The non-coding genome in early human development-Recent advancements

Paloviita, Pauliina

2022-11

Paloviita , P & Vuoristo , S 2022 , ' The non-coding genome in early human development-Recent advancements ' , Seminars in Cell and Developmental Biology , vol. 131 , pp. 4-13 . https://doi.org/10.1016/j.semcdb.2022.02.010

http://hdl.handle.net/10138/351337
https://doi.org/10.1016/j.semcdb.2022.02.010

Seminars in Cell and Developmental Biology 131 (2022) 4–13

Contents lists available at ScienceDirect

Seminars in Cell and Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

The non-coding genome in early human development – Recent advancements

Pauliina Paloviita, Sanna Vuoristo *

Department of Obstetrics and Gynaecology, University of Helsinki, 00014 Helsinki, Finland

ARTICLE INFO

Keywords:
Human embryo
Pre-implantation development
Non-coding genome
Non-coding RNA
Transposable element
Small RNA

ABSTRACT

Not that long ago, the human genome was discovered to be mainly non-coding, that is comprised of DNA sequences that do not code for proteins. The initial paradigm that non-coding is also non-functional was soon overturned and today the work to uncover the functions of non-coding DNA and RNA in human early embryogenesis has commenced. Early human development is characterized by large-scale changes in genomic activity and the transcriptome that are partly driven by the coordinated activation and repression of repetitive DNA elements scattered across the genome. Here we provide examples of recent novel discoveries of non-coding DNA and RNA interactions and mechanisms that ensure accurate non-coding activity during human maternal-to-zygotic transition and lineage segregation. These include studies on small and long non-coding RNAs, transposable element regulation, and RNA tailing in human oocytes and early embryos. High-throughput approaches to dissect the non-coding regulatory networks governing early human development are a foundation for functional studies of specific genomic elements and molecules that has only begun and will provide a wider understanding of early human embryogenesis and causes of infertility.

1. Introduction

Pre-implantation embryo development, including embryonic genome activation (EGA) and cellular lineage commitment, has been extensively investigated during the past decades and is reviewed in [1-5]. Studies in *Mus musculus*, *Drosophila melanogaster*, *Xaenopus laevis*, and other model organisms have provided fundamental mechanistic insights into early developmental processes [6-10], that are only partially recapitulated in humans, thus limiting the potential of using non-primate organisms to investigate human embryonic development. Considerable interspecies differences range from *bonafide* measures, such as distinct sets of genes, to more complex divergencies, like the timing of embryonic genome activation [8,11]. Enlightening the mechanisms of early embryo development in humans is of high scientific and clinical interest and will provide grounds for better understanding of reproduction-related failures and congenital defects, thus facilitating the development of improved diagnostics during infertility treatments.

Despite the potential and importance of human embryo research, a variety of challenges impede its progression. Human embryo research is typically performed using either supernumerary embryos or gametes that have been donated for research after termination of infertility treatments, which limits the number of available samples. Other challenges related to human embryo research include for instance differences in legislation and applicable funding between countries, and the paradigm of open sharing of genetic data while simultaneously protecting it according to strict interpretation of the General Data Projection Regulation. We propose that more uniform policies and legislation would likely facilitate collaborative approaches, open access publishing, and data sharing across the globe.

Recent improvements in single-cell technologies that enable studying the rare and unique embryo samples and retaining information on cellular heterogeneity have led to major breakthroughs in human embryo research. These landmarks include for instance the identification of the embryonic transcription start site regions that are upregulated during EGA [12], and recognition of the first protein-coding transcripts that are heterogeneously expressed during lineage commitment [13-17]. Development of the CRISPR-Cas9 genome editing technology has enabled the generation of knockout alleles at nucleotide-level accuracy [18], even in human embryos [19,20]. Although genome editing approaches in human germ cells and embryos are technically very challenging and their ethical justification in human embryo research must be meticulously evaluated, these techniques are unique in allowing

* Correspondence to: Functional Embryo Genomics Lab, University of Helsinki, Biomedicum Helsinki 1, Haartmaninkatu 8, A430B, 00014, Finland.
E-mail address: sanna.vuoristo@helsinki.fi (S. Vuoristo).

https://doi.org/10.1016/j.semcdb.2022.02.010
Received 1 December 2021; Received in revised form 8 February 2022; Accepted 8 February 2022
Available online 15 February 2022

scientists to investigate functions of specific genes during human early embryogenesis [19]. Embryonic transcripts have recently been knocked down in human zygotes and embryos using RNAi technologies that are expected to affect target mRNA stability and translation [21-23]. Manipulated embryos have typically been analysed for the presence (expression levels) of protein-coding marker genes either using immunofluorescence and microscopy in three dimensions or single-cell RNA-sequencing (scRNA-seq). Although some of the recent pivotal studies have shed light upon the non-coding genome [21,22,24,25] and transcripts [26-30] in human gametes and early embryos, the vast majority of human embryo research has focused on protein coding genes. The non-coding genome is widely accessible and transcriptionally active, producing a multitude of RNA classes reviewed in [31] (Fig. 1), during early human embryo development. Accessible genomic regions in early human embryos frequently comprise transposable elements, reviewed in [32,33]. The functions of the non-coding genome have been studied in human pluripotent stem cells (hPSCs) [34], however here we do not summarize this information but only refer to it, to exemplify mechanisms that may be extrapolatable to the human embryo context. In this review, we discuss recent advances in elucidating the non-coding regulatory genome in human embryo development and highlight prospects of this fundamental and challenging area of research.

2. Overview of human pre-implantation development

The maturation of oocytes occurs after ovulation, when the prophase I arrested germinal vesicle (GV) stage oocytes resume meiosis I, undergo asymmetric division, and exclude the first polar body, which is a haploid cell with a minimal amount of cytoplasm [35]. The oocytes arrest at meiosis II and final maturation takes place only at fertilization when the second polar body is extruded. Mature (MII) oocytes are transcriptionally quiescent but the composition of their stored RNAs and proteins changes during the maternal-to-zygotic transition (MZT) (Fig. 2). At fertilization the maternal and paternal pronuclei are assembled in a clustered configuration [36,37]. After DNA replication, the pronuclei membranes break and the first cleavage division can take place [38]. Maternal RNAs and proteins contribute to the programming of the totipotent fertilized oocyte, as elaborated by the classical cloning approach where the maternal cocktail in the oocyte cytoplasm can reprogram a somatic nucleus sufficiently enough to contribute to the production of live offspring [39-42]. Moreover, while maternal RNAs are essential for embryo development before the onset of zygotic RNA transcription, their clearance is equally important [43].

scRNA-seq of human oocytes and embryos has revealed that the human EGA essentially occurs by the 8-cell stage [44,45,17]. More recent findings indicate that the first genes with embryo-specific splicing patterns are activated soon after fertilisation [46] while other results reveal that the human EGA commences in a minor and major wave by the 4-cell, and 8-cell stage, respectively [12]. Some of the well-known EGA genes, such as *LEUTX* [47] that is upregulated at the 4-cell stage [12], were not upregulated in zygotes [46]. Therefore, further studies are needed to elaborate to what extent the zygotic gene expression[46] and the minor EGA wave [12] overlap and which mechanisms drive these early EGA process. Maternal RNAs are degraded along with the activation of transcription from the embryonic loci.

Acquirement of key morphological features such as blastomere polarization initiates in parallel with major epigenetic remodelling and transcriptional activation at the 8-cell stage. In mice, apicobasal polarity of the cells is driven by intracellular asymmetries in the placement of the cell adhesion protein E-cadherin and the partitioning defective (Par) complex, and cytoskeletal components, while the mechanism in humans remains to be elucidated. Blastomere polarization occurs concomitantly

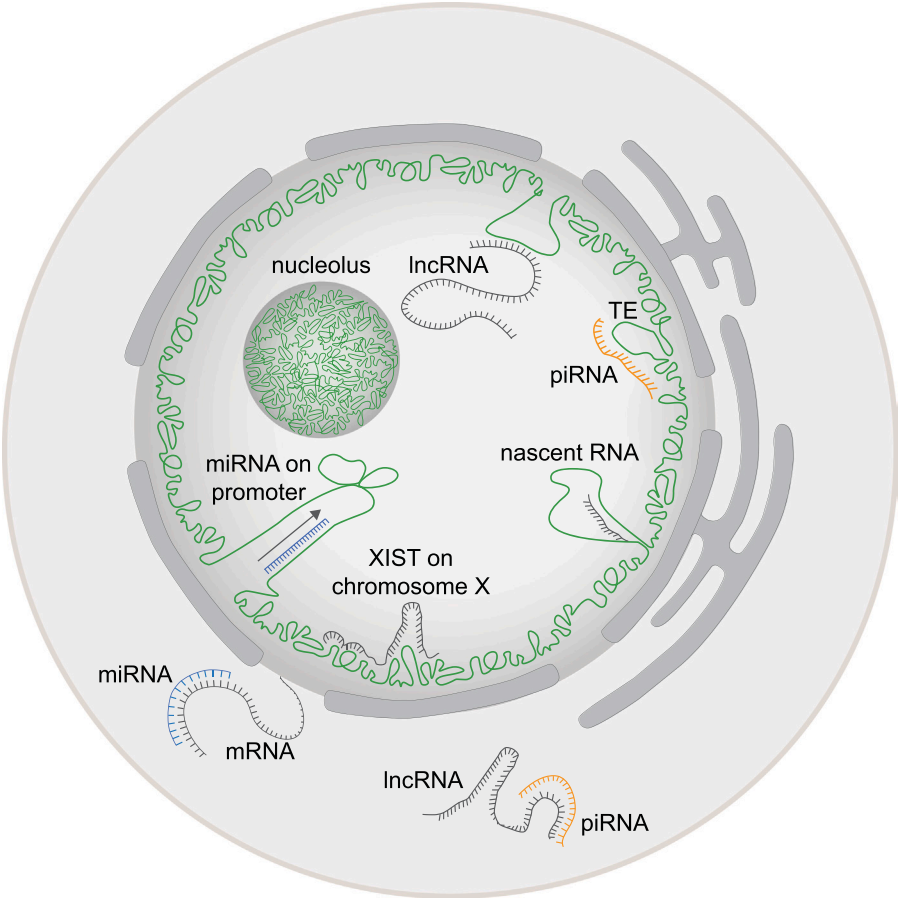

Fig. 1. Cellular compartmentalization of non-coding genomic elements and their RNA products.Genomic transposable elements provide cis-regulatory regions that can be bound both by non-coding RNAs, such as micro RNAs (miRNA), and proteins. Non-coding DNA produces both long and small non-coding transcripts (lncRNA and small RNA, respectively) that exist in the nuclear and cytoplasmic compartments. Piwi interacting RNAs (piRNA) are found both in the cytoplasm and nucleus where their canonical function is to silence transposable elements (TE) and transcripts generated from them. Canonical miRNA function, the repression of mRNAs, occurs in the cytoplasm but they may also contribute to gene regulation via binding to promoters or enhancers in the nucleus.

