dux in preimplantation development.

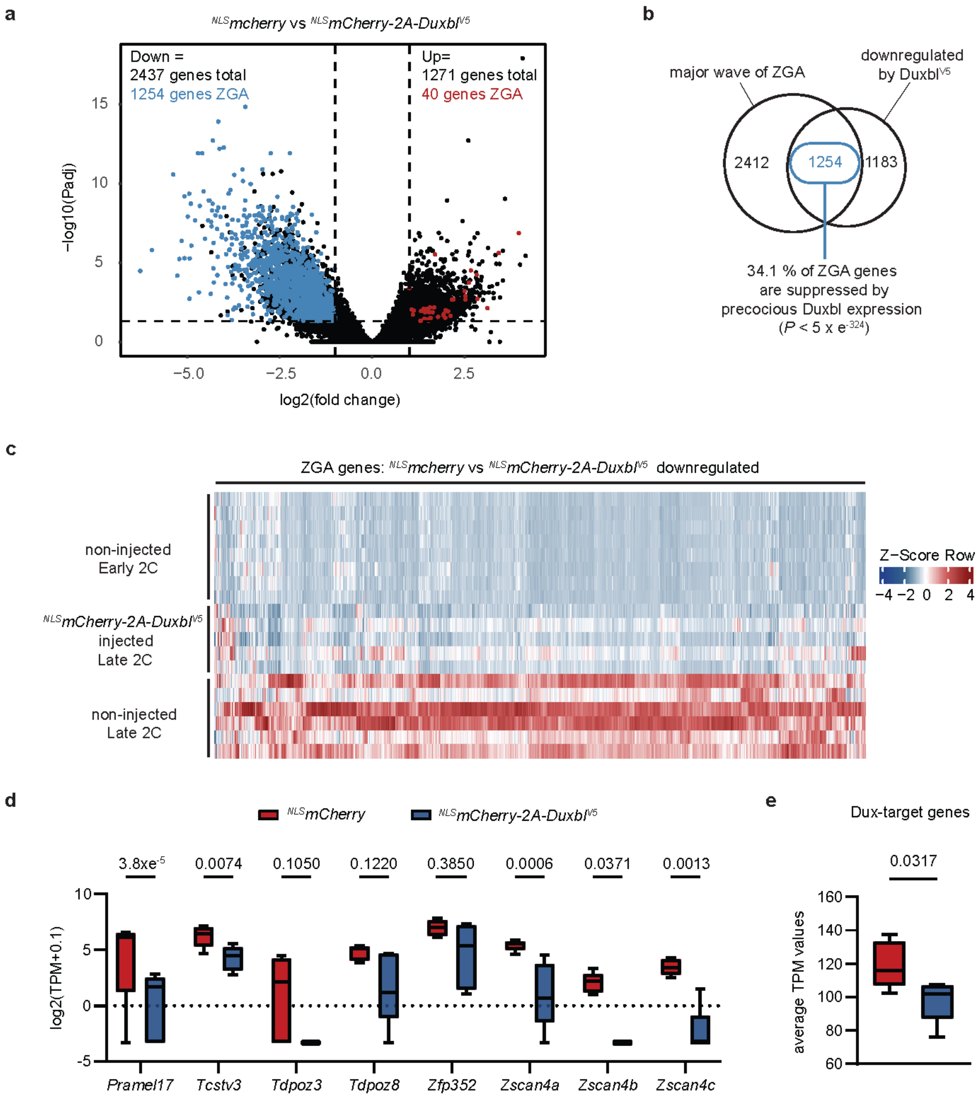

Figure 20: Duxbl suppresses ZGA *in vivo*:

(a) Volcano plot depicting differential gene expression at late 2C-stage (E1.5) between embryos that were microinjected at E0.5 with either ^{NLS}*mCherry-2A-Duxbl*^{V5}- or ^{NLS}*mCherry* (control)-encoding mRNAs. Blue and red dots label down- or upregulated ZGA genes, respectively. **(b)** Venn diagram showing overlap between genes downregulated by precocious Duxbl^{V5}-expression and ZGA-genes. Statistical significance was determined by hypergeometric test. **(c)** Heat map comparing expression levels of ZGA-genes downregulated by Duxbl^{V5} (blue circle in **(b)**) in non-injected control embryos at early and late 2C-stage to *Duxbl*^{V5}-mRNA injected embryos at late 2C-stage. **(d,e)** Box-and-whisker plots (minimum to maximum) showing normalized expression levels of individual ZGA-associated Dux-targets **(d)** or all Dux-targets **(e)**. Dux-targets were identified using data from[137]. FDRs were obtained from DESeq2 for **(d)**. Mann-Whitney-U test was used to determine statistical significance in **(e)**. ZGA genes = upregulated from early to mid 2C-stage, data from[159]. (DEGs: 2 >= FC <= -2; mean read count >= 5; FDR <= 0.05). ^{NLS}*mCherry*: n = 5. ^{NLS}*mCherry-2A-Duxbl*^{V5}: n = 5, non-injected early 2C: n = 8, non-injected late 2C: n = 6. 5 embryos per sample (n).

5.8 Overexpression of Duxbl after ZGA is compatible with preimplantation development

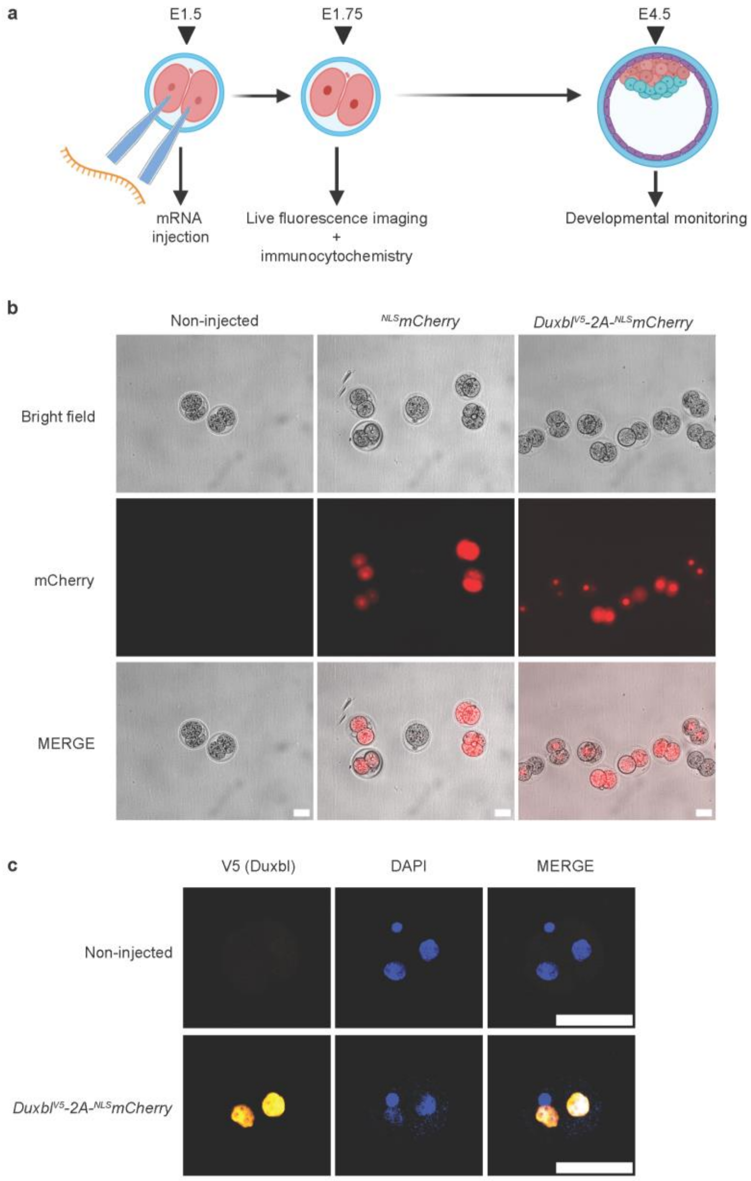

Figure 21: Overexpression of Duxbl post ZGA:

Both blastomeres of 2C-embryos were microinjected with either ^{NLS}*mCherry*- or ^{NLS}*mCherry-2A-Duxbl*^{V5} -encoding mRNA at E1.5 (48 h post HCG). **(a)** Schematic illustrating experimental workflow. **(b,c)** Images were taken 6 hours after microinjection. **(b)** Representative live images of embryos. **(c)** Confocal images of embryos after immunocytochemistry with anti-V5 epitope antibody. DAPI was used for nuclear staining. Scale bars = 100 µm. (**(a)** Created with BioRender.com)

Conceivably, the developmental arrest at 2C-stage that is mediated by overexpression of Duxbl (Figure 19) might be elicited by subsequent differential expression of genes that are not associated with ZGA. To test this idea, Duxbl-encoding mRNA was microinjected into both blastomeres of embryos at late 2C-stage, when the major wave of ZGA has already taken place and a potentially observed phenotype would therefore be independent of the initiation of ZGA (Figure 21a)[164]. To assess adequate timing of protein delivery, I performed live cell imaging and confocal microscopy. This confirmed expression of Duxbl at the transition from 2C- to 4C-stage (Figure 21b,c). Remarkably, Duxbl-overexpressing embryos developed into blastocysts with similar ratios as controls at E4.5 (Figure 22). These results further indicated that the observed developmental arrest upon preponed expression of Duxbl (Figure 19) was caused by suppression of ZGA and corroborated the idea that Duxbl acts as an antagonist towards the major wave of ZGA.

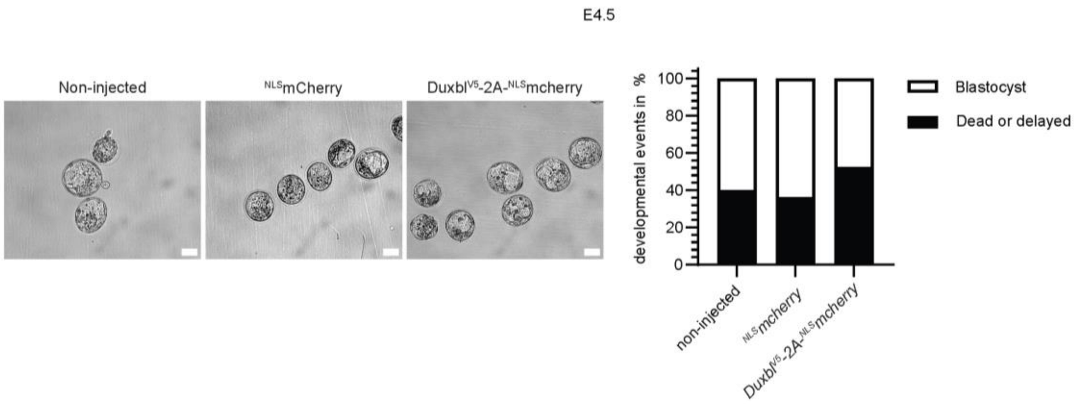

Figure 22: Overexpression of Duxbl post ZGA does not impede preimplantation development: Embryos were injected at E1.5 (24 post HCG) with either ^{NLS}*mCherry*- or *Duxbl*^{V5}*-2A-*^{NLS}*mCherry*-encoding mRNA. Representative bright field images showing embryos at E4.5. Staggered bar graphs depicting blastocyst formation rates at E4.5. Scale bars = 100 µm. Non-injected: n = 5, ^{NLS}*mCherry*: n = 11; *Duxbl*^{V5}*-2A-*^{NLS}*mCherry*: n = 21.

5.9 Loss of Duxbl results in 2C- to 4C-stage arrest

At the transition from 2C- to 4C-stage, the gene expression panel associated with ZGA is rapidly silenced in mice and accompanies the loss of totipotency[122,140,168,169]. This is necessary for development to progress and to enable the first rise of different cell lineages since sustained expression of ZGA-specific Dux target genes results in developmental arrest at the 4C-stage[170]. In this context, I set out to tackle the question whether Duxbl is necessary for downregulation of ZGA-associated transcription and hence for development.

