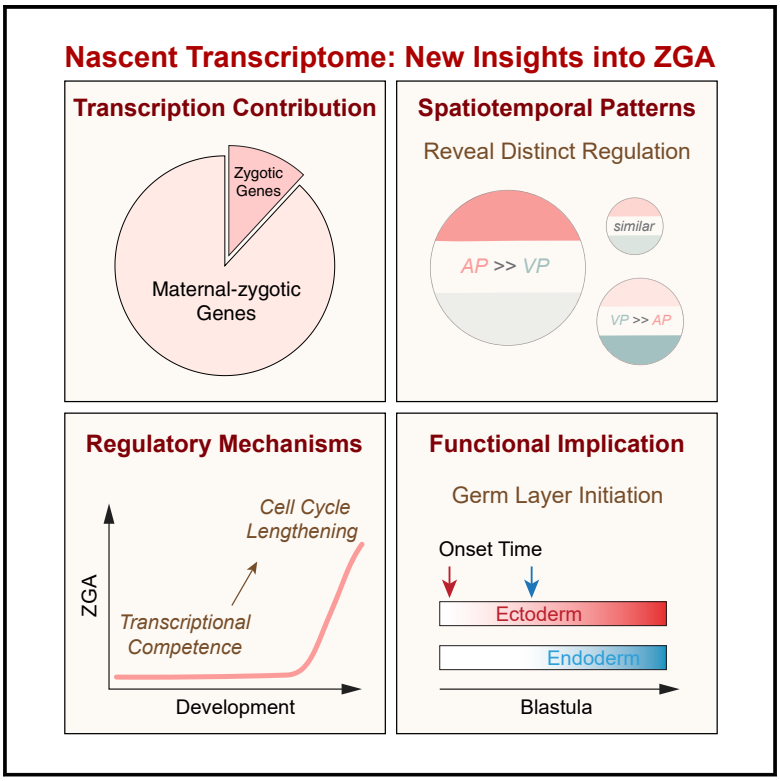
Report

Current Biology

Nascent transcriptome reveals orchestration of zygotic genome activation in early embryogenesis

Graphical abstract



Authors

Hui Chen, Matthew C. Good

Correspondence

mattgood@pennmedicine.upenn.edu

In brief

By profiling the nascent transcriptome during zygotic genome activation (ZGA) from whole *Xenopus* blastula and dissected subregions, Chen and Good unveil predominant transcription from maternal-zygotic genes and distinct spatial patterns, reconcile regulatory mechanisms of ZGA, and discover a link to sequential activation of germ-layer-specific genes.

Highlights

- Whole-embryo and regional EU-RNA-seq determines timing and spatial patterns of ZGA
- Maternal-zygotic genes dominate transcriptional output during ZGA
- Manipulation of translation and cell division reconciles regulatory mechanisms of ZGA
- Timing of germ-layer-specific expression appears sequential in the blastula

Chen & Good, 2022, Current Biology *32*, 4314–4324
October 10, 2022
https://doi.org/10.1016/j.cub.2022.07.078

CellPress

CellPress

Current Biology

Report

Nascent transcriptome reveals orchestration of zygotic genome activation in early embryogenesis

Hui Chen[1] and Matthew C. Good[1,2,3,*]
[1]Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]Lead contact
*Correspondence: mattgood@pennmedicine.upenn.edu
https://doi.org/10.1016/j.cub.2022.07.078

SUMMARY

Early embryo development requires maternal-to-zygotic transition, during which transcriptionally silent nuclei begin widespread gene expression during zygotic genome activation (ZGA).[1–3] ZGA is vital for early cell fating and germ-layer specification,[3,4] and ZGA timing is regulated by multiple mechanisms.[1–5] However, controversies remain about whether these mechanisms are interrelated and vary among species[6–10] and whether the timing of germ-layer-specific gene activation is temporally ordered.[11,12] In some embryonic models, widespread ZGA onset is spatiotemporally graded,[13,14] yet it is unclear whether the transcriptome follows this pattern. A major challenge in addressing these questions is to accurately measure the timing of each gene activation. Here, we metabolically label and identify the nascent transcriptome using 5-ethynyl uridine (5-EU) in *Xenopus* blastula embryos. We find that EU-RNA-seq outperforms total RNA-seq in detecting the ZGA transcriptome, which is dominated by transcription from maternal-zygotic genes, enabling improved ZGA timing determination. We uncover discrete spatiotemporal patterns for individual gene activation, a majority following a spatial pattern of ZGA that is correlated with a cell size gradient.[14] We further reveal that transcription necessitates a period of developmental progression and that ZGA can be precociously induced by cycloheximide, potentially through elongation of interphase. Finally, most ectodermal genes are activated earlier than endodermal genes, suggesting a temporal orchestration of germ-layer-specific genes, potentially linked to the spatially graded pattern of ZGA. Together, our study provides fundamental new insights into the composition and dynamics of the ZGA transcriptome, mechanisms regulating ZGA timing, and its role in the onset of early cell fating.

RESULTS AND DISCUSSION

Nascent EU-RNA-seq characterizes the composition and dynamics of ZGA with high sensitivity

Transcript levels in early embryos are dominated by maternal RNAs preloaded in the egg, whereas newly transcribed RNAs, including those from maternal-zygotic (MZ) and exclusively zygotic (Z) genes, constitute a small portion during zygotic genome activation (ZGA) (Figure 1A). Abundant maternal RNAs represent a major challenge in understanding the scope, timing, and underlying mechanisms regulating ZGA transcription. Recently, metabolic labeling of newly synthesized RNAs using uridine analogs, such as 4-thiol-UTP (4s-UTP)[15,16] and 5-ethynyl uridine (5-EU),[14,17,18] followed by physical separation of nascent and maternal RNA pools, has provided new tools to interrogate ZGA. We previously demonstrated that compared with conventional RNA sequencing (RNA-seq), EU-labeled nascent RNA-seq (EU-RNA-seq; Figure 1B) enriched biotinylated zygotic transcripts (EU-RNAs) from *Xenopus* late blastula on streptavidin beads,[14] suggesting its potential for detecting newly transcribed RNA. To characterize the nascent transcriptome in embryos at earlier stages, where transcript levels are much lower, we further optimized the EU-RNA-seq protocol (STAR Methods) and performed it on *Xenopus* embryos at 1-h intervals from 5 to 9 h post-fertilization (hpf), corresponding to Nieuwkoop and Faber (NF) stages 7–9, covering the period of pre-ZGA to widespread ZGA. In addition to sequencing the nascent transcripts (on "bead"), we sequenced the flowthrough after separation ("flowthrough," presumable maternal RNA) and the total RNA ("all") for comparison. We observed that from 5 to 9 hpf, nascent transcripts of an increasing number of genes enriched in the "bead" dataset compared with "flowthrough" (Figures 1C, S1A, and S1B), suggesting that EU-RNA-seq captures and enriches nascent transcripts. Although we were able to separate zygotic transcripts from maternal transcripts, we also noticed some maternal transcripts bound non-specifically to beads. Therefore, we chose to calculate nascent transcription based on the net increase of reads, treating those at 5 hpf as background. To characterize which genes were activated, we filtered the nascent transcriptome data (STAR Methods), generating a list of 2,577 genes (Figure S1C; Table S1).

Noticeably, at the time of ZGA widespread onset (7 hpf), over 44% genes were more highly detected using EU-RNA-seq than total RNA-seq (Figure 1D); an ~2- to 16-fold enrichment (Figure 1E), consistent with a previous observation.[14] More strikingly, activation of 240 genes was uniquely detected by

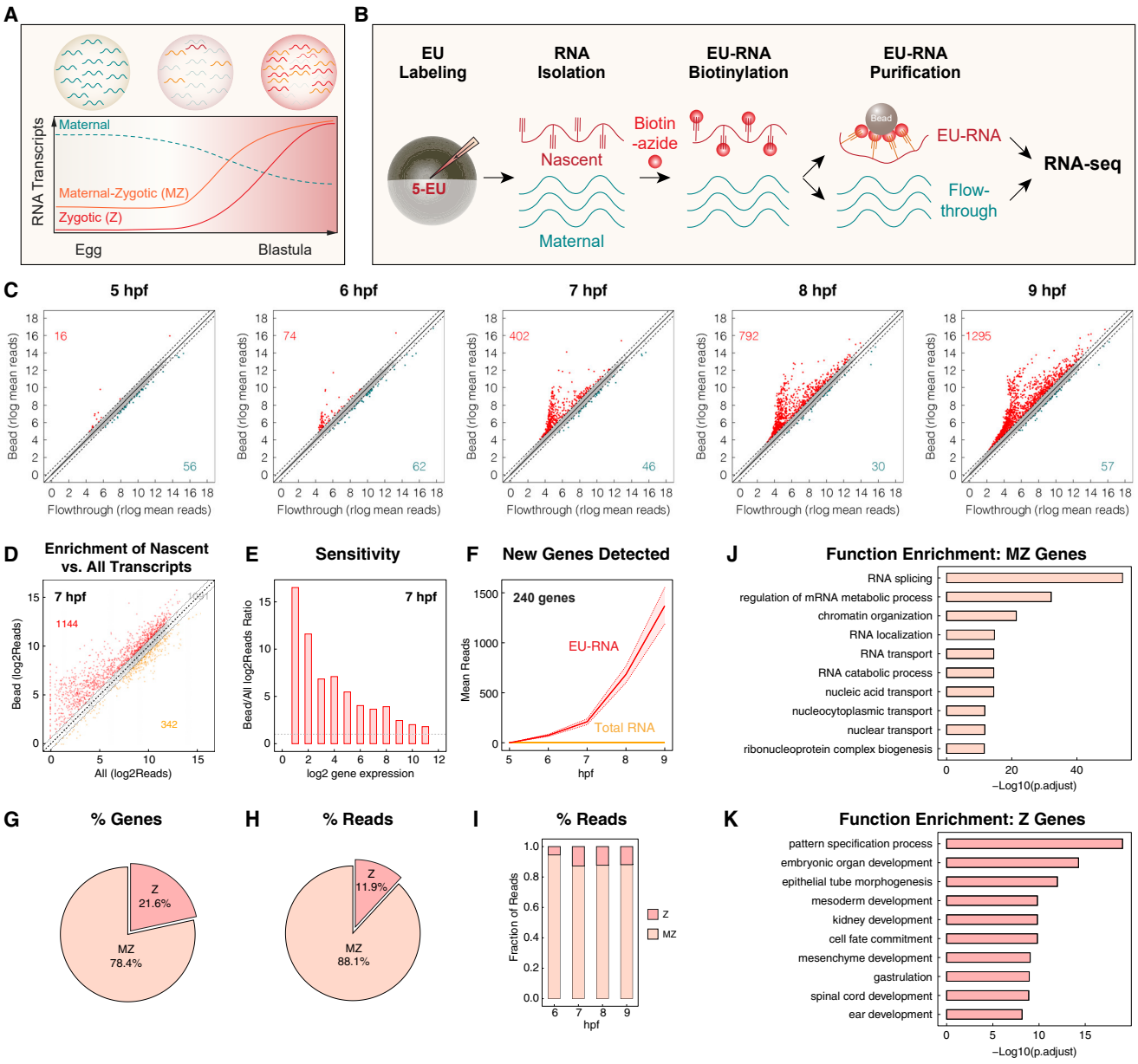


Figure 1. Nascent EU-RNA-seq to characterize the composition and dynamics of ZGA with high sensitivity

(A) Schematic of transcript composition during zygotic genome activation in early embryogenesis: from egg to late blastula. Red, transcripts of zygotic genes; orange, transcripts of maternal-zygotic genes; blue, transcripts of maternal genes.

(B) Schematic describing the EU-RNA-seq methodology. Nascent transcripts are metabolically labeled via 5-ethynyl uridine (5-EU) microinjected in 1-cell *Xenopus* embryos. Total RNAs are isolated for biotinylation via click reaction. The 5-EU-labeled nascent transcripts (red) are captured by streptavidin beads; flow-through contains maternal transcripts.

(C) Distinguishing nascent transcriptome ("bead") versus maternal transcriptome ("flowthrough") reads via RNA-seq from 5 to 9 hpf in blastula embryos. Each dot represents individual genes with rlog reads averaged from replicates quantified by DESeq2. Dashed lines: 1.5-fold threshold for enrichment.

(D) Nascent reads enrich in "bead" library versus traditional total transcriptome ("all") at 7 hpf from RNA-seq. Each dot represents individual genes with $\log_2$ reads averaged from replicates quantified by DESeq2. Dashed lines: 1.5-fold threshold for enrichment.

(E) Higher sensitivity for detection of zygotic expression in nascent transcriptome versus total transcriptome at all gene expression levels.

(F) Hundreds of transcripts are uniquely detected by nascent transcriptome. Mean reads from duplicates (mean ± SE) for 240 genes detected from nascent EU-RNA (red) and total RNA (orange), respectively.

(G) Percentage of genes expressed during ZGA that are classified as zygotic-only genes (Z) and maternal-zygotic gene (MZ).

(H and I) Percentage of total library reads from transcripts of zygotic-only genes (Z) and maternal-zygotic gene (MZ) from 5 to 9 hpf.

(J and K) Gene ontology (GO) analysis of MZ (J) and Z (K) genes.

See also Figure S1 and Tables S1 and S2.

EU-RNA-seq (Figure 1F; Table S2). Assay for transposase-accessible chromatin using sequencing (ATAC-seq) from early gastrula stage (NF stage 10) revealed that the promoter regions in most of these genes uniquely detected by EU-RNA-seq were accessible (Figures S1M and S1N), suggesting that these genes are likely actively transcribed. To further validate the sensitivity of EU-RNA-seq, we compared the levels of the most highly transcribed genes identified by Session et al.[19] and Yanai et al.[20] in our nascent versus total transcriptome data. Both comparisons showed an ~3- to 4-fold enrichment in the EU-RNA-seq (Figures S1D–S1F). This improved sensitivity also revealed earlier expression of a subset of genes in the zygotic transcriptome (Figure S1G). Together, the EU-RNA-seq enabled us to detect nascent transcription in early *Xenopus* embryos with unprecedented sensitivity and specificity.

Next, we wondered what fraction of the transcriptional output of large-scale ZGA is comprised of transcripts from MZ and Z genes. In traditional gene profiling, it is a challenge to characterize the activation of MZ genes—the presence of high maternal RNA levels can mask the onset of nascent transcription. Based on the presence of transcripts in the egg, we categorized the nascent list of 2,577 genes into Z genes that do not contain reads in the egg (≤5 reads) and MZ genes that contain transcripts in the egg (>5 reads) (Figure S1C). We found that among the genes activated from 5 to 9 hpf, MZ genes accounted for ~70% total reads and Z genes only ~10%, and both increased over time (Figures S1H–S1L). The ratio of MZ:Z genes was ~4:1, and the ratio of reads for MZ:Z genes was ~8:1, which was relatively constant over time (Figures 1G–1I). These data suggest that compositionally, ZGA transcriptional output is dominated by MZ gene expression, consistent with previous observations,[21] although zygotic-only transcripts are essential for development.[22–24] Gene ontology (GO) analysis revealed that MZ genes are mainly involved in RNA processing, splicing, and transport (Figure 1J), whereas Z genes are responsible for patterning, gastrulation, cell fate commitment, and germ-layer specification (Figure 1K). Thus, onset of widespread ZGA is dominated by a handoff from maternal-to-zygotic control of core regulatory genes, whereas zygotic-only factors that pattern later development represent a smaller portion of the transcriptional output. In summary, the highly sensitive EU-RNA-seq methodology reveals a greater depth to the composition and dynamics of ZGA, offering essential new insights on genome regulation in early embryo development.