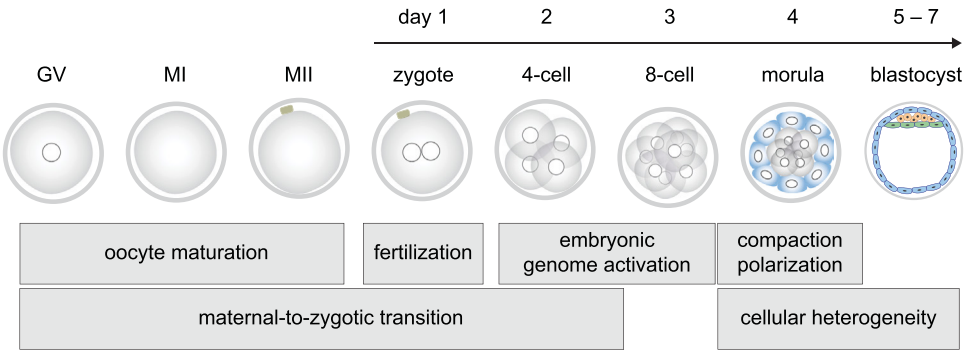

Fig. 2. Human early embryo developmental stages.In humans, the maternal-to-zygotic transition entails oocyte maturation, fertilisation, and embryonic genome activation at the 4- to 8-cell stages. This is followed by the morula stage during which cells acquire heterogeneity that leads to the establishment of the first embryonic lineages at the blastocyst stage. The developmental timeline is shown above the illustration.

with compaction, the process of blastomeres forming tight adhesions with each other, and contributes to cellular lineage segregation [23]. In the compacted embryo, the cell-cell contact area is maximized and a compressed morphology is attained. The human embryo reaches the morula stage typically on the fourth day post fertilization, when it is composed of approximately 16- to 32 cells. The morula to early blastocyst transition coincides with the opening of the blastocoelic cavity and the appearance of the first cellular lineages. Human embryos usually reach the blastocyst stage around five days post fertilization. At this stage the blastomeres segregate to form the inner cell mass (ICM) or trophectoderm lineages. During blastocyst maturation, the blastocyst expands, the trophectoderm forming trophoblast cells differentiate further and the ICM is divided into two lineages, the hypoblast (primitive endoderm in mice) and epiblast. Trophoblast cells form the embryonic part of the placenta, the pluripotent epiblast cells form the foetus, and the hypoblast cells form the yolk sac as development proceeds. The Zona pellucidae, a structure which encloses the blastocyst, becomes thinner allowing the blastocyst to finally hatch. Implantation in humans typically takes place around seven days post fertilization. The human embryo pre- to post-implantation morphology was recently coupled with embryonic and extra-embryonic cellular lineage segregation events using an in vitro pre-to post-implantation culture method [48,49]. These studies revealed that human embryos can self-organize through the implantation stages in vitro in the absence of maternal tissues [48,49]. Gastrulation that commences during the third week of human embryo development, determines the future body axes, reviewed in [50]. Tyser et al. used single cell RNA-sequencing to characterize cell populations in a 16–19-day post-fertilization human gastrula [51]. Thorough bioinformatics analyses revealed the specification of the three germ layers; ectoderm, endoderm, and mesoderm, as well as appearance of primordial germ cells and red blood cells, among the other further specified cell types [51]. Human early embryonic development seems to be far more variable than that of the typical model organisms, like the mouse, when it comes to for example timing and synchrony of the cleavage divisions, ratio of aneuploid cells, and cellular lineage commitment. The interindividual differences between human embryos call for repetition of central findings by incorporating samples from different donors and treatment facilities, and in sufficient quantities to ensure that the results are generalizable.

3. The interplay of transposable elements, DNA methylation and oocyte short piRNAs during human MZT

During MZT, the control of embryonic development is transferred from stored maternal RNAs and proteins to zygotic RNAs transcribed from the newly organized genome. Maternal protein-coding genes, including for instance *GROWTH AND DIFFERENTIATION FACTOR 9* [52] and the *ZONA PELLUCIDA* genes [53], are relatively well-characterized. Although most retrotransposons, reviewed in [54-56], in the human genome are non-mobile and transcriptionally silenced, some of these elements are transcribed and seem to have retained the ability to transpose in human oocytes [57,26,58-60,28]. Human GV oocytes express at least *LINE1*, *HERV-K10*, as well as *SINE-R*, *VNTR* and *ALU* (SVA) elements [58,61]. Long interspersed nuclear elements (LINE) have been suggested to regulate the expression of proximal protein coding genes, such as *AR* and *TMEFF2* [61]. Long terminal repeat (LTR) elements, particularly the endogenous retrovirus (ERV) ERV1 and the primate-specific LTR12C and LTR7 elements, are expressed and provide alternative transcription start sites in mouse and human oocytes [62,26,24,63]. Whether transcription from the retroelement loci and the generated transcripts are required for MZT and EGA in human oocytes is yet to be elucidated.

Oocyte maturation and MZT entail genome-wide changes in chromatin accessibility and epigenetic landscapes [24,64,25]. Although large-scale epigenetic changes seem to be conserved between mouse and human MZT, some characteristics differ between these two species, reviewed in [65]. Human oocyte maturation from the GV to MII stage is characterized by relatively stable and pre-established CpG methylation (occurring in CpG dinucleotides) levels, particularly in gene bodies [66]. In addition, increased non-CpG methylation explains the observed overall accumulation of methylation during oocyte maturation. Interestingly, most of the increasingly methylated regions [67] and non-CpG regions [61] in MII oocytes are located at transposable element (TE) loci. Moreover, increased methylation at several non-CpG distal regions that overlap with LINEs correlates negatively with the expression of nearby coding genes [61], while increased non-CpG methylation at gene bodies correlates positively with the expression of the corresponding genes in MII oocytes [67]. This suggests that non-CpG methylation levels during oocyte maturation may play a role in gene expression regulation and potentially involve TE-mediated mechanisms. After fertilization, the paternal pronucleus immediately initiates dramatic global demethylation, followed by that of the maternal pronucleus. Demethylation proceeds until the 2-cell stage [67]. At the 4-cell to 8-cell stage *de novo* DNA methylation commences and is enriched at evolutionary young transposable elements [68]. In conclusion, the methylation dynamics during human MZT seem to be extremely complex, involving abundant methylation level changes in transposable element regions.

Transposition may be deleterious to the genome and therefore human oocytes and embryos are likely to employ multiple mechanism to regulate TE silencing. Global re-methylation takes place at the post-implantation stage [67,68], but for instance LTR5-HS and LTR7B become accessible at the 8-cell stage and are rapidly downregulated at the morula stage [24], indicating efficient suppression. Small RNAs (sRNAs) contribute to the regulation of TEs in oocytes and the earliest embryonic stages. A novel sRNA class that is specifically expressed in primate oocytes, oocyte short Piwi-interacting RNAs (os-piRNAs), was

found to account for over 70% of the total sRNA-seq reads in human oocytes and are likely to contribute to the silencing of TEs, especially the L1 family members [30]. Conventional Piwi-interacting RNAs (piRNAs) are also found in human oocytes although at lower overall levels [29,30] and they preferably target ERV and LTR family members suggesting different regulatory roles for the two piRNA classes [30]. Further differences between the two piRNA classes are detailed in the glossary. The piRNA pathway is dispensable to mouse oogenesis and embryo development [69], where another sRNA class, endogenous short interfering RNAs (endo-siRNAs), is likely to be responsible for TE silencing. Endo-siRNAs are abundant in mammalian oocytes and early embryos [70] but lacking in human counterparts [30] and the contrary is true in the case of os-piRNAs [29,30]. This limits the potential to study human sRNA-TE interactions in model organisms. KRAB zinc-finger proteins have also been associated with TE silencing in the context of early embryonic genes [71] and future experimental work is needed to better elucidate how the co-evolved coding and non-coding transcripts and their products ensure controlled TE activity during MZT.

4. Modifications of maternal coding and non-coding transcripts in MZT

Transcriptomic heterogeneities can be accomplished in multiple ways, for instance through invocation of alternative transcription start sites, RNA-editing, and alternative splicing, resulting in transcript isoforms, reviewed in [72]. Differences in transcript isoform sequences, including the three prime untranslated region (3' UTR) regulatory elements, modulate transcript stability, degradation, and translation efficacy. Findings across species have emphasized selective regulation of maternal mRNAs to occur during MZT, reviewed in [73]. Maternal mRNAs can be roughly classified based on the timing of their degradation; RNAs belonging to the 'maternal RNA decay pathway' are downregulated (degraded) between the GV oocyte and zygote stage, while RNAs belonging to the 'zygotic decay pathway' are degraded in a delayed manner between the GV oocyte or zygote and 8-cell stage [74]. Recent data provide insights into these processes and show that embryonic transcription is necessary for the degradation of some maternal mRNAs in early human embryos, as alpha amanitin treatment causes accumulation of maternal mRNAs that belong to the 'zygotic decay pathway' [74]. One of the means to regulate maternal RNAs during MZT is through modifications in poly(A) tail sequence and length at the mRNA 3'UTR. Poly(A) tails of maternal mRNAs are deadenylated in mouse oocytes, where mRNA deadenylation coincides with mRNA storage and translational inactivity instead of leading to immediate mRNA degradation. [75]. Global poly(A) tail de-adenylation in human oocytes is suggested to be mediated by BTG4 that interacts with the CCR4-NOT de-adenylation complex [76-78] also shown to be active in the mouse [79,75,80].

mRNAs with short poly(A)tails are more likely to be either rapidly degraded or modified further by terminal nucleotidyltransferases (TENT), including terminal uridylyltransferases (TUT), that alter RNAs post-transcriptionally, reviewed in [81]. Recent methodological developments have allowed detailed investigation of poly(A) tail lengths and sequences also during human MZT [76-78]. Shortened poly(A) tails provide substrates for the incorporation of uridine or other nucleotides. TUT4/7 can add uridines to the 3' end of the poly(A) tails, suggesting that uridines located at the 5' end or internal parts of the poly(A) tail have been added prior to (re)-polyadenylation [76-78]. Addition of U residues seems to mark the transcript for degradation [82-84], however recent findings indicate that 3' uridylation may be temporally decoupled from degradation during human MZT and that these transcripts may be degraded only post EGA [76-78]. Transcripts can be further modified by TENT4 A and B-mediated addition of nucleotide residues. TENT4 proteins have somewhat relaxed nucleotide preferences and can incorporate also non-adenosine residues, especially guanosine, to poly (A) tails [81]. This incorporation of guanosine and adenosine residues may occur in a sequential manner [76-78]. Guanosine residues in the internal parts of the poly(A) sequences are suggested to stabilize re-adenylated mRNAs in human zygotes [76-78]. In conclusion, recent data indicate that poly (A) tail length and nucleotide composition are rigorously regulated during MZT in humans.