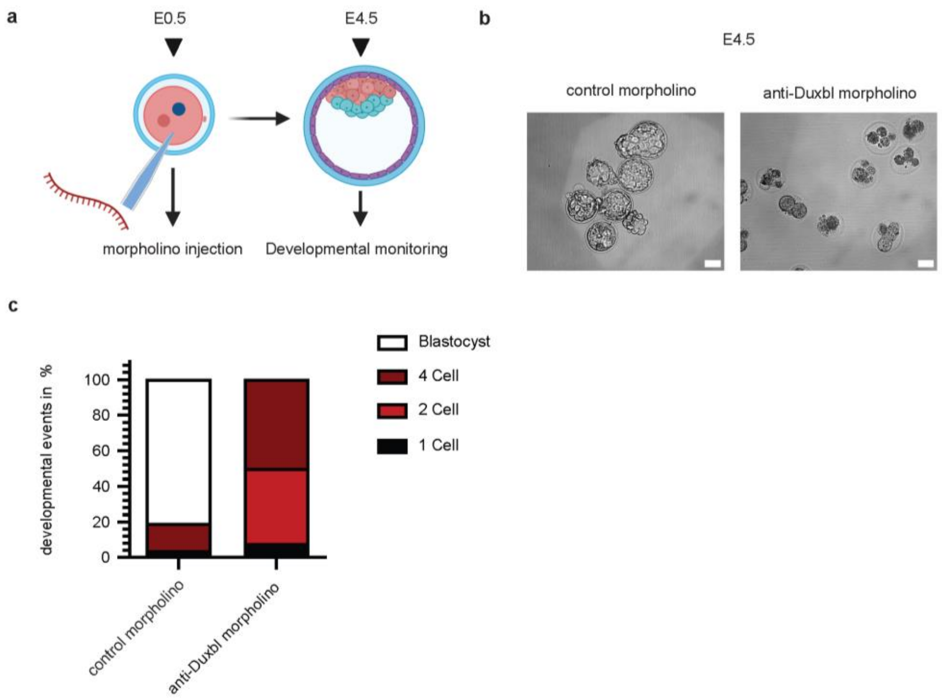

Figure 23: Duxbl-knockdown results in 2- to 4C-stage arrest:

a) Schematic illustrating experimental workflow. Embryos were microinjected at E0.5 with either control morpholino or a translation-blocking morpholino targeting the 5'-UTR of *Duxbl*. **(b)** Representative bright field images of injected embryos at E4.5. **(c)** Staggered Bar graphs depicting ratios of developmental events. Both groups n = 26. Scale bars = 100 µm. (**(a)** Created with BioRender.com)

To test this, embryos at the 1C-stage were microinjected with either a translation-blocking morpholino (MO) against Duxbl or a control MO and cultured until E4.5 (Figure 23a). While the control antisense oligo did not interfere with development to the blastocyst stage, the anti-Duxbl MO led to a complete developmental arrest at the 2C- to 4C-stage (Figure 23 b,c). Strikingly, these results phenocopied the effect of sustained expression of ZGA-genes[170]. These data strongly indicated that Duxbl is necessary for preimplantation development.

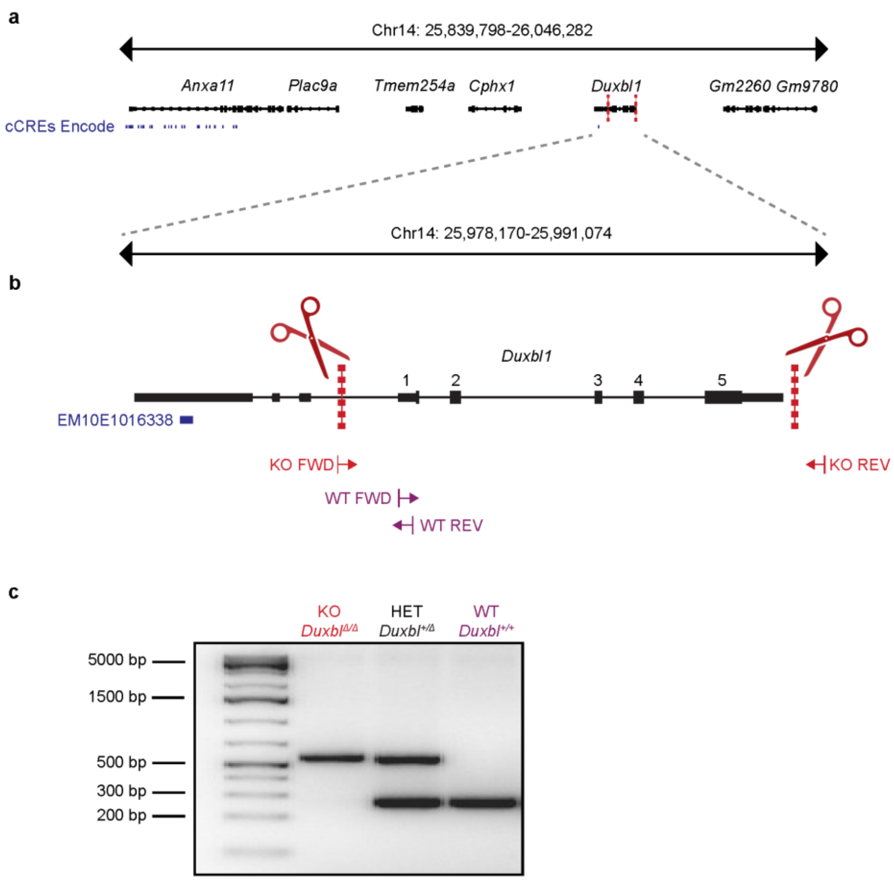

Figure 24: Generation and genotyping of *Duxbl*-KO mouse line:

(a) Schematic illustrating *Duxbl1* locus on chromosome 14 with neighboring genes and cCREs annotations from Encode[171]. **(b)** Schematic depicting the *Duxbl1* locus in detail with numbers of exons, cCREs Encode annotation (in blue), and locations of primers for genotyping. Red dashed lines indicate cutting sites for CRISPR-mediated knockout. **(c)** Gel picture showing result of PCR-based genotyping of 4C-stage embryos, obtained from heterozygous *Duxbl-KO* intercrosses. KO = knockout; WT = wild type; HET = heterozygous; FWD = forward; REV = reverse. Of note: Illustrations **(a)** and **(b)** were generated based on mm10. (**(b)** Created with BioRender.com)

Due to the lack of an IF-compatible anti-Duxbl antibody and the possibility of off-target effects of MOs[172], I further investigated the loss-of-function-effect of Duxbl using a mouse model wherein the Duxbl gene has been genetically inactivated. In this model, all five coding exons of the *Duxbl* gene were removed in mESCs via a CRISPR-Cas9 strategy (Figure 24 a,b). Importantly, no genetic regulatory elements were found within this deleted region according to the database for candidate cis-regulatory elements (cCREs) from the Encyclopedia of DNA Elements Project (ENCODE)[171] (Figure 24b). Moreover, successful deletion of the Duxbl-encoding region was verified by PCR using two independent primer sets, one detecting the deleted allele and the other one detecting the wild type (WT) allele (Figure 24c).

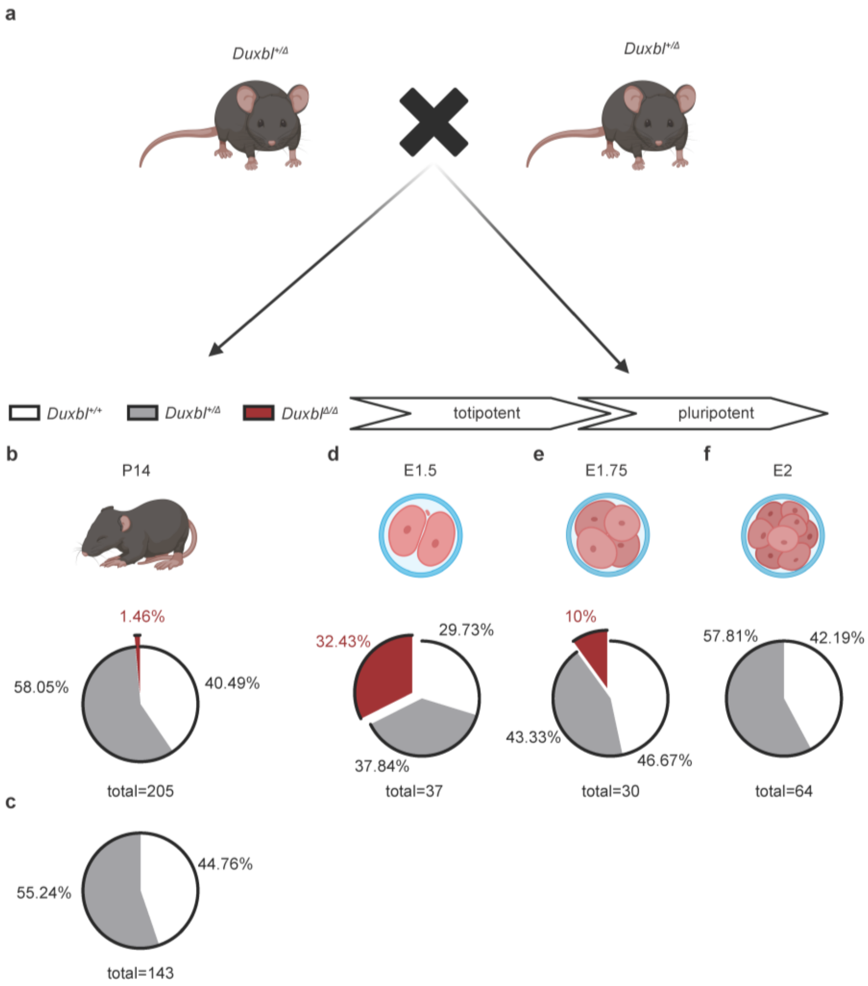

Figure 25: Duxbl is essential for exit from totipotency and preimplantation development:
(a) Schematic illustrating experimental workflow. Heterozygous *Duxbl*$^{+/Δ}$ mice were intercrossed and resulting offspring was genotyped via PCR. **(b,c)** Pie charts illustrating genotype ratios of all pups from heterozygous intercrosses **(b)** and from heterozygous intercrosses after F1 generation **(c)**. **(d-f)** Pie charts showing genotype ratios of single embryos at late 2C-stage (48 h post HCG) **(d)**, at 4C-stage (56 h post HCG) **(e)**, or 8C-stage (69 h post HCG) **(f)**. (Created with BioRender.com)

To generate chimeric founder mice, heterozygous *Duxbl*-knockout (*Duxbl*$^{+/Δ}$) mESCs were injected into blastocysts and implanted into foster mothers. Resulting *Duxbl*$^{+/Δ}$-mice were intercrossed, and obtained pups were genotyped at postnatal day 14 (P14) (Figure 25a-c). Strikingly, from 42 crosses and 205 pups, only three (1.45 %) were homozygous *Duxbl*-knockout mice (*Duxbl*$^{Δ/Δ}$) (Figure 25b). However, after the F2 generation, no *Duxbl*$^{Δ/Δ}$-pups could be obtained from 25 heterozygous intercrosses with 143 newborns at P14 (Figure 25c). These data clearly showed that

Duxbl is required for development. To determine the exact developmental stage at which *Duxbl*$^{\Delta/\Delta}$-embryos die, I genotyped embryos collected from heterozygous intercrosses specifically at late 2C-, 4C- and 8C-stages (Figure 25d-f). While Mendelian ratios of genotypes were obtained at the late 2C-stage (Figure 25d), the percentage of *Duxbl*$^{\Delta/\Delta}$-embryos was dramatically reduced at the 4C-stage (Figure 25e), and no *Duxbl*$^{\Delta/\Delta}$-embryos could be observed at the 8C-stage (Figure 25f). These results were consistent with the observed phenotype in the MO-experiments (Figure 23b,c). Taken together, these data demonstrated that *Duxbl* plays an essential role in preimplantation development during the transition from 2C- to 8C-stage, when the loss of totipotency occurs[120,121].

5.10 Duxbl enables exit from the totipotent stage by suppressing ZGA-associated transcription

To gain mechanistic insight into the role of Duxbl during the exit from totipotency and to assess to what extent the loss of Duxbl might deregulate gene expression, I collected single embryos from *Duxbl*$^{+/\Delta}$-intercrosses at the late 2C-stage (Figure 26a). RNA-seq of these embryos followed by differential gene expression analysis enabled the identification of five *Duxbl*$^{\Delta/\Delta}$-embryos (Figure 26a,b). Importantly, the expression level of the neighboring gene *Cphx* was not affected in *Duxbl*-depleted embryos, confirming the integrity of the *Duxbl*-KO allele (Figure 26b). Differential gene expression analysis revealed that the majority (74 %) of the top 50 DEGs were upregulated in *Duxbl*$^{\Delta/\Delta}$-embryos (Figure 26c). Of note, these upregulated genes comprised ZGA-associated genes like *Zscan4d* and *Zfp560*[137], as well as the *negative elongation factor A* (*NELFA*) (Figure 26c). This is of particular importance because NELFA was only recently identified as a major driver of a 2C-specific gene expression profile in mESCs[173]. These results revealed that Duxbl counteracts the ZGA-specific gene expression at the late 2C-stage *in vivo*.