EU-RNA-seq on segmented embryos uncovers spatial patterns of single-gene activation

We previously performed nascent EU-RNA imaging in single cells of whole-mount *Xenopus* early embryos and observed a stereotypical spatiotemporal pattern of large-scale ZGA, which initiates first in small cells at the animal pole (AP) and is delayed in large cells at the vegetal pole (VP), dependent on cells reaching a threshold size[13,14] (Figure 2A). However, an open question was whether the global pattern of ZGA held at the single-gene level for most genes. To address this question, we performed EU-RNA-seq on the dissected regions of AP and VP from embryos at 5–9 hpf (Figure 2A). We chose these regions because they display a striking phase shift in ZGA timing of ~90 min,[14] and they represent the presumptive ectoderm (the AP) and endoderm (the VP). Due to the lower quantity of material, we refined the AP-VP nascent transcriptome by selecting genes that undergo consistent continuous activation in the AP or VP regions at 5–9 hpf, which resulted in a list of 882 genes for further characterization of their spatial activation patterns (Figure S2A). To determine the onset time for each gene at the AP and the VP, respectively, we adapted a previously described method to fit the mean-normalized reads at 5–9 hpf with a smooth spline[25] (Figure S2B). By setting thresholds of expression (STAR Methods), we eventually classified 476 genes into five categories of activation patterns based on their spatial expression profiles (Figures 2B and S2C; Table S3). We verified the activation patterns of a subset of transcripts using RT-PCR and quantitative real-time PCR (Figures S2D and S2E). The existence of various patterns of gene activation (Figures S2F–S2I) is consistent with the view that the embryo is patterned during early embryogenesis,[23] and potentially distinct mechanisms regulate ZGA. As expected, the "VP regional" category is replete with genes involved in endoderm-related development (Figure S2G), including the bix1, mixer, sox17, and nodal family genes (Table S3), consistent with the VP as the physical location of the presumptive endoderm cells.[26,27] Genes uniquely activated or activated early in the VP may be regulated by region-specific maternal determinants, such as VegT,[28–30] that prepattern early embryos.

Intriguingly, among the five spatiotemporal patterns, the predominant category is the "AP early, VP delay" (~56% genes) (Figures 2B, 2C, and S2C), paralleling the global spatiotemporal patterning of ZGA from EU imaging, suggesting that these genes could be regulated by a cell-size- or DNA:cytoplasm-ratio-dependent mechanism.[14] GO analysis of this category revealed functional enrichment of ectoderm-related development such as epithelial tube morphogenesis and eye development (Figure S2F), consistent with the role of the AP as the physical location of the presumptive ectoderm. Separately, ~13% of genes show an activation pattern similar in space and time (Figures 2B, 2C, S2C, and S2I), potentially consistent with a timer mechanism. The observation of diverse patterns is also consistent with previous studies suggesting that ZGA includes expression of distinct subsets of genes regulated by distinct mechanisms.[6,7]

In summary, our regional nascent transcriptome analysis revealed patterning that is consistent with a classic developmental control view but also previously underappreciated major spatial and temporal pattern of gene activation, consistent with cell-size- or DNA:cytoplasm-ratio-dependent ZGA regulation. The results reveal that distinct modes of regulation likely control distinct subsets of gene expression patterns at the canonical onset of widespread ZGA and that more than half follow a pattern tied to the cell size gradient in the *Xenopus* blastula. Future studies are warranted to dissect the respective mechanisms underlying these distinct expression patterns.

Reconciling mechanisms that regulate ZGA timing

Multiple mechanisms have been suggested to regulate ZGA timing in various species,[1–5] including a timer,[6,17,21,31,32] DNA:cytoplasm ratio,[7,14,25,33–38] and cell-cycle elongation.[8–10,15,18,39–41] However, these mechanisms may work in concert or be interrelated, and their relative contribution to ZGA onset timing in embryogenesis is debated.[6–10] We reasoned that manipulating

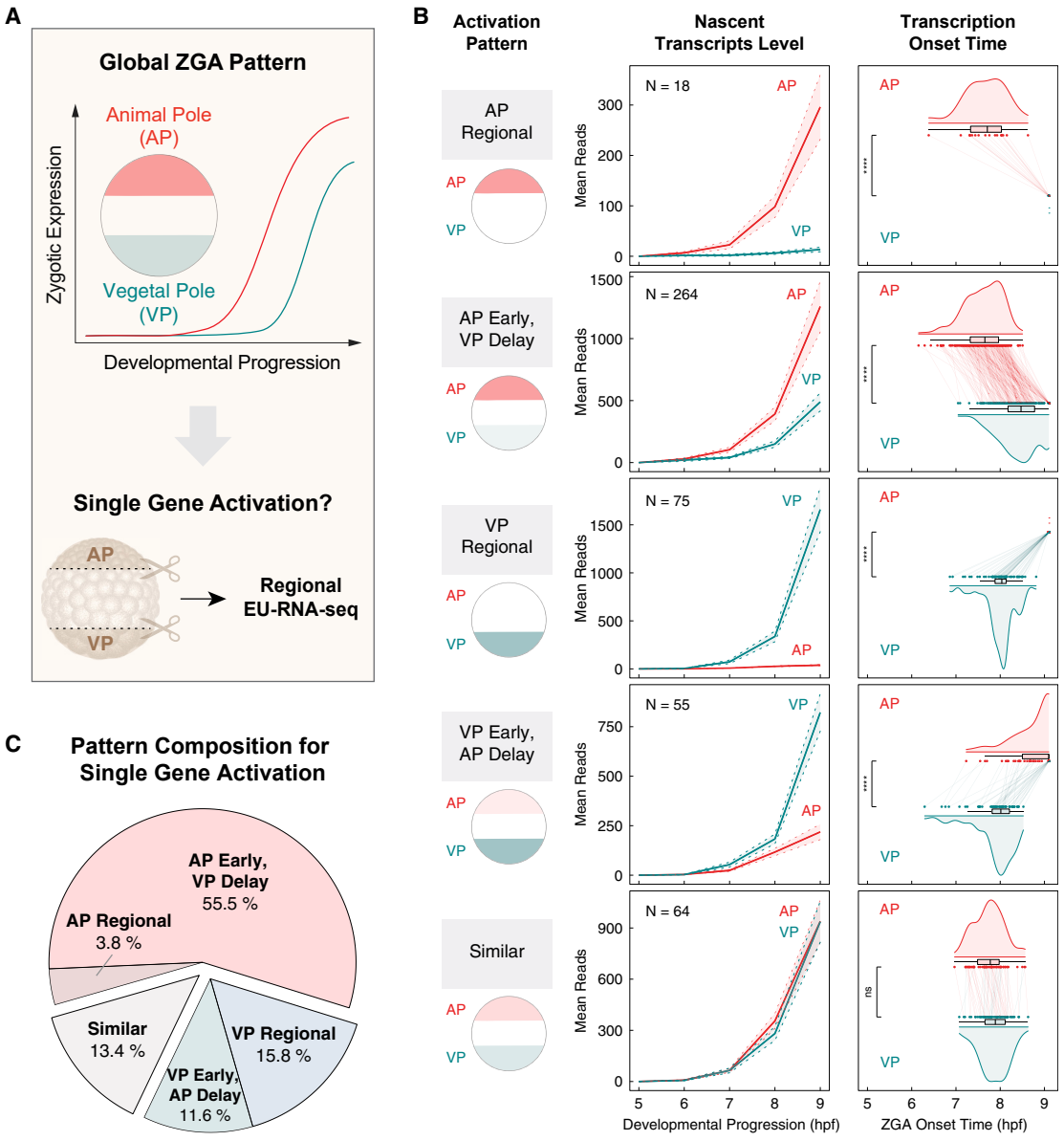


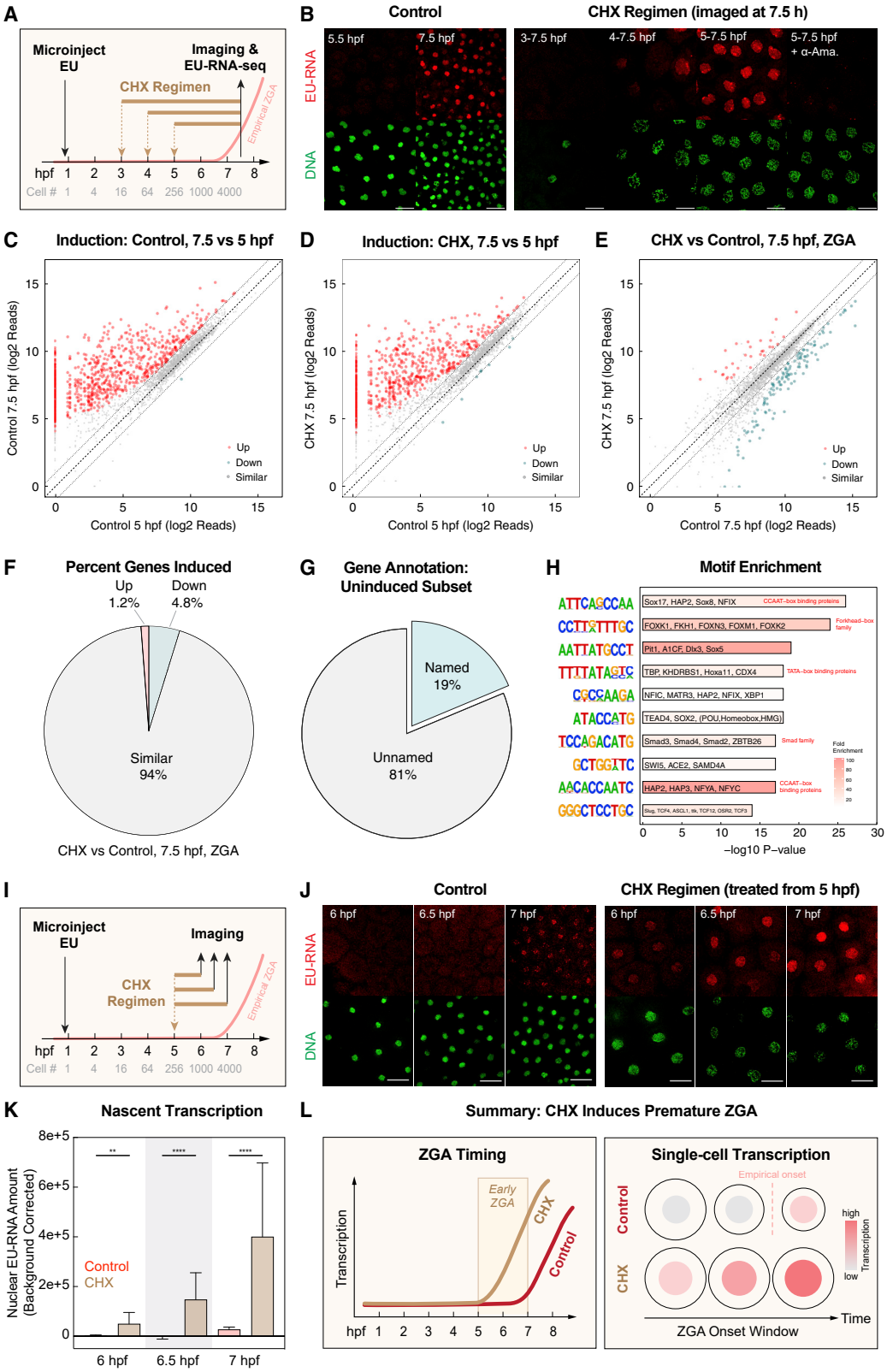
Figure 2. EU-RNA-seq on segmented embryos uncovers spatial patterns of single-gene activation

(A) Schematic showing spatial patterns of ZGA observed from EU-RNA imaging, and strategy for classifying gene expression profiles from segmented embryos using regional EU-RNA-seq. AP, animal pole; VP, vegetal pole.

(B) Categorization of distinct spatial patterns of activation for zygotic genes. Left: schematic of each pattern for gene activation. Middle: reads from EU-RNA-seq performed on the segmented AP (red) and VP (blue) regions from early embryo at 5–9 hpf. Transcript levels were thresholded to those at 5 hpf. Mean ± SE from indicated number of genes. Right: comparison of the activation onset time for transcription of each gene for AP (red) and VP (blue). ****$p < 0.0001$; ns, not significant. Note that to compare and plot the activation onset time for those genes that were not activated during 5–9 hpf at AP or VP, their activation onset time was set to 9.1.

(C) Percentage of zygotic genes that have distinct spatial patterns of expression during ZGA. A majority of genes (55%) show AP early and VP delayed expression. See also Figure S2 and Table S3.

the regulatory parameters at various phases of development and measuring nascent transcription would allow us to distinguish the contributions of distinct mechanisms. To this end, we treated EU-injected embryos at various pre-ZGA stages with cycloheximide (CHX), an inhibitor of translation that has been widely used for arresting *Xenopus* embryos in interphase by administration to the culture medium,[42–47] and analyzed its effect on nascent transcription by EU-RNA imaging (Figure 3A). Interestingly, we observed that CHX impacts nascent transcription in a stage-dependent manner—no transcription for 3–7.5 hpf arrest, modest transcription for 4–7.5 hpf arrest, and abundant transcription for 5–7.5 hpf arrest (Figure 3B). This result suggests that a developmental window is essential for the embryo gaining transcriptional competence before ZGA, consistent with a recent finding in zebrafish,[42] potentially by accumulating maternal translation of transcription activators such as the



(legend on next page)

pioneering pluripotency factors[21,32] and impacting chromatin remodeling.[17,48] Notably, the highly transcribing embryos arrested from 5 to 7.5 hpf only contain ~500 cells (versus ~5,000 cells in control embryos at 7.5 hpf; 3–4 cell cycles behind the control) in which DNA synthesis is arrested[42] (Figure S3A), and the DNA:-cytoplasm ratio is far below the threshold for ZGA onset.[14] This suggests that once the embryo gains transcriptional competence, prolonged arrest in interphase may enable nascent transcript accumulation, although the possibility of cell-cycle-independent activities, such as translational inhibition of potential transcriptional inhibitors, cannot be excluded.

We next wondered whether nascent transcription in CHX-arrested embryos represents bona fide ZGA. To test this, first, we co-microinjected EU with α-amanitin, an inhibitor of RNA polymerase II (RNAPII), and found that a majority of EU-RNA signal was abolished, suggesting RNAPII-dependent transcription in CHX-arrested embryos (Figure 3B). We then performed EU-RNA-seq to compare the nascent transcriptome between 7.5 hpf control and the CHX-arrested embryos from 5 to 7.5 hpf. Surprisingly, 94% of the genes were similarly transcribed, and only 4.8% and 1.2% of the genes were downregulated and upregulated by CHX, respectively (Figures 3C–3F and S3B–S3D). A majority of the CHX-downregulated genes are unnamed and unannotated (Figures 3G and S3E–S3G), although they seem to be involved developmental regulation (Figure 3H); in contrast, the majority of the CHX-upregulated genes are named and annotated (Figure S3F) and are enriched in the ectoderm-related development (Figures S4E and S4F). These data suggest that CHX arrest induces nearly full ZGA, despite its differential impact on subsets of the genome. Notably, this comparable level of transcription at 7.5 hpf was reached by the CHX-arrested embryo (~500 cells) that contains ~10 times fewer nuclei and DNA template than the control embryo (~5,000 cells), suggesting higher and possibly earlier transcriptional output per nucleus in CHX-arrested embryos.