Maternal micro RNAs (miRNAs) also undergo post-transcriptional adenylation, which is conserved with temporal differences from fly to human early development [70,85]. TENT2 has been suggested to carry out the majority of miRNA 3' adenylation in this context [86]. In humans, miRNA 3' adenylation peaks at the zygote stage and reduces in subsequent embryonic stages [29]. sRNA adenylation during early development seems to be class specific, as in mice endo-siRNAs are highly adenylated and piRNAs are not [70]. Additionally, 3' adenylation levels in humans and mice are highly miRNA species specific and both oligo- and polyadenylation has been reported [29,70]. Many maternal miRNAs seem to undergo degradation at the zygote stage in humans [29, 30] and other organisms [87,88,70] indicating that they belong to the 'maternal decay pathway'. Whether adenylation is linked to this process remains controversial as it may both prevent and promote miRNA degradation [89-91,85,70] and the outcome is likely to be context dependent. We speculate that RNA modifications and particularly adenylation introduce another significant layer of regulation in human MZT. Adenylation provides a rapid way to regulate RNA stability and mRNA translational outputs without requiring changes in chromatin accessibility or transcriptional activity. Some of the post-transcriptional modifications in mRNAs and miRNAs are mediated by the same factors, such as TENT4 A/B, and TUT4/7 [81]. Despite the valuable information gained from recent findings, further studies are needed to dissect how the modified target RNA species are selected spatially and temporally in the context of MZT. It is also interesting to speculate whether some long non-coding RNAs (lncRNAs), including transcripts originating from transposable element loci, might be modified by the factors that regulate mRNAs and miRNAs, and what might be the outcome.

5. Transposable elements may provide functionally important cis-regulatory regions for EGA

A recent study revealed that human EGA commences at the time of fertilization as some genes are upregulated although at relatively minor levels [46]. Comparison of the transcriptomes of human MII oocyte and cleavage stage embryos has revealed that while the absolute number of RNA molecules reduces dramatically between the transition from MII oocytes to 4-cell stage embryos [12,92], some transcripts are significantly upregulated [12]. Among the transcripts upregulated during this minor EGA stage are for example *ZSCAN4*, a well-known protein-coding EGA gene also expressed in mouse [93] and bovine early embryos [94]. In addition to protein-coding genes, several transcripts upregulated in zygotes [46] and 4-cell stage embryos [12] were annotated to the non-coding genome, especially at pseudogene regions and human ERV elements [46]. Some of the 4-cell stage non-coding transcripts are likely to be lncRNAs, often enriched at transposable element regions [12]. In addition to maternally expressed ERVs, expression of the ERV2 (LTR5HS, HERV-K) and ERV3 (LTR7Y) elements [28], which also harbour DNAse I hypersensitivity sites (DHS) indicating accessible chromatin loci [22], is obvious at the time of major EGA. LTR elements provide binding sites for transcription factors [95,96] and participate in the regulation of their corresponding ERVs. Human HERV-Ks are regulated by LTR5HSs, which bear binding sites for POU5F1/OCT4, a transcription factor that is expressed [28] and activates major EGA transcripts in human EGA stage embryos [22]. OCT4 has been suggested to bind at LTR5HS loci and recruit P300, which likely leads to the observed acquisition of H3K27 acetylation [97], H3K4 trimethylation and HERV-K expression [28]. Another ERV class, HERV-H, that is typically connected to the maintenance of pluripotency [98], is also expressed in 8-cell stage human embryos and beyond [22]. HERV-H elements are flanked by LTR7 elements that are accessible in early

human embryos [22]. In addition, SVA (SVA-A), L1 (L1 and L1-HS), and SINE (Alu) elements are expressed [12,28] and located in open chromatin [22] in human preimplantation embryos. Alu elements are the most abundant repeat class identified in the human genome [99]. Out of all Alu elements, particularly AluY and AluS are enriched at the transcription start sites of many human EGA genes [12], suggesting that they participate in the regulation of human EGA [12,100]. Alu elements frequently overlap with a 36-basepair DNA-binding motif, which resembles a known consensus binding sequences of PRD-like homeobox genes, T-box and bZIP [12]. Candidate binding factors of this Alu element overlapping *de novo* motif include protein products of both embryonically (e.g. *DUXA* and *LEUTX*) and maternally (e.g. *OTX2*, *PITX2*, and *GSC*) expressed genes [12]. Out of the maternal factors at least *GSC* and *OTX2* are highly enriched in human embryos at the 2-cell stage [25], indicating that their binding to Alu elements during the first days of development is possible.

Transient chromatin opening takes place in human zygotes, prior to the progressive increase in chromatin accessibility from the 2-cell stage onwards [24]. Recent studies have indicated that open chromatin does not automatically denote active transcription in human early embryos [25]. Alu loci seem to be inaccessible at the 2-cell stage but become accessible by the 8-cell stage when major EGA takes place in human embryos [25]. It appears possible that prior to EGA Alu loci might be accessible, even if not transcribed, and could provide *cis* regulatory elements with transcription factor binding sites for the EGA gene activators as suggested in human 4-cell and 8-cell stage embryos [12]. The reorganization of three-dimensional (3D) chromatin coincides with EGA across species, reviewed in [101]. The formation of topologically associated domains (TAD) seems to depend on polymerase II-mediated transcription in human embryos, given that the TADs appear at the 8-cell stage and beyond, and that alpha-amanitin treatment of human zygotes largely abolishes TAD formation at the 8-cell stage [21]. Interestingly, AluS elements were enriched around some of the very weak TAD boundaries already at the 2-cell stage and transcribed in later cleavage stage embryos [21]. Recent data from mouse embryos and 2-cell-embryo-like cells indicate that murine ERVL (MERVL) elements promote formation of insulating domains [102]. Therefore, it appears that TEs have a role in the establishment of TAD boundaries in early cleavage stage embryos.

6. Non-coding RNAs may contribute to cellular heterogeny preceding the establishment of the first embryonic lineages

The establishment of the first embryonic lineages requires the blastomeres to retain their capability to proliferate and simultaneously produce intercellular differences in their molecular composition, that is heterogeny. Classic protein markers of embryonic lineage segregation have been studied in humans both to assess the timing of the segregation events and to identify critical factors in this process. In human embryos, the depletion of OCT4 [19] and Phospholipase C (PLC) -Protein kinase C (PKC) [23,103] signalling pathway components has indicated that these factors are needed for the determination of the ICM and trophectoderm, respectively. Similarly restricted expression of either NANOG or GATA6 in the ICM cells demarcates the establishment of the human epiblast and hypoblast [104]. While some mechanism driving lineage determination are conserved between species, others have diverged. CDX2 drives the specification and maturation of the trophectoderm and ICM [105] in the mouse while it may not be necessary for this segregation event in human embryos [2]. Some of the observed interspecies differences in the regulation of lineage segregation have been explained by low conservation of the studied genes, highlighting again the limitations in the use of model organisms to study human embryogenesis, especially in the context of the non-coding genome and transcriptome that have greatly diverged between species.

In humans, cellular heterogeny between blastomeres is observed initially around the 8-celll to morula stage [13,103] in the compacting embryo. At that time blastomeres acquire apico-basal cell polarity that was recently reported to initiate the segregation of the trophectoderm and ICM via PLC – PKC activity, leading to restricted expression of trophectoderm-associated factors, such as GATA3, in the outer cells [23,103]. Morula cells have been described as heterogenous especially in terms of cell polarity determining proteins, yet it seems that also non-coding RNAs contribute to the establishment of cellular heterogeny. lncRNAs, especially those derived from human ERVS, HERV-H and HERV-K [27,28], are transcribed from the 8-cell stage onwards and could contribute to the priming of lineage segregation as has been proposed for the Heterogeneously expressed from the Intronic Plus Strand of the TFAP2A-locus RNA (HIPSTR) that is unevenly expressed between blastomeres of human 8-cell stage embryos [106].

In humans [13,15] and other mammals [107–109] the ICM and trophectoderm segregate prior to the formation of the epi- and hypoblast. Meistermann et al. [13] combined live imaging, which is optimal for human embryo staging [110], to single-cell transcriptomics and reported that the mRNA profiles of human trophectoderm and ICM diverge just before blastocyst expansion and prior to epiblast and hypoblast establishment in the expanded blastocyst. Previous studies of lineage segregation in humans rely heavily on marker mRNAs/proteins to distinguish the established lineages [92,111,2,14,104], while it is known that lncRNAs are more tissue specific and show greater heterogeny between blastomeres [112,17,106]. ERV-derived RNAs, especially HERV-H and HERV-K, have been associated with specific lineages of the human blastocyst [113,92,114,115,28]. Reijo Pera and colleagues [113] identified a set of long intergenic noncoding RNA (lincRNA) isoforms that typically incorporate ERVs, especially HERV-H. These Human pluripotency-associated transcripts (HPATs) are abundant in human pluripotent cells [113] and almost completely absent in differentiated adult tissues. In human blastocysts, HPAT-2, − 3 and − 5 expression was specific to the ICM/epiblast cells and depletion of these HPATs in single blastomeres excluded the cells from contributing to the ICM [92,114]. HPAT expression specific to also the trophectoderm and hypoblast lineages has been reported [114,115], altogether indicating that lncRNAs may be involved in lineage specification and could also be used for annotating the cell types in human embryos.

Factors associated with the human ICM resembling naïve hPSC state [34,116], OCT4, NANOG, KLF4, and LBP9, mediate transcriptional activation of HPATs and other HERV-H containing RNAs via binding to LTR7 elements in hPSCs, and this axis could function also in human embryos [92,117]. Conversely, in vitro knockdown of specific ERVs via RNA interference leads to a reduction in pluripotency markers [59,115]. Furthermore, HPAT5 was suggested to post-transcriptionally inhibit the let-7 miRNA and this controls the balance between pluripotency and differentiation in hPSCs [92]. While many discoveries on the function of non-coding RNAs in balancing pluripotency and differentiation have been made in hPSCs, it is likely that similar axes govern the segregation of the different lineages also in human blastocysts, where non-coding and coding RNAs and their products form elaborate circuitries that may provide developmental robustness to the early embryo.

The contribution of sRNAs to driving cellular heterogeny that precedes lineage segregation has not been studied in humans. Current knowledge on sRNA expression in human cleavage and morula stage embryos is at the single embryo level [29,30] and blastocysts have been profiled only for known miRNAs [118]. Differential miRNA expression has been reported in bovine [119,120] and mouse [87,121] embryos between the ICM and trophectoderm, and in mouse stem cell models of embryonic and extra-embryonic tissues [122], suggesting that miRNAs contribute to the formation of the first differentiated lineages in mammalian embryos. In humans and mice, miRNA levels increase as development proceeds [87,29,70]. Additionally, Dgcr8-lacking mouse oocytes and zygotes arrest at the blastocyst stage or around the time of gastrulation, respectively, indicating the requirement for an active miRNA biogenesis pathway in cellular lineage commitment [123-125]. In human oocytes, cleavage stage embryos, and blastocysts,

hsa-miR-372 and its miRNA cluster family members are detected in abundant levels [29,118]. Hsa-miR-372–3p promotes somatic cell reprogramming to pluripotency [126] and is found at higher levels in naive than primed hPSCs [127,128]. The homologous miR-290–295 cluster is an essential regulator of mouse embryonic stem cell pluripotency [129] and causes partial lethality and infertility upon depletion [130], altogether suggesting that this conserved miRNA family may participate in the regulation of mammalian embryogenesis. In addition to miRNAs and piRNAs, both ribosomal RNA (rRNA)- and transfer RNA (tRNA)-derived small RNAs are present in human sperm, oocytes, and early-stage embryos in biologically relevant abundancies [29,30]. Currently, the contribution of miRNAs and other sRNAs to cellular heterogeny of human blastomeres remains to be elucidated. However, it is likely that sRNAs may contribute to lineage segregation via the regulation of the pluripotency/differentiation axis, as has been described for miRNAs and rRNA- and tRNA-derived small RNAs in PSCs [131-133].

