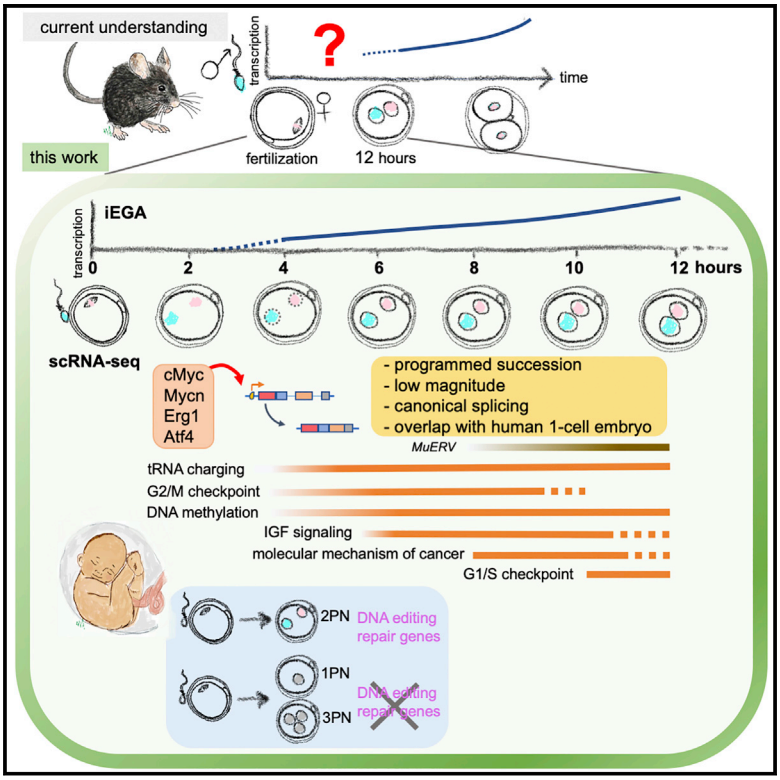
Cell Reports

Resource

A program of successive gene expression in mouse one-cell embryos

Graphical abstract



Highlights

- A gene expression program, iEGA, initiates within 4 h of mouse fertilization
- Upregulated genes are normatively spliced, protein coded, and soon downregulated
- iEGA genes predict cancer-associated pathways and transcription regulators
- Inhibiting predicted transcription regulators acutely disrupts iEGA and development

Authors

Maki Asami, Brian Y.H. Lam, Martin Hoffmann, ..., Giles S.H. Yeo, Christoph A. Klein, Anthony C.F. Perry

Correspondence

gshy2@cam.ac.uk (G.S.H.Y.), christoph.klein@klinik.uni-regensburg.de (C.A.K.), perry135@aol.com (A.C.F.P.)

In brief

How and when embryonic transcription begins following fertilization remains elusive. Asami et al. generate a high-resolution time course to reveal a transcription program initiating within 4 h of fertilization in mouse one-cell embryos. The program predicts embryonic and cancer-associated pathway expression, and inhibiting inferred regulators acutely blocks transcription and development.

Asami et al., 2023, Cell Reports *42*, 112023
February 28, 2023
https://doi.org/10.1016/j.celrep.2023.112023

CellPress

Cell Reports

CellPress
OPEN ACCESS

Resource

A program of successive gene expression in mouse one-cell embryos

Maki Asami,[1,7] Brian Y.H. Lam,[2,7] Martin Hoffmann,[3] Toru Suzuki,[1] Xin Lu,[4] Naoko Yoshida,[5] Marcella K. Ma,[2] Kara Rainbow,[2] Miodrag Gužvić,[4] Matthew D. VerMilyea,[6] Giles S.H. Yeo,[2,*] Christoph A. Klein,[3,4,*] and Anthony C.F. Perry[1,8,*]

[1]Laboratory of Mammalian Molecular Embryology, Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK
[2]Medical Research Council (MRC) Metabolic Diseases Unit, Wellcome-MRC Institute of Metabolic Science-Metabolic Research Laboratories, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 0QQ, UK
[3]Project Group Personalized Tumor Therapy, Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Regensburg, Germany
[4]Experimental Medicine and Therapy Research, University of Regensburg, Regensburg, Germany
[5]Department of Pathology, Kansai Medical University, Osaka 573-1010, Japan
[6]Embryology and Andrology Laboratories, Ovation Fertility Austin, Austin, TX 78731, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: gshy2@cam.ac.uk (G.S.H.Y.), christoph.klein@klinik.uni-regensburg.de (C.A.K.), perry135@aol.com (A.C.F.P.)
https://doi.org/10.1016/j.celrep.2023.112023

SUMMARY

At the moment of union in fertilization, sperm and oocyte are transcriptionally silent. The ensuing onset of embryonic transcription (embryonic genome activation [EGA]) is critical for development, yet its timing and profile remain elusive in any vertebrate species. We here dissect transcription during EGA by high-resolution single-cell RNA sequencing of precisely synchronized mouse one-cell embryos. This reveals a program of embryonic gene expression (immediate EGA [iEGA]) initiating within 4 h of fertilization. Expression during iEGA produces canonically spliced transcripts, occurs substantially from the maternal genome, and is mostly downregulated at the two-cell stage. Transcribed genes predict regulation by transcription factors (TFs) associated with cancer, including c-Myc. Blocking c-Myc or other predicted regulatory TF activities disrupts iEGA and induces acute developmental arrest. These findings illuminate intracellular mechanisms that regulate the onset of mammalian development and hold promise for the study of cancer.

INTRODUCTION

In fertilization, two heterotypic cells (the gametes, sperm and oocyte) combine to cause formation of a totipotent one-cell embryo.[1] This is a foundational developmental event that coincides with embryonic genome activation (EGA), in which transcription in the new embryo initiates from gamete-derived genomes that had been transcriptionally silent until fertilization.[2,3] EGA occurs in a milieu of complex biochemical and physical changes, many unique, within nascent one-cell embryos.[3–6] Phospho-relays with multiple targets, including the cytostatic factor Emi2, precipitate meiotic cell cycle progression[3,6–8] and are concurrent with specialized parental chromatin remodeling,[3,9] which presumptively regulates EGA. The major sperm nucleoprotein protamine[10] is removed by oocyte-derived nucleoplasmin activity and replaced by maternal histones prior to S phase, which begins ~8 h after fertilization in the mouse.[3,11] The new embryo undergoes atypical patterns of histone modification[12,13] and chromosome organisation,[14–16] but conclusions differ regarding the extent to which chromatin structure is inherited from the gametes[15,17] or assembled anew.[14,16] Parental genomes become bounded by pronuclear membranes that, in the mouse, are visible ~4.5 h after fertilization and remain for ~10 h until the first mitotic prometaphase.[3–5]

The dynamics of mouse EGA have been inferred from injected reporter gene expression,[18] bromodeoxyuridine (BrdU) labeling,[19] cDNA library construction,[20] microarray analysis,[21–23] and RNA sequencing (RNA-seq)[24,25] to initiate as an event referred to as "minor" EGA in late one-cell embryos, followed by "major" EGA at the two-cell stage. However, these studies have often relied on embryos for which the time of fertilization *in vivo* was indeterminate even though oocytes are fertilizable for more than 12 h after ovulation,[26] the time of coitus and duration of sperm passage and fusion at the fertilization site are unknown,[27] and one-cell embryo morphology and time since fertilization are not reliably correlated.[28] Studies have sometimes used hundreds or thousands of embryos,[20–22,24,29] potentially smoothing signals and obscuring biologically relevant differences.[30]

These challenges preclude the degree of inter-embryo synchrony necessary for accurate transcriptome profiling in one-cell embryos. Moreover, the models they produce do not account for how maternal factor activity required for early development is regulated in the absence of endogenous transcription or address the cue that instigates gene expression, which can

evidently be provided either *in vivo* or in distinct environments *in vitro*. Hints that transcription initiates at the early one-cell stage may also have been restricted by skewed library preparation protocols that potentially reflect mRNA polyadenylation (recruitment) rather than *de novo* gene expression.[21,25,31–33] There are conflicting views about whether gene expression in one-cell embryos produces spliced mRNAs, with evidence of efficient canonical splicing[32,34,35] and the suggestion that it does not occur.[29]

We recently employed polyadenylation-independent single-cell RNA-seq (scRNA-seq) of human embryos to address some of these issues, revealing that gene expression initiates at the one-cell stage.[34] However, the paucity of available healthy one-cell human embryos for research hampers characterization. It precludes precise synchronization and time course profiling and confounds orthogonal validation, including functional corroboration of gene expression. It has therefore not been possible to determine precisely when EGA initiates in human one-cell embryos or whether it initiates stochastically, as a monotonic burst, or as a succession. Indeed, no time courses within the one-cell stage have been reported in any species, and the onset of EGA has been ascribed to the earliest time point at which upregulation has been determined, not necessarily the earliest point at which it occurs. The degree to which the onset of EGA is conserved between species is unknown, and models of EGA are thus incomplete. To address this, we set out to delineate EGA based on multi-platform, single-cell transcriptome time course profiling and characterization of precisely staged mouse one-cell embryos.

RESULTS

Mouse fertilization rapidly triggers an embryonic transcription program

Synchronous one-cell mouse embryos were produced by precisely timed intracytoplasmic sperm injection (ICSI) of mature, metaphase II (mII) oocytes and collected at 2-h intervals for analysis by scRNA-seq. Precise embryo synchronization minimized noise to uncover hitherto inaccessible information about gene expression at the onset of development, and the scRNA-seq protocol avoided poly(A) capture and its attendant potential for library bias.[31,33] We performed scRNA-seq on *Mus musculus domesticus* F2 hybrid (F2) and genetically distinctive, *M. m. domesticus* × *M. m. castaneus* (B6*cast*) embryos that developed efficiently *in vitro* (Figure S1A) and treated F2-B6*cast* scRNA-seq as one dataset to account for strain-specific effects and achieve higher statistical power. This gave an average read depth (±SEM) of 30.4 ± 1.3 million per mII oocyte or one-cell embryo. Unsupervised t-distributed stochastic neighbor embedding (t-SNE) visualization allocated each cell in the F2-B6*cast* time course to its appropriate position, producing clusters at each time point (false discovery rate [FDR] < 0.05; Figure 1A). An analogous time course series of independent F2 embryos subjected to microarray analysis also appropriately allocated each oocyte and embryo to its corresponding time point, corroborating scRNA-seq (FDR < 0.05; Figures 1A and 1B). The F2-B6*cast* scRNA-seq series comprised 4,067 differentially expressed genes (DEGs) across the 12-h time course in one-cell embryos relative to mature mII oocytes (FDR < 0.05). Of the DEGs detected by scRNA-seq, 368 (55.2%) changed in the same direction as the 667 DEGs detected by microarrays (FDR < 0.05 vs. FDR < 0.05), representing a strong Pearson correlation (*r*) of 0.78 (Figures 1C and S1B). Unsupervised gene-by-gene analysis of the F2-B6*cast* scRNA-seq series (FDR < 0.05) assigned embryos to corresponding respective positions on the time course, visualized in the heatmap in Figure 1D.

Of the 4,067 F2-B6*cast* DEGs, 2,290 (56.3%) were downregulated, including canonical maternal transcripts such as *Mos*, *Plat*, *Dppa3* (*Stella*), and *Gdf9*, whose decay following fertilization has been documented[21,33] (Figure S2A). However, there was also a remarkable and systematic program of transcriptional upregulation (1,777 genes, 43.7%) that had initiated as early as within 4 h of fertilization (Figure 1D). Genome activation within 12 h of fertilization (referred to here as immediate EGA, iEGA) was distinctive but overlapped with canonical EGA[24] (Figures S2B–S2D). We assessed upregulation by qPCR, validated by "spike-in" standardization (Figure S1C), which corroborated transcript level changes identified by microarray and scRNA-seq platforms (n = 32 transcripts) (Figures S1D–S1F). These data reveal a program of gene expression that initiates in mouse one-cell embryos within 4 h of fertilization.

iEGA is accompanied by normative transcript processing

De novo transcriptome re-assembly revealed that 88.3% (1,235 of 1,398, FDR < 0.05) of iEGA transcripts possessed annotated 5′ ends, predicting canonical RNA polymerase II (RNA Pol II) transcription start sites. Most upregulated iEGA genes (1,398 of 1,777 [78.7%], FDR < 0.05) were protein coding, leading us to evaluate splicing in light of the suggestion that there are low levels of spliced, translatable mRNA in one-cell mouse embryos.[29] Transcript-level scRNA-seq data analyzed overall, and individually by Sashimi plots, revealed upregulation of mature RNA transcripts spliced at canonical exon junctions during iEGA (Figure S3). Increases in spliced transcript levels were corroborated by intron-flanking qPCR (i.e., qPCR across boundaries between different exons) of multiple targets in independent embryos (Figures 2A, 2B, and S4A). Genes for core spliceosomal components and the mRNA capping enzyme guanylyltransferase (*Rngtt*) were upregulated by 6 h (Figure S4B). Our scRNA-seq protocol avoided poly(A) capture, and random-primed qPCR further suggested that iEGA transcript level increases were not contingent upon polyadenylation (Figures S5A and S5B). Levels of some endogenous proteins apparently increased within 6 h of fertilization, and examples were consistent with classes corresponding to transcript upregulation (e.g., *Atf6* and *Bbs5*) and recruitment (e.g., *Cep44*), although the two classes are not necessarily exclusive (Figures S5B–S5D). We confirmed acute translational competence immediately after fertilization by coinjecting oocytes with sperm plus mCherry-encoding cRNA. Embryos expressed mCherry protein within 4 h of coinjection regardless of whether the injected cRNA had been polyadenylated *in vitro* (Figure 2C). These findings show that, in keeping with previous reports,[32,35] normative splicing predominates during iEGA and is accompanied by competence to translate new transcripts.