To determine whether CHX treatment induces early ZGA, we treated embryos with CHX starting from 5 hpf, at which embryos had gained transcriptional competence, and examined nascent transcription at time points before the canonical onset of widespread ZGA (Figure 3I). Remarkably, nascent transcription readily occurred upon CHX treatment from 5 to 6 or 5 to 6.5 hpf, when transcription is not detectable in control embryos even though they contain many fewer nuclei due to the arrest (Figures 3J, 3K, and S3H). Most strikingly, upon CHX treatment for 5 to 7 hpf transcriptional output per nucleus is 14.8-fold higher (Figure 3K), in embryos ~3 divisions behind control embryos (inferred from cell volume, Figure S3H). Furthermore, individual zygotic genes could be detected earlier and higher

Figure 3. Reconciling mechanisms that regulate ZGA timing

(A) Experimental design to probe effect of cycloheximide (CHX) treatment on ZGA regulation. Embryos were microinjected with EU at 1-cell stage; treated with CHX at 3, 4, and 5 hpf, respectively; and analyzed for nascent transcription by confocal imaging or EU-RNA-seq at 7.5 hpf. The red line indicates empirical ZGA timing in normal embryos. Estimated cell numbers at each hpf are shown for normal embryos.

(B) Imaging nascent transcript (EU-RNA) accumulation in blastomere nuclei of whole-mount embryos. Maximum-intensity Z-projections of several slices from individual embryo stacks are shown. Left: representative confocal imaging of control embryos, prior to ZGA (5.5 hpf) and at the canonical onset of widespread ZGA (7.5 hpf). Right: representative confocal imaging of embryos treated with CHX from 3 to 7.5, 4 to 7.5, and 5 to 7.5 hpf, respectively. DNA channel illustrates number and size of nuclei. Nuclei in CHX-treated embryos are large and round indicating prolonged arrest in interphase. Nascent transcription is absent from embryos arrested from 3 to 7.5 hpf. Note: nuclear morphology in 3–7.5-hpf-arrested embryos is more heterogenous, but no EU-RNA was observed in nuclei from multiple repeats of the experiment. Nuclei in blastomeres arrested from 5 to 7.5 hpf show strong levels of nascent transcription, similar to and exceeding control 7.5 hpf embryos. Signal is RNAPII dependent; inhibited by α-amanitin treatment (bottom). Scale bars, 50 μm.

(C–E) Nascent transcriptome reads for control (wild type) and embryos CHX-arrested from 5 to 7.5 hpf. Each dot represents individual genes from the selected 2,577 genes with $\log_2$ reads averaged from duplicates quantified by DESeq2. The differentially enriched genes in each group are labeled in red and blue, respectively, based on a 2-fold threshold (dashed lines) and adjusted $p < 0.05$. Accumulation of nascent reads from 5 to 7.5 hpf in control embryos (C) and from 5 to 7.5 hpf in embryos arrested in CHX from 5 to 7.5 hpf (D). (E) Comparison of reads during widespread ZGA (7.5 hpf) for control embryos versus embryos CHX-arrested from 5 to 7.5 hpf.

(F) Embryonic arrest in interphase via CHX treatment from 5 to 7.5 hpf leads to widespread induction of zygotic genes normally induced during ZGA. Percentage of 2,577 genes in nascent 7.5 hpf transcriptome of control and that in embryos subject to 5–7.5 hpf CHX treatment that are upregulated (red), downregulated genes (blue) and similarly expressed (gray); calculated from (E).

(G) Gene annotation for subset of the transcriptome that are not induced by arrest with CHX treatment from 5 to 7.5 hpf. Over 80% do not have assigned gene names in the *Xenopus* genome.

(H) Motifs strongly enriched in the promoter regions of genes that do not induce in response to CHX arrest and prolonged interphase.

(I) Experimental design to test whether CHX treatment induces premature ZGA. Embryos were microinjected with EU at 1-cell stage, treated with CHX from 5 hpf, and imaged for nascent transcription at 6, 6.5, and 7 hpf, respectively. The red line indicates ZGA timing in normal embryos at 22°C. Estimated cell numbers at each hpf are shown for control embryos.

(J) Imaging of nascent transcription (EU-RNA) in nuclei of blastomeres on the animal pole of whole-mount control and CHX-arrested embryos. Maximum intensity Z-projections from individual embryo substacks are shown. Observation of premature ZGA—nascent transcription—in CHX-arrested embryos treated from 5 to 6, 5 to 6.5, and 5 to 7 hpf. Nascent transcription is largely absent in control embryos at 6 and 6.5 hpf and beginning widespread accumulation in control embryos at 7 hpf. Scale bars, 50 μm.

(K) Quantification of nascent transcription in control and CHX-arrested embryos in (J). The nuclear EU-RNA amount in blastomere nuclei was calculated by multiplying the integrated nuclear EU-RNA signal after cytoplasmic background subtraction by the nucleus volume. A total of 80–160 cells from at least three embryos were analyzed for each group. Data are represented as mean + SD, and statistical significance was determined by one-way ANOVA (Fisher's LSD test). $^{**}p < 0.01$; $^{****}p < 0.0001$.

(L) Schematic summary of precocious ZGA initiated by CHX treatment. Left: CHX induces early ZGA from 5 hpf after embryos have gained transcriptional competence. Right: CHX treatment promotes nascent transcript accumulation to a higher level and at earlier times in individual nuclei. The color gradient indicates no or low (gray) to high (red) transcription in the nucleus. The red dashed vertical line indicates empirical widespread ZGA onset timing for control embryos.

See also Figure S3.

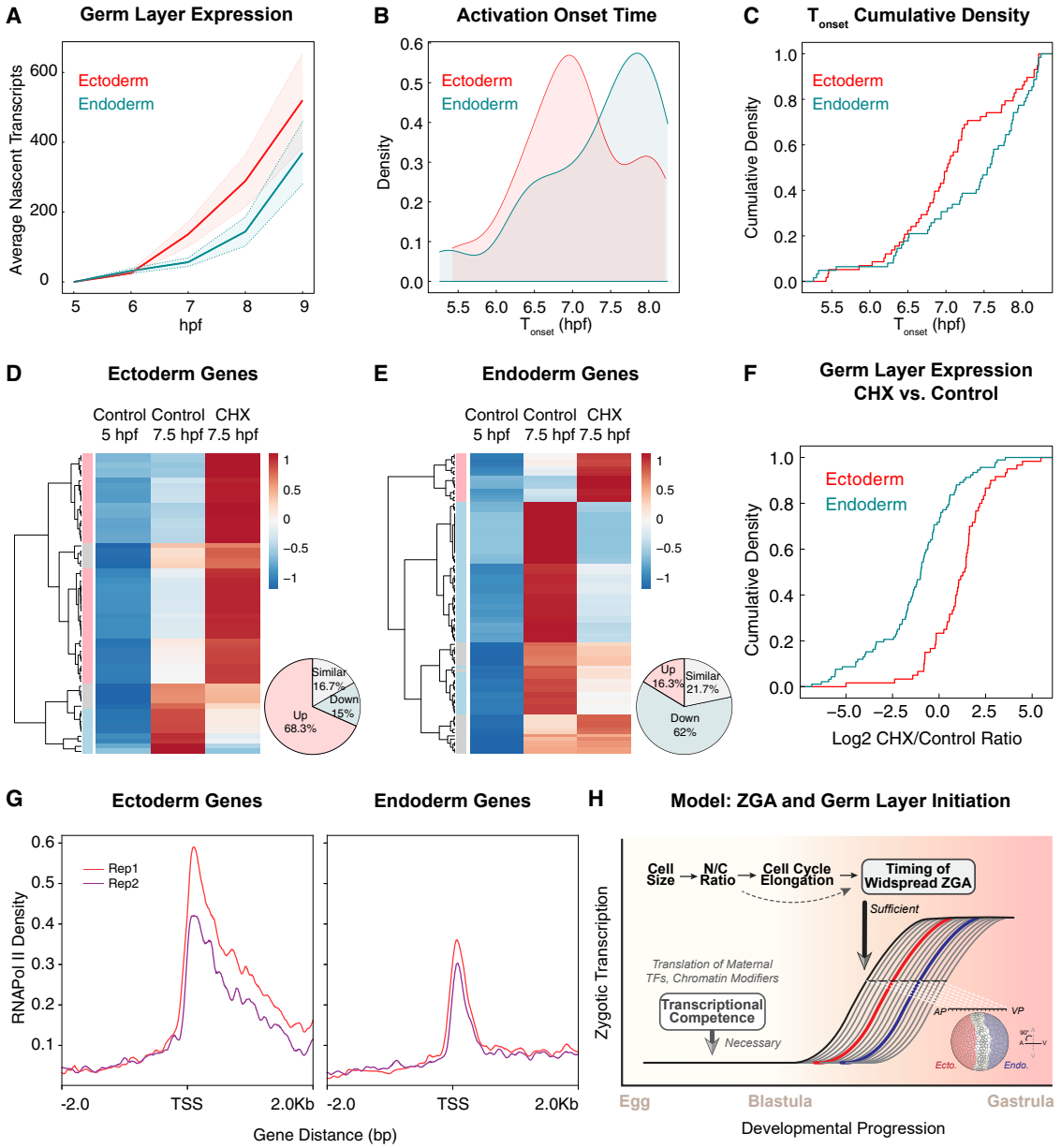


Figure 4. Timing of germ-layer initiation is correlated to spatially graded onset of ZGA

(A) Average transcript levels for ectoderm and endoderm genes in blastula embryos. Nascent reads averaged from replicate embryos. Red, set of ectoderm genes (N = 111 genes); blue, set of endoderm genes (N = 172 genes) at 5–9 hpf. Mean ± SE.

(B) Distribution of time of onset for transcriptional activation of individual genes within ectoderm (red) and endoderm (blue) sets.

(C) Cumulative density of time of onset for transcriptional activation for ectoderm (red) and endoderm (blue) genes.

(D and E) Heatmaps showing *Z* scores for induction of ectoderm (D) and endoderm (E) genes in control embryos at 7.5 hpf versus embryos treated with CHX from 5 to 7.5 hpf. Data are selected genes that have detectable transcriptional induction in control embryos from 5 to 7.5 hpf. Pie chart shows fraction of genes similar (gray) or upregulated (red) or downregulated (blue), comparing CHX-treated versus control embryos; threshold 1.5-fold difference.

(F) Cumulative density showing induction relative to control of germ-layer genes in CHX-treated embryos at 7.5 hpf. Red, ectoderm; blue, endoderm.

(G) Profile plots for ChIP-seq peaks of RNA Pol II for ectoderm and endoderm genes, respectively, in embryos at stage 10.5. Data from Session et al.[19] The regions of transcription start site (TSS) ±2kb from two replicates are shown. The level of RNA Pol II binding in the indicated regions is higher in ectoderm genes than in endoderm genes.

(H) Composite model and pathway describing regulated ZGA onset. Widespread ZGA is inhibited by rapid cell cycles in cleavage-stage embryos. As cell size reduces, due to cell division without cell growth, a DNA:cytoplasm ratio threshold is reached, promoting cell-cycle elongation and allowing for accumulation of nascent zygotic transcripts. Cell-cycle elongation is sufficient to promote large-scale ZGA in embryos that have achieved transcriptional competence. Translation of maternal TFs and histone acetyltransferases are necessary to generate transcriptionally competent embryos. *Xenopus* blastula embryos contain a gradient of cell sizes which achieve widespread ZGA onset at different times. A majority of the nascent transcriptome during ZGA displays early activation in smaller cells at

(legend continued on next page)

levels in embryos arrested with CHX starting at 5 hpf compared with control embryos by RT-PCR (Figure S3I). Together, these data suggest that CHX induces precocious ZGA onset by increasing transcriptional output in individual nuclei (Figure 3L). These findings agree with an interpretation that short cell cycle can repress nascent transcription and cell-cycle elongation can promote large-scale ZGA after an embryo gains transcriptional competence, although it cannot be excluded that separate translational inhibition by CHX may also contribute to transcription.[10,42] Noticeably, our finding is consistent with recent observations made in other embryonic systems, including zebrafish[17] and *Drosophila*,[7,8] that cell-cycle arrest increases zygotic transcription. Our study is limited by the inability of using similar regimens of specific cell-cycle inhibitors such as Cdk inhibitors which did not rapidly block embryo cleavages (Figure S3J).

Timing of germ-layer initiation is linked to spatially graded onset of ZGA

Many germ-layer-specific genes are transcribed in the blastula embryo during ZGA. In *Xenopus*, we previously found that cells of the AP—the presumptive ectoderm—initiate large-scale ZGA ~90 min earlier than cells of the VP—the presumptive endoderm,[14] which is linked to the AP cells reaching a threshold cell size for ZGA more quickly. We wondered whether the spatiotemporal patterning of ZGA might contribute to a temporal ordering of germ-layer-specific gene activation, the chronology of which is debated in different model embryos.[11,12] To determine whether a temporal order of germ-layer-specific gene activation exists in *Xenopus* embryos, we focused on a list of marker genes that had been experimentally defined and validated in gastrula embryos for the ectoderm[26] and the endoderm,[27] respectively. We discovered that on average, the ectodermal genes are more highly transcribed, and a majority are activated earlier, than the endodermal genes from 5 to 9 hpf (Figures 4A–4C and S4A–S4D). Chromatin immunoprecipitation sequencing (ChIP-seq) analysis for RNAPII and H3K4me3,[19] marks for active transcription, revealed higher transcriptional activity in the ectodermal genes than the endodermal genes that persist in early gastrula (Figures 4G and S4I). However, it should be noted that several endodermal genes are activated early (Figure 4B), consistent with previous findings that these genes are transcribed early in development, potentially regulated by the maternal T-box factor VegT.[49,50] These data, together with those from the AP/VP spatial patterns[14] (Figures 2 and S2), suggest that the timing of germ-layer initiation is largely correlated with the regional timing of ZGA onset and that distinct mechanisms may regulate the germ-layer-specific activation at the single-gene level.

Analysis of CHX-upregulated genes revealed enrichment in ectoderm-related development (Figures S4E and S4F). We wondered whether the ectodermal genes could be upregulated by CHX treatment from 5 to 7.5 hpf. Most of the ectodermal genes (68.3%) were hyperactivated, and only 15% genes were downregulated in CHX-arrested embryos (Figures 4D and S4A); in contrast, most of the endodermal genes (62%) were downregulated, and only 16.3% genes were upregulated in CHX-arrested embryos (Figures 4E and S4B). The striking difference in the impact of CHX on expression germ-layer genes (Figures 4D–4F and S4A–S4D) suggests that the ectoderm genes might be primed for activation and thus more susceptible to transcription via cell-cycle arrest. We did not observe a correlation between CHX-induced expression and time of onset (Figures S4G and S4H). In summary, our data suggest that the timing of germ-layer-specific gene activation may be linked to the timing of ZGA in different regions of the embryo. However, our study demonstrates only a correlative relationship between spatially patterned ZGA and germ-layer-specific gene activation, and future studies are required to directly probe the causal link between these two, ideally by manipulating cell size and ZGA onset gradient and measuring the regional timing of transcriptional initiation for germ-layer markers.

Composite model for ZGA

The combination of nascent imaging and transcriptome profiling coupled to embryo arrest at different times provided tools to link and parse regulatory mechanisms controlling ZGA timing. Our data are consistent with a model in which ZGA timing is regulated by the cell-cycle elongation once an embryo gains transcriptional competence and cells achieve a threshold size or DNA:cytoplasm ratio (Figures 4H, S4J, and S4K). In multiple species, including *Drosophila* and zebrafish, early rapid cell cycles block zygotic transcription, leading to abortive or short transcripts.[15,18] We interpret the inhibitory effect of short cell cycles to explain why a normal *Xenopus* embryo does not initiate widespread zygotic transcription at 5 hpf, even though it is likely transcriptional competent, based on our CHX arrest data (from 5 to 7.5 hpf). Importantly, artificial or natural lengthening of cell cycle promotes zygotic transcription.[7,8,17] For vertebrates, maturation and lengthening of the early cell cycle are linked to cells reaching a threshold size,[38,51] and the DNA:cytoplasm ratio and histone levels may also regulate the timing of the cell-cycle lengthening.[7,10,33,52,53] This logic helps explain how cells arrested at too low of DNA:cytoplasm ratio, in the CHX-arrested embryos from 5 to 7.5 hpf, nonetheless initiate ZGA concomitant with a longer time spent arrested in interphase. Linking these concepts, we propose a composite model in which *Xenopus* embryos must first achieve transcriptional competence, a necessary step, and then wait to initiate widespread ZGA until cell-cycle elongation occurs, coupled to when blastomeres achieve a threshold cell size and DNA:cytoplasm ratio.