7. Conclusions

The functions of the non-coding genome, both at the DNA and RNA level, are essential to early human development and so far, the number of studies in this context remains limited. The non-coding genome and transcriptome provide an additional layer of cellular regulation of which human embryo studies have focused on dissecting the chromatin conformation and epigenetic landscape at non-coding loci as well as stage-wise characterisation of the non-coding transcripts and their modifications. Other aspects such as the coding genome, cell organelles, mechanical ques [134,135], cellular polarization, and the 3D organization of the genome (reviewed in [101]), are also crucial to early (human) development and are likely to be connected to the non-coding genome with already some examples in literature. For instance, rRNA biogenesis and nucleolar maturation were recently found to contribute to the exit from the mouse 2-cell-embryo-like cell state [136,137], via alterations in the 3D organization of perinucleolar heterochromatin [137]. Cellular processes essential to early development and cell specification in general are regulated by extensive crosstalk and competition between different coding and non-coding molecules, with the highly tissue specific lncRNAs likely playing especially central roles [138]. This molecular dialogue is not limited to RNA-RNA/protein interactions but also includes DNA that is regulated by proteins, lncRNAs and sRNAs. The best known lncRNA-DNA interaction in the human developmental context is perhaps X chromosome (in)activation by *XIST* and *XACT* [139-142]. The hybridization of sRNAs to the genome has also been reported in this context for instance in the case of piRNA mediated epigenetic TE silencing [143]. Interestingly, TEs as well as other non-coding DNA sequences located at promoters and enhancers may also contribute to sRNA-mediated transcriptional activation. The contribution of sRNAs in regulating transcription of EGA genes by direct binding to the genome has only recently been discovered in mice [144] and so far, reports on sRNA-directed regulation of transcription in humans is limited to miRNAs in somatic tissues [145,146]. The functional annotation of non-coding DNA and RNA is only at its beginning, especially in human gametes and embryos, and future discoveries extending the canonical functions of different RNA classes or genomic elements is expected.

Recent and future novel discoveries on the contribution of the non-coding genome and transcriptome to human development rely heavily on sequencing-based methods appropriate for single-cell/embryo quantities such as PAIS-seq, Cas-seq and other sRNA-seq methods, Hi-C, and CUT&Tag [127,147,148,76,77,78,29,30]. While many sequencing-based methods requiring the homogenization of several gametes or embryos as starting material are feasible to study early development, the conclusions are limited as bulk analysis masks biologically relevant heterogeny between single blastomeres and interindividual differences between pooled human embryos may cause further noise in the data. Additionally, both current bulk and single cell methods to prepare sequencing libraries of non-coding RNAs are like to favour some RNA classes over others due to the used RNA capture method or the possible 3D structure and cellular compartmentalization of the RNAs. Methods allowing for more detailed investigation of the nuclear and cytoplasmic compartments, such as NET-CAGE [149] and RD-SPRITE [150] would provide new information on the diverse functions of nascent and nuclear non-coding RNAs, however, scaling these methods to the single cell/embryo level remains a challenge.

Challenges involved in bioinformatic analysis the non-coding genome and transcriptome are not limited to the context of embryology. Quantification of non-coding RNAs may be obscured by frequent multimapping of reads, especially short reads from sRNA-seq. Moreover, classification of non-coding RNAs is challenging as they have diverse characteristic features, are often lowly expressed and poorly conserved between species, and may resemble protein-coding RNAs [151]. Interestingly, translation has been observed beyond protein-coding transcripts and for instance translated lncRNAs are a potential source of peptides, some of which may have biological functions [152]. Furthermore, several mRNAs function on the RNA level in a manner resembling non-coding transcripts perturbing the division between coding and non-coding transcripts [152,153]. Consequently, non-coding RNA databases may include inadequate annotations and the recent discovery of os-piRNAs in human and primate oocytes exemplifies that the catalogue of the human developmental transcriptome as well as the genomic regulatome is far from complete.

Ex vivo human embryo research is also limited to methods that require a small number of cells. Thus, methods to locate specific molecules (antibodies for proteins and fluorescence in situ hybridization for RNA) or perturb a single genomic locus using CRISPR-Cas9 or transcript using small interfering RNAs (siRNAs) have been employed for functional studies. In comparison to the coding genome, the functions of the non-coding genome and transcriptome have been less studied not only in the field on human development but on a global scale. Recently, the FANTOM6 project piloted a high-throughput endeavour to functionally annotate lncRNAs using a cell culture platform coupled with antisense oligonucleotide (ASO)- and siRNA-mediated knockdown of lncRNAs [154]. The knockdown method was found in some cases to affect the outcome of the experiment, likely due to ASOs functioning primarily in the nuclear and siRNAs in the cytoplasmic compartment. Genetically engineered animals have also been crucial in functional dissection of lncRNAs [155] as well as sRNAs [156] and TEs [157], yet the number of these studies remains very small in comparison to mRNAs. Recent work characterized a conserved Cdk2ap1 protein isoform that originates from species-specific TEs and in the mouse has a developmentally reciprocal expression pattern and opposite biological function with the canonical Cdk2ap1 [158]. Expression of a human Cdk2ap1 splice isoform originating from a human-specific TE promoter was sufficient to restore the development of knockout mouse embryos that lacked the endogenous TE driving the expression of this isoform in mice [158]. Model organisms have also been indispensable for discoveries on the contribution of several non-coding RNA classes found in mammalian gametes, including miRNAs [159-161], piRNAs [162,163], tRNA-derived small RNAs [164-167], and lncRNAs [168], to intergenerational inheritance. While model organisms provide an important in vivo tool to understand the roles of the non-coding genome in human embryonic development their use is limited to studying the pathways and molecules that are conserved between species. Recently established cell models of human embryonic development, including naïve stem cells [116], blastoids/gastruloids [76-78,169,170-173], and 8-cell like cells [174,175], are crucial tools for mechanistic understanding of non-coding elements. However, meticulousness in the study designs employing these models is imperative to ensure that the non-coding RNA is studied under the prerequisites set by observations made in human oocytes and embryos regarding cellular compartmentalization and dominant transcript isoform (Fig. 1). As human gametes and embryos are known to employ

regulatory mechanisms elicited by the non-coding genome that are absent in somatic cell types and may differ in model organism, characterization of the human gamete and embryo transcriptome and genomic regulatome at the sequence and on the functional level is likely to be only at its beginning.

Acknowledgements

We acknowledge Dr Timo Tuuri for critical reading of the manuscript. This work was supported by the University of Helsinki Doctoral Programme in Biomedicine (P.P.), the Orion Research Foundation sr (P. P), the University of Helsinki Three-year grant (S.V.), the Sigrid Jusélius Young Investigator Grant (S.V.), and the Jane and Aatos Erkko Foundation (S.V.).