CellPress
OPEN ACCESS

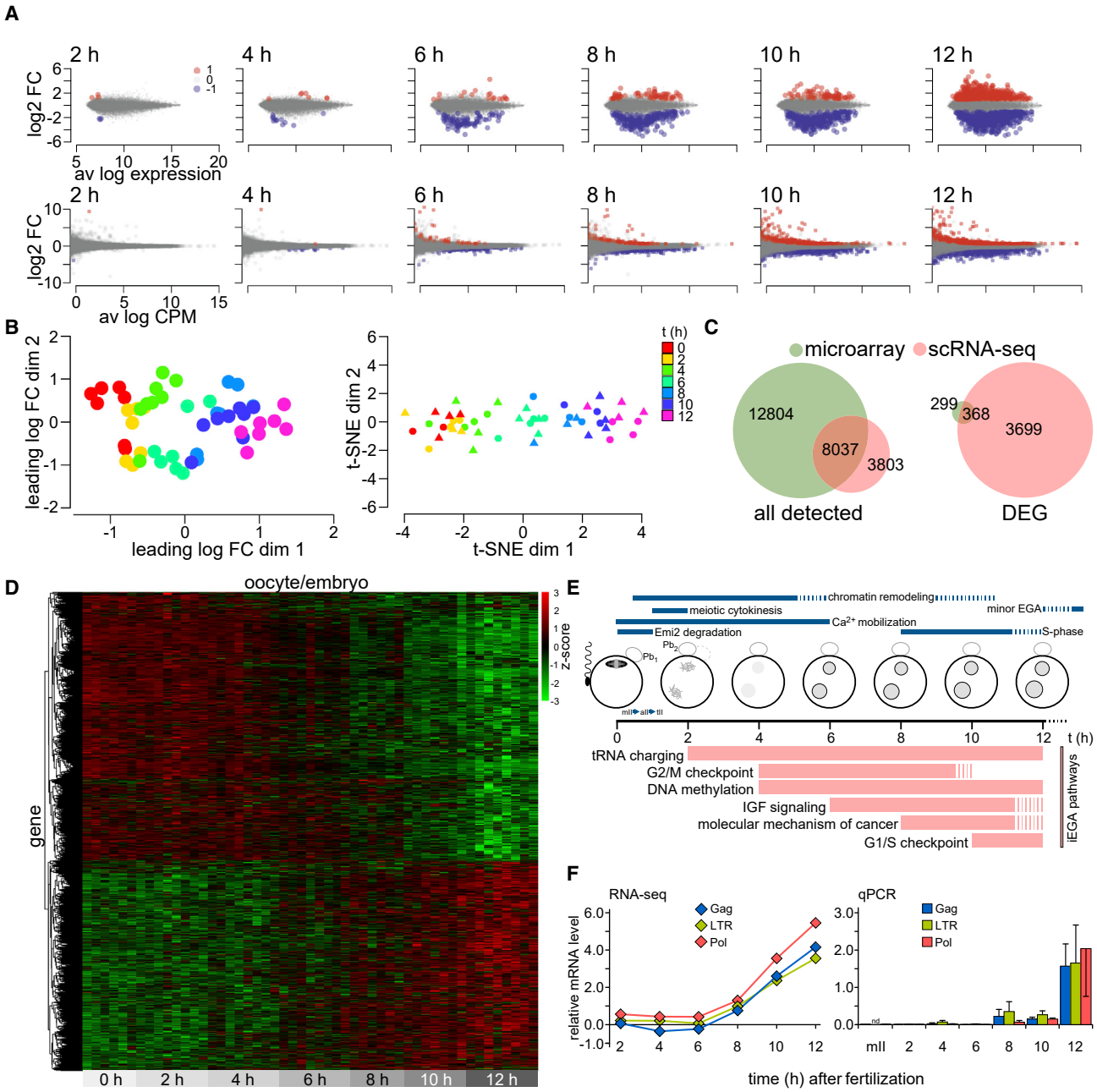


Figure 1. A mouse EGA program initiates rapidly after fertilization

(A) Map plots showing DEGs relative to levels in metaphase II (mII) oocytes from microarray (top) and F2-B6*cast* scRNA-seq series at separate time points.

(B) Multidimensional scaling (MDS) analysis of microarray data (left) and t-SNE visualization of F2 (dots) and B6*cast* (triangles) scRNA-seq data.

(C) Relationships of microarray and scRNA-seq data across the 12-h time course, comparing all detected transcripts (left) and differentially expressed genes (DEGs; FDR < 0.05 vs. FDR < 0.05, microarray vs. scRNA-seq).

(D) Heatmap showing mouse one-cell embryo gene expression changes (FDR < 0.05) from F2-B6*cast* scRNA-seq across the time course. mII oocytes correspond to 0 h. The plot indicates expression of 4,067 genes across $6 \leq n \leq 8$ embryos per time point.

(E) iEGA pathway succession and canonical embryonic events in the first 12 h. Pb_1, first polar body; Pb_2, second polar body; mII, all, and tII, second meiotic prophase, metaphase, anaphase, and telophase respectively.

(F) Relative levels (±SEM) of MuERV transcripts *LTR*, *Pol*, and *Gag* at the times shown after fertilization, from scRNA-seq (FDR $\leq$ 0.05, left) and qPCR of pools of *in vitro* fertilization (IVF) embryos (n = 6 pools/time point).

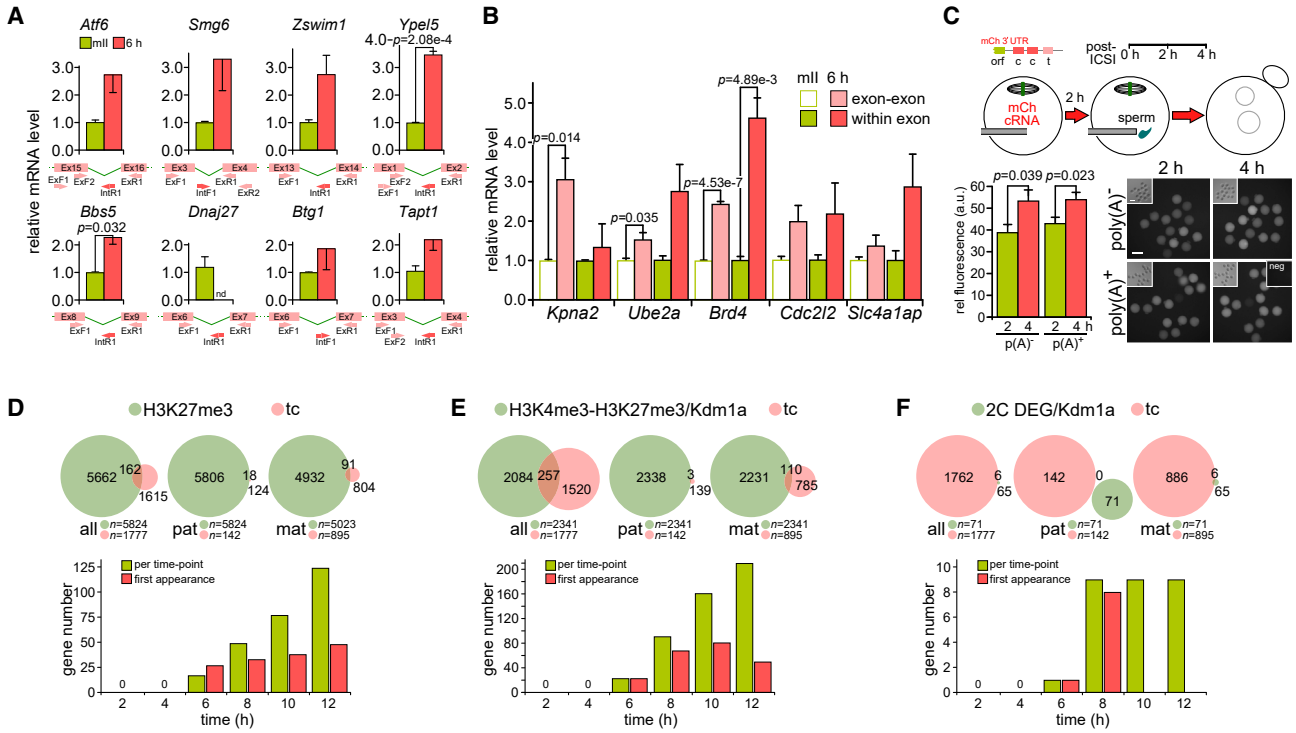


Figure 2. Normative transcript splicing and features of chromatin associated with iEGA

(A) Bar charts indicating transcript levels determined by qPCR using the intron-flanking primers indicated below. cDNA was random primed and derived from pools of 10 mII oocytes (mII) or IVF embryos 6 h after IVF (6 h) (n = 4). Intron-exon primer pairs gave products for genomic DNA but not cDNA. ExF1 and ExF2, exonic forward primers; ExR1 and ExR2, exonic reverse primers; IntF1, intronic forward primers; IntR1, intronic reverse primers; nd, not detected. Values are ± SEM. t test values are shown, where $p < 0.05$.

(B) qPCR analysis as in (A) with intra-exonic primer pairs (within exon) and primers flanking exon junctions (exon-exon) in pools of 10 embryos (n = 6) for up-regulated genes independently identified in previous work.[21]

(C) Injection of mCherry cRNA (mCh; top left: ORF, mCh open reading frame; c, cytoplasmic polyadenylation element; t, mRNA cleavage/polyadenylation signal). Fluorescence intensity quantification (bottom left) at the times shown after injection of mCh cRNA (0.6 ng/μL) polyadenylated *in vitro* (pA^+) or not (pA^-). Fluorescence micrographs show representative oocytes with corresponding bright-field images (insets, top left) and a non-injected control (neg; inset, top right). Scale bars, 100 μm. (A–C) Values are ± SEM. Unpaired t tests show $p < 0.05$.

(D) Venn diagrams (top) of upregulated iEGA genes in F2-B6*cast* (all) and B6*cast* paternal (pat) and maternal (mat) scRNA-seq datasets as they map to promoters marked with H3K27me3 in mII oocytes.[36] Bar charts for F2-B6*cast* scRNA-seq data show cumulative upregulated gene numbers and numbers upregulated for the first time. tc, F2-B6*cast* time course series scRNA-seq data.

(E) As in (D), showing overlaps with H3K4me3 or H3K27me3 promoter occupancy in sperm altered by exposure to transgene expression of Kdm1a.[37]

(F) As in (D) showing overlaps with DEGs in two-cell (2C) embryos following fertilization by transgenic *Kdm1a* sperm.[37]

Pathways predicted by the iEGA expression profile

Ingenuity Pathway Analysis (IPA)[38] of upregulated F2-B6*cast* series genes at each time point predicted a pathway succession that overlapped between microarray and scRNA-seq data and reflected features of one-cell embryonic development (Figures 1E, S2E, and S2F).[3] The pattern of embryonic developmental terms disappeared when IPA was performed on downregulated genes. Terms for upregulated genes included cell cycle regulation, exemplified by the meiotic-to-mitotic cell cycle transition and the first mitotic cytokinesis (e.g., cell cycle checkpoint terms Chek1, Cks1b, Ccne2, Cdc25a, Cul1, and E2F4), metabolism (e.g., the insulin growth factor [IGF] signaling terms Csnk2a1, Igfbp4, Jak1, Pdpk1, Pik3r1, and Pik3r3), and DNA methylation (e.g., Hist1h4a, Hist1h4f, Hist2h4a, Hist4h4, Rbbp4, and Sap30); the paternal genome (and, to a lesser extent, the maternal one) undergoes replication-independent ("active") DNA demethylation soon after fertilization.[9,39,40] Murine endogenous retrovirus (MuERV) *LTR*, *Pol*, and *Gag* transcripts[41] were upregulated at hundreds of loci from 8 h (Figure 1F), showing that MuERV gene activation is established in one-cell embryos as a feature of iEGA. The putative EGA regulator gene *Dux*[42] was upregulated by 6 h but Dux-responsive genes were not (Figure S1G), suggesting that iEGA is independent of Dux or its paralogs.[43] These results reflect coordination between iEGA and developmental processes immediately after fertilization.

Embryonic chromatin during iEGA

We next asked whether regulatory chromatin marks occupied iEGA genes. Only one iEGA gene contained a DNA methylation imprint (*Flt1*; FDR < 0.05; Table S1), and two (*Elavl2* and *Nav3*) were from recently described transient H3K27me3-associated

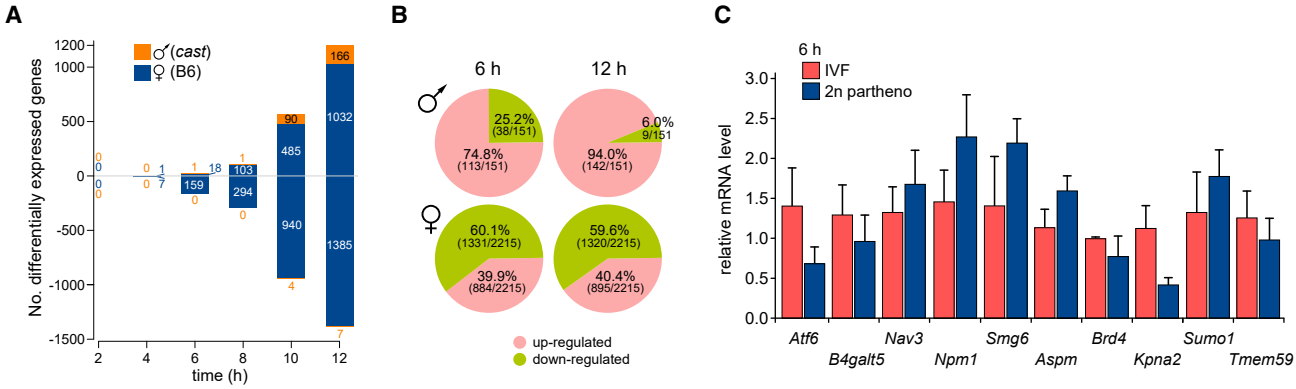


Figure 3. Parental genome contributions to gene expression during iEGA

(A) Bar charts indicating DEG numbers determined by scRNA-seq compared with mII oocytes (0 h) for maternal (B6) and paternal (*cast*) alleles in the B6*cast* series. Values reflect FDR < 0.05 for each time point separately.

(B) Proportions of DEGs determined by scRNA-seq allocated to each parental genome at the times indicated (6 or 12 h) relative to 2 h after fertilization. Values were calculated using FDR < 0.05 for the entire dataset and reveal an increase in paternal genome activity by 12 h.

(C) Bar charts indicating relative transcript levels (±SEM) determined by qPCR using random-primed cDNA from embryos produced by IVF or diploid (2n) parthenogenotes (n = 3 pools of 10 each) 6 h after activation of development. Pairwise t test p values for each target were p = 0.069 for *Kpna2* and p > 0.220 for all others.

mouse blastocyst imprints.[44] Some iEGA promoters contained H3K27me3 at the same promoter as in sperm (162, FDR < 0.05) or mII oocytes (91, FDR < 0.05)[36] or the active mark H3K4me2 in sperm[37] (p < 0.05; Figures 2D–2F; Table S1). Further data mining revealed 17,168 promoter regions associated with the active mark H3K4me3 in the sperm genome (FDR < 0.05),[45] of which 1,116 corresponded to iEGA genes (62.8% of the iEGA total, n = 1,777, FDR < 0.05). Co-occupancy by H3K4me3 and H3K27me3 constitutes a bivalent mark thought to poise promoters for transcription in mouse embryonic stem cells (ESCs),[46] but only 95 iEGA promoter regions (5.3%) were represented by the set of promoters co-occupied by H3K4me3 and H3K27me3 in mouse primordial germ cells (n = 3,498, FDR < 0.05).[47] This indicates that most iEGA occurs independent of bivalently marked gamete promoters.

EGA in non-mammalian vertebrates involves chromatin remodeling,[2,48] and because meiotic exit is a canonical feature of vertebrate fertilization, we evaluated dynamic changes to histone modifications in relation to cell cycle progression in the mouse. This was achieved by injecting developmentally competent, heated sperm heads that do not induce meiotic exit[9,49] into mII oocytes (Figure S6A) and comparing histone modifications (n = 33) in the resulting biparental mII oocytes (containing oocyte- and sperm-derived genomes) with those in age-matched one-cell embryos. In mII oocytes containing heated sperm heads, maternal chromatin shared only 54.5% of histone modifications with paternal chromatin, but in control embryos that had undergone meiotic exit, 78.8% of parental histone modifications were shared (p = 0.006; Figures S6B–S6D). The relative epigenetic uniformity ("epigenetic" here refers to chromatin composition, including DNA and histone modifications, that might affect gene expression) attained between parental genomes following meiotic exit reflected an increase in modifications associated with active transcription (e.g., the appearance of acetylated H3 lysine 27, H3K27ac; Figures S6B and S6C) from 6 to 14 active marks in paternal chromatin and from 14 to 16 in maternal chromatin (Figure S6D). These findings suggest that a transcriptionally favorable chromatin landscape in one-cell embryos is rapidly promoted by meiotic exit in response to fertilization.