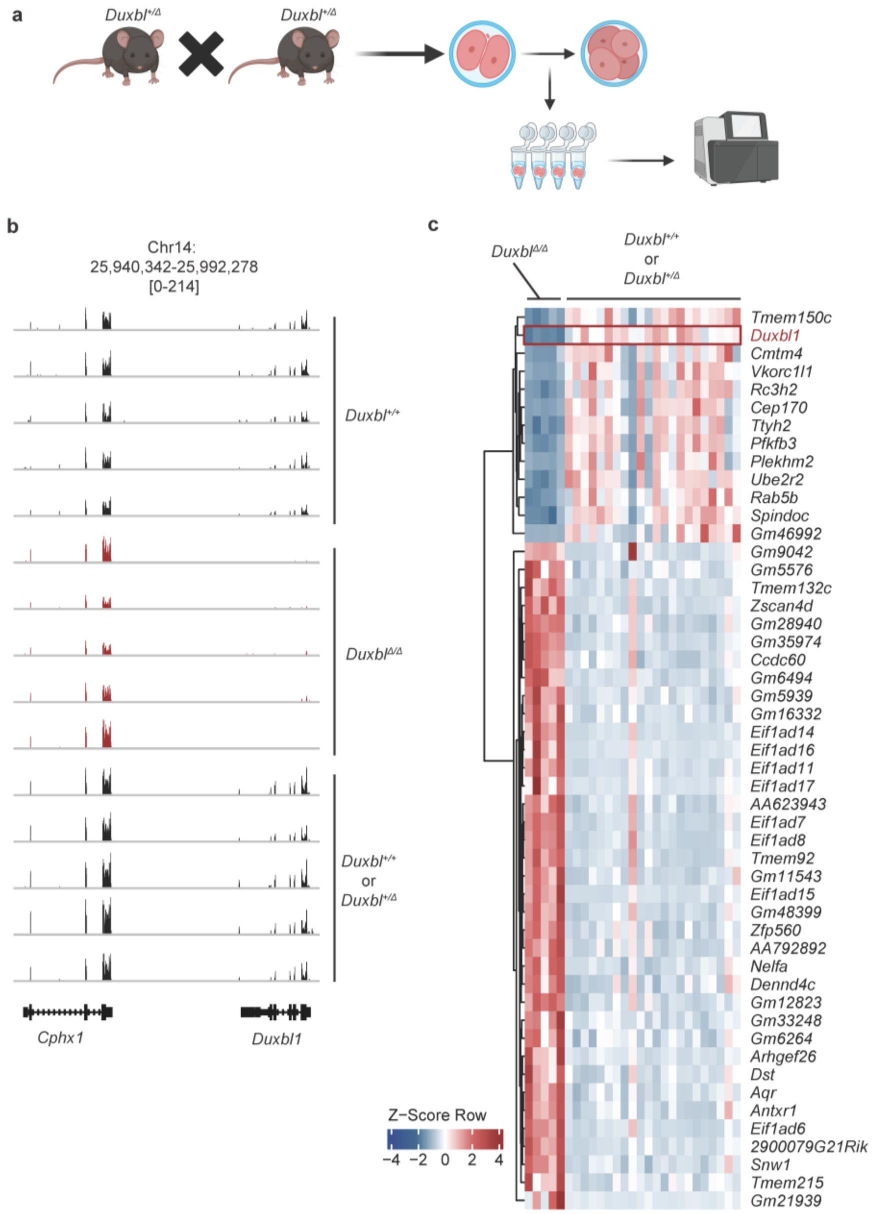

Figure 26: Identification and transcriptional profiling of *Duxbl*$^{Δ/Δ}$ embryos at late 2C-stage: **(a)** Schematic illustration of the experimental workflow. Single embryos from *Duxbl*$^{+/Δ}$- and *Duxbl*$^{+/+}$-intercrosses were collected at late 2C-stage (48 h post HCG) and submitted to RNA-seq. **(b)** Representative sequencing depth normalized genome browser tracks depicting expression levels of *Duxbl1* and *Cphx1.* Embryos labeled "*Duxbl*$^{Δ/Δ}$" and "*Duxbl*$^{+/+}$ or *Duxbl*$^{+/Δ}$" were obtained from the same crossings. **(c)** Heat map showing expression levels of the top 50 DEGs between *Duxbl*$^{Δ/Δ}$ vs. *Duxbl*$^{+/+}$ or *Duxbl*$^{+/Δ}$ at late 2C-stage. *Duxbl*$^{+/+}$: n = 5. "*Duxbl*$^{+/+}$ or *Duxbl*$^{+/Δ}$": n = 22. *Duxbl*$^{Δ/Δ}$: n = 5. (**(a)** Created with BioRender.com)

The idea of Duxbl acting as an antagonist of ZGA was further supported by the finding that substantially more ZGA-genes were upregulated (326) than downregulated (126) upon loss of Duxbl (Figure 27a). Indeed, two-thirds of the genes that were upregulated in *Duxbl*$^{Δ/Δ}$-embryos belonged to the major wave of

ZGA (Figure 27b). Along this line, expression levels of the ZGA-genes that were previously downregulated by precocious Duxbl expression via mRNA-injection (Figure 20b, blue circle) were noticeably elevated in *Duxbl*$^{\Delta/\Delta}$-embryos compared to WT-controls (Figure 27c). Thereby, both independent technical approaches the mRNA-injection and the knockout cross-validated each other, confirming a role for Duxbl as an antagonist of 2C-stage-specific ZGA-gene expression. Indeed, average expression of ZGA-genes (Figure 27d), as well as of Dux-target genes (Figure 27e), and of genes downregulated during the transition from 2C- to 4C-stage (late 2C to 4C, mean read count >= 5, FC <= -2, FDR <= 0.05, data[159]) (Figure 27f) was significantly upregulated in *Duxbl*$^{\Delta/\Delta}$-embryos at late 2C-stage. In line with these findings, genes upregulated during the transition from 2C- to 4C-stage (late 2C to 4C, mean read count >= 5, FC >= 2, FDR <= 0.05, data[159]) displayed significantly lower expression levels in average upon loss of Duxbl (Figure 27g).

Most importantly, all genes that were bound and suppressed by Duxbl in co-overexpression experiments with Dux in mESCs (Figure 16) were substantially downregulated in *Duxbl*$^{\Delta/\Delta}$-embryos (Figure 27h). Remarkably, this also included expression of *Zfp352*, which has been identified as a Dux-target gene *in vitro* that is not altered by loss of Dux *in vivo* and is therefore thought to be transcriptionally activated by a redundant factor (Figure 27h)[141]. This observation implied that Duxbl plays a more important role in the regulation of ZGA than Dux. In line with the developmental arrest at 2C- to 8C-stage (Figure 25 d-f), gene set enrichment analysis displayed strong downregulation of cell cycle-associated genes in *Duxbl*$^{\Delta/\Delta}$-embryos (Figure 27i).

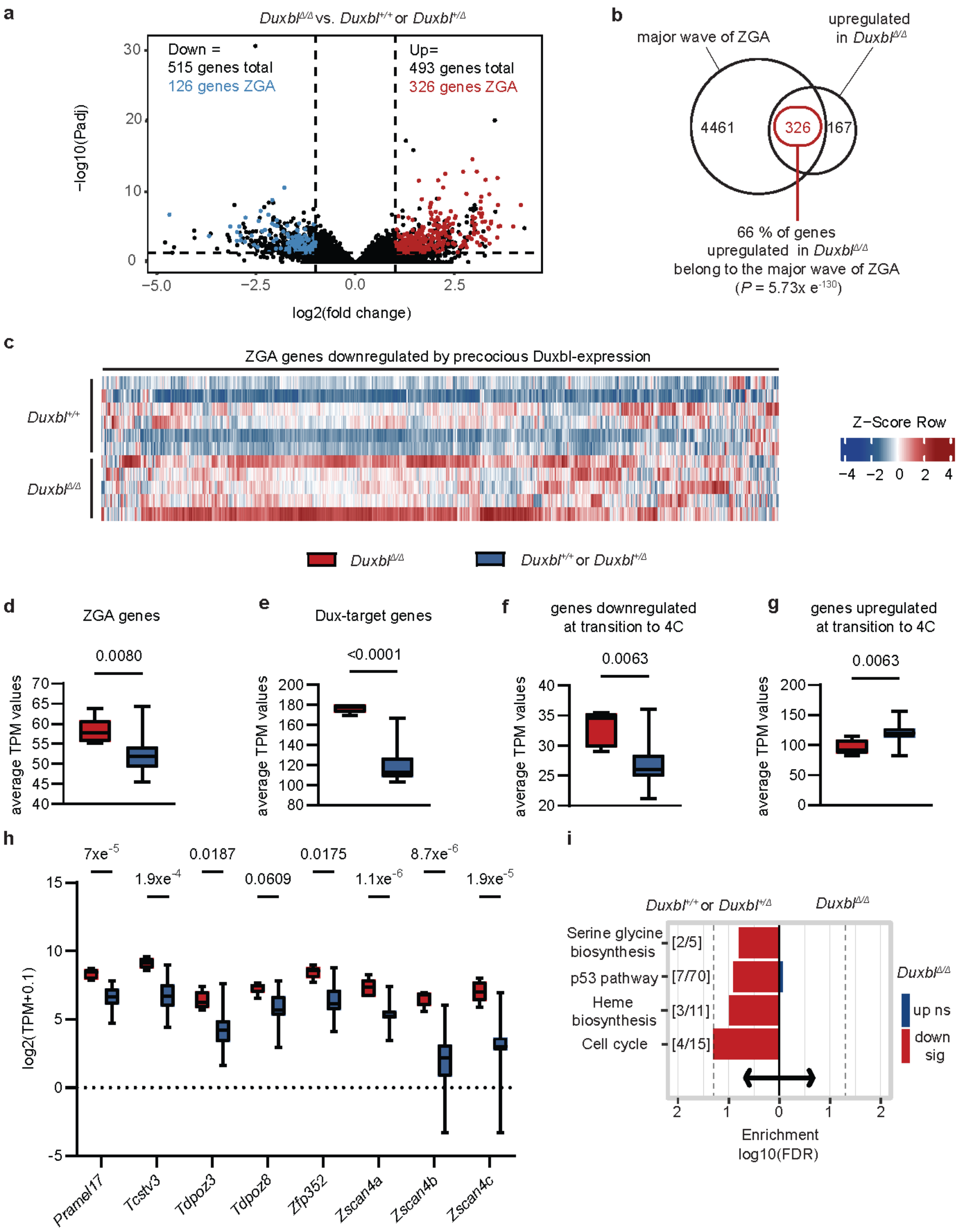

Figure 27: Loss of Duxbl results in insufficient silencing of 2C-specific gene expression:

(a) Volcano plot depicting differential gene expression between *Duxbl*$^{\Delta/\Delta}$-embryos and *Duxbl*$^{+/+}$- or *Duxbl*$^{+/\Delta}$-embryos at late 2C-stage (E1.5). Blue and red dots label down- or upregulated ZGA genes, respectively. **(b)** Venn diagram showing overlap between genes upregulated in *Duxbl*$^{\Delta/\Delta}$-embryos and ZGA-genes. Statistical significance was determined by hypergeometric test. **(c)** Heat map comparing expression levels of ZGA-genes downregulated by precocious Duxbl^{V5}-expression (blue circle in **(Figure 20b)**) in *Duxbl*$^{+/+}$- and *Duxbl*$^{\Delta/\Delta}$-embryos at late 2C-stage. **(d,e)** Box-and-whisker plots (minimum to maximum) comparing average gene expression values in *Duxbl*$^{+/+}$- and *Duxbl*$^{\Delta/\Delta}$-embryos at late 2C-stage for ZGA-genes **(d)**, Dux-target genes **(e)** and genes that are down- **(f)** or upregulated **(g)** upon transition from 2C- to 4C-stage. Mann-Whitney-U test was used to determine statistical significance. **(h)** Box-and-whisker plots (minimum to maximum) showing normalized expression levels of individual ZGA-associated Dux-targets in *Duxbl*$^{+/+}$- and *Duxbl*$^{\Delta/\Delta}$-embryos at late

2C-stage. FDRs were obtained from DESeq2**. (i)** Bar graph displaying results of gene set enrichment analysis for *Duxbl*$^{\Delta/\Delta}$-embryos vs. *Duxbl*$^{+/+}$- or *Duxbl*$^{+/\Delta}$-embryos at late 2C-stage. up ns = not significantly upregulated. down sig = significantly downregulated. Dux-targets were identified using data from[137]. ZGA genes (upregulated from early to mid 2C-stage) and genes differentially expressed at 2C- to 4C-transition were identified using data from[159]. (DEGs: 2 >= FC <= -2; mean read count >= 5; FDR <= 0.05). *Duxbl*$^{+/+}$: n = 6. "*Duxbl*$^{+/+}$ or *Duxbl*$^{+/\Delta}$": n = 22. *Duxbl*$^{\Delta/\Delta}$: n = 5.