References

[1] M.A. Mole, A. Weberling, M. Zernicka-Goetz, Comparative analysis of human and mouse development: from zygote to pre-gastrulation, Curr. Top. Dev. Biol. 136 (2020) 113–138.
[2] K.K. Niakan, K. Eggan, Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse, Dev. Biol. 375 (2013) 54–64.
[3] K.K. Niakan, J. Han, R.A. Pedersen, C. Simon, R.A. Pera, Human pre-implantation embryo development, Development 139 (2012) 829–841.
[4] M.N. Shahbazi, Mechanisms of human embryo development: from cell fate to tissue shape and back, Development (2020) *147*.
[5] S.E. Wamaitha, K.K. Niakan, Human pre-gastrulation development, Curr. Top. Dev. Biol. 128 (2018) 295–338.
[6] K. Cockburn, J. Rossant, Making the blastocyst: lessons from the mouse, J. Clin. Invest. 120 (2010) 995–1003.
[7] D.C. Hamm, M.M. Harrison, Regulatory principles governing the maternal-to-zygotic transition: insights from Drosophila melanogaster, Open Biol. 8 (2018), 180183.
[8] D. Jukam, S.A.M. Shariati, J.M. Skotheim, Zygotic genome activation in vertebrates, Dev. Cell 42 (2017) 316–332.
[9] J. Yang, T. Aguero, M.L. King, The xenopus maternal-to-zygotic transition from the perspective of the germline, Curr. Top. Dev. Biol. 113 (2015) 271–303.
[10] M. Zernicka-Goetz, S.A. Morris, A.W. Bruce, Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo, Nat. Rev. Genet 10 (2009) 467–477.
[11] K.N. Schulz, M.M. Harrison, Mechanisms regulating zygotic genome activation, Nat. Rev. Genet. 20 (2019) 221–234.
[12] V. Tohonen, S. Katayama, L. Vesterlund, E.M. Jouhilahti, M. Sheikhi, E. Madissoon, G. Filippini-Cattaneo, M. Jaconi, A. Johnsson, T.R. Burglin, et al., Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development, Nat. Commun. 6 (2015) 8207.
[13] D. Meistermann, A. Bruneau, S. Loubersac, A. Reignier, J. Firmin, V. Francois-Campion, S. Kilens, Y. Lelievre, J. Lammers, M. Feyeux, et al., Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification, Cell Stem Cell 28 (2021) 1625–1640, e1626.
[14] S. Petropoulos, D. Edsgard, B. Reinius, Q. Deng, S.P. Panula, S. Codeluppi, A. Plaza Reyes, S. Linnarsson, R. Sandberg, F. Lanner, Single-cell RNA-seq reveals lineage and x chromosome dynamics in human preimplantation embryos, Cell 165 (2016) 1012–1026.
[15] G.G. Stirparo, T. Boroviak, G. Guo, J. Nichols, A. Smith, P. Bertone, Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast, Development (2018) *145*.
[16] Z. Xue, K. Huang, C. Cai, L. Cai, C.Y. Jiang, Y. Feng, Z. Liu, Q. Zeng, L. Cheng, Y. E. Sun, et al., Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing, Nature 500 (2013) 593–597.
[17] L. Yan, M. Yang, H. Guo, L. Yang, J. Wu, R. Li, P. Liu, Y. Lian, X. Zheng, J. Yan, et al., Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells, Nat. Struct. Mol. Biol. 20 (2013) 1131–1139.
[18] E. Zuo, Y. Sun, W. Wei, T. Yuan, W. Ying, H. Sun, L. Yuan, L.M. Steinmetz, Y. Li, H. Yang, Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos, Science 364 (2019) 289–292.
[19] N.M.E. Fogarty, A. McCarthy, K.E. Snijders, B.E. Powell, N. Kubikova, P. Blakeley, R. Lea, K. Elder, S.E. Wamaitha, D. Kim, et al., Genome editing reveals a role for OCT4 in human embryogenesis, Nature 550 (2017) 67–73.
[20] H. Ma, N. Marti-Gutierrez, S.W. Park, J. Wu, Y. Lee, K. Suzuki, A. Koski, D. Ji, T. Hayama, R. Ahmed, et al., Correction of a pathogenic gene mutation in human embryos, Nature 548 (2017) 413–419.
[21] X. Chen, Y. Ke, K. Wu, H. Zhao, Y. Sun, L. Gao, Z. Liu, J. Zhang, W. Tao, Z. Hou, et al., Key role for CTCF in establishing chromatin structure in human embryos, Nature 576 (2019) 306–310.
[22] L. Gao, K. Wu, Z. Liu, X. Yao, S. Yuan, W. Tao, L. Yi, G. Yu, Z. Hou, D. Fan, et al., Chromatin accessibility landscape in human early embryos and its association with evolution, Cell 173 (2018) 248–259, e215.
[23] M. Zhu, M. Shahbazi, A. Martin, C. Zhang, B. Sozen, M. Borsos, R.S. Mandelbaum, R.J. Paulson, M.A. Mole, M. Esbert, et al., Human embryo polarization requires PLC signaling to mediate trophectoderm specification, Elife (2021) *10*.
[24] L. Liu, L. Leng, C. Liu, C. Lu, Y. Yuan, L. Wu, F. Gong, S. Zhang, X. Wei, M. Wang, et al., An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos, Nat. Commun. 10 (2019) 364.
[25] J. Wu, J. Xu, B. Liu, G. Yao, P. Wang, Z. Lin, B. Huang, X. Wang, T. Li, S. Shi, et al., Chromatin analysis in human early development reveals epigenetic transition during ZGA, Nature 557 (2018) 256–260.
[26] V. Franke, S. Ganesh, R. Karlic, R. Malik, J. Pasulka, F. Horvat, M. Kuzman, H. Fulka, M. Cernohorska, J. Urbanova, et al., Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes, Genome Res 27 (2017) 1384–1394.
[27] J. Goke, X. Lu, Y.S. Chan, H.H. Ng, L.H. Ly, F. Sachs, I. Szczerbinska, Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells, Cell Stem Cell 16 (2015) 135–141.
[28] E.J. Grow, R.A. Flynn, S.L. Chavez, N.L. Bayless, M. Wossidlo, D.J. Wesche, L. Martin, C.B. Ware, C.A. Blish, H.Y. Chang, et al., Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells, Nature 522 (2015) 221–225.
[29] P. Paloviita, C. Hyden-Granskog, D.A. Yohannes, P. Paluoja, J. Kere, J. S. Tapanainen, K. Krjutskov, T. Tuuri, U. Vosa, S. Vuoristo, Small RNA expression and miRNA modification dynamics in human oocytes and early embryos, Genome Res 31 (2021) 1474–1485.
[30] Q. Yang, R. Li, Q. Lyu, L. Hou, Z. Liu, Q. Sun, M. Liu, H. Shi, B. Xu, M. Yin, et al., Single-cell CAS-seq reveals a class of short PIWI-interacting RNAs in human oocytes, Nat. Commun. 10 (2019) 3389.
[31] T.R. Cech, J.A. Steitz, The noncoding RNA revolution-trashing old rules to forge new ones, Cell 157 (2014) 77–94.
[32] L.M. Payer, K.H. Burns, Transposable elements in human genetic disease, Nat. Rev. Genet. 20 (2019) 760–772.
[33] A.D. Senft, T.S. Macfarlan, Transposable elements shape the evolution of mammalian development, Nat. Rev. Genet 22 (2021) 691–711.
[34] T. Messmer, F. von Meyenn, A. Savino, F. Santos, H. Mohammed, A.T.L. Lun, J. C. Marioni, W. Reik, Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution, Cell Rep. 26 (2019) 815–824, e814.
[35] S.G. Hillier, Gonadotropic control of ovarian follicular growth and development, Mol. Cell Endocrinol. 179 (2001) 39–46.
[36] T. Cavazza, Y. Takeda, A.Z. Politi, M. Aushev, P. Aldag, C. Baker, M. Choudhary, J. Bucevicius, G. Lukinavicius, K. Elder, et al., Parental genome unification is highly error-prone in mammalian embryos, Cell 184 (2021) 2860–2877, e2822.
[37] J. Van Blerkom, P. Davis, J. Merriam, J. Sinclair, Nuclear and cytoplasmic dynamics of sperm penetration, pronuclear formation and microtubule organization during fertilization and early preimplantation development in the human, Hum. Reprod. Update 1 (1995) 429–461.
[38] G. Capmany, A. Taylor, P.R. Braude, V.N. Bolton, The timing of pronuclear formation, DNA synthesis and cleavage in the human 1-cell embryo, Mol. Hum. Reprod. 2 (1996) 299–306.
[39] K.H. Campbell, J. McWhir, W.A. Ritchie, I. Wilmut, Sheep cloned by nuclear transfer from a cultured cell line, Nature 380 (1996) 64–66.
[40] J.B. Gurdon, The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles, J. Embryol. Exp. Morphol. 10 (1962) 622–640.
[41] J.B. Gurdon, V. Uehlinger, "Fertile" intestine nuclei, Nature 210 (1966) 1240–1241.
[42] I. Wilmut, A.E. Schnieke, J. McWhir, A.J. Kind, K.H. Campbell, Viable offspring derived from fetal and adult mammalian cells, Nature 385 (1997) 810–813.
[43] W. Tadros, H.D. Lipshitz, The maternal-to-zygotic transition: a play in two acts, Development 136 (2009) 3033–3042.
[44] P. Braude, V. Bolton, S. Moore, Human gene expression first occurs between the four- and eight-cell stages of preimplantation development, Nature 332 (1988) 459–461.
[45] J. Tesarik, V. Kopecny, M. Plachot, J. Mandelbaum, High-resolution autoradiographic localization of DNA-containing sites and RNA synthesis in developing nucleoli of human preimplantation embryos: a new concept of embryonic nucleologenesis, Development 101 (1987) 777–791.
[46] M. Asami, B.Y.H. Lam, M.K. Ma, K. Rainbow, S. Braun, M.D. VerMilyea, G.S. H. Yeo, A.C.F. Perry, Human embryonic genome activation initiates at the one-cell stage, Cell Stem Cell 29 (2022) 209–216, e204.
[47] E.M. Jouhilahti, E. Madissoon, L. Vesterlund, V. Tohonen, K. Krjutskov, A. Plaza Reyes, S. Petropoulos, R. Mansson, S. Linnarsson, T. Burglin, et al., The human PRD-like homeobox gene LEUTX has a central role in embryo genome activation, Development 143 (2016) 3459–3469.
[48] A. Deglincerti, G.F. Croft, L.N. Pietila, M. Zernicka-Goetz, E.D. Siggia, A. H. Brivanlou, Self-organization of the in vitro attached human embryo, Nature 533 (2016) 251–254.
[49] M.N. Shahbazi, A. Jedrusik, S. Vuoristo, G. Recher, A. Hupalowska, V. Bolton, N. N.M. Fogarty, A. Campbell, L. Devito, D. Ilic, et al., Self-organization of the human embryo in the absence of maternal tissues, Nat. Cell Biol. 18 (2016) 700–708.
[50] J. Rossant, P.P.L. Tam, Early human embryonic development: blastocyst formation to gastrulation, Dev. Cell 57 (2022) 152–165.
[51] R.C.V. Tyser, E. Mahammadov, S. Nakanoh, L. Vallier, A. Scialdone, S. Srinivas, Single-cell transcriptomic characterization of a gastrulating human embryo, Nature 600 (2021) 285–289.