Distribution and regulation of iEGA genes

Upregulated iEGA genes were distributed throughout the genome (Figure S7A). To discriminate between contributions to iEGA from parental alleles, we focused on B6*cast* series DEGs to infer relative parental genomic contributions. Unsupervised cluster analysis placed B6*cast* embryos at their corresponding time course positions (Figure S7B) and permitted parent-specific expression assignment based on parental alleles with informative expressed SNPs. Changes reflected maternal transcript downregulation (in 2,785 of 4,695 [59.3%] of cases), including established precedents, *Mos*, *Zp2* and *Tle6*, that had initiated by 6 h and was followed by a progressive decline throughout iEGA (Figures 3A and 3B). Maternal allelic gene expression had initiated 2 h after onset of maternal transcript degradation (by 8 h after fertilization) and included *Zbed3*, *Zfp57*, *Spin1*, *Cnot6l*, and *Obox5*. This was supported by qPCR of diploid one-cell parthenogenotes (i.e., embryos lacking a paternal genome), in which expression of all (n = 10) assessed upregulated iEGA genes at 6 h was similar (p > 0.2) to that of controls (Figure 3C). Paternal iEGA lagged; its initiation was not detected until ~10 h but was clearly evident by 12 h (Figures 3A and 3B). The active mark H3K4me3 was strong in maternal chromatin at this time relative to that in paternal chromatin (Figures S6B–S6D). These findings suggest that initial gene expression in iEGA is predominantly driven by the maternal genome. We also investigated whether sperm-borne mRNA was introduced during sperm internalization and removal of the major sperm nucleoprotein protamine. In agreement with qualitative analysis,[49] we found that 50% of transgene-encoded protamine was removed within ~30 min of sperm entry (Figures S7C–S7F). We therefore produced an

independent B6*cast* scRNA-seq time course series (B6*cast*-HiRes) with high temporal resolution starting at 20 min. However, the B6*cast*-HiRes series detected zero upregulated DEGs (from 13,870 genes whose expression was detected) and no significant levels of sperm-borne mRNA,[50] even at a relatively relaxed stringency (FDR < 0.1). With the caveat that B6*cast* and B6*cast*-HiRes analyses were limited to informative reads following ICSI, these findings argue against delivery into oocytes of a stable submembrane cargo of persistent sperm-borne mRNA[51] and indicate that iEGA initiates from the maternal genome.

Analysis by qPCR of embryos treated with the RNA Pol II inhibitor α-amanitin were indicative of refractory and sensitive transcript levels (Figure S4C). "Minor" EGA has been shown previously to be refractory to α-amanitin,[21,23] and consistent with this, scRNA-seq of control and α-amanitin-treated embryos (Figure S4D) yielded statistically indistinguishable (FDR > 0.05) upregulated transcript profiles (Figure S4E). However, the heatmap in Figure S4E visually indicated a reduction in the α-amanitin treatment group, implying some α-amanitin sensitivity, consistent with qPCR. There was no evidence of splice variants corresponding to the α-amanitin target Rpb1 that might, in principle, mediate resistance (Figure S4F). Levels of some refractory transcripts (e.g., *Bbs5*), but not all (e.g., *Klf4*), were also reduced by the RNA Pol II elongation inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), which precipitated early developmental arrest (Figures S4G and S4H).

Mouse iEGA and the onset of EGA in human one-cell embryos

We recently reported gene expression in healthy human bipronuclear (2PN) one-cell embryos.[34] Of 874 upregulated human DEGs, 111 (12.7%) overlapped with mouse F2-B6*cast* iEGA series species orthologs (FDR < 0.05) (Figure 4A), supported by mouse embryo qPCR (n = 11 genes; Figure 4B). The F2-B6*cast* scRNA-seq time course revealed that recombination pathways are expressed during iEGA, leading us to focus on them in human one-cell embryos because their disruption in the mouse is associated with aneuploidy[52] and compromises prime editing.[53] Although heritable genome editing has been assessed in tripronuclear (3PN) human one-cell embryos,[54] single-cell qPCR showed that the normally upregulated DNA repair genes *MLH1*, *MSH4*, and *MSH6*[55] were dysregulated in 3PN and monopronuclear (1PN) human embryos (Figures 4C and 4D). Thus, DNA maintenance genes are expressed during iEGA, but DNA repair pathway expression is disrupted in human 1PN and 3PN one-cell embryos even though they often exhibit normal ploidy.[55,56]

Of 543 iEGA genes (FDR < 0.05, log2CPM ≥ 1 [counts per million, CPM], with at least one log2 fold change [log2FC] ≥ 0.5 across the 2- to 12 h time course), 302 overlapped with previous mouse data for mII oocytes and one-, two-, and four-cell embryos (Figure 4E).[24] In most cases (186 of 302; 61.6%), expression was sustained at the one-cell stage but markedly declined by the two-cell stage (clusters II and III; Figure 4E). Thus, part of the iEGA program is downregulated at the time of "major" EGA in mouse two-cell embryos,[21] a situation that mirrors the one in human embryos when allowance is made for the later occurrence of "major" EGA, around the eight-cell stage.[34] Expression of the remaining 302 upregulated iEGA genes (cluster I) remained elevated in two- and four-cell embryos (Figure 4E).[24]

Pathways predicted by mouse iEGA overlapped with terms predicted by upregulated DEGs of human one-cell embryos; the human dataset most closely corresponded to an early (4-h) time point in the mouse time course (r = 0.675; Figure 4F). The human embryos were heterogeneous and asynchronous and the precise time of fertilization was unknown (and any of these aspects might conceal human-mouse similarities); nevertheless, our findings are consistent with a conserved pattern of genome activation in mouse and human one-cell embryos, with slower genome activation in humans relative to that in mice.

Shared regulatory transcription factors (TFs) in iEGA and cancer

IPA of iEGA predicted cancer terms (Figure 1E) and upstream regulator effects (FDR < 0.05) mediated by c-Myc and additional cancer-associated TFs[58] (Figures 5A–5C, S8, and S9). Transcriptional networks regulated by MYC and MYCN were also predicted for human one-cell embryos,[34] leading us to investigate the functionality of cancer-associated TF pathways during mouse iEGA. We selected the TFs c-Myc, Mycn, Erg, and Atf4 because they are predicted to function within 8 h of fertilization and have corresponding inhibitors of high specificity, and each plays major and well-studied roles in cancer.[59–62] The four TFs were present in the cytoplasm of mII oocytes and in the cytoplasm and nuclei of one-cell embryos (Figures S9A–S9D). We did not detect their transcripts in either mII oocytes or one-cell embryos. c-Myc localized to spindles in mII oocytes and was present in immature oocytes[63,64] (Figures 5D and 5E). c-Myc and its canonical heterodimeric partner Max were present in one-, two-, and four-cell embryos (Figures 5E and S9E–S9G). Genes for SUMOylation and deSUMOylation, which modulate c-Myc activity,[64] were upregulated during iEGA (Figure S9H). We found evidence of the cytoplasmic c-Myc cleavage product Myc-nick[65] in one-cell embryos, and other isoforms were also readily detectable in cleavage-stage embryos (Figures 5E and S9I). These findings are consistent with roles of maternal c-Myc, Mycn, Erg, and Atf4 as iEGA regulators.

Abrogation of c-Myc acutely disrupts development and iEGA

Established gene targeting or knockdown (e.g., RNAi- or morpholino-based) strategies are not available to evaluate the function of TFs such as c-Myc, which pre-exist in mII oocytes as maternally inherited factors that might play key roles in oogenesis and maturation (Mos provides a precedent[6]). Alternative approaches, such as heterobifunctional chemical degraders (PROTACs) and antibody-dependent depletion (e.g., "Trim-Away"), assume efficient targeting and require additional manipulation that may impact iEGA. We thus evaluated c-Myc with structurally distinct small molecule inhibitors, 10,058-F4 (F4) and MYCi975 (i975), which specifically bind different regions within the c-Myc helix-loop-helix (HLH) domain to prevent c-Myc-Max heterodimerization and block gene-regulatory activity.[66,67] Treatment with either inhibitor induced developmental arrest at one- or two-cell stages (Figures 6A–6D), and i975 resulted in loss of c-Myc immunoreactivity, indicative of degradation (Figure S9J). We determined the transcriptomic impact of c-Myc inhibition by performing scRNA-seq of one-cell embryos (n = 8) 10 h after F4 treatment. This identified 1,500 DEGs (F4 DEGs, FDR < 0.05), of which 577 (38.5%)

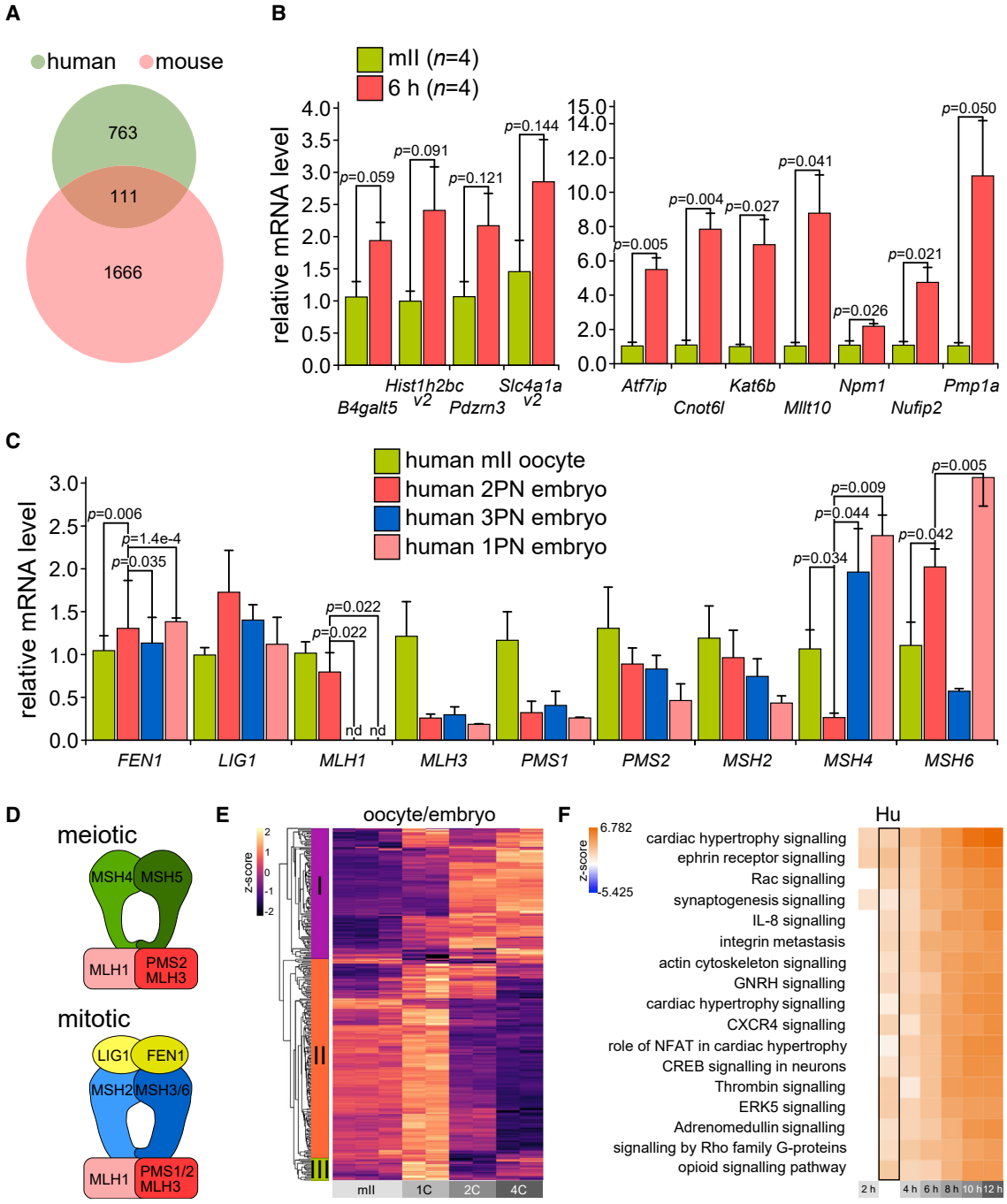


Figure 4. Initiation of human embryonic transcription: Conservation with iEGA and dysregulation of DNA repair gene expression

(A) Venn diagram showing upregulated gene overlap between mouse (full time course FDR < 0.05, slope > 0; n = 1,777) and human (FDR < 0.05, logFC > 0; n = 874) one-cell embryos (FDR < 0.1).[34]

(B) Bar charts for transcript levels (qPCR) of mouse orthologs of transcripts upregulated in human one-cell embryos.[34] Values (±SEM) are normalized against mII oocytes (~1.0). p values (unpaired t test) are indicated for pairwise comparisons.

(C) Single-cell qPCR for transcripts encoding DNA repair and recombination pathway components in human monopronuclear (1PN) and tripronuclear (3PN) one-cell embryos and healthy bipronuclear (2PN) and mII oocyte (mII) controls. Values (±SEM) are normalized against mII oocytes (~1.0). Unpaired t tests indicate $p < 0.05$.

(D) Schematic showing interactions between human mismatch repair proteins in mitotic and meiotic cells[57] whose transcripts were evaluated in (C).

(E) Heatmap indicating expression trajectories of 302 of 543 iEGA genes from mouse one-cell embryos (FDR < 0.05, log2CPM ≥ 1 and at least one log2FC ≥ 0.5 across the 2- to 12-h time course) relative to expression in mII oocytes or one-cell, 2C, and four-cell embryos.[24] Genes were sorted into three clusters (I, II, and III) as indicated.[34]

(F) Ingenuity Pathway Analysis (IPA) of upregulated genes in human (Hu) and mouse one-cell embryos.