In summary, these results demonstrated that *Duxbl* plays an essential role in preimplantation development by decommissioning the major wave of ZGA, enabling the exit from totipotency and the transition towards pluripotency.

6 Discussion

6.1 From ZGA to cell fate decisions

In this study, I discovered an essential role for the TF Duxbl in regulating the 2C-specific transcriptional burst during the major wave of ZGA. Thereby, I was able to shed light on a major unanswered question in preimplantation development: How is the totipotency-associated gene expression pattern so rapidly decommissioned after the 2C-stage to enable progression towards pluripotency and segregation into the very first cell lineages?

The emergence of totipotency and the ZGA-associated transcriptional program have been subject of extensive research[21,22]. Major recent achievements have been the identification of epigenetic modifiers and TFs that are involved in shaping the chromatin landscape and gene expression profile during ZGA, as for example Dux or more recently Nr5a2[137,174]. However, although it was shown that silencing of the 2C-specific expression profile per se is pivotal for development after the 4C-stage[170,175], the factors that enable this remained elusive. In this context, it is worth mentioning that the transposable element *long interspersed nuclear element 1* (*LINE1*) can act as an RNA scaffold to enable recruitment of a repressive protein complex consisting of Nucleolin and Trim28 that silences Dux expression[175]. While this mechanism can explain how Dux expression itself is rapidly decommissioned at the 2C-stage (Figure 11b), this is likely not sufficient to silence the Dux-induced transcriptional burst itself. This idea is based on the finding that the Dux-target Zscan4c can bind genomic loci of ZGA-genes and activate their expression[176]. These findings, together with the observation that overexpression of the zinc finger domains of Zscan4c in mESCs resulted in upregulation of endogenous *Zscan4*[176], suggest that some Dux-targets drive a positive feedback loop that enhances the Dux-initiated ZGA-cascade. Therefore, it is conceivable that repression of Dux alone would likely not be sufficient to terminate the 2C-specific ZGA-burst (Figure 28).

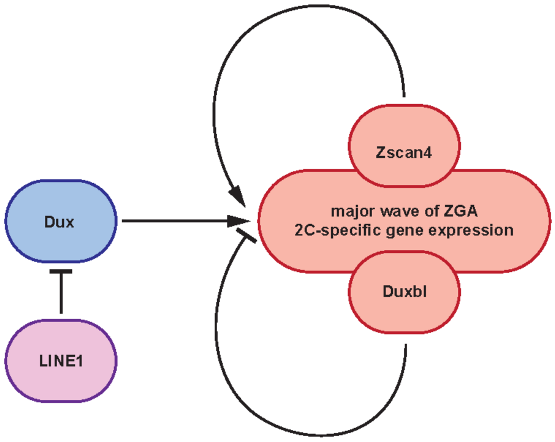

Figure 28: Duxbl as an antagonist of 2C-specific gene expression:

Schematic illustrating a model for the regulation of the major wave of ZGA in mice by Dux, LINE1, Zscan4 and Duxbl[137,175-177]. (Created with BioRender.com)

In this study, I discovered that this transcriptional program is antagonized by Duxbl, enabling the transition to the 4C- and 8C-stages in mice. At this time point, the developmental potential of the embryo changes as individual cells start to undergo the very first lineage segregation towards ICM and trophectoderm[7]. I demonstrated that Duxbl functions as an essential regulator that enables the exit from totipotency in a negative feedback loop.

6.2 Regulation of the 2C-program in a negative feedback loop

6.2.1 Splicing-dependent posttranscriptional regulation of Dux and Duxbl

Negative feedback loops play a significant role in cell fate decisions[178]. In this process of cellular differentiation, these loops manifest when an activator not only changes the fate of a cell but also concurrently elicits expression of an antagonist, thus restricting its own actions[178]. In the context of ZGA, Dux is expressed at late 1C- to early 2C-stage and is thought to drive expression of ZGA-specific genes at mid to late 2C-stage[137]. Among those genes is the Dux-paralog *Duxbl* (Figure 11). Notably, although *Duxbl*-expression is slightly deferred to Dux, *Duxbl*-transcripts can already be detected at the early 2C-stage and, thus immediately prior to the major wave of ZGA[164]. Interestingly, before the major wave of ZGA occurs, the murine embryo lacks an effective splicing machinery[98]. The impact of this phenomenon on cellular identity has been observed *in vitro* by artificially initiating a pluri- to totipotent transition[179]. Remarkably, administering the spliceosome inhibitor Pladienolide B elicits reconfiguration of gene expression dynamics in cultured mESCs including downregulation of pluripotency-associated genes like *Pou5f1* and *Nanog*, and concurrent upregulation of totipotency-related markers like the *Zscan4*-cluster, *Zfp352* and *MERVL*[179]. This phenomenon underscores a recurring pattern wherein genes related to early development exhibit a predilection for reduced or absent introns[114]. Thus, since Dux is an intron-less gene, one might expect Dux to be translated directly upon transcription at the late 1C-stage, while Duxbl, which is encoded by five exons, would only be functionally expressed as a protein at mid to late 2C-stage. Based on this idea, a reasoned conjecture is that Dux contributes to the initiation of the major wave of ZGA at late 1C- to early 2C-stage, while Duxbl counteracts the transcriptional program in a negative feedback loop from mid 2C-stage onwards to late 2C-stage. This suggests an additional layer of regulation, potentially involving a splicing-dependent posttranscriptional regulatory switch in preimplantation development that relies on the functional dynamics of Dux and Duxbl.

Future studies employing immunofluorescence with specific antibodies against Dux and Duxbl will be important to validate this hypothesis and to understand how and if uncoupling of transcription and translation might regulate development by titrating the expression of an activator and a repressor.

6.2.2 Antagonizing the major wave of ZGA *in vivo*

Almost five decades ago, Warner and Versteegh demonstrated that ZGA is an essential step in development[166]. They demonstrated that inhibiting RNA Pol II using the toxin α-aminitin in 2C-stage murine embyros halted developmental progression[166]. Building upon their findings, following work by Flach and colleagues using the same approach at different time points further unraveled the temporal dynamics of transcription in the embryo[164]. This investigation led to the separation of ZGA into two waves: the minor wave of ZGA at late 1C- to early 2C-stage and the major wave of ZGA at late 2C-stage[164]. While both studies clearly demonstrated the importance of the major wave of ZGA, the role of minor ZGA remained incompletely understood. More recent work employing the chemical 5,6-dichloro-1-β-d-ribofuranosyl-benzimidazole (DRB), which also inhibits RNA Pol II, revealed that the minor wave of ZGA is also necessary for development[115]. In fact, it was revealed that the minor wave of ZGA enables the preconfiguration of RNA Pol II for the major wave of ZGA[97]. Taking these findings into account, it is remarkable that prematurely inducing the expression of Duxbl alone at late 1C-stage is sufficient to induce complete developmental arrest, similar to chemically inhibiting the entire transcriptional machinery. Nevertheless, even though Duxbl suppresses a substantial portion of the major wave of ZGA (34.1 %), it cannot be completely ruled out that precociously expressing Duxbl via mRNA-injection at the late 1C-stage could interfere with the minor wave of ZGA. Conceivably, RNA-seq analysis of embryos at late 1C- and early 2C-stages, shortly after microinjection of Duxbl-encoding mRNA, may answer this question.

Indeed, similar mRNA-based microinjection methods were employed by Guo and colleagues to investigate the implications of prolonged Dux expression beyond the early 2C-stage[170]. This resulted in sustained expression of *Zscan4*, *Zfp352*, *MERVL,* as well as other Dux-targets, and was accompanied by a 4C-stage arrest[170]. Furthermore, the treatment of embryos with antisense oligonucleotides targeting LINE1 was shown to lead to sustained Dux expression at the 4-cell stage, resulting in elevated expression levels of 2-cell-specific ZGA targets and developmental arrest at the totipotent stage[175]. While these studies clearly demonstrated the necessity of rapidly switching off 2C-specific gene expression, they only focused on the transcriptional control of Dux. However, the posttranscriptional stability and, thus, the activity of Dux remained enigmatic. By reason, direct regulation of Dux-target gene expression is likely a pivotal factor for successful preimplantation stage development. Indeed, here I found that the developmental arrest at 4C-stage induced by loss of Duxbl correlated with upregulation of the same ZGA-specific genes *in vivo* that were bound by Duxbl *in vitro* (Figure 16, Figure 27h). In addition, I could show that *Duxbl*-KO embryos exhibit significantly elevated levels of *NELFA* gene expression (Figure 26c), which has recently been shown to drive 2C-like conversion in mESCs *in vitro*[173]. Taken together, these preliminary findings hint towards a potential role for NELFA in the regulation of the 2C-specific transcriptional profile *in vivo*.

Notably, while ectopic overexpression of Dux elicits strong expression of the ZGA-gene *Zfp352 in vitro* and *in vivo*, *Zfp352* expression is not significantly altered upon inactivation of Dux *in vivo*[137,141,170]. This underscores the notion of Dux as a non-essential factor in ZGA, which is reinforced by the non-lethal phenotype of *Dux*-KO mice[141]. This has led to the recent proposal that the initiation of ZGA upon loss of Dux might be redundantly compensated by another homeobox TF, called Obox4[180]. However, more recent work by Ji and colleagues challenged this concept, postulating that a combination of multiple Obox factors (Obox1/2/3/4/5/7) enables ZGA, independently of Dux[181]. Nevertheless, these studies collectively emphasize that ZGA is redundantly secured by several TFs. In contrast, my own findings have revealed that *Zfp352* expression is strongly upregulated upon loss of Duxbl, highlighting the essential role of Duxbl in the decommissioning of the ZGA-specific expression pattern.

It will be interesting to see whether Obox factors activate Duxbl in *Dux*-KO embryos and if Duxbl similarly counteracts Obox-mediated transcription as observed with Dux-mediated transcription.

6.2.3 Can a 2C-like state be stabilized by co-expression of Dux and Duxbl *in vitro*?

Beyond the importance of fully understanding preimplantation development *in vivo*, it is becoming increasingly important to understand the precise temporal regulation of the totipotentcy-associated transcriptional program *in vitro*. Indeed, a major goal of ongoing research is to find ways to stabilize the totipotent state[182]. This would open new avenues to better understand the developmental program during the first cleavages by using techniques that demand much larger amounts of input material that cannot be obtained from embryos[182]. In this regard, the establishment of a stable totipotent cell line would greatly facilitate research focused on determining the molecular basis of differentiation processes before, during and after implantation since such cells, in contrast to pluripotent stem cells, could possess the ability to differentiate into extraembryonic tissue[182]. For example, Yang and colleagues recently derived expanded potential stem cells (EPSCs) from 8C-stage murine embryos and were able to sustain the growth of those cells in a defined inhibitor cocktail *in vitro*[183]*.* In contrast to conventionally cultured mESCs, EPSCs exhibited elevated expression levels of early cleavage stage-specific genes and could contribute to embryonic and extraembryonic lineages when injected into morula-stage embryos[183]. However, EPSCs originate from 8C-stage embryos and individual blastomeres at this developmental stage do not support embryonic development on their own[123-125]. Consequently, EPSCs only fulfill a less strict definition of totipotency[1]. Notably, further work by Posfai et al. that compared EPSCs to embryos of different preimplantation stages revealed that the transcriptome of EPSCs appears to resemble that of pluripotent epiblast rather than that of totipotent cleavage stages[184]. In their work, EPSCs were only able to contribute to extraembryonic tissue in chimeric embryos *in vivo* to a marginal extent[184]. Thus, EPSCs appear not to be ideal for comprehensive biochemical assessments targeting totipotent features or physiological studies focused on segregation processes leading to distinct cell fates.