[52] S.A. McGrath, A.F. Esquela, S.J. Lee, Oocyte-specific expression of growth/differentiation factor-9, Mol. Endocrinol. 9 (1995) 131–136.
[53] S. Canosa, T. Adriaenssens, W. Coucke, P. Dalmasso, A. Revelli, C. Benedetto, J. Smitz, Zona pellucida gene mRNA expression in human oocytes is related to oocyte maturity, zona inner layer retardance and fertilization competence, Mol. Hum. Reprod. 23 (2017) 292–303.
[54] G. Bourque, K.H. Burns, M. Gehring, V. Gorbunova, A. Seluanov, M. Hammell, M. Imbeault, Z. Izsvak, H.L. Levin, T.S. Macfarlan, et al., Ten things you should know about transposable elements, Genome Biol. 19 (2018) 199.
[55] C. Feschotte, Transposable elements and the evolution of regulatory networks, Nat. Rev. Genet 9 (2008) 397–405.
[56] C. Hermant, M.E. Torres-Padilla, TFs for TEs: the transcription factor repertoire of mammalian transposable elements, Genes Dev. 35 (2021) 22–39.
[57] J. Bouckenheimer, P. Fauque, C.H. Lecellier, C. Bruno, T. Commes, J.M. Lemaitre, J. De Vos, S. Assou, Differential long non-coding RNA expression profiles in human oocytes and cumulus cells, Sci. Rep. 8 (2018) 2202.
[58] I. Georgiou, D. Noutsopoulos, E. Dimitriadou, G. Markopoulos, A. Apergi, L. Lazaros, T. Vaxevanoglou, K. Pantos, M. Syrrou, T. Tzavaras, Retrotransposon RNA expression and evidence for retrotransposition events in human oocytes, Hum. Mol. Genet 18 (2009) 1221–1228.
[59] P. Gerdes, S.R. Richardson, D.L. Mager, G.J. Faulkner, Transposable elements in the mammalian embryo: pioneers surviving through stealth and service, Genome Biol. 17 (2016) 100.
[60] N.K. Goto, T. Zor, M. Martinez-Yamout, H.J. Dyson, P.E. Wright, Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain, J. Biol. Chem. 277 (2002) 43168–43174.
[61] B. Yu, N. Doni Jayavelu, S.L. Battle, J.C. Mar, T. Schimmel, J. Cohen, R. D. Hawkins, Single-cell analysis of transcriptome and DNA methylome in human oocyte maturation, PLoS One 15 (2020), e0241698.
[62] J. Brind'Amour, H. Kobayashi, J. Richard Albert, K. Shirane, A. Sakashita, A. Kamio, A. Bogutz, T. Koike, M.M. Karimi, L. Lefebvre, et al., LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation, Nat. Commun. 9 (2018) 3331.
[63] A.E. Peaston, A.V. Evsikov, J.H. Graber, W.N. de Vries, A.E. Holbrook, D. Solter, B.B. Knowles, Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos, Dev. Cell 7 (2004) 597–606.
[64] J. Wu, B. Huang, H. Chen, Q. Yin, Y. Liu, Y. Xiang, B. Zhang, B. Liu, Q. Wang, W. Xia, et al., The landscape of accessible chromatin in mammalian preimplantation embryos, Nature 534 (2016) 652–657.
[65] C.W. Hanna, H. Demond, G. Kelsey, Epigenetic regulation in development: is the mouse a good model for the human? Hum. Reprod. Update 24 (2018) 556–576.
[66] B. Yu, X. Dong, S. Gravina, O. Kartal, T. Schimmel, J. Cohen, D. Tortoriello, R. Zody, R.D. Hawkins, J. Vijg, GenOME-WIDE, SINGLE-Cell DNA methylomics reveals increased non-CpG methylation during human oocyte maturation, Stem Cell Rep. 9 (2017) 397–407.
[67] H. Guo, P. Zhu, L. Yan, R. Li, B. Hu, Y. Lian, J. Yan, X. Ren, S. Lin, J. Li, et al., The DNA methylation landscape of human early embryos, Nature 511 (2014) 606–610.
[68] P. Zhu, H. Guo, Y. Ren, Y. Hou, J. Dong, R. Li, Y. Lian, X. Fan, B. Hu, Y. Gao, et al., Single-cell DNA methylome sequencing of human preimplantation embryos, Nat. Genet. 50 (2018) 12–19.
[69] Y.Q. Su, F. Sun, M.A. Handel, J.C. Schimenti, J.J. Eppig, Meiosis arrest female 1 (MARF1) has nuage-like function in mammalian oocytes, Proc. Natl. Acad. Sci. USA 109 (2012) 18653–18660.
[70] Q. Yang, J. Lin, M. Liu, R. Li, B. Tian, X. Zhang, B. Xu, M. Liu, X. Zhang, Y. Li, et al., Highly sensitive sequencing reveals dynamic modifications and activities of small RNAs in mouse oocytes and early embryos, Sci. Adv. 2 (2016), e1501482.
[71] J. Pontis, E. Planet, S. Offner, P. Turelli, J. Duc, A. Coudray, T.W. Theunissen, R. Jaenisch, D. Trono, Hominoid-specific transposable elements and KZFPs facilitate human embryonic genome activation and control transcription in naive human ESCs, Cell Stem Cell 24 (2019) 724–735, e725.
[72] E.M. Wissink, A. Vihervaara, N.D. Tippens, J.T. Lis, Nascent RNA analyses: tracking transcription and its regulation, Nat. Rev. Genet 20 (2019) 705–723.
[73] N.L. Vastenhouw, W.X. Cao, H.D. Lipshitz, The maternal-to-zygotic transition revisited, Development (2019) *146*.
[74] Q.Q. Sha, W. Zheng, Y.W. Wu, S. Li, L. Guo, S. Zhang, G. Lin, X.H. Ou, H.Y. Fan, Dynamics and clinical relevance of maternal mRNA clearance during the oocyte-to-embryo transition in humans, Nat. Commun. 11 (2020) 4917.
[75] R.M. Schultz, P. Stein, P. Svoboda, The oocyte-to-embryo transition in mouse: past, present, and future, Biol. Reprod. 99 (2018) 160–174.
[76] X. Liu, J.P. Tan, J. Schroder, A. Aberkane, J.F. Ouyang, M. Mohenska, S.M. Lim, Y.B.Y. Sun, J. Chen, G. Sun, et al., Modelling human blastocysts by reprogramming fibroblasts into iBlastoids, Nature 591 (2021) 627–632.
[77] Liu, Y., Wu, K., Shao, F., Nie, H., Zhang, J., Li, C., Hou, Z., Wang, J., Zhou, B., Zhao, H., et al. , 2021b. Dynamics of poly(A) tail length and non-A residues during the human oocyte-to-embryo transition. bioRxiv.
[78] Liu, Y., Zhang, Y., Nie, H., Liu, Z., Wang, J., Lu, F. (2021c). Re-polyadenylation occurs predominantly on maternal mRNA degradation intermediates during mammalian oocyte-to-embryo transition. bioRxiv.
[79] Y. Liu, X. Lu, J. Shi, X. Yu, X. Zhang, K. Zhu, Z. Yi, E. Duan, L. Li, BTG4 is a key regulator for maternal mRNA clearance during mouse early embryogenesis, J. Mol. Cell Biol. 8 (2016) 366–368.
[80] C. Yu, S.Y. Ji, Q.Q. Sha, Y. Dang, J.J. Zhou, Y.L. Zhang, Y. Liu, Z.W. Wang, B. Hu, Q.Y. Sun, et al., BTG4 is a meiotic cell cycle-coupled maternal-zygotic-transition licensing factor in oocytes, Nat. Struct. Mol. Biol. 23 (2016) 387–394.
[81] S. Yu, V.N. Kim, A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing, Nat. Rev. Mol. Cell Biol. 21 (2020) 542–556.
[82] C. De Almeida, H. Scheer, H. Zuber, D. Gagliardi, RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome, Wiley Inter. Rev. RNA (2018) 9.
[83] J. Lim, M. Ha, H. Chang, S.C. Kwon, D.K. Simanshu, D.J. Patel, V.N. Kim, Uridylation by TUT4 and TUT7 marks mRNA for degradation, Cell 159 (2014) 1365–1376.
[84] M. Morgan, C. Much, M. DiGiacomo, C. Azzi, I. Ivanova, D.M. Vitsios, J. Pistolic, P. Collier, P.N. Moreira, V. Benes, et al., mRNA 3' uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome, Nature 548 (2017) 347–351.
[85] M. Lee, Y. Choi, K. Kim, H. Jin, J. Lim, T.A. Nguyen, J. Yang, M. Jeong, A. J. Giraldez, H. Yang, et al., Adenylation of maternally inherited microRNAs by Wispy, Mol. Cell 56 (2014) 696–707.
[86] C.T. Neilsen, G.J. Goodall, C.P. Bracken, IsomiRs–the overlooked repertoire in the dynamic microRNAome, Trends Genet 28 (2012) 544–549.
[87] Y. Ohnishi, Y. Totoki, A. Toyoda, T. Watanabe, Y. Yamamoto, K. Tokunaga, Y. Sakaki, H. Sasaki, H. Hohjoh, Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development, Nucleic Acids Res 38 (2010) 5141–5151.
[88] F. Tang, M. Kaneda, D. O'Carroll, P. Hajkova, S.C. Barton, Y.A. Sun, C. Lee, A. Tarakhovsky, K. Lao, M.A. Surani, Maternal microRNAs are essential for mouse zygotic development, Genes Dev. 21 (2007) 644–648.
[89] A.M. Burroughs, Y. Ando, M.J. de Hoon, Y. Tomaru, T. Nishibu, R. Ukekawa, T. Funakoshi, T. Kurokawa, H. Suzuki, Y. Hayashizaki, et al., A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness, Genome Res. 20 (2010) 1398–1410.
[90] A. D'Ambrogio, W. Gu, T. Udagawa, C.C. Mello, J.D. Richter, Specific miRNA stabilization by Gld2-catalyzed monoadenylation, Cell Rep. 2 (2012) 1537–1545.
[91] T. Katoh, Y. Sakaguchi, K. Miyauchi, T. Suzuki, S. Kashiwabara, T. Baba, T. Suzuki, Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2, Genes Dev. 23 (2009) 433–438.
[92] J. Durruthy-Durruthy, V. Sebastiano, M. Wossidlo, D. Cepeda, J. Cui, E.J. Grow, J. Davila, M. Mall, W.H. Wong, J. Wysocka, et al., The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming, Nat. Genet. 48 (2016) 44–52.
[93] G. Falco, S.L. Lee, I. Stanghellini, U.C. Bassey, T. Hamatani, M.S. Ko, Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells, Dev. Biol. 307 (2007) 539–550.
[94] K. Takahashi, P.J. Ross, K. Sawai, The necessity of ZSCAN4 for preimplantation development and gene expression of bovine embryos, J. Reprod. Dev. 65 (2019) 319–326.
[95] Z. Izsvak, J. Wang, M. Singh, D.L. Mager, L.D. Hurst, Pluripotency and the endogenous retrovirus HERVH: conflict or serendipity? Bioessays 38 (2016) 109–117.
[96] D. Wolf, S.P. Goff, Embryonic stem cells use ZFP809 to silence retroviral DNAs, Nature 458 (2009) 1201–1204.
[97] S.H. Choi, M.D. Gearhart, Z. Cui, D. Bosnakovski, M. Kim, N. Schennum, M. Kyba, DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes, Nucleic Acids Res. 44 (2016) 5161–5173.
[98] Y. Zhang, T. Li, S. Preissl, M.L. Amaral, J.D. Grinstein, E.N. Farah, E. Destici, Y. Qiu, R. Hu, A.Y. Lee, et al., Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells, Nat. Genet. 51 (2019) 1380–1388.
[99] C.W. Schmid, Alu: structure, origin, evolution, significance and function of one-tenth of human DNA, Prog. Nucleic Acid. Res Mol. Biol. 53 (1996) 283–319.
[100] P. Polak, E. Domany, Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes, BMC Genom. 7 (2006) 133.
[101] E. Ing-Simmons, M. Rigau, J.M. Vaquerizas, Emerging mechanisms and dynamics of three-dimensional genome organisation at zygotic genome activation, Curr. Opin. Cell Biol. 74 (2022) 37–46.
[102] Kruse, K., Diaz, N., Enriquez-Gasca, R., Gaume, X., Torres-Padilla, M.-E., Vaquerizas, J.M. (2019). Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. bioRxiv.
[103] C. Gerri, A. McCarthy, G. Alanis-Lobato, A. Demtschenko, A. Bruneau, S. Loubersac, N.M.E. Fogarty, D. Hampshire, K. Elder, P. Snell, et al., Initiation of a conserved trophectoderm program in human, cow and mouse embryos, Nature 587 (2020) 443–447.
[104] M. Roode, K. Blair, P. Snell, K. Elder, S. Marchant, A. Smith, J. Nichols, Human hypoblast formation is not dependent on FGF signalling, Dev. Biol. 361 (2012) 358–363.
[105] H. Niwa, Y. Toyooka, D. Shimosato, D. Strumpf, K. Takahashi, R. Yagi, J. Rossant, Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation, Cell 123 (2005) 917–929.
[106] D. Yunusov, L. Anderson, L.F. DaSilva, J. Wysocka, T. Ezashi, R.M. Roberts, S. Verjovski-Almeida, HIPSTR and thousands of lncRNAs are heterogeneously expressed in human embryos, primordial germ cells and stable cell lines, Sci. Rep. 6 (2016) 32753.
[107] R. Alberio, Transcriptional and epigenetic control of cell fate decisions in early embryos, Reprod. Fertil. Dev. 30 (2017) 73–84.
[108] R. Alberio, Untangling early embryo development using single cell genomics, Theriogenology 150 (2020) 55–58.