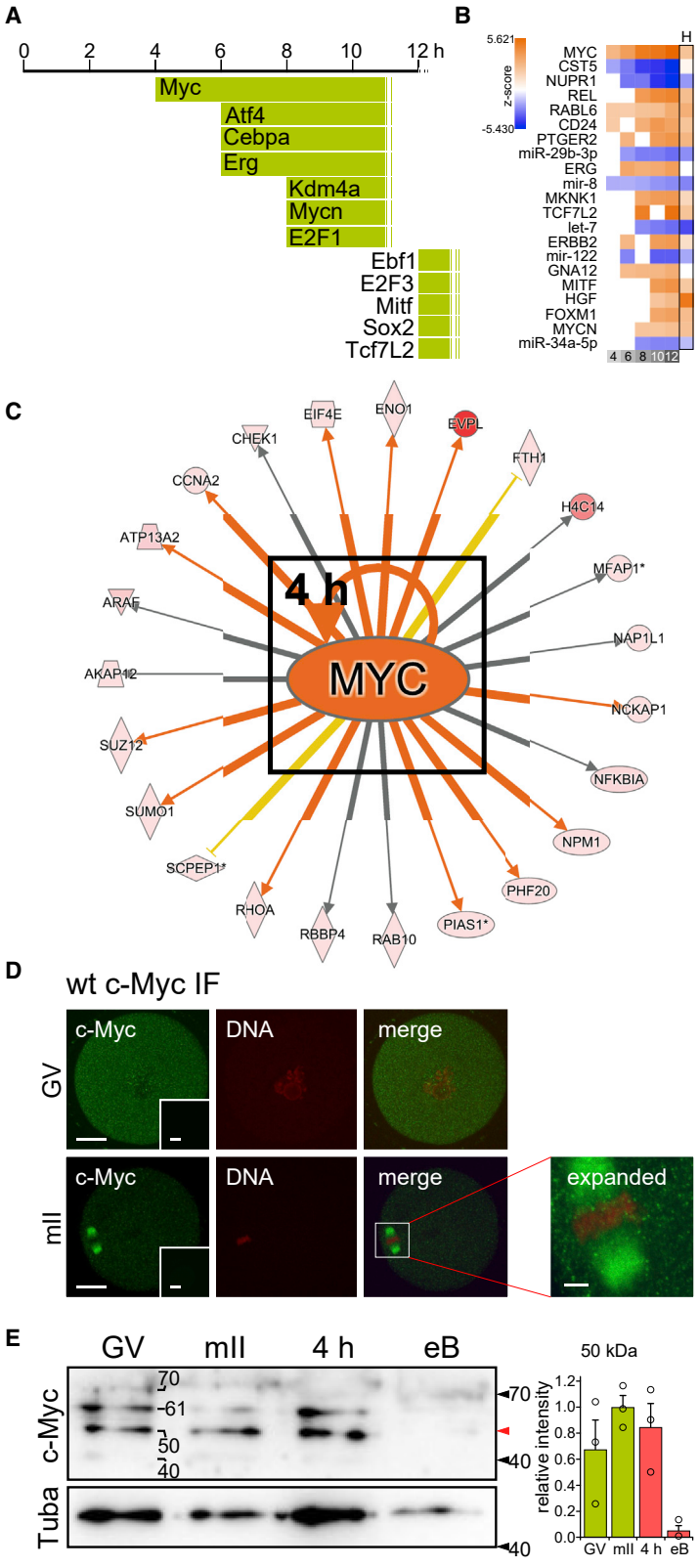


Figure 5. c-Myc in mouse oocytes and one-cell embryos

(A) Upstream transcription regulators inferred by IPA of mouse iEGA.

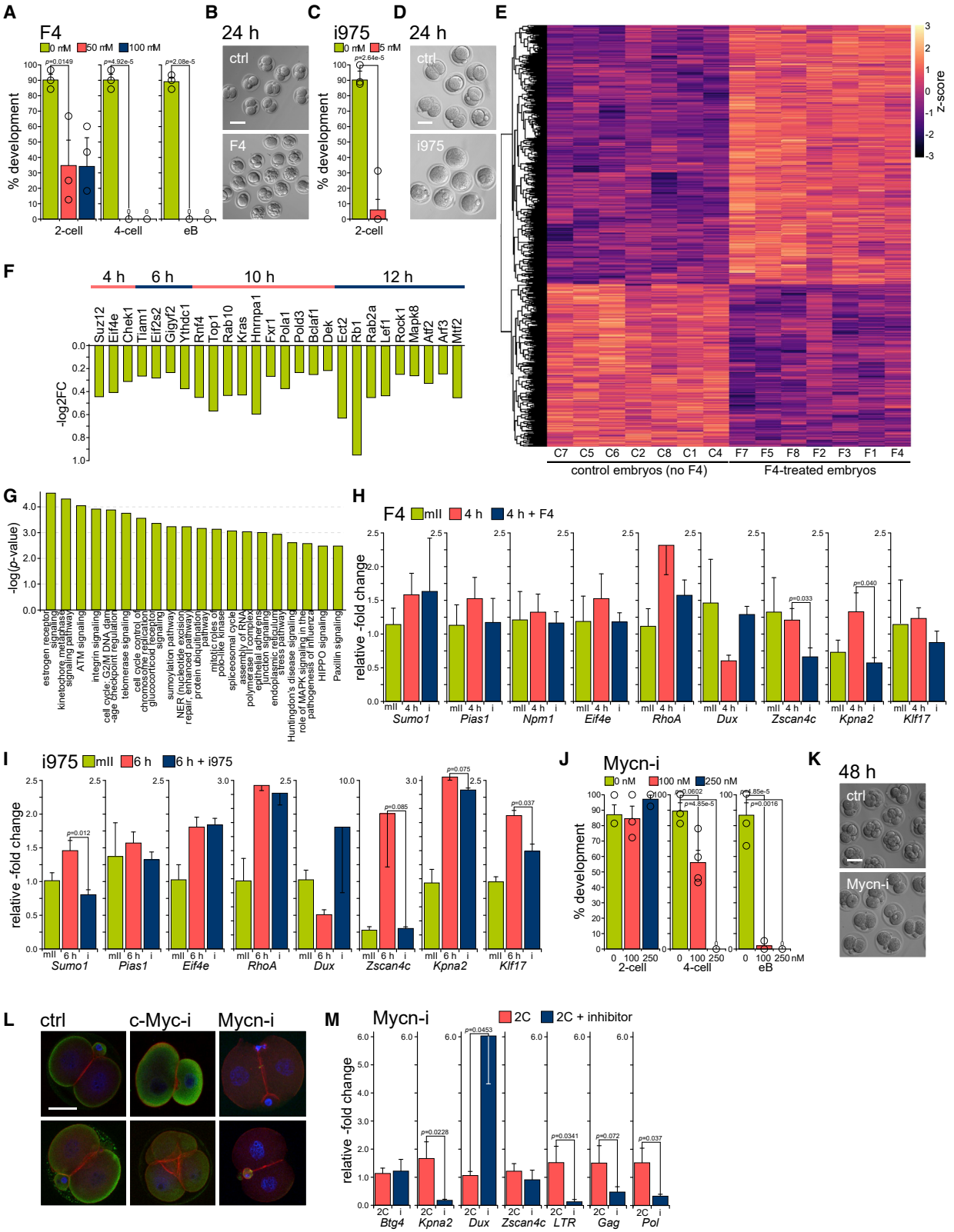
(B) Transcription regulators indicated for Hu and mouse one-cell embryos by IPA.

(C) The upstream network, with c-Myc at its center, predicted by IPA of upregulated F2-B6*cast* genes at 4 h.

(D) Representative immunofluorescence micrographs of a germinal vesicle (GV) oocyte (n = 6) and mII oocyte (n = 7) stained with anti-c-Myc antibodies (green) and propidium iodide (red, for DNA). Each row shows the same oocyte, with an expanded image showing the mII spindle (expanded). Insets show negative controls (primary antibody omitted) at the same stage. Scale bars, 20 μm and 2 μm (expanded).

(E) Representative (n = 3) immunoblot of GV oocytes (100/lane), mII oocytes (100/lane), one-cell embryos after 4 h (100/lane) and embryonic day 4.5 blastocysts (eB; 30/lane). Antibodies recognized c-Myc (Myc) or tubulin-α (Tuba) loading/transfer controls. Indicated sizes are in kilodaltons. A red arrowhead marks the position of canonical c-Myc (50 kDa). Quantification of canonical 50 kDa c-Myc in each lane relative to Tuba controls normalized against mII oocytes (set at 1.0) is plotted on the right.

CellPress
OPEN ACCESS



(legend on next page)

were expressed at reduced levels compared with untreated controls (Figure 6E), including most (229 of 240, 95.4%) of the F4 DEGs that overlapped with iEGA genes. The list contained known c-Myc targets (Figure 6F) and, when subjected to pathway analysis, predicted multiple relevant terms ($p < 0.01$), including G2/M DNA damage regulation, cell cycle control of chromosome replication, nucleotide excision repair, and the spliceosomal cycle (Figure 6G). These findings imply a role of c-Myc in promoting iEGA. c-Myc is also a repressor of transcription, including cell cycle arrest genes, consistent with upregulation of 923 F4 DEGs, including cell cycle-regulatory transforming growth factor β (TGF-β) signaling pathways. qPCR corroborated disruption of iEGA gene upregulation by F4 and i975 (Figures 6H and 6I). Analogous inhibition of the related basic HLH (bHLH) domain TF Mycn also blocked early development and disrupted iEGA gene upregulation (Figures 6J–6M), including MuERV gene expression ($p \leq 0.07$; Figure 6M). Treatment disrupted embryo morphology and impeded cytokinesis (Figure 6L). Abrogating the functions of additional cancer-associated TFs predicted to have iEGA-regulatory roles, Erg or Atf4, also impeded preimplantation development (Figures S9K and S9L). Taken together, these findings support the view that iEGA is predictive of TFs that initiate embryonic transcription and that iEGA shares features with gene-regulatory networks in cancer.

DISCUSSION

This work reveals a program of mouse embryonic gene expression that initiates within 4 h of fertilization. The trigger is likely fertilization itself; candidate iEGA-regulatory TFs such as c-Myc and Mycn are regulated by kinases, and fertilization triggers a phosphorylation cascade that might, in principle, activate them.[6,7,68–72] Moreover, each of the four tested TFs (c-Myc, Mycn, Erg, and Atf4) was present in oocytes and one-cell embryos (Figures 5D, 5E, and S9A–S9E). Selectively inhibiting c-Myc during iEGA led to acute developmental arrest and induced activation failure in ~95% of upregulated genes. c-Myc is a master regulator of gene transcription,[73] binding to many active genes as a c-Myc-Max heterodimer[59] and facilitating RNA Pol II promoter-proximal pause release.[74,75] It is also a transcriptional repressor,[59,73] and most F4 DEGs (61.5%) were upregulated following inhibition, including cell cycle arrest genes, suggesting that c-Myc represses some gene expression during iEGA. As well as developmental roles for c-Myc and Mycn, interfering with embryonic Erg and Atf4 also impaired early development (Figures S9K and S9L); ERG alters enhancer chromatin landscapes to promote cell proliferation,[76] while ATF4 modulates metabolic control to redirect nutrients into anabolic pathways in tumors, a role overlapping with that of c-Myc.[62,77] These findings imply major physiological gene-regulatory roles for c-Myc in one-cell embryos, and their collective force implicitly associates metabolic, cancer, and early embryonic pathways.

Onset of transcription immediately after gamete union has implications for models of epigenetic inheritance and reprogramming by providing a framework in which epigenetic marks in gametes can be rapidly decoded for transcription in early embryos. This may reflect maternal pioneer factor activity[12] or gametic bookmarking, in which epigenetic memory (e.g., histone modification or TF occupancy) persists into mitotic M phase.[78] Bookmarking TFs predicted by iEGA are exemplified by Brd4 and RNA Pol II general TFs, including TATA-binding protein and TFIID. In addition, c-Myc acts as a pioneer factor[79] and targets promoters during reprogramming of somatic cells to pluripotency.[80] We found little evidence of regulation of iEGA by canonical imprints or the bivalent mark H3K4me3/H3K27me3, although there was correspondence to the active mark H3K4me3 (Figures 2D–2F). It remains to be seen whether epigenetic marks from gametes are functional at the one-cell stage (i.e., whether they impact one-cell-stage gene expression), and if so, whether their regulation differs from that of later-acting epigenetic marks in preimplantation embryos.[44]

One interpretation of our data is that mouse iEGA is predominantly driven by the maternal genome (Figures 3A and 3B), which possessed a transcriptionally favorable global chromatin

Figure 6. Functional analysis of predicted iEGA regulators in early mouse embryos

(A) Developmental rates (percentage of one-cell embryos surviving IVF) following IVF (2 h) with continued culture in the presence or absence of 50 μM c-Myc inhibitor 10,058-F4 (F4) for 3 independent experiments ($10 \leq n \leq 30$ embryos each). 2-cell, 2C embryo (24 h after IVF); 4-cell, four-cell embryo (~48 h); eB, expanded blastocyst (~96 h).

(B) Hoffman micrographs after 24 h of embryos produced by IVF and incubated without (control, ctrl) or continuously with F4 (50 μM). Scale bar, 100 μm.

(C) Developmental rates per (A), except in the presence or absence of 5 μM c-Myc inhibitor MYCi975 (i975), for 3 independent experiments ($9 \leq n \leq 16$ embryos each).

(D) Hoffman micrographs after 24 h of embryos of (C) after incubation without (ctrl) or continuously with 5 μM i975. Scale bar, 50 μm.

(E) Heatmap showing mouse gene expression changes (FDR < 0.05) determined by scRNA-seq for mouse F2 one-cell embryos not treated (n = 7) or treated (n = 7) with F4 for 10 h after fertilization.

(F) IPA of scRNA-seq data to reveal c-Myc-regulated iEGA target genes ($p < 0.05$) that were not upregulated when embryos were treated with F4.

(G) Top pathway terms in IPA of genes inhibited by F4 as determined by scRNA-seq ($p < 0.01$).

(H) Ratiometric qPCR analysis of iEGA transcripts in mII oocytes or embryos following F4 treatment for 2 h per (A).

(I) Ratiometric qPCR analysis of iEGA transcripts in mII oocytes or embryos following i975 treatment for 4 h per (C).

(H and I) Transcript levels were normalized against mII oocytes (set at ~1.0) and values plotted ± SEM, indicating differences where $p < 0.10$ (unpaired t test).

(J) Developmental rates as in (A), except following embryo treatment with the Mycn inhibitor (Mycn-i) CD532 (n = 3 independent experiments; $10 \leq n \leq 48$ embryos each).

(K) Hoffman micrographs of embryos produced by IVF and incubated without (ctrl) or with CD532 (250 nM, Mycn-i) after 48 h. Scale bar, 100 μm.

(L) Representative immunofluorescence micrographs of 2C-stage embryos, 24 h after IVF, stained with propidium iodide (DNA, blue), phalloidin (F-actin, red), or for Tuba (green). Embryos were incubated with F4 (c-Myc-i), CD532 (Mycn-i), or without inhibitor (ctrl). Scale bars, 50 μm.

(M) Ratiometric qPCR analysis of iEGA transcripts in 2C embryos produced by IVF cultured in the absence (2C) or continuous presence (2C + inhibitor) of the Mycn inhibitor CD532. Values are plotted ± SEM and indicate differences where $p < 0.10$ (unpaired t test).

landscape (Figure S6D). This was in agreement with a correlation between H3K4me3 and the presence of maternal allele-specific RNA Pol II prior to "major" EGA.[81] It is also consistent with establishment of totipotency in mitotic parthenogenetic embryos that efficiently accommodate the paternal genome to support full-term development[5] and full-term development of parthenogenetic embryos, albeit at low efficiency, following epigenetic modification.[82] Maternal genome predominance at the onset of embryonic transcription is possibly an evolutionary corollary of the potential to develop parthenogenetically in non-mammalian vertebrates,[83] prevented in mammals by imprinting.[44]

Most (88.3%) iEGA transcripts predicted canonical RNA Pol II transcription start sites, and indeed, RNA Pol II chromatin binding exhibits widespread enrichment prior to "major" EGA.[74,81] However, early embryonic transcription exhibits anomalous features. Although the large catalytic subunit Rpb1 carboxy-terminal domain is canonically phosphorylated in active RNA Pol II, it is hypophosphorylated during "minor" EGA[84] and at genes activated after nuclear transfer.[85] Our analysis by qPCR and scRNA-seq of embryos treated with the RNA Pol II inhibitor α-amanitin (Figures S4C–S4E) indicated that gene expression was largely refractory; transcript levels in treated and untreated embryos were statistically indistinguishable (FDR > 0.05). This agrees with previous reports of α-amanitin-refractory transcription,[21,23] but although not statistically significant, transcript levels were apparently reduced by treatment (Figure S4E), suggesting that the effect is nuanced and non-binary. We speculate that steric impedance prevents α-amanitin from accessing Rpb1 in a manner that varies in one-cell embryos. Understanding the nature of this variation will benefit from profiling chromatin at iEGA gene-regulatory regions.