In contrast to this chemically induced approach, an alternative method for *in vitro* modeling of the totipotent state involves utilizing 2C-like cells[182]. Unlike EPSCs, 2C-like cells do not express markers of pluripotent epiblast or trophectoderm and closely resemble the transcriptome of 2C-embryos[185]. This characteristic appears to make them more suitable to study totipotency-associated features, such as chromatin landscape, protein complexes of epigenetic modifiers, or binding of TFs. However, the challenge with 2C-like cells is that they emerge from conventionally cultured mESCs at very low rates, with individual cells only transiently cycling through this state[112]. Consequently, prior enrichment steps are necessary to obtain sufficient amounts of 2C-like cells for biochemical experiments[186]. Importantly, the efficiency of the 2C-like conversion can be substantially increased by ectopic expression of Dux[137]. Nonetheless, overexpression of Dux can lead to increased DNA damage and induce cell death[187], potentially skewing the model towards an apoptotic state and hindering the stable culture of 2C-like cells. The cytotoxic effect of Dux is dependent on its ability to bind genomic targets and to induce global hyperacetylation of H3K27[138,139,147]. Considering that Duxbl can occupy the same genetic regulatory elements as Dux and counteract Dux-mediated transcription (Figure 16), one might speculate that Duxbl may be able to antagonize Dux-induced cell death. Conceivably, co-expression of Dux and Duxbl in mESCs using either two independent titratable systems or expressing both TFs in tandem utilizing a polycistronic 2A-self-cleaving peptide approach might create a viable equilibrium[188-190]. This approach might potentially enable stable expansion of a viable 2C-like state *in vitro*, which could facilitate gaining new insights into the molecular processes during preimplantation development. Moreover, adding other TFs that regulate differentiation processes to this model might promote biomedical research pertaining to implantation.

6.3 A Duxbl-mediated mechanism for the exit from totipotency

This study provides clear evidence that Duxbl plays an indispensable role in silencing the 2C-specific gene expression program. However, the precise molecular mechanisms that orchestrate this process still remains incompletely understood. The regulation of transcription during the transition from 1C- to 4C-stage involves a complex interplay of multiple mechanisms. During the switch from the minor to the major wave of ZGA in the 1C- to 2C-transition, a substantial reduction of broad H3K4me3- and H3K27ac-domains to their canonical forms occurs[73,76]. Mechanistically, previous establishment of the broad H3K27ac-domains relies on CBP/EP300, while their reduction is dependent on HDACs[76]. Interestingly, Dux has been shown to induce hyperacetylation of H3K27 via CBP/EP300 *in vitro*[139], thus potentially contributing to the emergence of the non-canonical H3K27ac at 1C-stage. Importantly, I could show that Duxbl does not interact with EP300 and binds to the same Dux-occupied ZGA-genes (Figure 13c, Figure 16). Therefore, it can be speculated that Duxbl might counteract the spread of H3K27ac from the 2C-stage onwards, potentially by directly competing with the Dux-CBP/EP300-complex for common DNA binding sites. Additionally, Duxbl might recruit HDACs to these sites and thereby actively reduce H3K27ac to its canonical pattern.

To explore the extent to which Duxbl contributes to constraining transcriptional activity during the 2C-stage, future studies are clearly warranted. These could involve low-input DNA-binding profiling techniques for Dux and Duxbl during 1C- to 2C-transition *in vivo* or ChIP-seq for Duxbl and H3K27ac upon co-overexpression of Dux and Duxbl in mESCs.

The transitioning from the 2C- to 4C-stage entails the deactivation of the 2C-specific gene expression pattern[140,170,175]. In this study, I demonstrated that Duxbl binds ZGA-associated Dux-targets and lacks transactivation capacity *in vitro.* Notably, I could show that precocious expression of Duxbl is sufficient to suppress ZGA and that Duxbl expression is necessary to downregulate ZGA as shown by the consequences of its loss-of-function *in vivo* (Figure 20, Figure 23, Figure 25, Figure

26, Figure 27). Taken together, Duxbl likely engages in direct competition with Dux for binding to ZGA-promoter sites, displacing the transcription machinery at these targets and contributing to the silencing of 2C-specific gene expression. Further elucidation of this model could be achieved by profiling the potential displacement of RNA Pol II at ZGA-promoters upon precocious Duxbl expression in 2C-embryos using the Stacc–seq technique developed by Liu and colleagues[97]. However, this will be challenging since this approach will require the microinjection of Duxbl-encoding mRNA into several hundred embryos.

Interestingly, the process of Duxbl-dependent silencing of 2C-specific gene expression appears to intersect with the deposition of H3K9me3 and subsequent heterochromatinization at the LTRs of MERVLs[85]. Since these LTRs serve as alternative promoters for the expression of 2C-specific genes, their silencing is presumably pivotal for the termination of the associated 2C-specific transcriptional program[112]. *In vitro* studies have indicated that TRIM28 and H3K9 methyltransferases, such as G9a, govern this process[191]. TRIM28 can act as a scaffold protein that recruits a complex machinery of repressive epigenetic modifiers, including another H3K9 methyltransferase, SETDB1, and HP1[192]. TRIM28 itself is normally recruited by other factors, providing sequence specificity[192]. Therefore, while Dux has been shown to drive expression of ZGA-genes from retrovirus-derived LTRs, Duxbl might recruit TRIM28 and, consequently, a repressive complex consisting of HP1, SETDB1 or G9a. To validate this hypothesis, future studies will need to elucidate the interaction partners of Duxbl and decipher the molecular mechanisms behind Duxbl-mediated silencing of the 2C-specific program. In this context, co-immunoprecipitation coupled with mass spectrometry analysis following overexpression of Duxbl in 2C-like cells could be a promising approach.

6.4 The DUX-family: Implications for ZGA in other species, disease and regeneration

It will be fascinating to explore whether other homologs of Duxbl function in similar ways in other species. Indeed, several DUX-factors have been identified in numerous placental mammals[143]. Besides humans and mice, this family of TFs has been found in other primates, such as gorillas or chimpanzees, as well as in glires, like rats and rabbits, or Laurasiatheria, like horses and dolphins[143]. Thus, this broad distribution prompts to consider extrapolating findings from this study and others to a diverse array of mammalian species in the context of preimplantation stage development and beyond. Notably, the genomes of several mammals do not encode for *Dux* or *DUX4*, but they do harbor an open reading frame for *DUXC*[143]. Unlike the intronless *Dux* and *DUX4*, *DUXC* contains several exons[143]. Despite this difference, the open reading frames of all three DUX factors encode for two homeobox domains that show substantial sequence similarities[143]. Moreover, the C-terminal domain of DUXC, exhibits substantial sequence conservation with the C-TAD of DUX4 *and* Dux[143]. Intriguingly, it has been shown that ectopic expression of DUXC in porcine and canine muscle cells induces a cleavage-stage associated transcriptional pattern, including expression of ZGA-specific genes and retrotransposons[193,194]. Thus, a similar role for DUXC in driving this program during preimplantation stage development can be assumed, indicating a cross-species conserved role for DUX TFs. Along this line, DUX-factors that appear to not contain a C-TAD, like Duxbl, also seem to be conserved across mammalian genomes[143]. Potential functional homologs are DUXA and DUXB, which have been identified in Laurasiatheria and several primates, including humans[143].

This raises the question whether the molecular mechanism of not only initiation but also of decommissioning of ZGA-specific gene expression might be evolutionary conserved in different mammalian species. Interestingly, DUXA and DUXB are direct transcriptional targets of DUX4 in human pluripotent stem cells and are both expressed during the human major wave of ZGA *in vivo*, which supports this idea[137,177,195]. Along this line, I was able to show in collaboration with Bosnakovski et al. that, similar to Duxbl, DUXA is not able to elicit expression of DUX4-target genes but is capable of binding to DUX4-occupied promoters and can suppress DUX4-mediated transcription when co-overexpressed *in vitro*[196]. These results strongly suggest that DUXA might play a significant role in the regulation of the ZGA-specific gene expression program during human early embryonic development, similar to Duxbl in murine preimplantation development. It will be interesting to investigate whether other functional homologs of Duxbl and DUXA may exist in species in which ZGA might be activated by DUXC[193,194].

Importantly, besides their role in preimplantation development, DUX factors have been linked to pathophysiological conditions[145,197,198]. For instance, DUX4 expression has been identified as the causative factor of facioscapulohumeral muscular dystrophy (FSHD) in humans[199,200]. This degenerative disease of skeletal muscle has been linked to insufficient silencing of the DUX4-encoding D4Z4-array and subsequent ectopic expression of DUX4[201-203]. Consequently, this misexpression of DUX4 causes cell death in the affected muscle tissue[204]. While the exact mechanism of DUX4-induced cell death still remains under discussion[205,206], intriguingly, FSHD-muscle tissue exhibits a ZGA-associated transcriptional profile[146,207]. To date, no cure for this disease exists, and approved drug treatments are lacking. However, a recent study demonstrated that an inhibitor blocking the interaction of the C-TAD of DUX4 with CBP/EP300 effectively suppresses DUX4-mediated transcription and the cytotoxic effect of DUX4 expression[139]. Since DUXA is able to suppress the DUX4-induced transcriptional program[196], it would be interesting to investigate whether DUXA is also capable of alleviating cell death.

Remarkably, a murine model of FSHD carrying the DUX4-encoding D4Z4 array displayed expression of DUX4 and, subsequently, Duxbl during muscle regeneration[208]. While it has been shown that Duxbl overexpression in a myogenic cell line inhibits differentiation and promotes proliferation *in vitro*[151], a role for Duxbl in regeneration *in vivo* has yet to be elucidated. In this context, future studies might reveal whether Dux is also physiologically expressed during murine muscle regeneration and if the Dux-Duxbl axis plays a role in other processes apart from preimplantation development.

7 Abbreviations

µg	microgram
16C	16-cell
2C	2-cell
32C	32-cell
3D	three-dimensional
4C	4-cell
5mC	methylation of cytosine
5'UTR	five prime untranslated region
64C	64-cell
8C	8-cell
aa	amino acids
ATP	adenosine triphosphate
ANOVA	analysis of variance
BSA	bovine serum albumin
CAG	cytomegalovirus enhancer fused to the chicken beta-actin promoter
CBP	CREB-binding protein
cCREs	database for candidate cis-regulatory elements
cDNA	complementary deoxyribonucleic acid
ChIP-qPCR	Chromatin immunoprecipitation coupled to quantitative polymerase chain reaction
ChIP-seq	Chromatin immunoprecipitation coupled to high-throughput DNA sequencing
CO_2	carbon dioxide
Co-IP	Co-immunoprecipitation
CpG	cytosine-guanine dinucleotide
CRISPR	clustered regularly interspaced short palindromic repeats
C-TAD	C-terminal activation domain
CTD	C-terminal domain
CT-value	cycle threshold
DAPI	4′,6-diamidino-2-phenylindole
DEGs	differentially expressed genes
DNA	deoxyribonucleic acid
DNMT1	deoxyribonucleic acid-methyltransferase 1
DRB	5,6-dichloro-1-β-d-ribofuranosyl-benzimidazole
DUX	double homeobox transcription factor
Duxbl$^{+/\Delta}$	heterozygous *Duxbl*-knockout
Duxbl^{V5}	V5-epitope tagged Duxbl
Duxbl$^{\Delta/\Delta}$	homozygous Duxbl-knockout
E0.5	embryonic day 0.5
E1.5	embryonic day 1.5
E2.5	embryonic day 2.5
E3.5	embryonic day 3.5