[109] R. Alberio, T. Kobayashi, M.A. Surani, Conserved features of non-primate bilaminar disc embryos and the germline, Stem Cell Rep. 16 (2021) 1078–1092.
[110] Apter, S., Ebner, T., Freour, T., Guns, Y., Kovacic, B., Le Clef, N., Marques, M., Meseguer, M., Montjean, D., Sfontouris, I., et al. , 2020. Good practice recommendations for the use of time-lapse technology (dagger). Hum Reprod Open 2020, hoaa008.
[111] E.W. Kuijk, L.T. van Tol, H. Van de Velde, R. Wubbolts, M. Welling, N. Geijsen, B. A. Roelen, The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos, Development 139 (2012) 871–882.
[112] J.J. Qiu, Z.R. Ren, J.B. Yan, Identification and functional analysis of long non-coding RNAs in human and mouse early embryos based on single-cell transcriptome data, Oncotarget 7 (2016) 61215–61228.
[113] K.F. Au, V. Sebastiano, P.T. Afshar, J.D. Durruthy, L. Lee, B.A. Williams, H. van Bakel, E.E. Schadt, R.A. Reijo-Pera, J.G. Underwood, et al., Characterization of the human ESC transcriptome by hybrid sequencing, Proc. Natl. Acad. Sci. USA 110 (2013) E4821–4830.
[114] J. Durruthy-Durruthy, M. Wossidlo, S. Pai, Y. Takahashi, G. Kang, L. Omberg, B. Chen, H. Nakauchi, R. Reijo Pera, V. Sebastiano,). Spatiotemporal reconstruction of the human blastocyst by single-cell gene-expression analysis informs induction of naive pluripotency, Dev. Cell 38 (2016) 100–115.
[115] G. Glinsky, J. Durruthy-Durruthy, M. Wossidlo, E.J. Grow, J.L. Weirather, K.F. Au, J. Wysocka, V. Sebastiano, Single cell expression analysis of primate-specific retroviruses-derived HPAT lincRNAs in viable human blastocysts identifies embryonic cells co-expressing genetic markers of multiple lineages, Heliyon 4 (2018), e00667.
[116] G. Guo, G.G. Stirparo, S.E. Strawbridge, D. Spindlow, J. Yang, J. Clarke, A. Dattani, A. Yanagida, M.A. Li, S. Myers, et al., Human naive epiblast cells possess unrestricted lineage potential, Cell Stem Cell 28 (1040–1056) (2021), e1046.
[117] J. Wang, G. Xie, M. Singh, A.T. Ghanbarian, T. Rasko, A. Szvetnik, H. Cai, D. Besser, A. Prigione, N.V. Fuchs, et al., Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells, Nature 516 (2014) 405–409.
[118] E.M. Rosenbluth, D.N. Shelton, A.E. Sparks, E. Devor, L. Christenson, B.J. Van Voorhis, MicroRNA expression in the human blastocyst, Fertil. Steril. 99 (2013) 855–861, e853.
[119] K. Goossens, W. De Spiegelaere, M. Stevens, C. Burvenich, B. De Spiegeleer, P. Cornillie, A. Van Zeveren, A. Van Soom, L. Peelman, Differential microRNA expression analysis in blastocysts by whole mount in situ hybridization and reverse transcription quantitative polymerase chain reaction on laser capture microdissection samples, Anal. Biochem. 423 (2012) 93–101.
[120] K. Goossens, P. Mestdagh, S. Lefever, M. Van Poucke, A. Van Zeveren, A. Van Soom, J. Vandesompele, L. Peelman, Regulatory microRNA network identification in bovine blastocyst development, Stem Cells Dev. 22 (2013) 1907–1920.
[121] S.R. Viswanathan, C.H. Mermel, J. Lu, C.W. Lu, T.R. Golub, G.Q. Daley, microRNA expression during trophectoderm specification, PLoS One 4 (2009), e6143.
[122] T. Spruce, B. Pernaute, A. Di-Gregorio, B.S. Cobb, M. Merkenschlager, M. Manzanares, T.A. Rodriguez, An early developmental role for miRNAs in the maintenance of extraembryonic stem cells in the mouse embryo, Dev. Cell 19 (2010) 207–219.
[123] E. Bernstein, S.Y. Kim, M.A. Carmell, E.P. Murchison, H. Alcorn, M.Z. Li, A. A. Mills, S.J. Elledge, K.V. Anderson, G.J. Hannon, Dicer is essential for mouse development, Nat. Genet. 35 (2003) 215–217.
[124] J. Ma, M. Flemr, P. Stein, P. Berninger, R. Malik, M. Zavolan, P. Svoboda, R. M. Schultz, MicroRNA activity is suppressed in mouse oocytes, Curr. Biol. 20 (2010) 265–270.
[125] N. Suh, L. Baehner, F. Moltzahn, C. Melton, A. Shenoy, J. Chen, R. Blelloch, MicroRNA function is globally suppressed in mouse oocytes and early embryos, Curr. Biol. 20 (2010) 271–277.
[126] D. Subramanyam, S. Lamouille, R.L. Judson, J.Y. Liu, N. Bucay, R. Derynck, R. Blelloch, Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells, Nat. Biotechnol. 29 (2011) 443–448.
[127] O.R. Faridani, I. Abdullayev, M. Hagemann-Jensen, J.P. Schell, F. Lanner, R. Sandberg, Single-cell sequencing of the small-RNA transcriptome, Nat. Biotechnol. 34 (2016) 1264–1266.
[128] H. Sperber, J. Mathieu, Y. Wang, A. Ferreccio, J. Hesson, Z. Xu, K.A. Fischer, A. Devi, D. Detraux, H. Gu, et al., The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition, Nat. Cell Biol. 17 (2015) 1523–1535.
[129] K. Yuan, W.B. Ai, L.Y. Wan, X. Tan, J.F. Wu, The miR-290-295 cluster as multi-faceted players in mouse embryonic stem cells. Cell Biosci. 38 (2017) https://doi.org/10.1186/s13578-017-0166-2.
[130] L.A Medeiros, L.M Dennis, M.E. Gill, H. Houbaviy, S. Markoulaki, D. Fu, A. C. White, O. Kirak, P.A. Sharp, D.C. Page, R. Jaenisch, Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. Proc Natl Acad Sci 108 (34) (2011) 14163-8, https://doi.org/10.1073/pnas.1111241108.
[131] S. Krishna, D.G. Yim, V. Lakshmanan, V. Tirumalai, J.L. Koh, J.E. Park, J. K. Cheong, J.L. Low, M.J. Lim, S.K. Sze, et al., Dynamic expression of tRNA-derived small RNAs define cellular states, EMBO Rep. 20 (2019), e47789.
[132] T.R. Leonardo, H.L. Schultheisz, J.F. Loring, L.C. Laurent, The functions of microRNAs in pluripotency and reprogramming, Nat. Cell Biol. 14 (2012) 1114–1121.
[133] J. Shi, Y. Zhang, D. Tan, X. Zhang, M. Yan, Y. Zhang, R. Franklin, M. Shahbazi, K. Mackinlay, S. Liu, et al., PANDORA-seq expands the repertoire of regulatory small RNAs by overcoming RNA modifications, Nat. Cell Biol. 23 (2021) 424–436.
[134] M.N. Shahbazi, A. Scialdone, N. Skorupska, A. Weberling, G. Recher, M. Zhu, A. Jedrusik, L.G. Devito, L. Noli, I.C. Macaulay, et al., Pluripotent state transitions coordinate morphogenesis in mouse and human embryos, Nature 552 (2017) 239–243.
[135] M. Zhu, M. Zernicka-Goetz, Principles of self-organization of the mammalian embryo, Cell 183 (2020) 1467–1478.
[136] Xie, S.Q., Leeke, B.J., Whidling, C., Wagner, R.T., Garcia-Llagostera, F., Chammas, P., Cheung, N.T.-F., Dormann, D., McManus, M.T., and Percharde, M. (2021). Nucleolar-based Dux repression is essential for 2-cell stage exit. bioRxiv.
[137] H. Yu, Z. Sun, T. Tan, H. Pan, J. Zhao, L. Zhang, J. Chen, A. Lei, Y. Zhu, L. Chen, et al., rRNA biogenesis regulates mouse 2C-like state by 3D structure reorganization of peri-nucleolar heterochromatin, Nat. Commun. 12 (2021) 6365.
[138] J. Xu, L. Feng, Z. Han, Y. Li, A. Wu, T. Shao, N. Ding, L. Li, W. Deng, X. Di, et al., Extensive ceRNA-ceRNA interaction networks mediated by miRNAs regulate development in multiple rhesus tissues, Nucleic Acids Res. 44 (2016) 9438–9451.
[139] I. Okamoto, C. Patrat, D. Thepot, N. Peynot, P. Fauque, N. Daniel, P. Diabangouaya, J.P. Wolf, J.P. Renard, V. Duranthon, et al., Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development, Nature 472 (2011) 370–374.
[140] C. Patrat, J.F. Ouimette, C. Rougeulle, X chromosome inactivation in human development, Development (2020) *147*.
[141] C. Vallot, C. Patrat, A.J. Collier, C. Huret, M. Casanova, T.M. Liyakat Ali, M. Tosolini, N. Frydman, E. Heard, P.J. Rugg-Gunn, et al., . XACT noncoding RNA competes with XIST in the control of X chromosome activity during human early development, Cell Stem Cell 20 (2017) 102–111.
[142] I.M. van den Berg, J.S. Laven, M. Stevens, I. Jonkers, R.J. Galjaard, J. Gribnau, J. H. van Doorninck, X chromosome inactivation is initiated in human preimplantation embryos, Am. J. Hum. Genet. 84 (2009) 771–779.
[143] C.D. Malone, G.J. Hannon, Small RNAs as guardians of the genome, Cell 136 (2009) 656–668.
[144] J.M. Zhang, W.B. Hou, J.W. Du, M. Zong, K.L. Zheng, W.J. Wang, J.Q. Wang, H. Zhang, Y.S. Mu, Z. Yin, et al., Argonaute 2 is a key regulator of maternal mRNA degradation in mouse early embryos, Cell Death Discov. 6 (2020) 133.
[145] R.F. Place, L.C. Li, D. Pookot, E.J. Noonan, R. Dahiya, MicroRNA-373 induces expression of genes with complementary promoter sequences, Proc. Natl. Acad. Sci. USA 105 (2008) 1608–1613.
[146] Y. Zhang, M. Fan, X. Zhang, F. Huang, K. Wu, J. Zhang, J. Liu, Z. Huang, H. Luo, L. Tao, et al., Cellular microRNAs up-regulate transcription via interaction with promoter TATA-box motifs, RNA 20 (2014) 1878–1889.
[147] I.M. Flyamer, J. Gassler, M. Imakaev, H.B. Brandao, S.V. Ulianov, N. Abdennur, S. V. Razin, L.A. Mirny, K. Tachibana-Konwalski, Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition, Nature 544 (2017) 110–114.
[148] H.S. Kaya-Okur, S.J. Wu, C.A. Codomo, E.S. Pledger, T.D. Bryson, J.G. Henikoff, K. Ahmad, S. Henikoff, CUT&Tag for efficient epigenomic profiling of small samples and single cells, Nat. Commun. 10 (2019) 1930.
[149] S. Hirabayashi, S. Bhagat, Y. Matsuki, Y. Takegami, T. Uehata, A. Kanemaru, M. Itoh, K. Shirakawa, A. Takaori-Kondo, O. Takeuchi, et al., NET-CAGE characterizes the dynamics and topology of human transcribed cis-regulatory elements, Nat. Genet. 51 (2019) 1369–1379.
[150] S.A. Quinodoz, J.W. Jachowicz, P. Bhat, N. Ollikainen, A.K. Banerjee, I. N. Goronzy, M.R. Blanco, P. Chovanec, A. Chow, Y. Markaki, et al., RNA promotes the formation of spatial compartments in the nucleus, Cell 184 (2021) 5775–5790, e5730.
[151] N. Amin, A. McGrath, Y.-P.P. Chen, Evaluation of deep learning in non-coding RNA classification, Nat. Mach. Learn. (2019).
[152] C.C.R. Hartford, A. Lal, When long noncoding becomes protein coding, Mol. Cell Biol. (2020) 40.
[153] J.W. Nam, S.W. Choi, B.H. You, Incredible RNA: dual functions of coding and noncoding, Mol. Cells 39 (2016) 367–374.
[154] J.A. Ramilowski, C.W. Yip, S. Agrawal, J.C. Chang, Y. Ciani, I.V. Kulakovskiy, M. Mendez, J.L.C. Ooi, J.F. Ouyang, N. Parkinson, et al., Functional annotation of human long noncoding RNAs via molecular phenotyping, Genome Res 30 (2020) 1060–1072.
[155] D. Andergassen, J.L. Rinn, From genotype to phenotype: genetics of mammalian long non-coding RNAs in vivo, Nat. Rev. Genet (2021).
[156] D.P. Bartel, Metazoan microRNAs, Cell 173 (2018) 20–51.
[157] M. Percharde, T. Sultana, M. Ramalho-Santos, What doesn't kill you makes you stronger: transposons as dual players in chromatin regulation and genomic variation, Bioessays 42 (2020), e1900232.
[158] A.J. Modzelewski, W. Shao, J. Chen, A. Lee, X. Qi, M. Noon, K. Tjokro, G. Sales, A. Biton, A. Anand, et al., A mouse-specific retrotransposon drives a conserved Cdk2ap1 isoform essential for development, Cell 184 (2021) 5541–5558, e5522.
[159] J.E. Cropley, S.A. Eaton, A. Aiken, P.E. Young, E. Giannoulatou, J.W.K. Ho, M. E. Buckland, S.P. Keam, G. Hutvagner, D.T. Humphreys, et al., Male-lineage transmission of an acquired metabolic phenotype induced by grand-paternal obesity, Mol. Metab. 5 (2016) 699–708.
[160] V. Grandjean, S. Fourre, D.A. De Abreu, M.A. Derieppe, J.J. Remy, M. Rassoulzadegan, RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders, Sci. Rep. 5 (2015) 18193.
[161] G. Sarker, W. Sun, D. Rosenkranz, P. Pelczar, L. Opitz, V. Efthymiou, C. Wolfrum, D. Peleg-Raibstein, Maternal overnutrition programs hedonic and metabolic