Overall, our work demonstrates that EU-RNA-seq is a highly sensitive method to characterize ZGA and determine the dynamics of transcription, useful for dissecting regulatory mechanisms underlying genome activation. Using this strategy, we identify distinct spatiotemporal gene expression patterns from segmented embryos, suggesting multiple modes of ZGA

the animal pole (AP)— the presumptive ectoderm—and later activation in larger cells at the vegetal pole (VP)—the presumptive endoderm. Chronological sequence of germ-layer induction may be linked to temporally graded onset of widespread ZGA in the blastula. Red, ectoderm genes; blue, endoderm genes. For clarity, the schematic of embryo is shown rotated 90° counterclockwise along the AV axis from the normal orientation of the embryo. Dotted arrow indicates that N/C ratio may affect ZGA transcription in a cell-cycle-independent manner.
See also Figure S4.

regulation in *Xenopus*, and unveil a potential link between ZGA patterning and germ-layer initiation in early development. The methodology is applicable to other embryonic systems, compatible with other high-throughput multi-omics technologies at the single-cell level, which will catalyze new insights into genome regulation and cell fating in development.

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
- METHOD DETAILS
  - *In vitro* fertilization (IVF)
  - Microinjection, cycloheximide treatment and collection of embryos
  - Cell cycle inhibitor incubation
  - RNA isolation, biotinylation and purification
  - EU-RNA-seq and analysis
  - Functional enrichment analysis
  - ATAC-seq and ChIP-seq analysis
  - RT-PCR and real-time PCR
  - Confocal imaging nascent transcripts in wholemount embryos ad image analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.07.078.

ACKNOWLEDGMENTS

We would like to thank the members of the Good, Klein, Mullins, Zaret, Berger, Bonasio, Lampson, Black, Greenberg, and Grishchuk labs at the University of Pennsylvania for helpful discussion and providing feedback. We particularly thank Dr. Peter Klein, Lily Einstein, Wenchao Qian, Boao Xia, Jorge Dabdoub, Haidar Ahmed, and Rachel Wells for providing help on frogs; Dr. Katherine Palozola (Ken Zaret Lab) for providing and discussing protocols for EU-RNA-seq; Dr. Jamie Kwasnieski (David Bartel Lab) for providing an alternative protocol for biotinylating EU-RNA; Dr. Dario Nicetto (Ken Zaret Lab) and Dr. Lihong Sheng (Roberto Bonasio Lab) for providing technical help on RNA-seq; and Dr. John Tobias (Penn Genomic Analysis Core) for providing help on RNA-seq data analyses. We also thank the Epigenetics Institute for training and providing instruments, the Cell and Developmental Biology Microscopy Core for imaging support, and National *Xenopus* Resource (NXR) for training and guidance on frogs. This work was supported in part by Burroughs Wellcome Fund, Charles E. Kaufman Foundation, the March of Dimes, and the National Institute of General Medical Sciences (R35GM128748) (M.C.G.) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R03HD105802) (H.C.).

AUTHOR CONTRIBUTIONS

Conceptualization, H.C. and M.C.G.; methodology, H.C. and M.C.G.; resources, H.C. and M.C.G.; investigation, H.C.; software, H.C.; formal analysis, H.C.; writing – original draft, H.C.; writing – review & editing, H.C. and M.C.G.; visualization, H.C. and M.C.G.; supervision, M.C.G.; funding acquisition, M.C.G.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 1, 2022
Revised: April 25, 2022
Accepted: July 29, 2022
Published: August 24, 2022

REFERENCES

1. Vastenhouw, N.L., Cao, W.X., and Lipshitz, H.D. (2019). The maternal-to-zygotic transition revisited. Development *146*, dev161471. https://doi.org/10.1242/dev.161471.
2. Jukam, D., Shariati, S.A.M., and Skotheim, J.M. (2017). Zygotic genome activation in vertebrates. Dev. Cell *42*, 316–332. https://doi.org/10.1016/j.devcel.2017.07.026.
3. Lee, M.T., Bonneau, A.R., and Giraldez, A.J. (2014). Zygotic genome activation during the maternal-to-zygotic transition. Annu. Rev. Cell Dev. Biol. *30*, 581–613. https://doi.org/10.1146/annurev-cellbio-100913-013027.
4. Schulz, K.N., and Harrison, M.M. (2019). Mechanisms regulating zygotic genome activation. Nat. Rev. Genet. *20*, 221–234. https://doi.org/10.1038/s41576-018-0087-x.
5. Pálfy, M., Joseph, S.R., and Vastenhouw, N.L. (2017). The timing of zygotic genome activation. Curr. Opin. Genet. Dev. *43*, 53–60. https://doi.org/10.1016/j.gde.2016.12.001.
6. Lu, X., Li, J.M., Elemento, O., Tavazoie, S., and Wieschaus, E.F. (2009). Coupling of zygotic transcription to mitotic control at the Drosophila mid-blastula transition. Development *136*, 2101–2110. https://doi.org/10.1242/dev.034421.
7. Syed, S., Wilky, H., Raimundo, J., Lim, B., and Amodeo, A.A. (2021). The nuclear to cytoplasmic ratio directly regulates zygotic transcription in Drosophila through multiple modalities. Proc. Natl. Acad. Sci. USA *118*. e2010210118. https://doi.org/10.1073/pnas.2010210118.
8. Strong, I.J.T., Lei, X., Chen, F., Yuan, K., and O'Farrell, P.H. (2020). Interphase-arrested Drosophila embryos activate zygotic gene expression and initiate mid-blastula transition events at a low nuclear-cytoplasmic ratio. PLoS Biol. *18*. e3000891. https://doi.org/10.1371/journal.pbio.3000891.
9. Edgar, B.A., Kiehle, C.P., and Schubiger, G. (1986). Cell cycle control by the nucleo-cytoplasmic ratio in early Drosophila development. Cell *44*, 365–372.
10. Collart, C., Allen, G.E., Bradshaw, C.R., Smith, J.C., and Zegerman, P. (2013). Titration of four replication factors is essential for the Xenopus laevis midblastula transition. Science *341*, 893–896. https://doi.org/10.1126/science.1241530.
11. Argelaguet, R., Clark, S.J., Mohammed, H., Stapel, L.C., Krueger, C., Kapourani, C.A., Imaz-Rosshandler, I., Lohoff, T., Xiang, Y., Hanna, C.W., et al. (2019). Multi-omics profiling of mouse gastrulation at single-cell resolution. Nature *576*, 487–491. https://doi.org/10.1038/s41586-019-1825-8.
12. Hashimshony, T., Feder, M., Levin, M., Hall, B.K., and Yanai, I. (2015). Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature *519*, 219–222. https://doi.org/10.1038/nature13996.
13. Chen, H., and Good, M.C. (2020). Imaging nascent transcription in wholemount vertebrate embryos to characterize zygotic genome activation. Methods Enzymol. *638*, 139–165. https://doi.org/10.1016/bs.mie.2020.03.002.
14. Chen, H., Einstein, L.C., Little, S.C., and Good, M.C. (2019). Spatiotemporal patterning of zygotic genome activation in a model

vertebrate embryo. Dev. Cell *49*, 852–866.e7. https://doi.org/10.1016/j.devcel.2019.05.036.

15. Heyn, P., Kircher, M., Dahl, A., Kelso, J., Tomancak, P., Kalinka, A.T., and Neugebauer, K.M. (2014). The earliest transcribed zygotic genes are short, newly evolved, and different across species. Cell Rep. *6*, 285–292. https://doi.org/10.1016/j.celrep.2013.12.030.

16. Holler, K., Neuschulz, A., Drewe-Boß, P., Mintcheva, J., Spanjaard, B., Arsiè, R., Ohler, U., Landthaler, M., and Junker, J.P. (2021). Spatio-temporal mRNA tracking in the early zebrafish embryo. Nat. Commun. *12*, 3358. https://doi.org/10.1038/s41467-021-23834-1.

17. Chan, S.H., Tang, Y., Miao, L., Darwich-Codore, H., Vejnar, C.E., Beaudoin, J.D., Musaev, D., Fernandez, J.P., Benitez, M.D.J., Bazzini, A.A., et al. (2019). Brd4 and P300 confer transcriptional competency during zygotic genome activation. Dev. Cell *49*, 867–881.e8. https://doi.org/10.1016/j.devcel.2019.05.037.

18. Kwasnieski, J.C., Orr-Weaver, T.L., and Bartel, D.P. (2019). Early genome activation in Drosophila is extensive with an initial tendency for aborted transcripts and retained introns. Genome Res. *29*, 1188–1197. https://doi.org/10.1101/gr.242164.118.

19. Session, A.M., Uno, Y., Kwon, T., Chapman, J.A., Toyoda, A., Takahashi, S., Fukui, A., Hikosaka, A., Suzuki, A., Kondo, M., et al. (2016). Genome evolution in the allotetraploid frog *Xenopus laevis*. Nature *538*, 336–343. https://doi.org/10.1038/nature19840.

20. Yanai, I., Peshkin, L., Jorgensen, P., and Kirschner, M.W. (2011). Mapping gene expression in two Xenopus species: evolutionary constraints and developmental flexibility. Dev. Cell *20*, 483–496. https://doi.org/10.1016/j.devcel.2011.03.015.

21. Lee, M.T., Bonneau, A.R., Takacs, C.M., Bazzini, A.A., DiVito, K.R., Fleming, E.S., and Giraldez, A.J. (2013). Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature *503*, 360–364. https://doi.org/10.1038/nature12632.

22. Zernicka-Goetz, M., Morris, S.A., and Bruce, A.W. (2009). Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. Nat. Rev. Genet. *10*, 467–477. https://doi.org/10.1038/nrg2564.

23. Heasman, J. (2006). Patterning the early Xenopus embryo. Development *133*, 1205–1217. https://doi.org/10.1242/dev.02304.

24. Langdon, Y.G., and Mullins, M.C. (2011). Maternal and zygotic control of zebrafish dorsoventral axial patterning. Annu. Rev. Genet. *45*, 357–377. https://doi.org/10.1146/annurev-genet-110410-132517.

25. Jukam, D., Kapoor, R.R., Straight, A.F., and Skotheim, J.M. (2021). The DNA-to-cytoplasm ratio broadly activates zygotic gene expression in Xenopus. Curr. Biol. *31*, 4269–4281.e8. https://doi.org/10.1016/j.cub.2021.07.035.

26. Blitz, I.L., Paraiso, K.D., Patrushev, I., Chiu, W.T.Y., Cho, K.W.Y., and Gilchrist, M.J. (2017). A catalog of Xenopus tropicalis transcription factors and their regional expression in the early gastrula stage embryo. Dev. Biol. *426*, 409–417. https://doi.org/10.1016/j.ydbio.2016.07.002.

27. Sinner, D., Kirilenko, P., Rankin, S., Wei, E., Howard, L., Kofron, M., Heasman, J., Woodland, H.R., and Zorn, A.M. (2006). Global analysis of the transcriptional network controlling Xenopus endoderm formation. Development *133*, 1955–1966. https://doi.org/10.1242/dev.02358.

28. Zhang, J., Houston, D.W., King, M.L., Payne, C., Wylie, C., and Heasman, J. (1998). The role of maternal VegT in establishing the primary germ layers in Xenopus embryos. Cell *94*, 515–524. https://doi.org/10.1016/s0092-8674(00)81592-5.

29. Xanthos, J.B., Kofron, M., Wylie, C., and Heasman, J. (2001). Maternal VegT is the initiator of a molecular network specifying endoderm in *Xenopus laevis*. Development *128*, 167–180. https://doi.org/10.1242/dev.128.2.167.

30. Zhang, J., and King, M.L. (1996). Xenopus VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. Development *122*, 4119–4129.

31. Blythe, S.A., and Wieschaus, E.F. (2016). Establishment and maintenance of heritable chromatin structure during early Drosophila embryogenesis. eLife *5*, e20148. https://doi.org/10.7554/eLife.20148.

32. Leichsenring, M., Maes, J., Mössner, R., Driever, W., and Onichtchouk, D. (2013). Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science *341*, 1005–1009. https://doi.org/10.1126/science.1242527.

33. Amodeo, A.A., Jukam, D., Straight, A.F., and Skotheim, J.M. (2015). Histone titration against the genome sets the DNA-to-cytoplasm threshold for the Xenopus midblastula transition. Proc. Natl. Acad. Sci. USA *112*, E1086–E1095. https://doi.org/10.1073/pnas.1413990112.

34. Newport, J., and Kirschner, M. (1982). A major developmental transition in early Xenopus embryos. II: control of the onset of transcription. Cell *30*, 687–696.

35. Newport, J., and Kirschner, M. (1982). A major developmental transition in early Xenopus embryos. I: Characterization and timing of cellular changes at the midblastula stage. Cell *30*, 675–686.

36. Dekens, M.P., Pelegri, F.J., Maischein, H.M., and Nüsslein-Volhard, C. (2003). The maternal-effect gene futile cycle is essential for pronuclear congression and mitotic spindle assembly in the zebrafish zygote. Development *130*, 3907–3916.

37. Prioleau, M.N., Huet, J., Sentenac, A., and Méchali, M. (1994). Competition between chromatin and transcription complex assembly regulates gene expression during early development. Cell *77*, 439–449.

38. Kane, D.A., and Kimmel, C.B. (1993). The zebrafish midblastula transition. Development *119*, 447–456.

39. Zhang, M., Skirkanich, J., Lampson, M.A., and Klein, P.S. (2017). Cell cycle remodeling and zygotic gene activation at the midblastula transition. Adv. Exp. Med. Biol. *953*, 441–487. https://doi.org/10.1007/978-3-319-46095-6_9.

40. Shermoen, A.W., and O'Farrell, P.H. (1991). Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. Cell *67*, 303–310. https://doi.org/10.1016/0092-8674(91)90182-x.

41. Rothe, M., Pehl, M., Taubert, H., and Jäckle, H. (1992). Loss of gene function through rapid mitotic cycles in the Drosophila embryo. Nature *359*, 156–159. https://doi.org/10.1038/359156a0.

42. Kimelman, D., Kirschner, M., and Scherson, T. (1987). The events of the midblastula transition in Xenopus are regulated by changes in the cell cycle. Cell *48*, 399–407.

43. Gard, D.L., Hafezi, S., Zhang, T., and Doxsey, S.J. (1990). Centrosome duplication continues in cycloheximide-treated Xenopus blastulae in the absence of a detectable cell cycle. J. Cell Biol. *110*, 2033–2042. https://doi.org/10.1083/jcb.110.6.2033.

44. Gerhart, J., Wu, M., and Kirschner, M. (1984). Cell cycle dynamics of an M-phase-specific cytoplasmic factor in Xenopus laevis oocytes and eggs. J. Cell Biol. *98*, 1247–1255. https://doi.org/10.1083/jcb.98.4.1247.

45. Miake-Lye, R., Newport, J., and Kirschner, M. (1983). Maturation-promoting factor induces nuclear envelope breakdown in cycloheximide-arrested embryos of Xenopus laevis. J. Cell Biol. *97*, 81–91. https://doi.org/10.1083/jcb.97.1.81.

46. Newport, J.W., and Kirschner, M.W. (1984). Regulation of the cell cycle during early Xenopus development. Cell *37*, 731–742. https://doi.org/10.1016/0092-8674(84)90409-4.

47. Levy, D.L., and Heald, R. (2010). Nuclear size is regulated by importin alpha and Ntf2 in Xenopus. Cell *143*, 288–298. https://doi.org/10.1016/j.cell.2010.09.012.