Multiple strands of evidence demonstrate that normative splicing occurred throughout iEGA (Figure S3A), contrasting with the suggestion that transcription is promiscuous and uncoupled from splicing.[29] Upregulated spliced transcripts were detected by scRNA-seq (Figure S3) or intron-flanking qPCR after 6 or 12 h (Figures 2A, 2B, and S4A). Human one-cell embryos[34] provide orthogonal evidence of splicing, corroborated by expression trajectory analysis in the mouse (Figure 4E) and other reports of upregulated spliced transcripts.[32,35] Pathway analysis was predictive of embryonic processes (Figures 1E and S2E), early embryonic translational competence (Figure 2C), and upregulation of protein-coding mRNAs (Figure S5C). These data argue that canonical transcript splicing occurs efficiently in mammalian one-cell embryos.

Gene regulation that drives or sustains totipotency in one-cell embryos[1] and pluripotency in ESCs possesses distinct yet overlapping features. For example, 1%–5% of pluripotent ESCs exhibit characteristics of two-cell embryos, including MuERV gene expression,[86] and MuERV genes are activated by Dux[42] and expressed during "major" EGA at mid- to late two-cell stages.[21] Our work demonstrates that MuERV *Gag*, *Pol*, and *LTR* expression initiates as part of the iEGA program within 8 h of fertilization and ~10 h prior to cell division[5] (Figures 1E and 1F). However, Dux is dispensable in early embryos,[43] and we show that *Dux* upregulation in iEGA is preceded by MuERV gene expression and not accompanied by upregulation of other Dux-responsive genes (Figure S1G). Although inhibiting c-Myc or Mycn led to increased Dux expression, it reduced MuERV gene expression (Figures 6I and 6M); c-Myc and Mycn may thus activate MuERV expression while repressing "major" EGA pathways during iEGA. Gene expression in iEGA is indicative of a succession of iEGA-regulatory TFs that segues to "minor" and then "major" EGA; this is supported by stepwise RNA Pol II loading of promoters in one- and two-cell mouse embryos[81] in concert with activation of other regulatory components and mechanisms (Figures 1E and 5A).

The mouse iEGA profile overlapped with gene expression in human one-cell embryos.[34] Differences may reflect the slower kinetics of human development,[87] and predicted pathways for human one-cell embryos were most conserved in early (4 h) mouse embryos (Figure 4F). Indeed, mouse and human lineage specification TFs are shared but differ in their respective timing, roles, and localization.[25] As in human one-cell embryos, the amplitude of upregulation in iEGA was small (mean log2FC = 0.77 ± 0.03, FDR < 0.05). This is consistent induction by MYC of target genes less than 2-fold,[59] modest expression increases of genes specific to ESCs during pluripotency induction,[80] and an average negative effect of the key transcription regulator, Polycomb, of approximately 2-fold.[88] The present work coheres with the idea that small-magnitude gene expression changes are a cardinal feature of cellular potency transitions.

Genes expressed in human one-cell embryos were precipitously deactivated between four- and eight-cell stages, commensurate with "major" EGA.[34] We here show that this gene expression pattern is conserved in the mouse; many iEGA genes were similarly deactivated prior to "major" EGA, which, in the mouse, occurs at the two-cell stage[21,24,25] (Figure 4E). One inference from this is that mammalian embryonic transcription is initiated by maternal factors immediately after fertilization but subsequently yields to a program of higher amplitude expression driven by distinctive TFs, leading to "major" EGA.

In sum, this work reports a programmed succession of gene expression triggered following fertilization in mammals: iEGA. iEGA employs canonical promoters and splicing. Predicted iEGA pathways and regulators overlap with cancer-associated terms and are critical determinants in early development. This model of transcriptional instatement in embryos opens doors to understanding the mechanistic basis by which cellular potency changes are effected, in particular the emergence of embryonic totipotency, with the promise of revealing targets for cancer therapy.

Limitations of the study

This study predicts that maternal factors instigate iEGA, but they are difficult to interfere with experimentally; the factors may play earlier roles in oocyte development and/or cannot be acutely removed after fertilization. We addressed this limitation using inhibitors, but it is possible that they were not completely specific to their intended targets. This limitation also applies to PROTAC- and antibody-based strategies, which also require additional, potentially disruptive manipulation. We mitigated this orthogonally via structurally distinct specific inhibitors; future studies may employ a broader inhibitor panel or use genomic knockins of the sensitive (non-toxic), reversible, auxin-inducible degron AID2 to generate rapidly degradable fusions.[89] Our study employed different mouse strains, but some are related, and it is possible that they behave idiosyncratically with respect to

iEGA. Future work may also employ reciprocal crosses to confirm parent-of-origin allelic expression. To recapitulate acrosome-reacted sperm-membrane depletion, we treated sperm with detergent, but treatment may have been too harsh and depleted some sperm-borne transcripts that would otherwise be delivered into the oocyte. It should also be noted that our scRNA-seq protocol mostly eliminated short RNAs. Finally, refinements that are robustly able to characterize upregulated iEGA promoter chromatin at high resolution in single cells would be highly desirable.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Human oocyte and embryo sample collection
- METHOD DETAILS
  - Human metaphase II oocytes and one-cell embryos (zygotes)
  - Animals
  - Collection and culture of oocytes
  - Sperm preparation and microinjection (ICSI)
  - Fertilization *in vitro* (IVF)
  - Embryo culture
  - Direct fluorescence imaging and image analysis
  - Immunocytochemistry
  - Immunoblotting
  - Preparation and injection of cRNA
  - Ratiometric RT-PCR (qPCR)
  - Single-cell RNA-sequencing (scRNA-seq)
  - RNA-sequencing for F4 and α-amanitin experiments
  - Microarrays
  - Sequencing alignment and gene and transcript counts
  - Microarray data analysis
  - Differential expression analysis
  - iEGA gene trajectory analysis
  - Bioinformatics and differential gene expression analysis for F4 and α-amanitin experiments
  - Sperm H3K4me3 ChIP-seq analysis
  - Pathway analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistics and reproducibility

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112023.

ACKNOWLEDGMENTS

We thank the University of Bath Animal Facility support staff and C. Tickle and M. Leeb for comments during manuscript preparation. We acknowledge support to A.C.F.P. from the UK Medical Research Council (G1000839, MR/N000080/1, MR/N020294/1, and MR/W024845/1) and Biotechnology and Biological Sciences Research Council (BB/P009506/1) and to C.A.K. from the Josef Steiner Foundation and ERC grant 322602. B.Y.H.L. is supported by BBSRC project grant BB/S017593/1 and B.Y.H.L. and G.S.H.Y. by the MRC Metabolic Diseases Unit (MC_UU_00014/1). Next-generation sequencing was performed by the IMS Genomics and transcriptomics core facility, supported by the MRC (MC_UU_00014/5), the Wellcome Trust (208363/Z/17/Z), and the Cancer Research UK Cambridge Institute Genomics Core.

AUTHOR CONTRIBUTIONS

A.C.F.P. conceived the core idea. A.C.F.P. and M.A. conceived molecular and imaging experiments, which were performed by M.A. and N.Y. Embryo micromanipulation was performed by T.S. and A.C.F.P., and additional embryology was performed with M.A. Single-cell whole-transcriptome amplification was performed by M.G. (microarray) and B.Y.H.L., K.R., and M.K.M. (scRNA-seq). C.A.K. designed microarray analyses, which were performed and evaluated by M.H., X.L., B.Y.H.L., and C.A.K. B.Y.H.L. designed scRNA-seq analyses, which were performed and evaluated by B.Y.H.L., M.K.M., and G.S.H.Y. Data analysis was performed by M.A., B.Y.H.L., G.S.H.Y., and A.C.F.P. A.C.F.P. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no conflicting interests.

Received: May 16, 2022
Revised: July 26, 2022
Accepted: January 9, 2023

REFERENCES

1. Condic, M.L. (2014). Totipotency: what it is and what it is not. Stem Cells Dev. *23*, 796–812.
2. Jukam, D., Shariati, S.A.M., and Skotheim, J.M. (2017). Zygotic genome activation in vertebrates. Dev. Cell *42*, 316–332.
3. Zhou, L.Q., and Dean, J. (2015). Reprogramming the genome to totipotency in mouse embryos. Trends Cell Biol. *25*, 82–91.
4. Duch, M., Torras, N., Asami, M., Suzuki, T., Arjona, M.I., Gómez-Martínez, R., VerMilyea, M.D., Castilla, R., Plaza, J.A., and Perry, A.C.F. (2020). Tracking intracellular forces and mechanical property changes in mouse one-cell embryo development. Nat. Mater. *19*, 1114–1123.
5. Suzuki, T., Asami, M., Hoffmann, M., Lu, X., Gužvić, M., Klein, C.A., and Perry, A.C.F. (2016). Mice produced by mitotic reprogramming of sperm injected into haploid parthenogenotes. Nat. Commun. *7*, 12676.
6. Suzuki, T., Suzuki, E., Yoshida, N., Kubo, A., Li, H., Okuda, E., Amanai, M., and Perry, A.C.F. (2010). Mouse Emi2 as a distinctive regulatory hub in second meiotic metaphase. Development *137*, 3281–3291.
7. Perry, A.C.F., and Verlhac, M.-H. (2008). Second meiotic arrest and exit in frogs and mice. EMBO Rep. *9*, 246–251.
8. Shoji, S., Yoshida, N., Amanai, M., Ohgishi, M., Fukui, T., Fujimoto, S., Nakano, Y., Kajikawa, E., and Perry, A.C.F. (2006). Mammalian Emi2 mediates cytostatic arrest and transduces the signal for meiotic exit *via* Cdc20. EMBO J. *25*, 834–845.
9. Yoshida, N., Brahmajosyula, M., Shoji, S., Amanai, M., and Perry, A.C.F. (2007). Epigenetic discrimination by mouse metaphase II oocytes mediates asymmetric chromatin remodeling independently of meiotic exit. Dev. Biol. *301*, 464–477.
10. Balhorn, R. (1982). A model for the structure of chromatin in mammalian sperm. J. Cell Biol. *93*, 298–305.
11. Moore, G.D., Ayabe, T., Kopf, G.S., and Schultz, R.M. (1996). Temporal patterns of gene expression of G1-S cyclins and cdks during the first and second mitotic cell cycles in mouse embryos. Mol. Reprod. Dev. *45*, 264–275.

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

13. Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., Xia, W., et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652–657.

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

15. Flyamer, I.M., Gassler, J., Imakaev, M., Brandão, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110–114.

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

17. Gassler, J., Brandão, H.B., Imakaev, M., Flyamer, I.M., Ladstätter, S., Bickmore, W.A., Peters, J.M., Mirny, L.A., and Tachibana, K. (2017). A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. EMBO J. *36*, 3600–3618.

18. Ram, P.T., and Schultz, R.M. (1993). Reporter gene expression in G2 of the 1-cell mouse embryo. Dev. Biol. *156*, 552–556.

19. Aoki, F., Worrad, D.M., and Schultz, R.M. (1997). Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. Dev. Biol. *181*, 296–307.

20. Ko, M.S., Kitchen, J.R., Wang, X., Threat, T.A., Wang, X., Hasegawa, A., Sun, T., Grahovac, M.J., Kargul, G.J., Lim, M.K., et al. (2000). Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. Development *127*, 1737–1749.

21. Hamatani, T., Carter, M.G., Sharov, A.A., and Ko, M.S.H. (2004). Dynamics of global gene expression changes during mouse preimplantation development. Dev. Cell *6*, 117–131.

22. Wang, Q.T., Piotrowska, K., Ciemerych, M.A., Milenkovic, L., Scott, M.P., Davis, R.W., and Zernicka-Goetz, M. (2004). A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. Dev. Cell *6*, 133–144.

23. Zeng, F., and Schultz, R.M. (2005). RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. Dev. Biol. *283*, 40–57.

24. Park, S.J., Komata, M., Inoue, F., Yamada, K., Nakai, K., Ohsugi, M., and Shirahige, K. (2013). Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. Genes Dev. *27*, 2736–2748.

25. Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C.Y., Feng, Y., Liu, Z., Zeng, Q., Cheng, L., Sun, Y.E., et al. (2013). Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature *500*, 593–597.

26. Marston, J.H., and Chang, M.C. (1964). The fertilizable life of ova and their morphology following delayed insemination in mature and immature mice. J. Exp. Zool. *155*, 237–251.

27. Suarez, S.S. (1987). Sperm transport and motility in the mouse oviduct: observations *in situ*. Biol. Reprod. *36*, 203–210.

28. Adenot, P.G., Mercier, Y., Renard, J.P., and Thompson, E.M. (1997). Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. Development *124*, 4615–4625.

29. Abe, K.I., Yamamoto, R., Franke, V., Cao, M., Suzuki, Y., Suzuki, M.G., Vlahovicek, K., Svoboda, P., Schultz, R.M., and Aoki, F. (2015). The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3' processing. EMBO J. *34*, 1523–1537.

30. Olsen, T.K., and Baryawno, N. (2018). Introduction to single-cell RNA sequencing. Curr. Protoc. Mol. Biol. *122*, e57.

31. Blower, M.D., Jambhekar, A., Schwarz, D.S., and Toombs, J.A. (2013). Combining different mRNA capture methods to analyze the transcriptome: analysis of the *Xenopus laevis* transcriptome. PLoS One *8*, e77700.

32. Daniels, R., Kinis, T., Serhal, P., and Monk, M. (1995). Expression of the myotonin protein kinase gene in preimplantation human embryos. Hum. Mol. Genet. *4*, 389–393.

33. Temeles, G.L., and Schultz, R.M. (1997). Transient polyadenylation of a maternal mRNA following fertilization of mouse eggs. J. Reprod. Fertil. *109*, 223–228.

34. Asami, M., Lam, B.Y.H., Ma, M.K., Rainbow, K., Braun, S., VerMilyea, M.D., Yeo, G.S.H., and Perry, A.C.F. (2022). Human embryonic transcription initiates at the one-cell stage. Cell Stem Cell *29*, 209–216.e4.

35. Qiao, Y., Ren, C., Huang, S., Yuan, J., Liu, X., Fan, J., Lin, J., Wu, S., Chen, Q., Bo, X., et al. (2021). High-resolution annotation of the mouse preimplantation embryo transcriptome using long-read sequencing. Nat. Commun. *12*, 1767.

36. Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016). Resetting epigenetic memory by reprogramming of histone modifications in mammals. Mol. Cell *63*, 1066–1079.

37. Siklenka, K., Erkek, S., Godmann, M., Lambrot, R., McGraw, S., Lafleur, C., Cohen, T., Xia, J., Suderman, M., Hallett, M., et al. (2015). Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science *350*, aab2006.

38. Krämer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in ingenuity pathway analysis. Bioinformatics *30*, 523–530.

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

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

41. Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. Virology *411*, 273–287.

42. Hendrickson, P.G., Doráis, J.A., Grow, E.J., Whiddon, J.L., Lim, J.W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R., Wilcox, A.L., et al. (2017). Conserved roles of mouse Dux and human Dux4 in activating cleavage-stage genes and MERVL/HERVL retro- transposons. Nat. Genet. *49*, 925–934.

43. De Iaco, A., Verp, S., Offner, S., Grun, D., and Trono, D. (2020). DUX is a non-essential synchronizer of zygotic genome activation. Development *147*, dev177725.