E4.5	embryonic day 4.5
EDTA	ethylenediaminetetraacetic acid
EF1a	eukaryotic translation elongation factor 1 alpha
ENCODE	Encyclopedia of DNA Elements Project
EPSCs	expanded potential stem cells
ESCs	embryonic stem cells
ES-E14TG2a	blastocyst-derived mouse embryonic stem cell line
FC	fold change
FDR	false discovery rate
FSHD	facioscapulohumeral muscular dystrophy
G2 phase	gap 2 phase
GC-rich	guanine-cytosine-rich
h	hour
H_2O	water
H3K27ac	acetylation of histone H3 at lysine 27
H3K27me3	trimethylation of histone H3 at lysine 27
H3K4me1	monomethylation of histone H3 at lysine 4
H3K4me3	trimethylation of histone H3 at lysine 4
H3K9me3	trimethylaiton of histone H3 at lysine 9
HA	human influenza hemagglutinin
HAT	histone acetyltransferase
HCG	human chorionic gonadotropin
HDAC	histone deacetylase
HEK293	human embryonic kidney cells 293
HERVL	human endogenous retrovirus with leucine tRNA primer
Hi-C	high-resolution chromosome conformation capture
HP1	heterochromatin protein 1
ICM	inner cell mass
IF	immunofluorescence
IgG	immunoglobulin G
IP	immunoprecipitation
IU	international unit
IVT	*in vitro* transcription
kDa	kilodalton
KDM	lysine demethylase
KO	knockout
LiCl	lithium chloride
LINE1	long interspersed nuclear element 1
log2FC	binary logarithm of fold change
LTR	long terminal repeat
M	molar
MERVL	murine endogenous retrovirus with leucine tRNA primer
MES	2-(N-morpholino)ethanesulfonic acid
mESCs	murine embryonic stem cells
mg	milligram

min	minute
ml	milliliter
mm	millimeter
mM	millimoles
MO	morpholino
mRNA	messenger ribonucleic acid
MW	molecular weight
MZT	maternal-to-zygotic transition
n	number of replicates
NaCl	sodium chloride
NELFA	negative elongation factor A
ng	nanogram
NLS	nuclear localization signal
^{NLS}mCherry	nuclear localization signal tagged mCherry
nt	nucleotide
Oct4	octamer-binding transcription factor 4
OE	overexpression
P14	postnatal day 14
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pou5f1	POU domain class 5 transcription factor 1
PTM	posttranslational modification
RIN value	ribonucleic acid integrity number value
RIPA buffer	radioimmunoprecipitation assay buffer
RNA	ribonucleic acid
RNA Pol II	ribonucleic acid polymerase II
RNA-seq	ribonucleic acid sequencing / sequencing for whole transcriptome profiling
rtTA	reverse tetracycline-controlled transactivator
S phase	synthesis phase
SDS	sodium dodecyl sulfate
sgRNA	single guide ribonucleic acid
Stacc-seq	small scale transpoase5-assisted chromatin cleavage with sequencing
TAD	topology associated domain
TAE	tris-acetate ethylenediaminetetraacetic acid
TBS	tris buffered saline
TE	transposable element
TE	tris ethylenediaminetetraacetic acid
Tet	Ten-eleven translocation
TF	transcription factor
Tn5	transposase 5
TRE	tetracycline response element
TSS	transcription start site
5‘-UTR	five prime untranslated region

V5	epitope tag derived from a small epitope present on the P and V proteins of the paramyxovirus of simian virus
WT	wild type
ZGA	zygotic genome activation

8 Literature

1 Condic, M. L. Totipotency: what it is and what it is not. *Stem Cells Dev* **23**, 796-812, doi:10.1089/scd.2013.0364 (2014).

2 Jungnickel, M. K., Sutton, K. A. & Florman, H. M. In the beginning: Lessons from fertilization in mice and worms. *Cell* **114**, 401-404, doi:Doi 10.1016/S0092-8674(03)00648-2 (2003).

3 Perez-Cerezales, S. *et al.* The oviduct: from sperm selection to the epigenetic landscape of the embryo. *Biol Reprod* **98**, 262-276, doi:10.1093/biolre/iox173 (2018).

4 Li, S. & Winuthayanon, W. Oviduct: roles in fertilization and early embryo development. *J Endocrinol* **232**, R1-R26, doi:10.1530/JOE-16-0302 (2017).

5 Wang, H. & Dey, S. K. Roadmap to embryo implantation: clues from mouse models. *Nat Rev Genet* **7**, 185-199, doi:10.1038/nrg1808 (2006).

6 Piliszek, A., Grabarek, J. B., Frankenberg, S. R. & Plusa, B. Cell fate in animal and human blastocysts and the determination of viability. *Mol Hum Reprod* **22**, 681-690, doi:10.1093/molehr/gaw002 (2016).

7 Mihajlovic, A. I. & Bruce, A. W. The first cell-fate decision of mouse preimplantation embryo development: integrating cell position and polarity. *Open Biol* **7**, doi:10.1098/rsob.170210 (2017).

8 Vastenhouw, N. L., Cao, W. X. & Lipshitz, H. D. The maternal-to-zygotic transition revisited. *Development* **146**, doi:10.1242/dev.161471 (2019).

9 Xia, W. & Xie, W. Rebooting the Epigenomes during Mammalian Early Embryogenesis. *Stem Cell Reports* **15**, 1158-1175, doi:10.1016/j.stemcr.2020.09.005 (2020).

10 Xu, R., Li, C., Liu, X. & Gao, S. Insights into epigenetic patterns in mammalian early embryos. *Protein Cell* **12**, 7-28, doi:10.1007/s13238-020-00757-z (2021).

11 Roze, D. Disentangling the benefits of sex. *PLoS Biol* **10**, e1001321, doi:10.1371/journal.pbio.1001321 (2012).

12 Petronczki, M., Siomos, M. F. & Nasmyth, K. Un menage a quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**, 423-440, doi:10.1016/s0092-8674(03)00083-7 (2003).

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

14 Wang, Q. T. *et al.* A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev Cell* **6**, 133-144, doi:10.1016/s1534-5807(03)00404-0 (2004).

15 Medvedev, S., Pan, H. & Schultz, R. M. Absence of MSY2 in mouse oocytes perturbs oocyte growth and maturation, RNA stability, and the transcriptome. *Biol Reprod* **85**, 575-583, doi:10.1095/biolreprod.111.091710 (2011).

16 Yu, J., Hecht, N. B. & Schultz, R. M. Expression of MSY2 in mouse oocytes and preimplantation embryos. *Biol Reprod* **65**, 1260-1270, doi:10.1095/biolreprod65.4.1260 (2001).

17 Yu, J., Hecht, N. B. & Schultz, R. M. RNA-binding properties and translation repression in vitro by germ cell-specific MSY2 protein. *Biol Reprod* **67**, 1093-1098, doi:10.1095/biolreprod67.4.1093 (2002).

18 Oh, B., Hwang, S., McLaughlin, J., Solter, D. & Knowles, B. B. Timely translation during the mouse oocyte-to-embryo transition. *Development* **127**, 3795-3803, doi:10.1242/dev.127.17.3795 (2000).

19 Oh, B., Hwang, S. Y., Solter, D. & Knowles, B. B. Spindlin, a major maternal transcript expressed in the mouse during the transition from oocyte to embryo. *Development* **124**, 493-503, doi:10.1242/dev.124.2.493 (1997).

20 Temeles, G. L. & Schultz, R. M. Transient polyadenylation of a maternal mRNA following fertilization of mouse eggs. *J Reprod Fertil* **109**, 223-228, doi:10.1530/jrf.0.1090223 (1997).

21 Vallot, A. & Tachibana, K. The emergence of genome architecture and zygotic genome activation. *Curr Opin Cell Biol* **64**, 50-57, doi:10.1016/j.ceb.2020.02.002 (2020).

22 Ladstatter, S. & Tachibana, K. Genomic insights into chromatin reprogramming to totipotency in embryos. *J Cell Biol* **218**, 70-82, doi:10.1083/jcb.201807044 (2019).

23 Bickmore, W. A. & van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell* **152**, 1270-1284, doi:10.1016/j.cell.2013.02.001 (2013).
24 Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harb Perspect Biol* **2**, a003889, doi:10.1101/cshperspect.a003889 (2010).
25 Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293, doi:10.1126/science.1181369 (2009).
26 Fritz, A. J., Sehgal, N., Pliss, A., Xu, J. & Berezney, R. Chromosome territories and the global regulation of the genome. *Genes Chromosomes Cancer* **58**, 407-426, doi:10.1002/gcc.22732 (2019).
27 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).
28 Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341-1345, doi:10.1126/science.aau0320 (2018).
29 Gassler, J. *et al.* A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J* **36**, 3600-3618, doi:10.15252/embj.201798083 (2017).
30 Battulin, N. *et al.* Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol* **16**, 77, doi:10.1186/s13059-015-0642-0 (2015).
31 Du, Z. *et al.* Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232-235, doi:10.1038/nature23263 (2017).
32 Flyamer, I. M. *et al.* Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110-114, doi:10.1038/nature21711 (2017).
33 Jung, Y. H. *et al.* Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep* **18**, 1366-1382, doi:10.1016/j.celrep.2017.01.034 (2017).
34 Ke, Y. *et al.* 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell* **170**, 367-381 e320, doi:10.1016/j.cell.2017.06.029 (2017).

35 Wang, Y. *et al.* Reprogramming of Meiotic Chromatin Architecture during Spermatogenesis. *Mol Cell* **73**, 547-561 e546, doi:10.1016/j.molcel.2018.11.019 (2019).
36 Di Pierro, M., Cheng, R. R., Lieberman Aiden, E., Wolynes, P. G. & Onuchic, J. N. De novo prediction of human chromosome structures: Epigenetic marking patterns encode genome architecture. *Proc Natl Acad Sci U S A* **114**, 12126-12131, doi:10.1073/pnas.1714980114 (2017).
37 Sun, X. Y., Zhang, J. & Cao, C. W. CTCF and Its Partners: Shaper of 3D Genome during Development. *Genes-Basel* **13**, doi:ARTN 1383

10.3390/genes13081383 (2022).
38 Li, F. F., An, Z. Y. & Zhang, Z. H. The Dynamic 3D Genome in Gametogenesis and Early Embryonic Development. *Cells-Basel* **8**, doi:ARTN 788

10.3390/cells8080788 (2019).
39 Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260, doi:10.1038/38444 (1997).
40 Kornberg, R. D. & Klug, A. The nucleosome. *Sci Am* **244**, 52-64, doi:10.1038/scientificamerican0281-52 (1981).
41 Rathke, C., Baarends, W. M., Awe, S. & Renkawitz-Pohl, R. Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta* **1839**, 155-168, doi:10.1016/j.bbagrm.2013.08.004 (2014).
42 Rodman, T. C., Pruslin, F. H., Hoffmann, H. P. & Allfrey, V. G. Turnover of Basic Chromosomal-Proteins in Fertilized-Eggs - a Cytoimmunochemical Study of Events Invivo. *J Cell Biol* **90**, 351-361, doi:DOI 10.1083/jcb.90.2.351 (1981).
43 Jenkins, T. G. & Carrell, D. T. Dynamic alterations in the paternal epigenetic landscape following fertilization. *Front Genet* **3**, 143, doi:10.3389/fgene.2012.00143 (2012).
44 Kawamura, M., Funaya, S., Sugie, K., Suzuki, M. G. & Aoki, F. Asymmetrical deposition and modification of histone H3 variants are essential for zygote development. *Life Sci Alliance* **4**, doi:10.26508/lsa.202101102 (2021).

45 Nashun, B., Yukawa, M., Liu, H., Akiyama, T. & Aoki, F. Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. *Development* **137**, 3785-3794, doi:10.1242/dev.051805 (2010).

46 Funaya, S. & Aoki, F. Regulation of zygotic gene activation by chromatin structure and epigenetic factors. *J Reprod Dev* **63**, 359-363, doi:10.1262/jrd.2017-058 (2017).

47 Li, A. *et al.* Phosphorylation of histone H2A.X by DNA-dependent protein kinase is not affected by core histone acetylation, but it alters nucleosome stability and histone H1 binding. *J Biol Chem* **285**, 17778-17788, doi:10.1074/jbc.M110.116426 (2010).

48 Shinagawa, T. *et al.* Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell* **14**, 217-227, doi:10.1016/j.stem.2013.12.015 (2014).

49 Burton, A. & Torres-Padilla, M. E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat Rev Mol Cell Biol* **15**, 723-734, doi:10.1038/nrm3885 (2014).

50 Bowman, G. D. & Poirier, M. G. Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. *Chem Rev* **115**, 2274-2295, doi:10.1021/cr500350x (2015).

51 Morrison, O. & Thakur, J. Molecular Complexes at Euchromatin, Heterochromatin and Centromeric Chromatin. *Int J Mol Sci* **22**, doi:ARTN 6922

10.3390/ijms22136922 (2021).

52 Grewal, S. I. S. & Jia, S. T. Heterochromatin revisited. *Nature Reviews Genetics* **8**, 35-46, doi:10.1038/nrg2008 (2007).

53 Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994-1004, doi:10.1016/j.cell.2012.09.045 (2012).

54 Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707-719, doi:10.1016/j.cell.2007.01.015 (2007).

55 Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-395, doi:10.1038/cr.2011.22 (2011).

56 Eberharter, A. & Becker, P. B. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep* **3**, 224-229, doi:10.1093/embo-reports/kvf053 (2002).