48. Hontelez, S., van Kruijsbergen, I., Georgiou, G., van Heeringen, S.J., Bogdanovic, O., Lister, R., and Veenstra, G.J.C. (2015). Embryonic transcription is controlled by maternally defined chromatin state. Nat. Commun. *6*, 10148. https://doi.org/10.1038/ncomms10148.

49. Skirkanich, J., Luxardi, G., Yang, J., Kodjabachian, L., and Klein, P.S. (2011). An essential role for transcription before the MBT in Xenopus laevis. Dev. Biol. *357*, 478–491. https://doi.org/10.1016/j.ydbio.2011.06.010.

50. Yang, J., Tan, C., Darken, R.S., Wilson, P.A., and Klein, P.S. (2002). Beta-catenin/Tcf-regulated transcription prior to the midblastula transition. Development *129*, 5743–5752.

51. Wang, P., Hayden, S., and Masui, Y. (2000). Transition of the blastomere cell cycle from cell size-independent to size-dependent control at the mid-blastula stage in Xenopus laevis. J. Exp. Zool. *287*, 128–144.

52. Chari, S., Wilky, H., Govindan, J., and Amodeo, A.A. (2019). Histone concentration regulates the cell cycle and transcription in early development. Development *146*, dev177402. https://doi.org/10.1242/dev.177402.

53. Shindo, Y., and Amodeo, A.A. (2021). Excess histone H3 is a competitive Chk1 inhibitor that controls cell-cycle remodeling in the early *Drosophila* embryo. Curr. Biol. *31*, 2633–2642.e6. https://doi.org/10.1016/j.cub.2021.03.035.

54. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419. https://doi.org/10.1038/nmeth.4197.

55. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

56. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

57. 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. https://doi.org/10.1038/nbt.1754.

58. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

59. Duttke, S.H., Chang, M.W., Heinz, S., and Benner, C. (2019). Identification and dynamic quantification of regulatory elements using total RNA. Genome Res. *29*, 1836–1846. https://doi.org/10.1101/gr.253492.119.

60. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 3. https://doi.org/10.14806/ej.17.1.200.

61. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

62. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

63. 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. https://doi.org/10.1186/gb-2008-9-9-r137.

64. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.

65. Palozola, K.C., Donahue, G., and Zaret, K.S. (2021). EU-RNA-seq for in vivo labeling and high throughput sequencing of nascent transcripts. Star Protoc. *2*, 100651. https://doi.org/10.1016/j.xpro.2021.100651.

66. Esmaeili, M., Blythe, S.A., Tobias, J.W., Zhang, K., Yang, J., and Klein, P.S. (2020). Chromatin accessibility and histone acetylation in the regulation of competence in early development. Dev. Biol. *462*, 20–35. https://doi.org/10.1016/j.ydbio.2020.02.013.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Xenopus laevis embryos	This paper	N/A
Chemicals, peptides, and recombinant proteins		
5-Ethynyl Uridine (5-EU)	Thermo Fisher Scientific	Cat# E10345
Cycloheximide	Sigma	Cat# C1988
Tetramethylrhodamine (TAMRA)-azide	Abcam	Cat# ab146486
Disulfide Biotin Azide	Click Chemistry Tools	Cat# 1168-5
THPTA	Click Chemistry Tools	Cat# 1010-100
$CuSO_4$	Sigma	Cat# 61230
Ascorbic acid	Sigma	Cat# A7506
TO-PRO-3	Thermo Fisher Scientific	Cat# T3605
Paraformaldehyde	EMS	Cat# 15710-S
Ficoll 400	Sigma	Cat# F4375
Hydrogen peroxide	Sigma	Cat# H1009
Formamide	Thermo Fisher Scientific	Cat# AC181090010
Benzyl alcohol	Sigma	Cat# 305197
Benzyl benzoate	ACROS Organics	Cat# 105860010
RO-3306	Sigma	Cat# SML0569
JNJ-7706621	Selleck Chemicals	Cat# S1249
AZD5438	Selleck Chemicals	Cat# S2621
BMS-265246	Selleck Chemicals	Cat# S2014
SKPin C1	Selleck Chemicals	Cat# S8652
Critical commercial assays		
RNeasy Mini Kit	Qiagen	Cat# 74104
DNase I recombinant, RNase-free	Roche	Cat# 04716728001
SuperScript III First-Strand Synthesis System for RT-PCR	Invitrogen	Cat# 18080-051
DreamTaq Green PCR Master Mix (2×)	Thermo Fisher Scientific	Cat# K1081
PowerUP SYBR Green Master Mix (2×)	Thermo Fisher Scientific	Cat# A25742
Click-iT Nascent RNA Capture Kit	Thermo Fisher Scientific	Cat# C10365
RNaseOUT Recombinant Ribonuclease Inhibitor	Thermo Fisher Scientific	Cat# 10777019
Universal RNA-seq with NuQuant	NuGEN	Cat# 0364
Agilent High Sensitivity DNA Kit	Agilent	Cat# 5067-4626
SPRIselect Reagent	Beckman Coulter	Cat# B23317
NEBNext Library Quant Kit for Illumina	NEB	Cat# E7630
NSQ 500/550 Hi Output KT v2.5 (75 CYS)	Illumina	Cat# 20024906
Deposited data		
Superseries of all RNA-seq datasets from this study	This paper	GEO: GSE201874; https://www.ncbi.nlm.nih.gov/geo/
Total and nascent transcriptome of *Xenopus laevis* embryos at 5-9 hpf (stage 7-9)	This paper	GEO: GSE201835; https://www.ncbi.nlm.nih.gov/geo/
Nascent transcriptome of the animal pole (AP) and the vegetal pole (VP) from *Xenopus laevis* embryos at 5-9 hpf (stage 7-9)	This paper	GEO: GSE201833; https://www.ncbi.nlm.nih.gov/geo/
Nascent transcriptome of cycloheximide (CHX)-arrested *Xenopus laevis* embryos at 5-7.5 hpf	This paper	GEO: GSE201834; https://www.ncbi.nlm.nih.gov/geo/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Xenopus laevis embryos	This paper	N/A
Oligonucleotides		
has1.S Forward: 5'-GTGGCATTCCAGCCTATTGT-3'	This paper	https://www.idtdna.com/
has1.S Reverse: 5'-TCAGGAATCTCCATTGTTTCTGC-3'	This paper	https://www.idtdna.com/
foxi1.L Forward: 5'-TGAAGATGATCCAGGCAAGGG-3'	This paper	https://www.idtdna.com/
foxi1.L Reverse: 5'-TAGGGCTCTCACTTAGCGGG-3'	This paper	https://www.idtdna.com/
grhl3.S Forward: 5'-CAGACTTAGCCAAGGCACCA-3'	This paper	https://www.idtdna.com/
grhl3.S Reverse: 5'-GGTCTGTAGCTGTTAATTCTGTCAA-3'	This paper	https://www.idtdna.com/
rgcc.L Forward: 5'-GTGGATACCCCTCATAAAGCAAG-3'	This paper	https://www.idtdna.com/
rgcc.L Reverse: 5'-TCGGTGTCACAGCATATCACT-3'	This paper	https://www.idtdna.com/
crx.L (Figure S2) Forward: 5'-AGTTGAGCACAAAGTACCATCCT-3'	This paper	https://www.idtdna.com/
crx.L (Figure S2) Reverse: 5'-TCAGTGACCAAGTATCAAGGGAC-3'	This paper	https://www.idtdna.com/
ventx1.2.L Forward: 5'-AAGCCTTCCTCAGCAGTGTT-3'	This paper	https://www.idtdna.com/
ventx1.2.L Reverse: 5'-GGGGGTGAATGCTGTTCTCA-3'	This paper	https://www.idtdna.com/
odc.S Forward: 5'-TTCATTCAGGCAGTCGTCGC-3'	This paper	https://www.idtdna.com/
odc.S Reverse: 5'-GCGCTGTTCTGCTGTTTGTA-3'	This paper	https://www.idtdna.com/
crx.L (Figure S3) Forward: 5'-CAGTTGAGCACAAAGTACCATCC-3'	This paper	https://www.idtdna.com/
crx.L (Figure S3) Reverse: 5'-CAGTGACCAAGTATCAAGGGACT-3'	This paper	https://www.idtdna.com/
dlc.L Forward: 5'-CGGACATGCGAATGGTCTCA-3'	This paper	https://www.idtdna.com/
dlc.L Reverse: 5'-GCTGGATACACCAGCGGCA-3'	This paper	https://www.idtdna.com/
pcdh18.L Forward: 5'-TGATTTGGGCAGAGATTCGC-3'	This paper	https://www.idtdna.com/
pcdh18.L Reverse: 5'-GTGCAGAGCCTCATAGCTGAA-3'	This paper	https://www.idtdna.com/
Software and algorithms		
salmon (v0.12.0)	Patro et al.[54]	https://github.com/COMBINE-lab/salmon
DESeq2 (bioconductor v3.8)	Love et al.[55]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
STAR (v 2.7.0)	Dobin et al.[56]	https://github.com/alexdobin/STAR
IGV (v2.8.0)	Robinson et al.[57]	https://software.broadinstitute.org/software/igv
clusterProfiler (v4.2.0)	Wu et al.[58]	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
HOMER (v4.11)	Duttke et al.[59]	http://homer.ucsd.edu/homer
Cutadapt (v3.7)	Martin[60]	https://cutadapt.readthedocs.io
Bowtie 2 (v2.3.4.1)	Langmead et al.[61]	http://bowtie-bio.sourceforge.net/bowtie2
Samtools (v1.1)	Li et al.[62]	https://www.htslib.org
MACS2 (v2.2.7.1)	Zhang et al.[63]	https://github.com/macs3-project/MACS
deepTools (v3.5.1)	Ramírez et al.[64]	https://deeptools.readthedocs.io
Fiji	NIH	https://imagej.net/software/fiji
GraphPad Prism 9	GraphPad	https://www.graphpad.com
Excel 2021	Microsoft	https://www.microsoft.com/EN-US/microsoft-365
RStudio	RStudio	https://www.rstudio.com
Adobe Illustrator CC 2021	Adobe	https://www.adobe.com/products/illustrator.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Matthew Good (mattgood@pennmedicine.upenn.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The RNA-seq data generated in this study have been deposited at Gene Expression Omnibus (GEO) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- The 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

Experiments in this study were performed using the African clawed frog *Xenopus laevis* according to the Animal Use Protocol approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Pennsylvania. Mature *Xenopus laevis* females and males were purchased from Nasco, and they were maintained at 20°C in tanks of a recirculating aquatic system. The females were used for procuring eggs and males were used for preparing sperms, the procedures of which have been described previously.[14] Briefly, to induce ovulation, 100 U of pregnant mare serum gonadotropin (PMSG) and 500 U of human chorionic gonadotropin (HCG) were sequentially injected into the dorsal lymph sac of female adult frogs at 3-7 days and 14-15 h, respectively, before experiment. Before egg collection, the females were temporarily kept at 16 °C in 1× Marc's Modified Ringer's (1× MMR was diluted from the stock 20× MMR that includes 100 mM HEPES pH 7.8, 2 mM EDTA, 2 M NaCl, 40 mM KCl, 20 mM $MgCl_2$, and 40 mM $CaCl_2$). Eggs were obtained by gently squeezing the female frogs and collected in glass dishes. After egg procurement, the females were quarantined in high marine salt for at least one week before returning to the recirculating aquatic system. The ovulated females were not used until they rested for at least 3 months. To prepare sperms, adult males were euthanized with 0.2% benzocaine for at least 20 min before dissection for isolating the testes. The isolated testes were kept in L-15 medium on ice and used within one week. Sperm slurry was prepared by crushing 1/2 of a testis using a plastic pestle in 1 ml of deionized water inside a 1.5-ml microfuge tube.

METHOD DETAILS

In vitro fertilization (IVF)

All IVF in this study were performed at room temperature (22 °C ± 0.5 °C). The procedures for IVF have been described previously.[14] Briefly, 1 ml of sperm slurry was evenly added onto a monolayer of eggs inside a glass dish collected as described above. The sperms and eggs were mixed by gently sliding a plastic pestle on the surface of the glass dish. Five minutes after adding sperms, the glass dish was flooded with ~20 ml of 0.1× MMR, with all eggs submerged. At ~30 minutes post-fertilization (mpf), the fertilized eggs were incubated with 20 ml of 2% L-cysteine in 0.1× MMR for 2-5 min and the jelly coats were removed by washing with 0.1× MMR for multiple times. The embryos were sorted and kept in 0.1× MMR for further use.

Microinjection, cycloheximide treatment and collection of embryos

The procedures for microinjection have been described previously.[13,14] Briefly, embryos at 1-cell stage (~35-40 mpf) were transferred into a microinjection chamber containing 3% Ficoll in 0.5 × MMR. Embryos were microinjected with 10 nl of 50 mM 5-ethynyl uridine (EU) using a PLI-100 picoliter microinjector (Medical Systems, Greenvale, NY). The final concentration of EU inside embryos is ~ 0.5 mM, as detailed previously.[13] After microinjection, the embryos were transferred into a glass dish containing 3% Ficoll in 0.5 × MMR for 1-2 h before being transferred in 0.1× MMR to continue embryo development.

To prepare embryos for characterizing the nascent transcriptome in whole embryos (Figures 1 and S1), embryos from two clutches (i.e., two individual frogs) were microinjected with EU as described above. EU-microinjected embryos developed to 5, 6, 7, 8 and 9 hours post-fertilization (hpf), respectively, were collected in microcentrifuge tubes (N = 20 embryos each), followed by complete removal of residual medium before snap-freezing in liquid nitrogen. The biological replicates of samples were stored at -80 °C before use. To distinguish maternal-zygotic (MZ) genes and zygotic-only (Z) genes during ZGA (Figures 1G–1K and S1H–S1L) based on the presence of their transcripts in eggs (see below), eggs from two clutches were collected and the egg transcriptome was directly compared to the nascent transcriptome from embryos at 5-9 hpf. The nascent transcriptome at 5-9 hpf from both experiments (a total of four replicates) were used for analysis (see below).

To prepare embryos to spatially characterize the nascent transcriptome in the animal pole (AP) and the vegetal pole (VP) regions (Figures 2 and S2), embryos from a total of four clutches (i.e., four frogs) were microinjected with EU as described above. Embryos were segmented via dissection using a hair knife to collect the AP (~ top 1/3 region of embryo) and the VP (~ bottom 1/3 region of embryo). For one of the four clutches, EU-microinjected embryos developed to 6, 7, 8 and 9 hpf, respectively, and were then dissected and collected in in microcentrifuge tubes (N = 30 segments each), followed by complete removal of residual medium before snap-freezing in liquid nitrogen. Two technical replicates of samples were stored at -80 °C before use. For the remainder of the three clutches, developing EU-microinjected embryos were collected at 5, 6, 7, 8 and 9 hpf and dissected in the same manner as above; AP and VP regions were collected (N = 50 segments each). A total of five replicates (four biological replicates with one technical replicate) were used for the spatial nascent transcriptome analysis.

To prepare embryos to characterize the effect of cell cycle lengthening on nascent zygotic transcription (Figures 3 and S3), embryos from one clutch (e.g., one frog) were microinjected with EU as described above. Two replicates of EU-microinjected embryos developed to 3, 4 and 5 hpf, respectively, were incubated in 0.1× MMR (control) or 0.1× MMR containing 0.2 mg/ml of cycloheximide (CHX) to block embryonic divisions and maintain the cells in the interphase. Control embryos and CHX-treated embryos developed to 7.5 hpf were collected as described above and stored at -80 °C before use. For confocal imaging of nascent zygotic transcription in single cells, control embryos and CHX-treated embryos at 7.5 hpf were fixed in 4% paraformaldehyde / 1× MEM (100 mM MOPS pH 7.4, 2 mM EGTA, and 1 mM $MgSO_4$) solution in 2-ml scintillation vials by rotating for 2 hours at room temperature. Embryos were completely dehydrated with methanol before being stored at -20 °C. To test whether CHX regulated nascent transcription is RNA polymerase II dependent, 0.1 ng of α-amanitin was co-microinjected with 5-EU into embryos at 1-cell stage and the embryos were treated with CHX and fixed as described above.