44. Santini, L., Halbritter, F., Titz-Teixeira, F., Suzuki, T., Asami, M., Ma, X., Ramesmayer, J., Lackner, A., Warr, N., Pauler, F., et al. (2021). Genomic imprinting in mouse blastocysts is predominantly associated with H3K27me3. Nat. Commun. *12*, 3804.

45. Erkek, S., Hisano, M., Liang, C.Y., Gill, M., Murr, R., Dieker, J., Schübeler, D., van der Vlag, J., Stadler, M.B., and Peters, A.H.F.M. (2013). Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. Nat. Struct. Mol. Biol. *20*, 868–875.

46. Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

47. Sachs, M., Onodera, C., Blaschke, K., Ebata, K.T., Song, J.S., and Ramalho-Santos, M. (2013). Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline *in vivo*. Cell Rep. *3*, 1777–1784.

48. Vastenhouw, N.L., Zhang, Y., Woods, I.G., Imam, F., Regev, A., Liu, X.S., Rinn, J., and Schier, A.F. (2010). Chromatin signature of embryonic pluripotency is established during genome activation. Nature *464*, 922–926.

49. Perry, A.C.F., Wakayama, T., and Yanagimachi, R. (1999). A novel *trans*-complementation assay suggests full mammalian oocyte activation is coordinately initiated by multiple, submembrane sperm compartments. Biol. Reprod. *60*, 747–755.

50. Sun, Y.H., Wang, A., Song, C., Shankar, G., Srivastava, R.K., Au, K.F., and Li, X.Z. (2021). Single-molecule long-read sequencing reveals a conserved intact long RNA profile in sperm. Nat. Commun. *12*, 1361.

51. Zhou, D., Suzuki, T., Asami, M., and Perry, A.C.F. (2019). Caput epididymal mouse sperm support full development. Dev. Cell *50*, 5–6.

52. Singh, P., Fragoza, R., Blengini, C.S., Tran, T.N., Pannafino, G., Al-Sweel, N., Schimenti, K.J., Schindler, K., Alani, E.A., Yu, H., and Schimenti, J.C. (2021). Human MLH1/3 variants causing aneuploidy, pregnancy loss, and premature reproductive aging. Nat. Commun. *12*, 5005.

53. Chen, P.J., Hussmann, J.A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.F., Chen, C., Nelson, J.W., Newby, G.A., Sahin, M., et al. (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell *184*, 5635–5652.e29.

54. Liang, P., Xu, Y., Zhang, X., Ding, C., Huang, R., Zhang, Z., Lv, J., Xie, X., Chen, Y., Li, Y., et al. (2015). CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. Protein Cell *6*, 363–372.

55. Capalbo, A., Treff, N., Cimadomo, D., Tao, X., Ferrero, S., Vaiarelli, A., Colamaria, S., Maggiulli, R., Orlando, G., Scarica, C., et al. (2017). Abnormally fertilized oocytes can result in healthy live births: improved genetic technologies for preimplantation genetic testing can be used to rescue viable embryos in in vitro fertilization cycles. Fertil. Steril. *108*, 1007–1015.e3.

56. Mutia, K., Wiweko, B., Iffanolida, P.A., Febri, R.R., Muna, N., Riayati, O., Jasirwan, S.O., Yuningsih, T., Mansyur, E., and Hestiantoro, A. (2019). The frequency of chromosomal euploidy among 3PN embryos. J. Reproduction Infertil. *20*, 127–131.

57. Bellacosa, A. (2001). Functional interactions and signaling properties of mammalian DNA mismatch repair proteins. Cell Death Differ. *8*, 1076–1092.

58. Dang, C.V. (2012). MYC on the path to cancer. Cell *149*, 22–35.

59. Baluapuri, A., Wolf, E., and Eilers, M. (2020). Target gene-independent functions of MYC oncoproteins. Nat. Rev. Mol. Cell Biol. *21*, 255–267.

60. Papas, T.S., Watson, D.K., Sacchi, N., Fujiwara, S., Seth, A.K., Fisher, R.J., Bhat, N.K., Mavrothalassitis, G., Koizumi, S., Jorcyk, C.L., et al. (1990). ETS family of genes in leukemia and Down syndrome. Am. J. Med. Genet. Suppl. *7*, 251–261.

61. Shen-Li, H., O'Hagan, R.C., Hou, H., Jr., Horner, J.W., 2nd, Lee, H.W., and DePinho, R.A. (2000). Essential role for Max in early embryonic growth and development. Genes Dev. *14*, 17–22.

62. Wortel, I.M.N., van der Meer, L.T., Kilberg, M.S., and van Leeuwen, F.N. (2017). Surviving stress: modulation of ATF4-mediated stress responses in normal and malignant cells. Trends Endocrinol. Metabol. *28*, 794–806.

63. Alexandrova, N., Niklinski, J., Bliskovsky, V., Otterson, G.A., Blake, M., Kaye, F.J., and Zajac-Kaye, M. (1995). The N-terminal domain of c-Myc associates with alpha-tubulin and microtubules *in vivo* and *in vitro*. Mol. Cell Biol. *15*, 5188–5195.

64. González-Prieto, R., Cuijpers, S.A., Kumar, R., Hendriks, I.A., and Vertegaal, A.C. (2015). c-Myc is targeted to the proteasome for degradation in a SUMOylation-dependent manner, regulated by PIAS1, SENP7 and RNF4. Cell Cycle *14*, 1859–1872.

65. Conacci-Sorrell, M., Ngouenet, C., and Eisenman, R.N. (2010). Myc-nick: a cytoplasmic cleavage product of Myc that promotes alpha-tubulin acetylation and cell differentiation. Cell *142*, 480–493.

66. Han, H., Jain, A.D., Truica, M.I., Izquierdo-Ferrer, J., Anker, J.F., Lysy, B., Sagar, V., Luan, Y., Chalmers, Z.R., Unno, K., et al. (2019). Small-molecule MYC inhibitors suppress tumor growth and enhance immunotherapy. Cancer Cell *36*, 483–497.e15.

67. Huang, M.J., Cheng, Y.C., Liu, C.R., Lin, S., and Liu, H.E. (2006). A small-molecule c-Myc inhibitor, 10058-F4, induces cell-cycle arrest, apoptosis, and myeloid differentiation of human acute myeloid leukemia. Exp. Hematol. *34*, 1480–1489.

68. Joshi, K., Banasavadi-Siddegowda, Y., Mo, X., Kim, S.H., Mao, P., Kig, C., Nardini, D., Sobol, R.W., Chow, L.M.L., Kornblum, H.I., et al. (2013). MELK-dependent FOXM1 phosphorylation is essential for proliferation of glioma stem cells. Stem Cell. *31*, 1051–1063.

69. Morillo, S.M., Abanto, E.P., Román, M.J., and Frade, J.M. (2012). Nerve growth factor-induced cell cycle reentry in newborn neurons is triggered by $p38^{MAPK}$-dependent E2F4 phosphorylation. Mol. Cell Biol. *32*, 2722–2737.

70. Sjostrom, S.K., Finn, G., Hahn, W.C., Rowitch, D.H., and Kenney, A.M. (2005). The Cdk1 complex plays a prime role in regulating N-Myc phosphorylation and turnover in neural precursors. Dev. Cell *9*, 327–338.

71. Vervoorts, J., Lüscher-Firzlaff, J., and Lüscher, B. (2006). The ins and outs of MYC regulation by posttranslational mechanisms. J. Biol. Chem. *281*, 34725–34729.

72. Zhao, B., Wei, X., Li, W., Udan, R.S., Yang, Q., Kim, J., Xie, J., Ikenoue, T., Yu, J., Li, L., et al. (2007). Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. Genes Dev. *21*, 2747–2761.

73. Lourenco, C., Resetca, D., Redel, C., Lin, P., MacDonald, A.S., Ciaccio, R., Kenney, T.M.G., Wei, Y., Andrews, D.W., Sunnerhagen, M., et al. (2021). MYC protein interactors in gene transcription and cancer. Nat. Rev. Cancer *21*, 579–591.

74. Abuhashem, A., Garg, V., and Hadjantonakis, A.-K. (2022). RNA polymerase II pausing in development: orchestrating transcription. Open Biol. *12*, 210220.

75. Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432–445.

76. Adamo, P., and Ladomery, M.R. (2016). The oncogene ERG: a key factor in prostate cancer. Oncogene *35*, 403–414.

77. Luo, Y., Yang, S., Wu, X., Takahashi, S., Sun, L., Cai, J., Krausz, K.W., Guo, X., Dias, H.B., Gavrilova, O., et al. (2021). Intestinal MYC modulates obesity-related metabolic dysfunction. Nat. Metab. *3*, 923–939.

78. Palozola, K.C., Lerner, J., and Zaret, K.S. (2019). A changing paradigm of transcriptional memory propagation through mitosis. Nat. Rev. Mol. Cell Biol. *20*, 55–64.

79. Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K.S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell *161*, 555–568.

80. Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative binding of transcription factors orchestrates reprogramming. Cell *168*, 442–459.e20.

81. Liu, B., Xu, Q., Wang, Q., Feng, S., Lai, F., Wang, P., Zheng, F., Xiang, Y., Wu, J., Nie, J., et al. (2020). The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. Nature *587*, 139–144.

82. Wei, Y., Yang, C.R., and Zhao, Z.A. (2022). Viable offspring derived from single unfertilized mammalian oocytes. Proc. Natl. Acad. Sci. USA *119*, e2115248119.

83. Sarvella, P. (1973). Adult parthenogenetic chickens. Nature *243*, 171.

84. Bellier, S., Chastant, S., Adenot, P., Vincent, M., Renard, J.P., and Bensaude, O. (1997). Nuclear translocation and carboxyl-terminal domain phosphorylation of RNA polymerase II delineate the two phases of zygotic gene activation in mammalian embryos. EMBO J. *16*, 6250–6262.

85. Miyamoto, K., Nguyen, K.T., Allen, G.E., Jullien, J., Kumar, D., Otani, T., Bradshaw, C.R., Livesey, F.J., Kellis, M., and Gurdon, J.B. (2018). Chromatin accessibility impacts transcriptional reprogramming in oocytes. Cell Rep. *24*, 304–311.

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

87. Niakan, K.K., Han, J., Pedersen, R.A., Simon, C., and Pera, R.A.R. (2012). Human pre-implantation embryo development. Development *139*, 829–841.

88. Berrozpe, G., Bryant, G.O., Warpinski, K., Spagna, D., Narayan, S., Shah, S., and Ptashne, M. (2017). Polycomb responds to low levels of transcription. Cell Rep. *20*, 785–793.

89. Yesbolatova, A., Saito, Y., Kitamoto, N., Makino-Itou, H., Ajima, R., Nakano, R., Nakaoka, H., Fukui, K., Gamo, K., Tominari, Y., et al. (2020). The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. Nat. Commun. *11*, 5701.

90. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. j. *17*, 10. https://doi.org/10.14806/ej.17.1.200.

91. Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

92. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

93. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

94. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

95. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

96. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. *11*, R25.

97. Van der Maaten, L.J. (2014). Accelerating t-SNE using tree-based algorithms. J. Mach. Learn. Res. *15*, 3221–3245.

98. Turro, E., Su, S.Y., Gonçalves, Â., Coin, L.J.M., Richardson, S., and Lewin, A. (2011). Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. Genome Biol. *12*, R13.

99. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595.

100. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.

101. Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: database for annotation, visualization, and integrated discovery. Genome Biol. *4*, P3.

102. Yoshida, N., and Perry, A.C.F. (2007). Piezo-actuated mouse intracytoplasmic sperm injection (ICSI). Nat. Protoc. *2*, 296–304.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Primary antibodies used in this work	see Table S2	Table S2
Biological samples		
Mouse (*Mus musculus domesticus*) metaphase II (mII) oocytes	In-house	C57BL/6 × DBA/2 F1 (B6D2F1)
Mouse (*Mus musculus domesticus*) one-cell embryos	In-house	B6D2F1 x B6D2F1
Mouse (*Mus musculus castaneous*) sperm	MRC Harwell	CAST/EiJ
Mouse (*Mus musculus domesticus* x *Mus musculus castaneous*) one-cell embryos	This study	C57BL/6 × CAST/EiJ
Human metaphase II (mII) oocytes	Ovation Fertility Austin, Embryology and Andrology Laboratories, Austin, TX 78731, USA	https://www.ovationfertility.com/
Human bipronuclear (2PN) embryos	O Ovation Fertility Austin, Embryology and Andrology Laboratories, Austin, TX 78731, USA	https://www.ovationfertility.com/
Human monopronuclear (1PN) embryos	Ovation Fertility Austin, Embryology and Andrology Laboratories, Austin, TX 78731, USA	https://www.ovationfertility.com/
Human tripronuclear (3PN) embryos	Ovation Fertility Austin, Embryology and Andrology Laboratories, Austin, TX 78731, USA	https://www.ovationfertility.com/
Chemicals, peptides, and recombinant proteins		
α-amanitin	Merck	CAS No: 23109-05-9
5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole	Merck	CAS No: 53-85-0
5-[(4-Ethylphenyl)methylene]-2-thioxo-4-thiazolidinone (10058-F4)	Merck	CAS No: 403811-55-2
NUCC-0200975 (MYCi975)	Selleckchem	NUCC-0200975
CD532	Calbiochem	Cat # 532605
protein kinase RNA-like endoplasmic reticulum kinase (PERK)	Selleckchem	Cat # GSK2656157
ERGi-USU	R&D Systems	Cat # 6632/50
Critical commercial assays		
Clontech SMARTer Total RNA-Seq Kit Pico Input (V1) and (V2) system	Takara Clontech	Cat # 635006, Cat # 634412
T7 mScript™ Standard mRNA Production System	Cellscript	Cat # C-MSC100625
Deposited data		
Single-cell RNA-sequencing datasets. See Table S1	Asami et al.[34] This paper.	GEO: GSE157834 and GSE222130
DNA microarray data	This paper	SE64648, GSE64649 and GSE64650
RNA-sequencing datasets	Park et al.[24]	

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Database of transcriptome in mouse early embryos	DBTMEE	http://dbtmee.hgc.jp
A sperm H3K4me3 ChIP-seq dataset	Erkek et al.[45]	GSE42629
ChIP-seq data: raw fastq reads for H3K4me3-bound DNA	Erkek et al.[45]	GSM1046833
ChIP-seq data: sonicated genomic DNA from sperm	Input control: GEO	GSM1046836
H3K27me3 ChIP-seq, H3K4me3 ChIP-seq and RNA-sequencing	Zheng et al.[36]	GSE76687 and GSE71434
Experimental models: Organisms/strains		
Human (*Homo sapiens*)	N/A	N/A
Mouse (*Mus musculus domesticus* and *Mus musculus castaneus*)	N/A	N/A
Oligonucleotides		
Primers for qPCR. See Table S3	This paper	N/A
Software and algorithms		
Cutadapt (1.7.1)	Martin[90]	https://cutadapt.readthedocs.io/en/stable/
Tophat (2.0.11)	Trapnell et al.[91]	https://ccb.jhu.edu/software/tophat/index.shtml
Cufflinks (2.2.1)	Trapnell et al.[92]	http://cole-trapnell-lab.github.io/cufflinks/
ht-seq-count (0.6.1p1)	Anders et al.[93]	https://htseq.readthedocs.io/en/release_0.11.1/count.html
Limma-voom	Ritchie et al.[94]	https://bioconductor.org/packages/release/bioc/html/limma.html
Igv (2.4.19)	Robinson et al.[95]	https://software.broadinstitute.org/software/igv/
edgeR	Robinson et al.[95] Robinson and Oshlack[96]	https://bioconductor.org/packages/release/bioc/html/edgeR.html
Rtsne	Van Der Maaten[97]	https://cran.r-project.org/web/packages/Rtsne/index.html
mmseq 1.0.9	Turro et al.[98]	https://github.com/eturro/mmseq
fastx-Toolkit (version 0.0.13)	Hannon lab	http://hannonlab.cshl.edu/fastx_toolkit/download.html
BWA MEM (v0.7.12)	Li et al.[99]	https://arxiv.org/abs/1303.3997
MACS2	Zhang et al.[100]	https://hbctraining.github.io/Intro-to-ChIPseq/lessons/05_peak_calling_macs.html
Qiagen Ingenuity Pathway Analysis (IPA)	QIAGEN	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/
DAVID	Dennis et al.[101]	https://david.ncifcrf.gov/
Other		
Agilent Bioanalyzer 2100 Agilent, USA	Agilent, USA	N/A
Metamorph software 7.8.12	Molecular Devices, LLC, USA	N/A
ABI 7500 Real Time PCR System	Applied Biosystems, CA	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Tony Perry (perry135@aol.com).