57 Chen, Y. C., Koutelou, E. & Dent, S. Y. R. Now open: Evolving insights to the roles of lysine acetylation in chromatin organization and function. *Mol Cell* **82**, 716-727, doi:10.1016/j.molcel.2021.12.004 (2022).

58 Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**, 1025-1040, doi:10.1038/nsmb1338 (2007).

59 Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-112, doi:10.1038/nature07829 (2009).

60 Lauberth, S. M. *et al.* H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021-1036, doi:10.1016/j.cell.2013.01.052 (2013).

61 Wang, H. *et al.* Publisher Correction: H3K4me3 regulates RNA polymerase II promoter-proximal pause-release (vol 615, pg 339, 2023). *Nature*, doi:10.1038/s41586-023-05958-0 (2023).

62 Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**, 21931-21936, doi:10.1073/pnas.1016071107 (2010).

63 Zentner, G. E., Tesar, P. J. & Scacheri, P. C. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res* **21**, 1273-1283, doi:10.1101/gr.122382.111 (2011).

64 Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279-283, doi:10.1038/nature09692 (2011).

65 Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet* **32**, 29-41, doi:10.1016/j.tig.2015.11.001 (2016).

66 Saksouk, N., Simboeck, E. & Dejardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3, doi:10.1186/1756-8935-8-3 (2015).

67 Dellino, G. I. *et al.* Polycomb silencing blocks transcription initiation. *Mol Cell* **13**, 887-893, doi:10.1016/s1097-2765(04)00128-5 (2004).

68 Breiling, A., Turner, B. M., Bianchi, M. E. & Orlando, V. General transcription factors bind promoters repressed by Polycomb group proteins. *Nature* **412**, 651-655, doi:10.1038/35088090 (2001).

69 Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326, doi:10.1016/j.cell.2006.02.041 (2006).

70 Vastenhouw, N. L. & Schier, A. F. Bivalent histone modifications in early embryogenesis. *Curr Opin Cell Biol* **24**, 374-386, doi:10.1016/j.ceb.2012.03.009 (2012).

71 Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124, doi:10.1038/35065138 (2001).

72 Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120, doi:10.1038/35065132 (2001).

73 Dahl, J. A. *et al.* Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548-+, doi:10.1038/nature19360 (2016).

74 Zhang, B. J. *et al.* Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553-+, doi:10.1038/nature19361 (2016).

75 Liu, X. *et al.* Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558-562, doi:10.1038/nature19362 (2016).

76 Wang, M., Chen, Z. & Zhang, Y. CBP/p300 and HDAC activities regulate H3K27 acetylation dynamics and zygotic genome activation in mouse preimplantation embryos. *EMBO J* **41**, e112012, doi:10.15252/embj.2022112012 (2022).

77 Jin, Q. *et al.* Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J* **30**, 249-262, doi:10.1038/emboj.2010.318 (2011).

78 Trotter, K. W. & Archer, T. K. The BRG1 transcriptional coregulator. *Nucl Recept Signal* **6**, e004, doi:10.1621/nrs.06004 (2008).

79 Narita, T. *et al.* Enhancers are activated by p300/CBP activity-dependent PIC assembly, RNAPII recruitment, and pause release. *Mol Cell* **81**, 2166-2182 e2166, doi:10.1016/j.molcel.2021.03.008 (2021).

80 Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343-349, doi:10.1038/nature09784 (2011).

81 Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* **10**, 697-708, doi:10.1038/nrm2763 (2009).

82 Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28**, 1-13, doi:10.1016/j.molcel.2007.09.011 (2007).

83 Zheng, H. *et al.* Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol Cell* **63**, 1066-1079, doi:10.1016/j.molcel.2016.08.032 (2016).

84 Matoba, S. *et al.* Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. *Cell* **159**, 884-895, doi:10.1016/j.cell.2014.09.055 (2014).

85 Wang, C. *et al.* Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat Cell Biol* **20**, 620-631, doi:10.1038/s41556-018-0093-4 (2018).

86 Cohen, C. J., Lock, W. M. & Mager, D. L. Endogenous retroviral LTRs as promoters for human genes: A critical assessment. *Gene* **448**, 105-114, doi:10.1016/j.gene.2009.06.020 (2009).

87 Bhat, A. *et al.* Role of Transposable Elements in Genome Stability: Implications for Health and Disease. *Int J Mol Sci* **23**, doi:ARTN 7802

10.3390/ijms23147802 (2022).

88 Canzio, D. *et al.* Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol Cell* **41**, 67-81, doi:10.1016/j.molcel.2010.12.016 (2011).

89 Smallwood, A., Esteve, P. O., Pradhan, S. & Carey, M. Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev* **21**, 1169-1178, doi:10.1101/gad.1536807 (2007).

90 Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490-495, doi:10.1038/nature10716 (2011).

91 Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-1093, doi:DOI 10.1126/science.1063443 (2001).

92 Guo, F. *et al.* Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**, 447-459, doi:10.1016/j.stem.2014.08.003 (2014).

93 Shen, L. *et al.* Tet3 and DNA Replication Mediate Demethylation of Both the Maternal and Paternal Genomes in Mouse Zygotes. *Cell Stem Cell* **15**, 459-470, doi:10.1016/j.stem.2014.09.002 (2014).

94 Smith, Z. D. *et al.* Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature* **549**, 543-547, doi:10.1038/nature23891 (2017).

95 Zhang, Y. *et al.* Dynamic epigenomic landscapes during early lineage specification in mouse embryos. *Nat Genet* **50**, 96-105, doi:10.1038/s41588-017-0003-x (2018).

96 Zhou, F. *et al.* Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* **572**, 660-664, doi:10.1038/s41586-019-1500-0 (2019).

97 Liu, B. *et al.* The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature* **587**, 139-144, doi:10.1038/s41586-020-2847-y (2020).

98 Abe, K. *et al.* The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3' processing. *EMBO J* **34**, 1523-1537, doi:10.15252/embj.201490648 (2015).

99 Hamamoto, G., Suzuki, T., Suzuki, M. G. & Aoki, F. Regulation of transketolase like 1 gene expression in the murine one-cell stage embryos. *PLoS One* **9**, e82087, doi:10.1371/journal.pone.0082087 (2014).

100 Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell* **6**, 117-131, doi:10.1016/s1534-5807(03)00373-3 (2004).

101 Park, S. J. *et al.* Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. *Genes Dev* **27**, 2736-2748, doi:10.1101/gad.227926.113 (2013).

102 Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593-597, doi:10.1038/nature12364 (2013).

103 Zeng, F., Baldwin, D. A. & Schultz, R. M. Transcript profiling during preimplantation mouse development. *Dev Biol* **272**, 483-496, doi:10.1016/j.ydbio.2004.05.018 (2004).

104 Majumder, S. & Depamphilis, M. L. A Unique Role for Enhancers Is Revealed during Early Mouse Development. *Bioessays* **17**, 879-889, doi:DOI 10.1002/bies.950171010 (1995).

105 Schultz, R. M. Regulation of zygotic gene activation in the mouse. *Bioessays* **15**, 531-538, doi:10.1002/bies.950150806 (1993).

106 Telford, N. A., Watson, A. J. & Schultz, G. A. Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Mol Reprod Dev* **26**, 90-100, doi:10.1002/mrd.1080260113 (1990).

107 Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204-220, doi:10.1038/nrg3354 (2013).

108 Bouniol, C., Nguyen, E. & Debey, P. Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Exp Cell Res* **218**, 57-62, doi:10.1006/excr.1995.1130 (1995).

109 Peaston, A. E. *et al.* Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* **7**, 597-606, doi:10.1016/j.devcel.2004.09.004 (2004).

110 Kigami, D., Minami, N., Takayama, H. & Imai, H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biology of Reproduction* **68**, 651-654, doi:10.1095/biolreprod.102.007906 (2003).

111 Wu, J. *et al.* The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652-657, doi:10.1038/nature18606 (2016).

112 Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57-63, doi:10.1038/nature11244 (2012).

113 Nothias, J. Y., Miranda, M. & DePamphilis, M. L. Uncoupling of transcription and translation during zygotic gene activation in the mouse. *EMBO J* **15**, 5715-5725 (1996).

114 Svoboda, P. Mammalian zygotic genome activation. *Semin Cell Dev Biol* **84**, 118-126, doi:10.1016/j.semcdb.2017.12.006 (2018).

115 Abe, K. *et al.* Minor zygotic gene activation is essential for mouse preimplantation development. *P Natl Acad Sci USA* **115**, E6780-E6788, doi:10.1073/pnas.1804309115 (2018).

116 Aoki, F. Zygotic gene activation in mice: profile and regulation. *J Reprod Dev* **68**, 79-84, doi:10.1262/jrd.2021-129 (2022).

117 Falco, G. *et al.* Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev Biol* **307**, 539-550, doi:10.1016/j.ydbio.2007.05.003 (2007).

118 Srinivasan, R. *et al.* Zscan4 binds nucleosomal microsatellite DNA and protects mouse two-cell embryos from DNA damage. *Sci Adv* **6**, eaaz9115, doi:10.1126/sciadv.aaz9115 (2020).

119 Ishiuchi, T. *et al.* Reprogramming of the histone H3.3 landscape in the early mouse embryo. *Nat Struct Mol Biol* **28**, 38-49, doi:10.1038/s41594-020-00521-1 (2021).

120 Rossant, J. Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *J Embryol Exp Morphol* **36**, 283-290 (1976).

121 Tarkowski, A. K. & Wroblewska, J. Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J Embryol Exp Morphol* **18**, 155-180 (1967).

122 Tarkowski, A. K. Experiments on the Development of Isolated Blastomeres of Mouse Eggs. *Nature* **184**, 1286-1287, doi:DOI 10.1038/1841286a0 (1959).

123 Kelly, S. J. Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *J Exp Zool* **200**, 365-376, doi:10.1002/jez.1402000307 (1977).

124 Pinyopummin, A., Takahashi, Y., Hishinuma, M. & Kanagawa, H. Development of single blastomeres from 4-cell stage embryos after aggregation with parthenogenones in mice. *Jpn J Vet Res* **42**, 119-126 (1994).

125 Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S. & Zernicka-Goetz, M. Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479-490, doi:10.1242/dev.01602 (2005).

126 Chazaud, C. & Yamanaka, Y. Lineage specification in the mouse preimplantation embryo. *Development* **143**, 1063-1074, doi:10.1242/dev.128314 (2016).

127 Johnson, M. H. & Ziomek, C. A. Induction of Polarity in Mouse 8-Cell Blastomeres - Specificity, Geometry, and Stability. *J Cell Biol* **91**, 303-308, doi:DOI 10.1083/jcb.91.1.303 (1981).

128 Balakier, H. & Pedersen, R. A. Allocation of Cells to Inner Cell Mass and Trophectoderm Lineages in Pre-Implantation Mouse Embryos. *Developmental Biology* **90**, 352-362, doi:Doi 10.1016/0012-1606(82)90384-0 (1982).

129 Fleming, T. P. A Quantitative-Analysis of Cell Allocation to Trophectoderm and Inner Cell Mass in the Mouse Blastocyst. *Developmental Biology* **119**, 520-531, doi:Doi 10.1016/0012-1606(87)90055-8 (1987).

130 Posfai, E. *et al.* Position-and Hippo signaling -dependent plasticity during lineage segregation in the early mouse embryo. *Elife* **6**, doi:ARTN e22906

10.7554/eLife.22906 (2017).

131 Manejwala, F. M., Cragoe, E. J., Jr. & Schultz, R. M. Blastocoel expansion in the preimplantation mouse embryo: role of extracellular sodium and chloride and possible apical routes of their entry. *Dev Biol* **133**, 210-220, doi:10.1016/0012-1606(89)90312-6 (1989).

132 Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat Rev Genet* **20**, 221-234, doi:10.1038/s41576-018-0087-x (2019).

133 Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med* **50**, 1-14, doi:10.1038/s12276-018-0071-8 (2018).

134 Gilfillan, G. D. *et al.* Limitations and possibilities of low cell number ChIP-seq. *BMC Genomics* **13**, 645, doi:10.1186/1471-2164-13-645 (2012).

135 Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**, 613-626, doi:10.1038/nrg3207 (2012).

136 Kidder, B. L., Hu, G. & Zhao, K. ChIP-Seq: technical considerations for obtaining high-quality data. *Nat Immunol* **12**, 918-922, doi:10.1038/ni.2117 (2011).

137 Hendrickson, P. G. *et al.* Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat Genet* **49**, 925-934, doi:10.1038/ng.3844 (2017).