Cell cycle inhibitor incubation

To assess the effect of specific cell cycle inhibitors on arresting blastula cell cycles (Figure S3J), normal embryos were incubated with 100 μM of Cdk inhibitors, including RO-3306 (Sigma, Cat# SML0569), JNJ-7706621 (Selleck Chemicals, Cat# S1249), AZD5438 (Selleck Chemicals, Cat# S2621), BMS-265246 (Selleck Chemicals, Cat# S2014), and SKPin C1 (Selleck Chemicals, Cat# S8652), respectively, from 5 hpf to 7.5 hpf. Untreated and DMSO-treated embryos were used as negative controls, and CHX (0.2 mg/ml) treated embryos were used as positive controls. Live embryos at 7.5 hpf were imaged under a stereomicroscope using Leica Application Suite X (LAS X) (Leica Microsystems, Germany).

RNA isolation, biotinylation and purification

Total RNAs were isolated using the RNeasy Mini Kit (Qiagen), following the instructions provided by the manufacturer. Briefly, eggs or EU-microinjected whole embryos or segmented AP and VP regions were added with 700 μl of Buffer RLT and homogenized by gentle pipetting the samples up and down for multiple times until all embryos were completely dissolved. The homogenates were added with 700 μl of 70% ethanol and the mixtures were transferred into the columns used for binding RNA. The columns were span at 13,000 rpm for 1 min. The columns were washed with 700 μl of Buffer RW1 and incubated with 80 μl of DNase I for 15 min at room temperature. After DNase I incubation, the columns were added with 600 μl of Buffer RW1 and centrifuged at 13,000 rpm for 1 min. The columns were washed twice with Buffer RPE and completely dried by centrifugation at 13,000 rpm for 2 min. The total RNAs were finally eluted in 20 μl of RNase-free water.

To biotinylate RNA, 2.5-10 μg of total RNAs were incubated with a 20 μl reaction that contains 2 mM disulfide biotin azide, 50 mM Hepes (pH 7.5), 1.25 mM $CuSO_4$/THPTA mix and 10 mM ascorbic acid for 1 h at room temperature. The reaction was stopped by adding 1 μl of 50 mM EDTA. To precipitate the RNA, the reaction was added with 1μl of glycogen, 1 volume of 5 M ammonium acetate and 700 μl of chilled ethanol, incubated at -80 °C for overnight, and centrifuged at 13, 000×g for 20 min at 4 °C. The supernatant was removed, and the pellet was washed twice with 700 μl of chilled 75% ethanol by centrifugation at 13, 000×g for 5 min at 4 °C. The pellet was air dried and resuspended in 10 μl RNase-free H_2O.

The nascent EU-RNA was purified using streptavidin beads following the instructions provided by the Click-iT Nascent RNA Capture Kit (Thermo Fisher Scientific, Cat# C10365), with minor modifications.[65] The 10 μl biotinylated RNA from above was added with a 15-μl reaction mix that contains 12.5 μl of Click-iT RNA binding buffer, 0.2 μl of RNaseOUT Recombinant Ribonuclease Inhibitor and 2.3 μl of RNase-free water. The reaction was incubated at 69 °C for 5 min and added with 5 μl of Dynabeads MyOne Streptavidin T1 that were pre-washed with Click-iT reaction wash buffer 2 for three times. The reaction was incubated for 30 min at room temperature. The beads were concentrated using a magnetic separator (Permagen) and sequentially washed with 50 μl of Click-iT reaction wash buffer 1 for five times and wash buffer 2 for five times. The beads were resuspended in 5 μl of for Click-iT reaction wash buffer 2 and used directly for first-strand cDNA synthesis and subsequent library prep (see below).

EU-RNA-seq and analysis

To perform RNA-seq, cDNA libraries were prepared for total RNA isolated from embryos ('All'), purified nascent EU-RNA ('Bead') and the flowthrough that contains maternal, non-nascent RNA ('Flowthrough'). cDNA libraries were prepared using the Universal RNA-seq with NuQuant kit (NuGEN, Cat# 0364), following the manual provided by the manufacturer. Ribosomal RNAs were depleted using the custom designed AnyDeplete Probe Mix for *Xenopus laevis* provided by the kit. The cDNA libraries were further analyzed following the instructions specified in respective kits below. The quality of cDNA libraries was analyzed using the Agilent High Sensitivity DNA Kit (Agilent, Cat# 5067-4626) in the Agilent 2100 Bioanalyzer System (Agilent Technologies, CA). The cDNA libraries were subjected to

size selection using SPRIselect beads (Beckman Coulter, Cat# B23317). The concentrations of cDNA libraries were quantified using the NEBNext Library Quant Kit (NEB, Cat# E7630). The individual cDNA libraries were pooled at equal molar ratios and the pooled cDNA libraries were sequenced using the NSQ 500/550 Hi Output KT v2.5 (75 CYS) (Illumina, Cat# 20024906) in the NextSeq 500 sequencer (Illumina, CA). To quantify transcripts, raw sequence data (fastq files) were aligned to *Xenopus laevis* genome build 9.2 using salmon v0.12.0.[54] Data were normalized for sequencing depth using DESeq2 (bioconductor v3.8).[55] To map the transcripts to the genome, the STAR (v 2.7.0) aligner[56] was used and the peaks were visualized in the Integrative Genomics Viewer (IGV, v2.8.0).[57]

To circumvent the issue of potential nonspecific binding of maternal transcripts to beads, we decided to use the net increase of reads at each blastula timepoint, subtracting as background the 5 hpf reads. When defining the list of nascent transcribing genes from 5-9 hpf, which we termed the nascent list, the nascent transcriptome data from whole embryos was filtered using the following criteria for each replicate: (1) the gene is continuously transcribed from 5 hpf to 9 hpf, and (2) at least with an increase of 50 reads and 1.5-fold induction from 5 hpf to 9 hpf (using average reads of 8-9 hpf vs. average reads of 5-6 hpf). The final list was determined for the genes meeting these criteria in all replicates, which included 2577 genes (used in Figures 1D–1F, S1D, and 3C–3G). To categorize the nascent list into subgroups of MZ vs Z genes (Figures 1G–1K and S1H–S1L), the presence of transcripts detected in the egg was used to determine their identities: Z genes were those with $\leq$ 5 reads in the eggs and MZ genes were those with > 5 reads. The maternal genes in the rest of all genes were those with the presence of transcripts > 100 reads in the egg but with no transcription in the Bead. To select the genes most highly transcribed at the MBT from previous studies, the transcriptome data from Session et al.[19] (Figure S1E) and Yanai et al.[20] (Figure S1F), respectively, were filtered using the following criteria: at least with an increase of 20 reads and 1.5-fold induction from Stage 8 (st08) to Stage 10 (st10).

To select genes for categorizing their spatial patterns of activation (Figures 2 and S2), the AP-VP nascent transcriptome data was filtered using the following criteria for each replicate: (1) the gene is continuously transcribed from 5-9 hpf at either the AP or the VP, and (2) at least with an increase of 10 reads from 5 or 6 hpf to 9 hpf (using average reads of 8-9 hpf vs. average reads of 5-6 hpf). The final list was determined for the genes meeting these criteria in at least three out of five replicates, which included 882 genes. To determine the activation patterns for each gene, the reads at both AP and VP from 5 hpf to 9 hpf were first normalized to their mean reads and then categorized using the following criteria for each pattern, respectively: (1) AP regional: both the total reads and the reads at 8-9 hpf were at least 10-fold higher in the AP than the VP; (2) AP early, VP delay: both the total reads and the reads at 8-9 hpf were 1.5-10-fold higher in the AP than the VP; (3) VP regional: both the total reads and the reads at 8-9 hpf were at least 10-fold higher in the VP than the AP; (4) VP early, AP delay: both the total reads and the reads at 8-9 hpf were 1.5-10-fold higher in the VP than the AP; (5) Similar: both the total reads and the reads at 8-9 hpf were within 1.5-fold difference between the VP and the AP. The resulting categorized lists of genes were further manually inspected to remove a small portion of genes with inconsistent profiles between replicates (at least three out of five replicates) or to correct their categorization based on expression profiles. To determine the activation onset time for each gene (Figures 2B and S2B), we adapted the method described in Jukam et al.[25] by fitting the mean-normalized reads with a smooth spline function and used the time reaching 20% of the maximum reads (the maximum of AP and VP combined) as the onset time. The fittings were manually inspected and corrected for some genes by fitting with an exponential or sigmoidal function optimal for them. To determine the average activation onset time for each pattern (Figure S2C), the same fitting of a smooth spline function was performed except for using the average of the mean-normalized reads for the AP and VP in each pattern. In Figure 2B, to compare and plot the activation onset time for those genes that were not activated during 5-9 hpf at AP or VP, their activation onset time was set to 9.1.

Lists of ectoderm and endoderm genes (Figure 4) were defined from previous studies. For ectoderm genes, we used the data from Blitz et al.[26] by selecting the animally enriched genes in gastrula of *Xenopus tropcalis* and matching their names in *Xenopus laevis*, which generated a list of 111 genes. For endoderm genes, we used the data from Sinner et al.[27] and matched their names in Affymetrix microarray with the ones in *Xenopus laevis* genome build 9.2 at Xenbase, which generated a list of 172 genes. Because many germ-layer specific genes are not expressed or very lowly expressed in blastula embryos, to characterize the effect of CHX on germ layer expression by 7.5 hpf (Figures 4D–4F and S4A–S4D), the ectoderm and endoderm genes were filtered for those genes with an increase of at least 10 reads from 5 hpf to 7.5 hpf in the control embryos.

Functional enrichment analysis

Gene ontology (GO) analysis for genes with functional enrichment in biological processes was performed using clusterProfiler (v4.2.0).[58] The top 10 of the most significantly enriched GO terms were selected and the -$\log_{10}$(p.adjust) was used as the proxy of enrichment. The enrichment of motifs at the promoter regions of CHX down-regulated genes or up-regulated endoderm genes were performed using HOMER (v4.11).[59] The names of transcription factors that bind the enriched motifs were manually inspected and confirmed their presence in *Xenopus laevis*.

ATAC-seq and ChIP-seq analysis

To validate the chromatin accessibility of nascent transcripts uniquely detected by EU-RNA-seq (Figures S2M and S2N), the public ATAC-seq data (GEO accession number: GSE138905) from the animal caps of *Xenopus laevis* embryos at stages 10 and 12 were analyzed.[66] Three replicates of each stage were included in the analysis. The Illumina Nextera adapter sequences were trimmed using Cutadapt (v3.7)[60] before the ATAC-seq sequences were aligned to the *Xenopus laevis* genome build 9.2 using Bowtie 2 (v2.3.4.1)[61] and BAM files were generated using Samtools (v1.1).[62] The peak calling was made using MACS2 (v2.2.7.1),[63] and the heatmap and profile plots for the ATAC-seq peaks were generated using deepTools (v3.5.1).[64]

To compare the RNA Pol II binding and H3K4me3 mark between ectoderm and endoderm genes (Figures 4G and S4I), the public ChIP-seq data (GEO: GSE76059) from stage 10.5 *Xenopus laevis* embryos were analyzed.[19] The ChIP-seq sequences were similarly processed to ATAC-seq sequences as described above, except for without adaptor removal.

RT-PCR and real-time PCR

Total RNAs were isolated as described above. cDNA was generated by using the SuperScript III First-Strand Synthesis System (Invitrogen, Cat# 18080-051), following the instructions provided by the product. Briefly, for each sample 2.5 μg of total RNA was mixed with 1 μl of 50 ng/μl random hexamers and 1 μl of 10 mM dNTP mix to make a reaction of 10 μl, which was incubated at 65 °C for 5 min and placed on ice for at least 1 min. Each reaction was added with 10 μl of cDNA Synthesis Mix, which was composed of 2 μl of 10× RT buffer, 4 μl of 25 mM $MgCl_2$, 2 μl of 0.1 M DTT, 1 μl of RNaseOUT (40 U/μl) and 1 μl of SuperScript III RT (200 U/μl), and it was incubated at 25 °C for 10 min followed by at 50 °C for 50 min. The reaction was terminated by incubation at 85 °C for 5 min. To eliminate RNA contamination, 1 μl of RNase H was added to each reaction by incubation at 37 °C for 20 min.

RT-PCR was performed by mixing 100 ng of cDNA with 10 μl of the 2× DreamTaq Green PCR Master Mix (Thermo Fisher Scientific, Cat# K1081) and 0.25 μM of each gene-specific Forward and Reverse primers (see key resources table) to make a total reaction of 20 μl, followed by performing PCR in a Bio-Rad C1000 Touch thermal cycler: 95 °C for 3 min; 25-35 cycles of 95 °C for 30 s, 51-55 °C for 30 s (note that the choose of annealing temperature is primer dependent) and 72 °C for 30 s; and 72 °C for 10 min. The PCR products were subjected to electrophoresis in 2% agarose gel containing 0.5 μg/ml of ethidium bromide and the gene-specific bands were visualized under ultraviolet light in a Bio-Rad Gel Doc EZ Imager.

Real-time PCR was performed by mixing 100 ng of cDNA with 10 μl of the 2× PowerUP SYBR Green Master Mix (Thermo Fisher Scientific, Cat# A25742) and 0.5 μM of each gene-specific Forward and Reverse primers (see key resources table) to make a total reaction of 20 μl (in triplicates), followed by performing PCR in a QuantStudio 3 Real-Time PCR System (Applied Biosystems) with the standard cycling mode: 50 °C for 2 min; 95 °C for 2 min; 40 cycles of 95 °C for 15 s, 56 °C for 15 s and 72 °C for 1 min. The melt curve stage was performed by the following conditions: 95 °C for 15 s (1.6 °C/s), 60 °C for 1 min (1.6 °C/s) and 95 °C for 15 s (0.15 °C/s). The fold difference in expression level between AP and VP was calculated by $2^{-\Delta Ct(AP-VP)}$, where ΔCt(AP-VP) was the average Ct(AP) of triplicates – average Ct(VP) of triplicates.

Confocal imaging nascent transcripts in wholemount embryos ad image analysis

Confocal imaging of nascent transcripts in wholemount embryos have been described previously.[13,14] Briefly, the fixed EU-microinjected control and CHX-treated embryos were sequentially rehydrated with 75%, 50% and 25% methanol in 0.5× SSC (75 mM NaCl and 7.5 mM sodium citrate) for 10 min each, followed by washing with 0.5× SSC for three times. Embryos were bleached in the solution of 5% formamide/2% H_2O_2/0.5× SSC for 6 h under light. Embryos were briefly rinsed with 0.5× SSC for three times, followed by washing with 1× TBST (containing 0.1% vol/vol Triton X-100) for 30 min each of six times and with 1× TBS for 10 min each of three times. Embryos were incubated with 25 μM TAMRA-azide, 100 mM Tris-HCl pH 8.5, 1 mM $CuSO_4$, and 100 mM ascorbic acid for 12 h at room temperature. Embryos were extensively washed with 1× TBST at room temperature for 1 day by changing buffer every 2 h. Embryos were incubated with TO-PRO-3 (1:500 dilution) overnight at 4°C, followed by washing with 1× TBST for 1 day at room temperature by changing buffer every 2 h. Embryos were completely dehydrated in anhydrous methanol by changing it for several times. Embryos were cleared in a mixture of 1 part of benzyl alcohol and 2 parts of benzyl benzoate (BABB) for 24 h before confocal imaging. Confocal imaging was performed with the ZEN software on a Zeiss LSM710 confocal microscope. EU-RNA and TO-PRO-3 were imaged with a frame size of 1,024 pixels × 1,024 pixels using lasers 561 nm (0.15% power) and 633 nm (10% power), respectively, without saturating signals. Z-stacks were collected at an interval of 2 μm using the Plan-Apochromat 25× / 0.8 immersion oil objective. Images were processed in Fiji (NIH) and presented as Z-projections with maximum intensity for several selected slices.