phenotypes across generations through sperm tsRNAs, Proc. Natl. Acad. Sci. USA 116 (2019) 10547–10556.
[162] A. Ashe, A. Sapetschnig, E.M. Weick, J. Mitchell, M.P. Bagijn, A.C. Cording, A. L. Doebley, L.D. Goldstein, N.J. Lehrbach, J. Le Pen, et al., piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans, Cell 150 (2012) 88–99.
[163] T. Grentzinger, C. Armenise, C. Brun, B. Mugat, V. Serrano, A. Pelisson, S. Chambeyron, piRNA-mediated transgenerational inheritance of an acquired trait, Genome Res. 22 (2012) 1877–1888.
[164] Q. Chen, M. Yan, Z. Cao, X. Li, Y. Zhang, J. Shi, G.H. Feng, H. Peng, X. Zhang, Y. Zhang, et al., Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder, Science 351 (2016) 397–400.
[165] D. Natt, U. Kugelberg, E. Casas, E. Nedstrand, S. Zalavary, P. Henriksson, C. Nijm, J. Jaderquist, J. Sandborg, E. Flinke, et al., Human sperm displays rapid responses to diet, PLoS Biol. 17 (2019), e3000559.
[166] H. Peng, J. Shi, Y. Zhang, H. Zhang, S. Liao, W. Li, L. Lei, C. Han, L. Ning, Y. Cao, et al., A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm, Cell Res. 22 (2012) 1609–1612.
[167] Y. Zhang, X. Zhang, J. Shi, F. Tuorto, X. Li, Y. Liu, R. Liebers, L. Zhang, Y. Qu, J. Qian, et al., Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs, Nat. Cell Biol. 20 (2018) 535–540.
[168] X. Zhang, F. Gao, J. Fu, P. Zhang, Y. Wang, X. Zeng, Systematic identification and characterization of long non-coding RNAs in mouse mature sperm, PLoS One 12 (2017), e0173402.
[169] N. Moris, K. Anlas, S.C. van den Brink, A. Alemany, J. Schroder, S. Ghimire, T. Balayo, A. van Oudenaarden, A. Martinez Arias, An in vitro model of early anteroposterior organization during human development, Nature 582 (2020) 410–415.
[170] B. Sozen, V. Jorgensen, B.A.T. Weatherbee, S. Chen, M. Zhu, M. Zernicka-Goetz, Reconstructing aspects of human embryogenesis with pluripotent stem cells, Nat. Commun. 12 (2021) 5550.
[171] S.C. van den Brink, A. van Oudenaarden, 3D gastruloids: a novel frontier in stem cell-based in vitro modeling of mammalian gastrulation, Trends Cell Biol. 31 (2021) 747–759.
[172] A. Yanagida, D. Spindlow, J. Nichols, A. Dattani, A. Smith, G. Guo, Naive stem cell blastocyst model captures human embryo lineage segregation, Cell Stem Cell 28 (1016–1022) (2021), e1014.
[173] L. Yu, Y. Wei, J. Duan, D.A. Schmitz, M. Sakurai, L. Wang, K. Wang, S. Zhao, G. C. Hon, J. Wu, Blastocyst-like structures generated from human pluripotent stem cells, Nature 591 (2021) 620–626.
[174] Taubenschmid-Stowers, J., Rostovskaya, M., Santos, F., Ljung, S., Argelaguet, R., Krueger, F., Nichols, J., and Reik, W. (2021). Modelling human zygotic genome activation in 8C-like cells in vitro. bioRxiv.
[175] Yoshihara, M.K., I. Nykänen, S. Sokka, J. Weltner, J. Lundin, K. Gawriyski, L. Jouhilahti, E.-M. Varjosalo, M. Tervaniemi MH. Otonkoski, T. Trokovic, R. Katayama, S. Vuoristo, S. Kere, J. (2021). Transient DUX4 expression induces blastomere-like expression program that is marked by SLC34A2. In bioRxiv.

Glossary

Maternal-to-zygotic transition: The maternal-to-zygotic transition entails the switch in control from the maternal to the zygotic genome during embryonic development. This requires both the transcriptional activation of the embryonic genome that has assembled from the maternally and paternally inherited haploid genomes and the degradation of maternal products stored in the oocyte.

Embryonic genome activation: The onset of transcription from the embryo genome, is called embryonic genome activation. Transcription is initiated in a stepwise manner with temporal differences between species. In human embryos, the first phase of transcription, a "minor" wave occurs by the 4-cell stage. The "major" wave occurs at the 8-cell stage and transcriptional activity is increased.

Non-coding genome: The non-coding genome refers to DNA sequences that do not encode for proteins. The non-coding genome functions in transcriptional regulation, chromatin organization, DNA replication, maintenance of chromosomal stability and segregation, and is also transcribed producing functional non-coding RNA molecules. Of the human genome 98–99% is considered non-coding and the ratio of coding to non-coding DNA varies greatly among species.

Transposable element: Transposable elements are mobile DNA sequence that account for approximately 50% of human DNA. Transposable elements are divided into Class I and II, that is retrotransposon and DNA transposons, respectively. The former move via reverse transcription while the latter encode the protein transposase enabling their independent transposition. Transposition often results in the duplication of the original moving sequence. Transposable elements provide regulatory sequences and an efficient means of genomic regulation as they exist in multiple copies across the genome.

Retrotransposon: Retrotransposons are divided into two groups, distinguished by the presence or absence of long terminal repeats. The vast majority of human transposable elements are non-long terminal repeat retrotransposons, mainly long interspersed nuclear element 1, Alu, and SINE-R, VNTR and ALU (SVA) elements that together make up approximately one third of the human genome. Human long terminal repeat elements are endogenous retroviruses and they account for approximately 8% of the genome.

HERV: Human endogenous retroviruses are DNA sequences of retroviral origin that have been acquired over the last 100 million years as the result of multiple integrations by ancient viruses. Most Human endogenous retroviruses have lost their coding potential due to mutations and produce mainly non-coding RNAs, however some have retained residual protein coding capacity. Human endogenous retroviruses are typically composed of two flanking long terminal repeats and an internal portion with viral genes gag, pro-pol and env.

Non-coding RNA: Non-coding RNAs are molecules that are not transcribed into proteins but can be functional. Non-coding RNA comprises multiple subclasses with the main division done according to sequence length into long and small non-coding RNAs. Non-coding RNAs create networks that include also coding transcripts and proteins to regulate chromatin conformation, epigenetic state, transcription, translation and RNA modification and turnover.

Long non-coding RNA: Long non-coding RNAs are over 200 nucleotides long and possess very little protein-coding potential. The FANTOM5 project reported 27,919 long non-coding RNAs in various human samples. Most long non-coding RNAs are transcribed as networks of protein-coding gene overlapping sense and antisense transcripts or are intergenic.

Small non-coding RNA: Small non-coding RNAs are a class of approximately 30 nucleotides long molecules with diverse regulatory functions during germ cell and early embryonic development as well as in somatic tissues. Mammalian oocytes and embryos primarily express Piwi-interacting RNAs, endogenous small interfering RNAs, micro RNAs, and ribosomal RNA- and transfer RNA-derived small RNAs.

Micro RNA: Micro RNAs are ~22 nucleotide long RNAs that mediate gene silencing via binding to partially complementary elements in the 3′ untranslated region of target mRNAs, resulting in mRNA deadenylation and/or inhibition of translation. Micro RNAs are generated by sequential DROSHA/DGCR8 and DICER1 cleavage of hairpin structures embedded in long transcripts and associate with Argonaute proteins to form the RNA-induced silencing complex.

PIWI-interacting RNA: Piwi-interacting RNAs are generated from primary transcripts that are cleaved into intermediates by the nuclease Zucchini, subsequently bound by PIWI proteins, further trimmed to lengths of ~26–30 nt and 2′-O-methylated at their 3′ termini by Hua enhancer 1. Most mammalian genomes encode four PIWI proteins (PIWIL1–4). The PIWI proteins form a complex with the primary piwi-interacting RNAs and cleave target RNAs between positions 10 and 11 of the complementary sequence. Piwi-interacting RNAs are also generated via a secondary mechanism upon binding to their target RNAs and this is called the "ping-pong" model for biogenesis. Primary piwi-interacting RNAs often have an uracil at the 5′ end however, secondary Piwi-interacting RNAs usually also have an A residue at position 10 and a 10 nucleotide 5′- end overlap with the primary Piwi-interacting RNAs. Piwi-interacting RNAs contribute to the silencing of transposable elements primarily in germ line cells.

Oocyte short piRNAs: Oocyte short piwi-interacting RNAs are ~20 nucleotides long, which is shorter than the piwi-interacting RNAs found in the testes and ovaries of other species. Oocyte short piwi-interacting RNAs specifically associate with PIWIL3 and lack 2'-O-methylation at the 3' end. They also have the 1 U/10 A sequence feature like the long piwi-interacting RNAs in human oocytes however, their overall sequence characteristics show closer resemblance to the mouse endogenous short interfering RNAs. Oocyte short piwi-interacting RNAs are primarily expressed in primate oocytes, lacking in mouse oocytes, and are likely to mediate the silencing of recently evolved transposable elements.