Materials availability

All unique reagents generated in this study are available via the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- The sequencing data generated in this study have been deposited in the NCBI GEO database under the accession number listed in the key resources table. Microarray data generated for this study have been deposited in the Gene Expression Omnibus (GEO) repository under accession numbers GEO: GSE64648, GEO: GSE64649, and GEO: GSE64650. Single-cell RNA-seq data generated for this study have been deposited under the accession numbers GEO: GSE157834 and GEO: GSE222130. This paper analyzes existing, publicly available data. The accession numbers for the datasets are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

Oocytes were from a *Mus musculus domesticus* C57BL/6 (B6) x DBA/2 F1 hybrid (B6D2F$_1$) and *M. m. castaneus*/EiJ (*cast*). Embryos were from (B6 x DBA/2) x (B6 x DBA/2) F2 hybrids (F2), or B6 x *cast* (B6*cast*) crosses. Mice were housed in specific pathogen-free facilities. Animal experiments were reviewed and approved by the UK Home Office (PPL PP5839163) and the University of Bath Animal Welfare Ethical Review Body. Female and male mice were used at 8-12 weeks of age.

Human oocyte and embryo sample collection

Human oocytes and single monopronuclear (1PN), bipronuclear (2PN) and tripronuclear (3PN) one-cell embryos were supplied anonymously subject to informed consent for use in research by couples who had finished family building or decided for other reasons to discontinue fertility treatment. Consents strictly adhered to guidelines of the Ethics Committee of the American Society for Reproductive Medicine. Embryos were cryopreserved and lysed on site at *Ovation Fertility* before being anonymized and shipped for analysis.

METHOD DETAILS

Human metaphase II oocytes and one-cell embryos (zygotes)

All human sample collection, processing and analytical protocols were as described previously.[34] Informed consent strictly adhered to guidelines of the Ethics Committee of the American Society for Reproductive Medicine. Embryos were lysed on site at *Ovation Fertility* before being anonymized and shipped for analysis.

Animals

Experimental animals in this study were all non-wild laboratory mice. Experiments were performed in accordance with local and national statutes including the University of Bath Animal Welfare Ethical Review Body and complied with the UK Animals (Scientific Procedures) Act, 1986 and its embodiments. Mice were housed in cages of up to five animals in ventilated cabinets with *ad libitum* access to food and water. Where appropriate, they were allowed to acclimate for seven days prior to treatment. Cage temperatures and relative humidity were recorded daily and ranged from 22-24°C and 40-55%, respectively.

Collection and culture of oocytes

For the B6 x *cast* (B6*cast*) cross, sperm from *M. m. castaneus* males were injected into mII oocytes from B6 females. Oviductal metaphase II (mII) oocyte complexes were typically collected in M2 medium (EMD Millipore, UK) from 8-12-week-old C57BL/6 or B6D2F$_1$ females (produced by crossing C57BL/6 females with DBA/2 males in-house or otherwise supplied by Charles River; L'Arbresle, France) 12 to 15 h after standard superovulation by serial injection of equine and human chorionic gonadotropin (PMSG and hCG). Complexes were then either used in IVF (below), or cumulus cells were removed by hyaluronidase treatment and after multiple washing in M2 medium, denuded oocytes incubated in kalium simplex optimized medium (KSOM; Millipore) under mineral oil in humidified 5% CO_2 (v/v air) at 37°C, until required.

Sperm preparation and microinjection (ICSI)

Preparation of cauda epididymidal sperm from 8- to 12-week-old B6D2F$_1$ or *cast* males for ICSI was essentially as previously described.[102] Injection was completed in ± ≤2.5 min for each timepoint. B6*cast* and B6*cast*-HiRes series were generated by injecting *cast* sperm into C57BL/6 oocytes (in ± ≤1.0 min per timepoint for the latter). Sperm were prepared by trituration for 45 sec in nuclear isolation medium (NIM; 125 mM KCl, 2.6 mM NaCl, 7.8 mM Na_2HPO_4, 1.4 mM KH_2PO_4, 3.0 mM EDTA; pH 7.0) containing 1.0% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) at room temperature (25°C). They were then washed twice in NIM and pelleted at ambient temperature; head-tail detachment was enhanced by trituration during pellet resuspension. Finally, sperm were resuspended in ice-cold NIM (~0.5 mL per epididymis equivalent) and stored at 4°C for up to 3 h until required, but typically injected immediately after preparation. This gentle protocol liberates sperm membrane and other (eg acrosomal) components that do not enter the oocyte during fertilization; the sperm support normal, healthy full-term development[5]; ICSI, *in vitro* fertilization (IVF) and natural mating generate indistinguishable offspring at similar rates.

Immediately before microinjection, ~50 μL of each sperm suspension was mixed with 20 μL of polyvinylpyrrolidone (PVP, average $M_r \approx 360{,}000$; Sigma-Aldrich) solution (15% [w/v]) and sperm injected (ICSI) within ~60 min of PVP mixing into oocytes in a droplet of M2 as described.[5,102] After a brief recovery period (~5 min), injected oocytes were transferred to KSOM under mineral oil equilibrated in humidified 5% CO_2 (v/v air) at 37°C and cultured until required.

Fertilization *in vitro* (IVF)

Where appropriate, embryos were generated by standard B6D2F$_1$ × B6D2F$_1$ IVF. Sperm were collected from mature males by epididymidal puncture followed by dispersal for 5 min in pre-warmed human tubal fluid (HTF; Millipore) in humidified CO_2 (5% [v/v] in air) at 37°C. Of the 400 μL dispersal droplet, 10 μL were transferred to a fresh fertilization dish containing 200 μL HTF and incubation continued for 1 h before placing ~30 cumulus oophorous complexes freshly-isolated from superovulated 8-week-old B6D2F$_1$ females and incubating in the CO_2 incubator at 37°C. The resultant embryos were washed in fresh HTF and dead and clearly unfertilized oocytes removed. Embryos were then washed 5× in KSOM and incubated until required in KSOM droplets equilibrated under mineral oil in humidified 5% CO_2 (v/v air) at 37°C.

Embryo culture

Embryos were typically cultured in KSOM droplets equilibrated under mineral oil in humidified 5% CO_2 ([v/v] in air) at 37°C as previously described.[5,102] Where appropriate for the inhibition of Pol II-mediated transcription in qPCR, mII oocytes were incubated in the presence of 100 μg/mL α-amanitin (Sigma-Aldrich) for 40 min prior to ICSI and post-ICSI incubation continued in the presence of α-amanitin. For the α-amanitin time course (Figure S4C), samples were collected 5 min post-ICSI (t = 0) or at the indicated times after continued culture in α-amanitin-supplemented media. For scRNA-seq, embryos were generated by sperm-oocyte mixing for 1 h and then incubated in KSOM containing α-amanitin (100 μg/mL) for a further 10 h, when bipronuclear embryo samples (with or without exposure to α-amanitin) were collected. Culture of some embryos was continued to confirm developmental attenuation. There were $n = 2$ biological replicates, performed on different days, with untreated mII oocytes collected in parallel. After removal of a single technical drop-out, data for all ($n = 11$) scRNA-seq samples are presented (Figure S4E). For inhibition of Pol II-mediated transcript elongation by interference with Cdk9, embryos were incubated in the presence of 120 μM 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB; Merck) beginning 1 h after the start of IVF (when most fertilization has occurred) and collected 6 h after the start of DRB treatment (Figure S4G); control embryos were incubated in parallel in media lacking DRB. To inhibit the *trans*-activation of c-Myc target gene expression, embryos produced by IVF 2 h after sperm-oocyte mixing were washed and transferred to KSOM supplemented with 5-[(4-Ethylphenyl)methylene]-2-thioxo-4-thiazolidinone (10058-F4; Merck) or NUCC-0200975 (MYCi975; Selleckchem, UK) and incubation continued until sample collection. 10058-F4 and MYCi975 are structurally distinct inhibitors of c-Myc-Max interaction, preventing *trans*-activation of c-Myc target gene expression; MYCi975 promotes MYC phosphorylation and degradation. 10058-F4 was used within commonly-used range of working concentrations at 50 or 100 μM and MYCi975 at 5 μM, solubilized in dimethylsulphoxide (DMSO); the same amount of DMSO was included in media without inhibitor for negative controls. For single-cell RNA-sequencing (scRNA-seq; below), 10058-F4-treated samples were collected after 10 h. Mycn, Atf4 and Erg were inhibited using an analogous protocol with their respective inhibitors. Inhibition of Mycn was achieved with CD532 (Calbiochem, US), which disrupts the Mycn-Aurora A complex. Inhibition of Atf4 was achieved by blocking its upstream kinase, protein kinase RNA-like endoplasmic reticulum kinase (PERK), with GSK2656157 (Selleckchem, UK) and of Erg by inhibiting its upstream kinase, Rio Kinase 2 (RIOK2) with ERGi-USU (Cat. No. 6632/50, R&D Systems, US). Experiments in all cases were performed on at least two experimental days and included controls performed in parallel using media lacking inhibitor.

Direct fluorescence imaging and image analysis

Differential interference contrast microscopy (DIC) and epifluorescence imaging have been described previously.[4–6] Images of live oocytes or embryos following cRNA injection were captured on an Olympus IX71 equipped with an Andro Zyla sCMOS camera and OptoLED illumination system (Cairn Research Ltd., UK) and processed using Metamorph software (Molecular Devices, LLC, USA). Excitation at 587 nm in combination with an ET-mCherry filter system was used for mCherry fluorescence detection and at 484 nm with an ET-EYFP filter system to detect Venus epifluorescence. Confocal images were obtained using LaserSharp 2000 6.0 Build 846 software on an Eclipse E600 (Nikon, Japan) microscope equipped with a Radiance 2100 laser scanning system

(BioRad, USA; LSM Technical Service, UK). Image J (http://rsbweb.nih.gov/ij) was used in image data analysis. The parental provenance of pronuclei in mouse one-cell embryos was assigned according to size and position; the female pronucleus is consistently smaller and closer to the second polar body than the paternal pronucleus.

Immunocytochemistry

Oocytes, embryos and cultured cells were fixed in 4% (w/v) paraformaldehyde and either processed immediately or stored at 4°C for up to two weeks until required. Fixed cells were permeabilized by incubation in PBS supplemented with 0.5% (v/v) triton X-100 and 0.1% (w/v) BSA for 30 min at 37°C, followed by blocking in PBS supplemented with 3% (v/v) normal goat serum and 0.1% (w/v) BSA for 30 min at room temperature. Labelling was by incubating samples overnight at 4°C in primary antibody (Table S2) followed by incubation for 1 h at 37°C with the appropriate secondary antibody (1:250 [v/v]; Life Technologies Ltd., UK) conjugated to Alexa 488, Alexa 594 and Alexa 647. DNA was stained by incubating samples at 37°C for 20 min in propidium iodide (1:200 [v/v]; Sigma, USA) or Hoechst 33342 (1:1000 [v/v]; Sigma). Chromatin epitopes in mII oocytes and embryos are accessible to their cognate antibodies[9] and most (27/33) here resided on solvent-exposed N-terminal histone tails. In 28/33 cases (84.8%), samples contained in-built positive controls in which one or both parental chromatin sets stained within a single given cell and only one situation corresponded to an epitope (the core modification, H4K91ac) that was unrecognised in both parental genomes at mII; however, both were stained in one-cell embryos. Antibodies with no reactivity in either mII oocytes or embryos were excluded from the analysis. For c-Myc immunocytochemistry, anti-c-Myc rabbit polyclonal antibody 10828-1-AP (Protein Tech) was used at 1:200 (v/v). Anti-α-tubulin (Sigma) was used at 1:1000 (v/v), and for F-actin staining, phalloidin 488 (Invitrogen) was used at 1:100 (v/v). DNA was stained by incubating samples at 37°C for 20 min in propidium iodide (1:200 [v/v]; Sigma, USA).

Immunoblotting

To characterize protein expression, mII oocytes or one-cell embryos were washed in BSA-free medium, transferred with minimum extraneous medium to a 1.5 mL tube, collected at the bottom of the tube by a centrifugation pulse, immediately flash-frozen in liquid N_2, and stored at ¯80°C until required. Immediately prior to electrophoresis, samples were thawed by vortexing in 5-10 μL ice-cold lysis buffer containing 50 mM Tris-Cl (pH7.0), 0.5 mM EDTA, 150 mM NaCl, 1% (v/v) Triton X-100 and a protease inhibitor cocktail (Thermo Scientific, UK). Samples were supplemented by mixing with an equal volume of loading buffer (2× Laemmli Lysis-buffer, Sigma), heated at 95°C for 5 min and immediately placed on ice. Polyacrylamide gel electrophoresis was through a 4% (w/v) stacking gel and a 10 or 12% (w/v) separating gel (ProtoGel, National Diagnostics, UK) under reducing conditions. Each track contained 100 mII oocytes, 100 one-cell embryos, or prestained protein size markers (Thermo Fisher Scientific, Inc., US, and Cleaver Scientific, Ltd., UK). Separated proteins were transferred onto PVDF-transfer membrane (GE Healthcare, USA) *via* a Trans-Blot Turbo (BioRad, USA), followed by blocking with 3% (w/v) non-fat dry milk powder in standard TBST buffer at room temperature for 1 h. Membranes were then incubated at 37°C for 2 h or 4°C overnight in 1% (v/v) blocking solution (1% [w/v] non-fat dry milk powder in TBS buffer) containing the following primary antibodies at the following dilutions (v/v): rabbit polyclonal c-Myc (Protein Tech, 10828-1-AP), 1:2000; monoclonal β-actin (Abcam), 1:2000; monoclonal α-tubulin (Sigma), 1:1000 (Table S2). After washing 3 × 15 min in TBST buffer (room temperature), membranes were incubated with secondary antibody (anti-mouse IgG horseradish peroxidase conjugate, 1:5,000 [Millipore, USA] or anti-rabbit IgG horseradish peroxidase conjugate, 1:10,000 [Abcam]) in TBS buffer at 37°C for 2 h or 4°C overnight with agitation. Membranes were washed 5× with TBST for 15 min each and immunoreactivity detected using Clarity Western ECL Substrate (BioRad). Where appropriate, re-blotting was performed after removing primary and secondary antibodies with Restore Western Blot Stripping Buffer (Thermo Fisher Scientific, UK) according to the recommendations of the manufacturer. Signal intensity quantification was performed with ImageJ (http://rsbweb.nih.gov/ij) by subtracting mean background signals from mean signals in areas defined with the selection tool.