To assess the effect of CHX on DNA synthesis (Figure S3A), DNA-integrated TO-PRO-3 signal in the nucleus was quantified using confocal image stacks collected using the 25× objective. A total number of 50-70 cells from five embryos were analyzed for each group. The boundaries of nuclei and cell of individual blastomeres were manually demarcated in Fiji (NIH) using the slice with the highest signal of the TO-PRO-3 channel by specific DNA signal and non-specific background signal, respectively. TO-PRO-3 signal in the nucleus as well as in the cell was measured and the net nuclear TO-PRO-3 signal was calculated by subtracting the cytoplasmic background signal from the nucleus signal. The nuclear TO-PRO-3 amount was calculated by multiplying the net nuclear TO-PRO-3 signal with the nucleus volume assuming a spherical shape of the nucleus.

To quantify the nascent EU-RNA intensity (Figure 3K) and cell size (Figure S3H) of single blastomeres, confocal image stacks collected using the 25× objective were used. For proper comparison between groups, only blastomeres in the animal pole regions within 100 μm depth of the image stacks were analyzed. For each group, a total number of 80-160 cells from at least three embryos were analyzed. The nucleus and cell boundaries of individual blastomeres were manually demarcated in Fiji (NIH) using the slice with the highest signal of the TO-PRO-3 channel by specific DNA signal and non-specific background signal, respectively. EU-RNA signal in the nucleus as well as in the cell was measured using the EU-RNA channel and the net nuclear EU-RNA signal was calculated by subtracting the cytoplasmic background signal from the nucleus signal. The nuclear EU-RNA amount was calculated by multiplying the net nuclear EU-RNA signal with the nucleus volume assuming a spherical shape of the nucleus. Cell size, represented as cell diameter or cell volume, was also calculated by assuming a spherical shape of blastomeres.

QUANTIFICATION AND STATISTICAL ANALYSIS

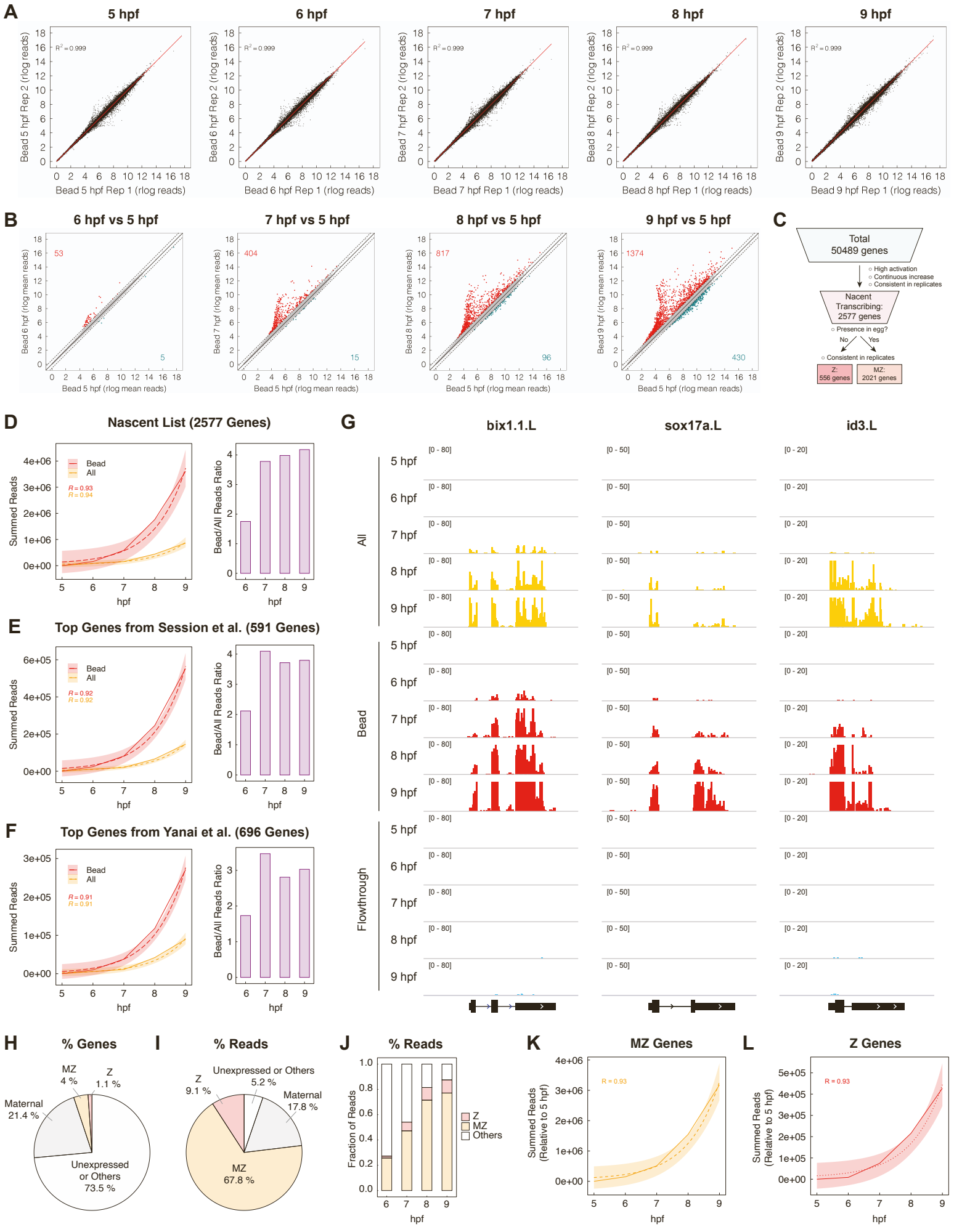
All statistical parameters, including sample numbers, mean and standard deviation or error, were included in Figures and Figure legends. In Figure 2B, the statistical significance was determined by Student's t test to compare the activation onset time between the AP and the VP. In Figures 3K and S3A, the statistical significance was determined by one-way ANOVA (Fisher's LSD test) to compare the nuclear EU-RNA level and cell size between control and CHX treatment at various developmental stages. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns, not significant. In other plots where p values or adjusted p values (p.adjust) were used, e.g. for functional enrichment and motif enrichment, the p values were calculated and provided by individual packages in R.

Current Biology, Volume *32*

Supplemental Information

Nascent transcriptome reveals orchestration of zygotic genome activation in early embryogenesis

Hui Chen and Matthew C. Good


A
5 hpf
6 hpf
7 hpf
8 hpf
9 hpf
Bead 5 hpf Rep 2 (rlog reads)
Bead 5 hpf Rep 1 (rlog reads)
Bead 6 hpf Rep 2 (rlog reads)
Bead 6 hpf Rep 1 (rlog reads)
Bead 7 hpf Rep 2 (rlog reads)
Bead 7 hpf Rep 1 (rlog reads)
Bead 8 hpf Rep 2 (rlog reads)
Bead 8 hpf Rep 1 (rlog reads)
Bead 9 hpf Rep 2 (rlog reads)
Bead 9 hpf Rep 1 (rlog reads)
R² = 0.999
B
6 hpf vs 5 hpf
7 hpf vs 5 hpf
8 hpf vs 5 hpf
9 hpf vs 5 hpf
Bead 6 hpf (rlog mean reads)
Bead 7 hpf (rlog mean reads)
Bead 8 hpf (rlog mean reads)
Bead 9 hpf (rlog mean reads)
Bead 5 hpf (rlog mean reads)
53
5
404
15
817
96
1374
430
C
Total
50489 genes
High activation
Continuous increase
Consistent in replicates
Nacent
Transcribing:
2577 genes
Presence in egg?
No
Yes
Consistent in replicates
Z:
556 genes
MZ:
2021 genes
D
Nascent List (2577 Genes)
Summed Reads
Bead
All
R = 0.93
R = 0.94
Bead/All Reads Ratio
hpf
E
Top Genes from Session et al. (591 Genes)
R = 0.92
R = 0.92
F
Top Genes from Yanai et al. (696 Genes)
R = 0.91
R = 0.91
G
bix1.1.L
sox17a.L
id3.L
All
Bead
Flowthrough
5 hpf
6 hpf
7 hpf
8 hpf
9 hpf
H
% Genes
MZ
4 %
Z
1.1 %
Maternal
21.4 %
Unexpressed
or Others
73.5 %
I
% Reads
Unexpressed or Others
5.2 %
Z
9.1 %
Maternal
17.8 %
MZ
67.8 %
J
% Reads
Fraction of Reads
Z
MZ
Others
K
MZ Genes
Summed Reads
(Relative to 5 hpf)
R = 0.93
L
Z Genes
R = 0.93

(Figure S1 Continued)

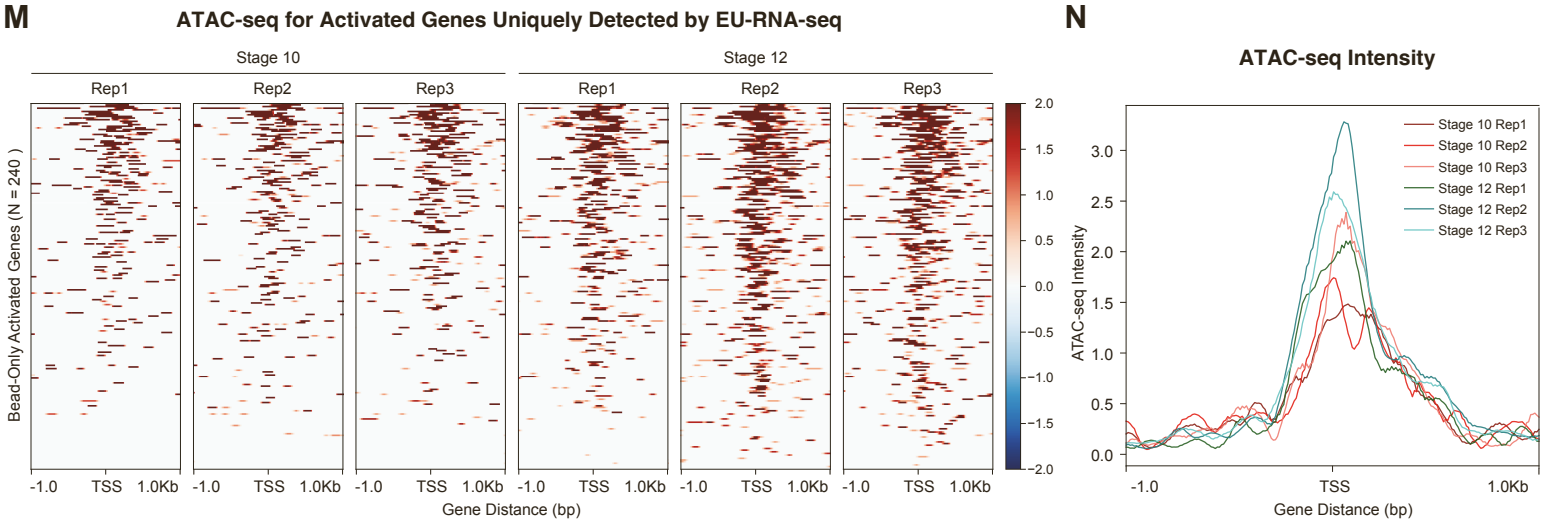


Figure S1. EU-RNA-seq to characterize the nascent transcriptome in early *Xenopus* embryogenesis. Related to Figure 1.

(A) Reproducibility of nascent transcriptome library reads among replicates. Nascent EU-RNA-seq was performed on two clutches of early embryos (Rep 1 and Rep 2) from 5-9 hpf. Each dot represents individual genes with rlog reads quantified by DESeq2. Red line: linear fitting of the data in each plot; R squared values for fit.

(B) Beads biochemically capture the nascent transcriptome. Increase in nascent EU-RNA-seq reads each hpf compared to pre-ZGA, 5 hpf. Each dot represents individual genes with rlog reads averaged from duplicates quantified by DESeq2. The numbers are genes enriched above 1.5-fold threshold (dashed lines). Red and blue dots indicate increased and decreased nascent transcripts at each hpf as compared with 5 hpf, respectively.

(C) Schematic describing filtering the transcription data to define the bona fide nascent transcriptome for whole embryos at 5-9 hpf. The nascent list is further split into zygotic-only (Z) and maternal-zygotic (MZ) genes based on absence or presence of reads in the egg.

(D-F) Enhanced sensitivity: comparison of reads levels from nascent EU-RNA-seq versus total RNA-seq. (Left) Summed reads for selected nascently transcribing genes (D, N = 2577), top genes from Session et al. (E, N = 591) and top genes from Yanai et al. (F, N = 696) from 5-9 hpf. Red, EU-RNA-seq for bead-purified nascent RNAs; Orange, total RNA-seq for all RNAs. Exponential fitting is shown with 95% CI. (Right) Enrichment of reads in nascent transcriptome as compared with total transcriptome over time. The ratio of summed Bead/All reads was calculated for each hpf from 6-9 hpf.

(G) Genome browser view of reads for representative zygotic genes. Comparing total RNA-seq (top, All; orange), nascent EU-RNA-seq (middle, Bead; red) and maternal RNA-seq (bottom, Flowthrough; blue). The scales are set to the same level for each gene. Reads for these zygotic genes are higher in nascent (bead) transcriptome and absent from the maternal transcriptome (flowthrough), indicating biochemical separation of nascent from maternal.

(H-J) Composition of genes or reads expressed during ZGA.

(H) Percentage of genes that are unexpressed, maternal, or induced: maternal-zygotic (MZ) vs. zygotic (Z).