138 Choi, S. H. *et al.* DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res* **44**, 5161-5173, doi:10.1093/nar/gkw141 (2016).

139 Bosnakovski, D. *et al.* A novel P300 inhibitor reverses DUX4-mediated global histone H3 hyperacetylation, target gene expression, and cell death. *Sci Adv* **5**, eaaw7781, doi:10.1126/sciadv.aaw7781 (2019).

140 Lu, F. & Zhang, Y. Cell totipotency: molecular features, induction, and maintenance. *Natl Sci Rev* **2**, 217-225, doi:10.1093/nsr/nwv009 (2015).

141 Chen, Z. Y. & Zhang, Y. Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development. *Nature Genetics* **51**, 947-+, doi:10.1038/s41588-019-0418-7 (2019).

142 To, B. & Andrechek, E. R. Transcription factor compensation during mammary gland development in E2F knockout mice. *PLoS One* **13**, e0194937, doi:10.1371/journal.pone.0194937 (2018).

143 Leidenroth, A. & Hewitt, J. E. A family history of DUX4: phylogenetic analysis of DUXA, B, C and Duxbl reveals the ancestral DUX gene. *BMC Evol Biol* **10**, 364, doi:10.1186/1471-2148-10-364 (2010).

144 UniProt, C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523-D531, doi:10.1093/nar/gkac1052 (2023).

145 Himeda, C. L. & Jones, P. L. The Genetics and Epigenetics of Facioscapulohumeral Muscular Dystrophy. *Annu Rev Genomics Hum Genet* **20**, 265-291, doi:10.1146/annurev-genom-083118-014933 (2019).

146 Campbell, A. E., Belleville, A. E., Resnick, R., Shadle, S. C. & Tapscott, S. J. Facioscapulohumeral dystrophy: activating an early embryonic transcriptional program in human skeletal muscle. *Hum Mol Genet* **27**, R153-R162, doi:10.1093/hmg/ddy162 (2018).

147 Eidahl, J. O. *et al.* Mouse Dux is myotoxic and shares partial functional homology with its human paralog DUX4. *Hum Mol Genet* **25**, 4577-4589, doi:10.1093/hmg/ddw287 (2016).

148 Lee, J. K. *et al.* Crystal Structure of the Double Homeodomain of DUX4 in Complex with DNA. *Cell Reports* **25**, 2955-+, doi:10.1016/j.celrep.2018.11.060 (2018).

149 Wu, S. L. *et al.* Characterization of genomic structures and expression profiles of three tandem repeats of a mouse double homeobox gene: Duxbl. *Dev Dyn* **239**, 927-940, doi:10.1002/dvdy.22210 (2010).

150 Klein, F. *et al.* The transcription factor Duxbl mediates elimination of pre-T cells that fail beta-selection. *J Exp Med* **216**, 638-655, doi:10.1084/jem.20181444 (2019).

151 Wu, S. L. *et al.* Double homeobox gene, Duxbl, promotes myoblast proliferation and abolishes myoblast differentiation by blocking MyoD transactivation. *Cell Tissue Res* **358**, 551-566, doi:10.1007/s00441-014-1974-x (2014).

152 De Iaco, A., Verp, S., Offner, S., Grun, D. & Trono, D. DUX is a non-essential synchronizer of zygotic genome activation. *Development* **147**, doi:ARTN dev177725

10.1242/dev.177725 (2020).

153 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

154 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

155 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

156 Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204-2207, doi:10.1093/bioinformatics/btq351 (2010).

157 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, doi:ARTN 550

10.1186/s13059-014-0550-8 (2014).

158 Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26, doi:10.1038/nbt.1754 (2011).

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

160 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

161 Xie, C. *et al.* KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* **39**, W316-W322, doi:10.1093/nar/gkr483 (2011).

162 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* **57**, 289-300, doi:DOI 10.1111/j.2517-6161.1995.tb02031.x (1995).

163 Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191, doi:10.1093/bioinformatics/btp033 (2009).

164 Flach, G., Johnson, M. H., Braude, P. R., Taylor, R. A. & Bolton, V. N. The transition from maternal to embryonic control in the 2-cell mouse embryo. *EMBO J* **1**, 681-686, doi:10.1002/j.1460-2075.1982.tb01230.x (1982).

165 Lu, F. *et al.* Establishing Chromatin Regulatory Landscape during Mouse Preimplantation Development. *Cell* **165**, 1375-1388, doi:10.1016/j.cell.2016.05.050 (2016).

166 Warner, C. M. & Versteegh, L. R. In vivo and in vitro effect of alpha-amanitin on preimplantation mouse embryo RNA polymerase. *Nature* **248**, 678-680, doi:10.1038/248678a0 (1974).

167 Ahier, A. & Jarriault, S. Simultaneous expression of multiple proteins under a single promoter in Caenorhabditis elegans via a versatile 2A-based toolkit. *Genetics* **196**, 605-613, doi:10.1534/genetics.113.160846 (2014).

168 Papaioannou, V. E., Mkandawire, J. & Biggers, J. D. Development and phenotypic variability of genetically identical half mouse embryos. *Development* **106**, 817-827, doi:10.1242/dev.106.4.817 (1989).

169 Tsunoda, Y. & Mclaren, A. Effect of Various Procedures on the Viability of Mouse Embryos Containing Half the Normal Number of Blastomeres. *J Reprod Fertil* **69**, 315-322 (1983).

170 Guo, M. Y. *et al.* Precise temporal regulation of Dux is important for embryo development. *Cell Res* **29**, 956-959, doi:10.1038/s41422-019-0238-4 (2019).

171 Moore, J. E. *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699-+, doi:10.1038/s41586-020-2493-4 (2020).

172 Eisen, J. S. & Smith, J. C. Controlling morpholino experiments: don't stop making antisense. *Development* **135**, 1735-1743, doi:10.1242/dev.001115 (2008).

173 Hu, Z. H. *et al.* Maternal factor NELFA drives a 2C-like state in mouse embryonic stem cells. *Nature Cell Biology* **22**, 175-+, doi:10.1038/s41556-019-0453-8 (2020).

174 Gassler, J. *et al.* Zygotic genome activation by the totipotency pioneer factor Nr5a2. *Science* **378**, doi:ARTN eabn7478

10.1126/science.abn7478 (2022).

175 Percharde, M. *et al.* A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell* **174**, 391-405 e319, doi:10.1016/j.cell.2018.05.043 (2018).

176 Zhang, W. *et al.* Zscan4c activates endogenous retrovirus MERVL and cleavage embryo genes. *Nucleic Acids Res* **47**, 8485-8501, doi:10.1093/nar/gkz594 (2019).

177 De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet* **49**, 941-945, doi:10.1038/ng.3858 (2017).

178 Bornstein, C. *et al.* A negative feedback loop of transcription factors specifies alternative dendritic cell chromatin States. *Mol Cell* **56**, 749-762, doi:10.1016/j.molcel.2014.10.014 (2014).

179 Shen, H. *et al.* Mouse totipotent stem cells captured and maintained through spliceosomal repression. *Cell* **184**, 2843-2859 e2820, doi:10.1016/j.cell.2021.04.020 (2021).

180 Guo, Y. *et al.* Obox4 secures zygotic genome activation upon loss of Dux. *bioRxiv*, 2022.2007.2004.498763, doi:10.1101/2022.07.04.498763 (2022).

181 Ji, S. *et al.* OBOX regulates mouse zygotic genome activation and early development. *Nature* **620**, 1047-1053, doi:10.1038/s41586-023-06428-3 (2023).

182 Ghazimoradi, M. H., Khalafizadeh, A. & Babashah, S. A critical review on induced totipotent stem cells: Types and methods. *Stem Cell Res* **63**, 102857, doi:10.1016/j.scr.2022.102857 (2022).

183 Yang, J. *et al.* Establishment of mouse expanded potential stem cells. *Nature* **550**, 393-397, doi:10.1038/nature24052 (2017).

184 Posfai, E. *et al.* Evaluating totipotency using criteria of increasing stringency. *Nat Cell Biol* **23**, 49-60, doi:10.1038/s41556-020-00609-2 (2021).

185 Genet, M. & Torres-Padilla, M. E. The molecular and cellular features of 2-cell-like cells: a reference guide. *Development* **147**, doi:10.1242/dev.189688 (2020).

186 Xie, S. Q. *et al.* Nucleolar-based Dux repression is essential for embryonic two-cell stage exit. *Gene Dev* **36**, 331-347, doi:10.1101/gad.349172.121 (2022).

187 Olbrich, T. *et al.* CTCF is a barrier for 2C-like reprogramming. *Nat Commun* **12**, doi:ARTN 4856

10.1038/s41467-021-25072-x (2021).

188 Gossen, M. & Bujard, H. Tight Control of Gene-Expression in Mammalian-Cells by Tetracycline-Responsive Promoters. *P Natl Acad Sci USA* **89**, 5547-5551, doi:DOI 10.1073/pnas.89.12.5547 (1992).

189 Mullick, A. *et al.* The cumate gene-switch: a system for regulated expression in mammalian cells. *Bmc Biotechnol* **6**, doi:Artn 43

10.1186/1472-6750-6-43 (2006).

190 Liu, Z. Q. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep-Uk* **7**, doi:ARTN 2193

10.1038/s41598-017-02460-2 (2017).

191 Maksakova, I. A. *et al.* Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells. *Epigenet Chromatin* **6**, doi:Artn 15

10.1186/1756-8935-6-15 (2013).

192 Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Gene Dev* **16**, 919-932, doi:DOI 10.1101/gad.973302 (2002).

193 Nip, Y. *et al.* Human DUX4 and porcine DUXC activate similar early embryonic programs in pig muscle cells: implications for preclinical models of FSHD. *Hum Mol Genet* **32**, 1864-1874, doi:10.1093/hmg/ddad021 (2023).

194 Wong, C. J., Whiddon, J. L., Langford, A. T., Belleville, A. E. & Tapscott, S. J. Canine DUXC: implications for DUX4 retrotransposition and preclinical models of FSHD. *Hum Mol Genet* **31**, 1694-1704, doi:10.1093/hmg/ddab352 (2022).

195 Tohonen, V. *et al.* Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nat Commun* **6**, 8207, doi:10.1038/ncomms9207 (2015).

196 Bosnakovski, D. *et al.* Antagonism among DUX family members evolved from an ancestral toxic single homeodomain protein. *iScience* **26**, 107823, doi:10.1016/j.isci.2023.107823 (2023).

197 Yasuda, T. *et al.* Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. *Nat Genet* **48**, 569-574, doi:10.1038/ng.3535 (2016).

198 Brcic, I. *et al.* Undifferentiated round cell sarcomas with CIC-DUX4 gene fusion: expanding the clinical spectrum. *Pathology* **52**, 236-242, doi:10.1016/j.pathol.2019.09.015 (2020).

199 Clapp, J. *et al.* Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am J Hum Genet* **81**, 264-279, doi:10.1086/519311 (2007).

200 Bosnakovski, D. *et al.* An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *EMBO J* **27**, 2766-2779, doi:10.1038/emboj.2008.201 (2008).

201 Wijmenga, C. *et al.* Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet* **2**, 26-30, doi:10.1038/ng0992-26 (1992).

202 van Overveld, P. G. *et al.* Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet* **35**, 315-317, doi:10.1038/ng1262 (2003).

203 Gabellini, D., Green, M. R. & Tupler, R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* **110**, 339-348, doi:10.1016/s0092-8674(02)00826-7 (2002).

204 Rickard, A. M., Petek, L. M. & Miller, D. G. Endogenous DUX4 expression in FSHD myotubes is sufficient to cause cell death and disrupts RNA splicing and cell migration pathways. *Hum Mol Genet* **24**, 5901-5914, doi:10.1093/hmg/ddv315 (2015).

205 Wallace, L. M. *et al.* DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann Neurol* **69**, 540-552, doi:10.1002/ana.22275 (2011).

206 Bosnakovski, D. *et al.* p53-independent DUX4 pathology in cell and animal models of facioscapulohumeral muscular dystrophy. *Dis Model Mech* **10**, 1211-1216, doi:10.1242/dmm.030064 (2017).

207 Ferreboeuf, M. *et al.* DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum Mol Genet* **23**, 171-181, doi:10.1093/hmg/ddt409 (2014).

208 Knopp, P. *et al.* DUX4 induces a transcriptome more characteristic of a less-differentiated cell state and inhibits myogenesis. *J Cell Sci* **129**, 3816-3831, doi:10.1242/jcs.180372 (2016).