Preparation and injection of cRNA

5′-capped cRNA was synthesized *in vitro* from linearized plasmid template DNA in a T7 mScript™ Standard mRNA Production System (Cellscript, USA) according to the recommendations of the manufacturer, as previously described.[5] Where appropriate, the polyadenylation step following synthesis was omitted (Figure 2C). cRNA was dissolved in nuclease-free water, quantified on a Nanophotometer and stored in aliquots at ¯80°C until required. cRNA solutions were diluted as appropriate with sterile water and injected (typically at concentrations of 0.01–1.00 μg/μL) within 1 h of thawing *via* a piezo-actuated micropipette into mII oocytes or embryos in M2 medium.

Ratiometric RT-PCR (qPCR)

For the analysis of discrete mRNAs, metaphase II (mII) oocytes or embryos were subjected to ratiometric transcript quantification by PCR (qPCR). qPCR was selected as an orthogonal analytical technique relative to scRNA-seq (below): it employed non-fragmented RNA, lacked a subtraction step, had a distinctive amplification schedule and used different embryos. The dynamic range of qPCR spans 8∼9 orders of magnitude, so detection is less a function of target starting concentrations. Mouse embryos for qPCR were produced either by IVF or ICSI and incubated in humidified 5% CO_2 (v/v air) at 37°C until required. Embryos at the appropriate time post-injection (6 h unless otherwise stated) were examined to confirm morphology (*eg* the presence of a second polar body and two pronuclei) and 5∼10 oocytes or embryos transferred in a minimal volume (<0.5 μL) to 1 μL 0.2% (w/v) sarkosyl (Merck) containing

10 ng tRNA (Hoffmann-La Roche Ltd., Basel, Ch) in a thin-walled 0.2 mL PCR tube, for use either directly or for flash-freezing in liquid nitrogen followed by storage at ¯80°C until required. Cells were heated at 65°C for 5 min and used to program cDNA synthesis primed with random 8-mers plus oligo(dT)$_{20}$ (each at 30 μM) in a 21 μL reaction volume containing 200 U SuperScript IV reverse transcriptase (Thermo Fisher Scientific, UK). qPCR reactions were performed in an ABI 7500 Real Time PCR System (Applied Biosystems, CA) in reactions (20 μL) containing 1-2 μL template cDNA, forward and reverse primers (100 nM each) and 12.5 μL of Power SYBR (ABI), using the parameters: 10 min at 95°C, followed by cycles of (15 sec at 95°C, 1 min at 58°C and 35 sec at 72°C). Each experiment was performed with biological triplicates collected on at least two days and included technical duplicates of each sample. Primer sets (Eurofins Genomics, Germany, and Merck) were non-dimerizing under the conditions employed. Primers corresponded to exonic (forward: ExF1, ExF2; reverse: ExR1 and ExR2) and intronic (forward: IntF1; reverse: IntR1) sequences in Figure 2A. Primer pairs (Table S3) were intra-exonic (iE) and in splicing analysis straddled exon-exon (EE) or intron-exon (IE) boundaries. Reactions lacking input cDNA were used to verify absence of contamination in cocktail components. Steady state transcript levels were normalized with respect to the internal reference, *H3f3a*, which is robustly expressed in mouse oocytes and preimplantation embryos.[6,20] The CT value for *H3f3a* corrected for cell number (CT + log2 embryo cell number) is constant during preimplantation stages, giving a mean value of 29.56 ± 0.23 (for $25 \leq n \leq 29$ independent replicates). Comparative quantification of RNA extraction and cDNA synthesis efficiencies was achieved by adding control 'spike-in' RNA to mII oocyte ($n = 5$ pooled cells) or 6 h embryo ($n = 6$ pooled cells) lysates prior to cDNA synthesis, according to the vendor's recommendation (Primerdesign Ltd, UK) (Figure 1B). For human mII oocytes and one-cell embryos, single cells were subjected to qPCR as described previously.[34] PCR primer sequences are given in Table S3. Statistical differences between pairs of datasets were analyzed by two-tailed unpaired *t*-tests and p values ≤ 0.05 considered statistically significant unless stated otherwise. All qPCR employed $n \geq 3$ independent biological samples per target.

Single-cell RNA-sequencing (scRNA-seq)

RNA-sequencing libraries were prepared from single mII oocytes and one-cell embryos 2, 4, 6, 8, 10 and 12 h post-ICSI using the Clontech SMARTer Total RNA-Seq Kit Pico Input (V1) system (Takara Clontech). Briefly, total RNA was liberated by lysis of single oocytes or one-embryos in 8.0 μL lysis buffer supplemented with RNase inhibitor (both from Takara Clontech). The RNA was incubated with SMART Pico Oligo mix at 94°C for 3 min and then subjected to first-strand synthesis with SMARTScribe reverse transcription using a template-switching oligo (TSO). After first strand synthesis, cDNA amplification was performed using SeqAmp DNA polymerase with Illumina barcoded adapters for 5 cycles of 15 sec at 98°C; 15 sec at 55°C; 30 sec at 68°C, followed by final extension for 2 min at 68°C. Ribosomal cDNA was removed using ZapR and R-Probes, after which there was a second round of cDNA amplification with SeqAmp DNA Polymerase for 15 cycles of 15 sec at 98°C, 15 sec at 55°C, and 30 sec at 68°C, to generate the final sequencing libraries. Libraries were analyzed using a High-sensitivity DNA kit and Agilent Bioanalyzer 2100 (Agilent, USA). For next-generation sequencing, libraries were combined at equimolar concentrations before loading onto an Illumina HiSeq 4000 instrument (Illumina, USA) for paired-end 100 (PE100) sequencing.

Quality control (QC) of scRNA-seq is reported in: Table S1. F2-B6cast scRNA-sequencing data QC and analyses.

RNA-sequencing for F4 and α-amanitin experiments

For each experiment, sequencing libraries were constructed with the SMARTer Stranded Total RNA-Seq Kit v2 – Pico Input Mammalian (Takara), as previously described.[34] Briefly, total RNA was liberated from one-cell embryos in 8 μL of lysis buffer containing RNAse inhibitor (Takara). RNA was then fragmented for 4 min and subjected to reverse transcription with SMART Pico N6 primers and Pico v2 SMART adaptor (template-switching oligo). Second strand cDNA synthesis was performed by subjecting the samples to five cycles of PCR with Illumina-compatible barcoded libraries. This was followed by removal of rRNA and an addition 15 cycles of PCR amplification. Indexed libraries were QC-validated using high-sensitivity D1000 ScreenTape and Agilent TapeStation 4200. Libraries were then normalised, pooled, and sequenced on an Illumina NovaSeq 6000 instrument to generate ∼500 million paired-end reads (PE100 and PE50 for F4 and α-amanitin experiments, respectively).

Microarrays

For time course transcriptomic microarray analysis, oocyte and embryo transcriptomes were prepared as for scRNA-seq (above) and preparation and whole transcriptome amplification fidelity assessed as previously described.[20] Gene expression data were quality-assessed by inspection of chip raw images and gene expression frequency distributions. Only high quality data were approved for further bioinformatic analysis, producing 51 samples. Raw gene expression data were background-corrected (limma R-package, normexp method) and normalized by quantile normalization. Technically replicated probes (identical Agilent IDs) were replaced by their median per sample. The original standard deviation was 2.07. For clustering and functional annotation, probes targeting the same gene were disambiguated by retaining only the probe with the lowest p value. This reduced the 41,000 non-control probes to 29,078, of which 21,391 were annotated by proper gene symbols. The number of exons was retrieved from UCSC known genes (mm9) by matching Agilent probe sequence locations to UCSC transcript coordinates. If different transcripts for the same probe sequence were available the number of exons was averaged across these transcripts.

Sequencing alignment and gene and transcript counts

Raw sequence reads (fastq) from mouse oocytes and embryos were trimmed using the fastx_trimmer command from fastx-Toolkit (version 0.0.13; respectively 5 nt from 5′ and 6 nt from 3′ ends), and sequencing adapter sequences were removed by Cutadapt (1.7.1).[90] This removed terminal primer (6mer) sequences that might have introduced mismatches. Mapping of the trimmed reads was by Tophat (2.0.11)[91] using the GRCm38 genome and Ensembl 54 reference transcriptome. After sequence alignment, the transcriptome was remodeled *via de novo* re-assembly of transcripts based on empirical data from mapped reads and incorporation into the original reference (Cufflinks 2.2.1).[92] Gene-level counts based on the updated transcriptome were then performed using ht-seq-count (0.6.1p1)[93] and transcript level counts *via* cuffnorm command from Cufflinks (2.2.1).[92] The final average read-pairs per sample (±s.e.m.) was 30.4 ± 1.3 million. Transcript classification was according to Cufflinks canonical splicing identifiers ' = ' (complete match of intron chain) and 'j' (potentially novel isoform).[92] For visualization in Figure S3A, maternal genes were subtracted from F2-B6*cast* series DEGs (FDR<0.05) at the time-points indicated. Sashimi plots were generated using mapped reads from oocytes and one-cell embryos using Integrative Genome Viewer (IGV) version 2.4.19. This analysis corroborated de novo gene expression, as distinct from extant maternal transcripts and genomic DNA.

Microarray data analysis

For the microarray dataset, raw gene expression values were background-corrected and normalized by quantile normalization using the Limma package.[94] As with RNA sequencing analysis, a GLM was fitted and eBayes used to detect differential gene expression over the time series. Genes with an FDR <0.05 were considered to be differentially expressed.

Differential expression analysis

Gene- and transcript-level count tables were imported into edgeR[95] for downstream differential gene expression analysis. For gene-level analysis, genes expressed at low levels (<1 count per million, >45 samples) were filtered, retaining a final total of 19,184 genes. Samples were subjected to trim-mean of M values (TMM) normalization from the edgeR package,[96] a common normalization method in which a weighted trimmed mean of the log expression ratios is used to normalize sequencing depth, here with a normalization factor range of 0.93–1.08. This is demonstrably among the most robust methods for RNA-seq differential expression analysis, including single-cell studies where global gene expression differences are large. Normalized gene abundance was listed as read-counts per million of mapped reads (CPM) and determined using the cpm() function in edgeR.

A visual overview of transcriptome profiles was generated in the Rtsne package (https://cran.r-project.org/web/packages/Rtsne/Rtsne.pdf) by performing dimensionality reduction *via* the principal component analysis (PCA, ndims = 50) coupled with t-distributed stochastic neighbor embedding (t-SNE), using the CPM of all genes as input with the default parameters.[97] The advantage of using t-SNE over PCA is that it provides superior resolution with which to differentiate samples in fewer dimensions (two projections were used here) and it is commonly used in scRNA-seq studies. For differential gene expression, a generalized linear model (GLM) was applied to determine the common, trend and gene-wise dispersions, and likelihood-ratio tests were employed to detect differential gene expression. Genes with a false discovery rate (FDR) of <0.05 were considered differentially expressed.

For transcript-level analysis, transcripts with low expression were first removed and estimated transcript level counts then TMM-normalized (normalization factor: 0.85–1.08) and CPM determined by edgeR + limma-voom.[94] The total number of transcripts that remained in the analysis was 53,150. We then calculated a gene-variability statistic to adjust for the mean-variance relationship using limma-voom, and differential expression was determined using empirical Bayesian *t* test (eBayes) from the limma package. Transcripts with an FDR of <0.05 were considered to be differentially expressed.

For parent-of-origin analysis, gffread utility from Tophat (2.0.11) was used to generate the reference fasta for the transcriptomes of C57/BL6 and Cast/EiJ based on their corresponding reference genomes (GRCm38, Ensembl 70) and remodeled transcriptome data as described above. The transcriptomes were merged to form a single 'F1′ reference transcriptome for mapping of trimmed sequence reads using Bowtie 1.1.0. Mapped reads were subsequently analyzed using MMSEQ 1.0.9 to estimate the transcript level and aggregated gene level abundance originating from the genome of each strain.[98] Abundance tables were generated using mmseq.R R script accompanied by the MMSEQ package for downstream analyses.[98] Estimated counts were then imported to limma-voom (normalization factor: 0.91–1.08) for downstream differential gene expression analysis similar to the method described above.

iEGA gene trajectory analysis

To determine the expression trajectories of iEGA genes beyond the iEGA time course, we performed bioinformatic analysis of data from Park et al.[24] using normalised FPKM values for oocytes (n = 3), and one-, two- and four-cell embryos (n = 2) extracted from DBTMEE (database of transcriptome in mouse early embryos, http://dbtmee.hgc.jp). The gene expression profiles of 302 iEGA genes (FDR<0.05 throughout the time course series, $\geq$0.5 log2FC in at least one time point compared to mII oocytes) were extracted from this dataset and used to create a heatmap using the pHeatmap package in R. Hierarchical clustering was performed on genes using the 'average' clustering algorithm. Data are presented in Figure 4E, whose scale bar indicates expression z-scores derived from Log2FPKM values.

Bioinformatics and differential gene expression analysis for F4 and α-amanitin experiments

Sequence reads were mapped to the mouse genome (GRCm38), Ensembl Version 100, and gene-level expression was determined, both using STAR (version 2.5.0a) with the parameters (–outFilterScoreMinOverLread 0.3 –outFilterMatchNminOverLread 0.3 –outSAMstrandField intronMotif –outSAMtype BAM SortedByCoordinate). Count tables were then TMM-normalised using EdgeR (version 3.36.0) and Limma (Version 3.50.1). A gene-wise Negative Binomial GLM model was then fitted, and a quasi-likelihood test used to determine DEGs.

Sperm H3K4me3 ChIP-seq analysis

A sperm H3K4me3 ChIP-seq dataset[45] was downloaded from Gene Expression Omnibus (Accession GSE42629). Raw fastq reads for H3K4me3-bound DNA (GSM1046833, ChIP) and sonicated genomic DNA from sperm (GSM1046836, input control) were aligned to the mouse GRCm38 genome using BWA MEM (v0.7.12).[99] Mapped reads were sorted by their genomic coordinates and used as input for peak detection using MACS2,[100] which performed PCR de-duplication, library normalization, ChIP peak modeling and calling by comparing reads from the H3K4me3 bound library to background controls (sperm genomic DNA). FDR<0.05 and peak fold-enrichment >5 were used to filter peak calls (total 44,209 peaks spanning a total of 62 Mb, or ~1.78% of the genome). Filtered peaks were annotated using ChIPSeeker R package with Ensembl mouse reference gene model GTF version 100. Of the peaks, 28,190 (63.8%) were labeled within the promotor (±3 kb from the transcriptional start site [TSS]) and 5′ untranslated regions (UTRs). Only peaks within the proximal promotor region (1.5 kb upstream or downstream of the TSS) were used for downstream comparisons.

Pathway analysis

DAVID[101] (https://david.ncifcrf.gov/) was used for some of the analysis of DEGs in F4-treated embryos ($p < 0.1$). Other pathway and upstream regulator analyses were performed using Qiagen Ingenuity Pathway Analysis (IPA) software with MGI gene symbols and a nominal, uncorrected $p < 0.05$ cut-off. Loci encompassing multiple genes were split in the analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics and reproducibility

Statistical differences between pairs of datasets were analyzed by χ^2 or, more typically, two-tailed unpaired *t*-tests. Values of $p < 0.05$ were considered statistically significant unless stated otherwise. Data were collected on at least two experimental days and reflect ≥ 2 independent biological replicates.