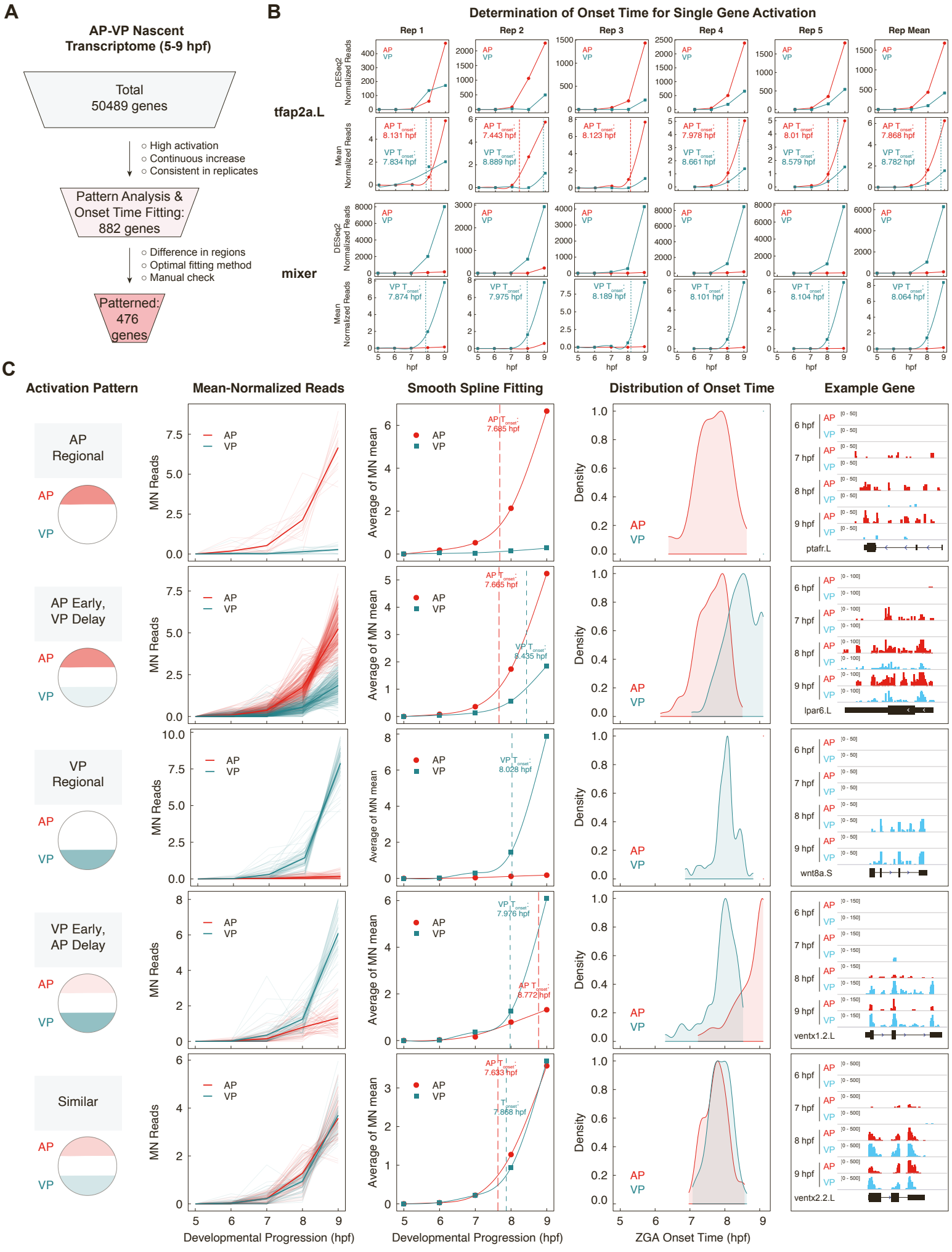
(I) Percentage of total reads from nascent transcriptome (Bead) that can be categorized.

(J) Fraction of assigned reads from nascent transcriptome (Bead) as development progresses.

(K and L) Summed reads for MZ (K) and Z (L) genes from 5-9 hpf. Showing increase of reads, subtracting background (5 hpf). Exponential fitting is shown with 95% CI.

(M) Heatmap for ATAC-seq peaks for genes whose activation was uniquely detected by EU-RNA-seq during 5-9 hpf in embryos at stages 10 and 12, respectively. Three replicates from each stage are shown (data source: Esmaeili et al., *Dev Biol* 2020). The regions of transcription start site (TSS) +/- 1kb are shown for each gene (each row represents each gene; N = 240 in total). The indicated regions of majority genes are accessible by stage 10 and further increased at stage 12.

(N) Profile plot for ATAC-seq intensity in M.


A
AP-VP Nascent Transcriptome (5-9 hpf)
Total 50489 genes
High activation
Continuous increase
Consistent in replicates
Pattern Analysis & Onset Time Fitting: 882 genes
Difference in regions
Optimal fitting method
Manual check
Patterned: 476 genes
B
Determination of Onset Time for Single Gene Activation
Rep 1
Rep 2
Rep 3
Rep 4
Rep 5
Rep Mean
tfap2a.L
mixer
DESeq2 Normalized Reads
Mean Normalized Reads
AP
VP
hpf
C
Activation Pattern
Mean-Normalized Reads
Smooth Spline Fitting
Distribution of Onset Time
Example Gene
AP Regional
AP Early, VP Delay
VP Regional
VP Early, AP Delay
Similar
MN Reads
Average of MN mean
Density
Developmental Progression (hpf)
ZGA Onset Time (hpf)
ptafr.L
lpar6.L
wnt8a.S
ventx1.2.L
ventx2.2.L

(Figure S2 Continued)

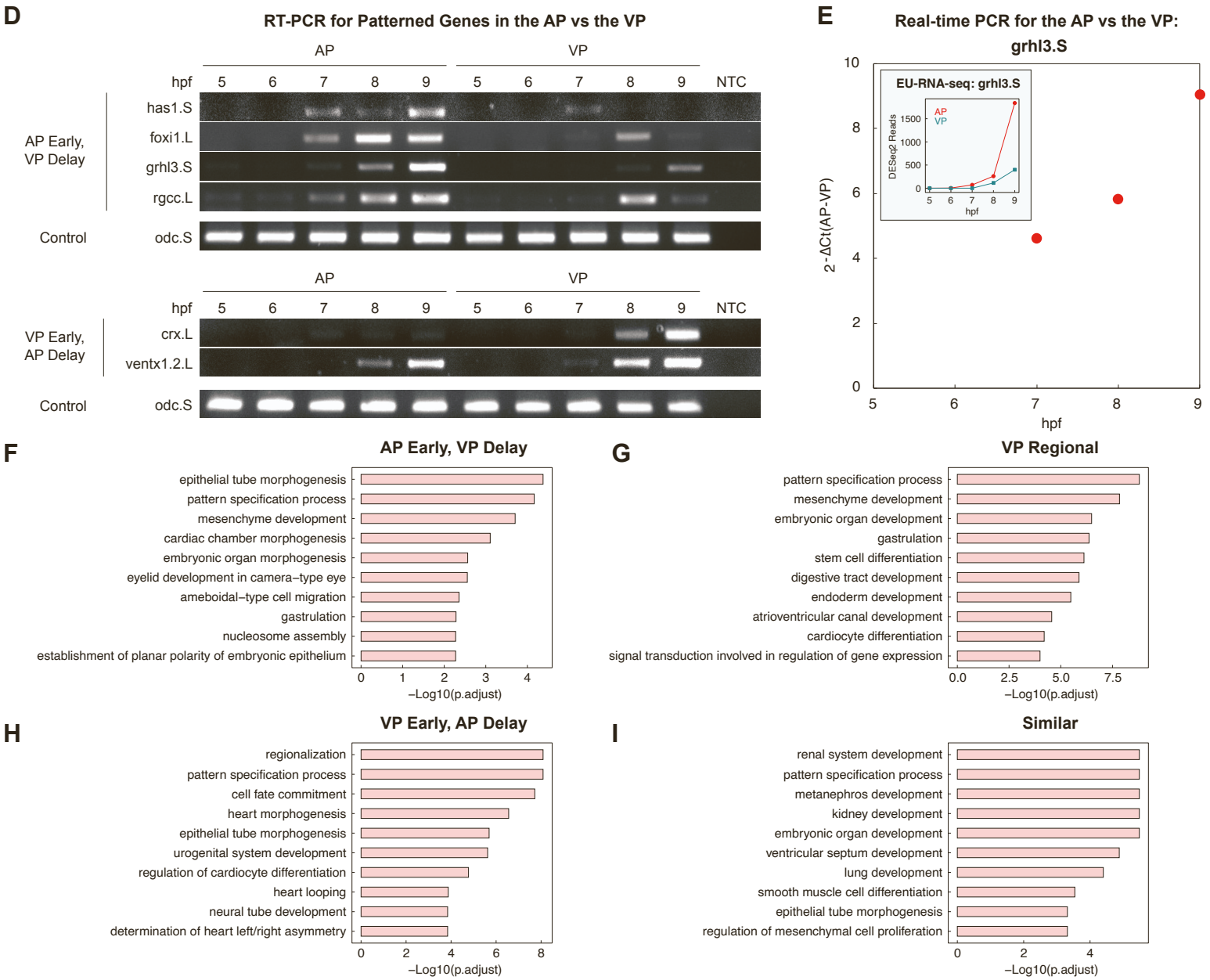


Figure S2. Characterizing spatial patterns of single gene activation in early *Xenopus* embryogenesis. Related to Figure 2.

(A) Scheme for filtering genes used for analysis of activation patterns at the animal pole (AP) and the vegetal pole (VP).

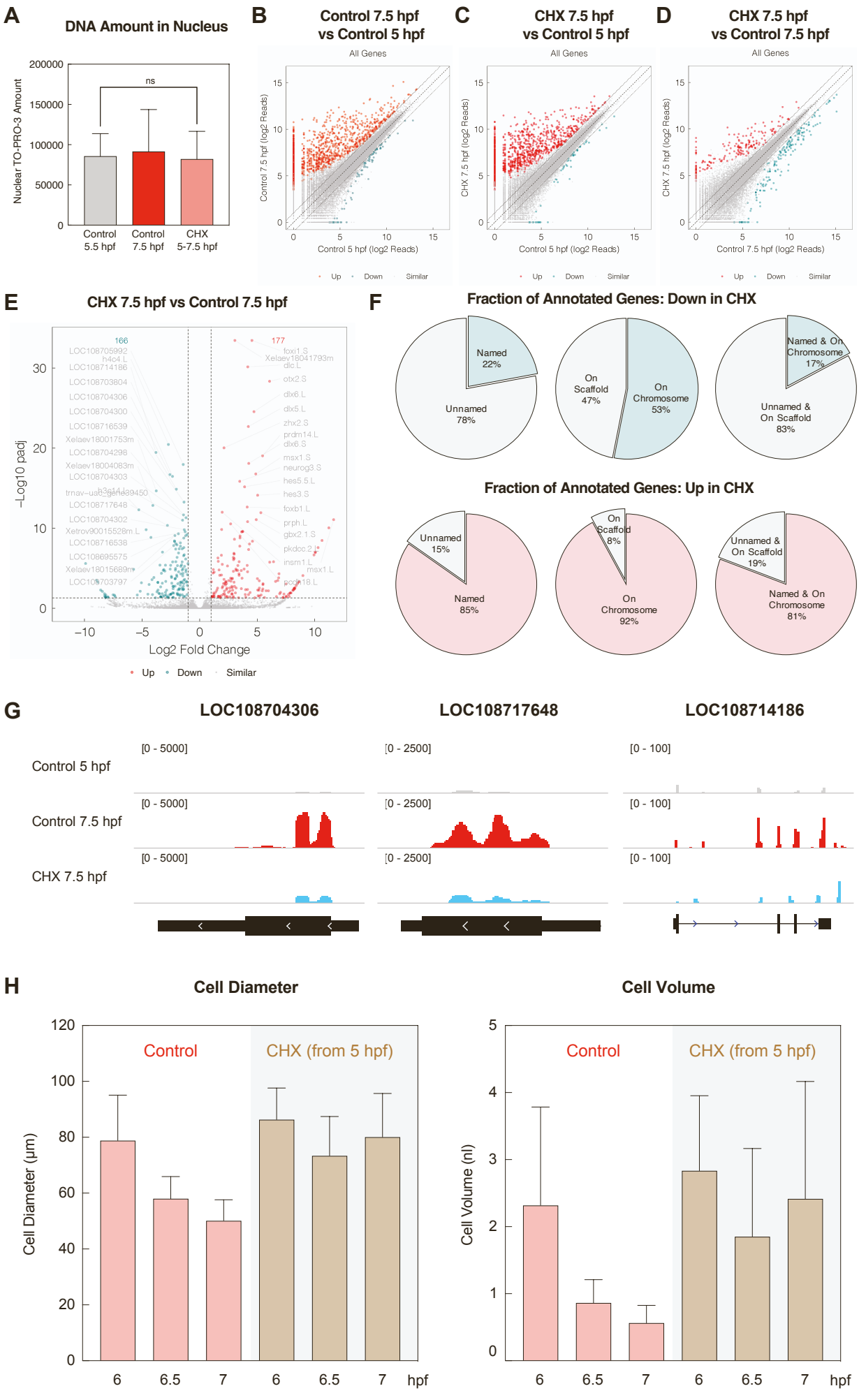
(B) Representative gene profiles to demonstrate the determination of activation onset time for individual genes. Genes tfap2a.L (top panels) and mixer (bottom panels) are shown as examples. DESeq2-normalized reads for 5 replicates (Rep 1-5) and their average (Rep Mean) from the nascent transcriptome at the AP (red) and the VP (blue) at 5-9 or 6-9 hpf. Mean-normalized reads were used for fitting with a smooth spline to determine the activation onset time at the AP and VP, respectively, based on reaching 20% of the maximum reads.

(C) Categorization of 5 distinct spatial patterns for single gene activation. From left to right: schematic; mean-normalized reads for individual genes and their average; smooth spline fitting for the average mean-normalized reads to determine the activation onset time in AP and VP; histogram showing distribution of activation onset time for each gene within a category for the AP (red) and the VP (blue); genome browser view of example genes for each pattern. AP, red; VP, blue.

(D) RT-PCR for example genes in the patterns of 'AP Early, VP Delay' (top panels) and 'VP Early, AP Delay' (bottom panels). Internal control was used for normalization. NTC, no-template control.

(E) Real-time PCR for an example gene grhl3.S in the AP and VP. The expression level was measured from triplicates and the fold difference between AP and VP is represented as $2^{-\Delta Ct(AP-VP)}$, where Ct(AP-VP) was calculated from the average Ct values of triplicates for AP and VP, respectively. Note that the expression was not detected at 5 and 6 hpf, and therefore the differences between AP and VP at these time points are not shown. As a reference for the expression pattern, the inset plot shows the average DESeq2 Reads by EU-RNA-seq in the AP and VP regions, respectively.

(F-I) Top 10 enriched GO terms for patterned genes: 'AP Early, VP Delay' (F), 'VP Regional' (G), 'VP Early, AP Delay' (H) and 'Similar' (I). The GO terms for 'AP Regional' genes could not be enriched due to small number of genes in this list.


A
DNA Amount in Nucleus
Nuclear TO-PRO-3 Amount
200000
150000
100000
50000
0
ns
Control 5.5 hpf
Control 7.5 hpf
CHX 5-7.5 hpf
B
Control 7.5 hpf vs Control 5 hpf
All Genes
Control 7.5 hpf (log2 Reads)
Control 5 hpf (log2 Reads)
Up
Down
Similar
C
CHX 7.5 hpf vs Control 5 hpf
All Genes
CHX 7.5 hpf (log2 Reads)
Control 5 hpf (log2 Reads)
Up
Down
Similar
D
CHX 7.5 hpf vs Control 7.5 hpf
All Genes
CHX 7.5 hpf (log2 Reads)
Control 7.5 hpf (log2 Reads)
Up
Down
Similar
E
CHX 7.5 hpf vs Control 7.5 hpf
166
177
foxi1.S
Xelaev18041793m
dlc.L
otx2.S
dlx6.L
dlx5.L
zhx2.S
prdm14.L
dlx6.S
msx1.S
neurog3.S
hes5.5.L
hes3.S
foxb1.L
prph.L
gbx2.1.S
pkdcc.2.L
insm1.L
msx1.L
pcdh18.L
LOC108705992
h4c4.L
LOC108714186
LOC108703804
LOC108704306
LOC108704300
LOC108716539
Xelaev18001753m
LOC108704298
Xelaev18004083m
LOC108704303
trnav-uac-gene39450
LOC108717648
LOC108704302
Xetrov90015528m.L
LOC108716538
LOC108695575
Xelaev18015689m
LOC108703797
−Log10 padj
Log2 Fold Change
Up
Down
Similar
F
Fraction of Annotated Genes: Down in CHX
Named 22%
Unnamed 78%
On Scaffold 47%
On Chromosome 53%
Named & On Chromosome 17%
Unnamed & On Scaffold 83%
Fraction of Annotated Genes: Up in CHX
Unnamed 15%
Named 85%
On Scaffold 8%
On Chromosome 92%
Unnamed & On Scaffold 19%
Named & On Chromosome 81%
G
LOC108704306
LOC108717648
LOC108714186
Control 5 hpf
Control 7.5 hpf
CHX 7.5 hpf
[0 - 5000]
[0 - 2500]
[0 - 100]
H
Cell Diameter
Cell Diameter (µm)
Control
CHX (from 5 hpf)
6
6.5
7
hpf
Cell Volume
Cell Volume (nl)

(Figure S3 Continued)

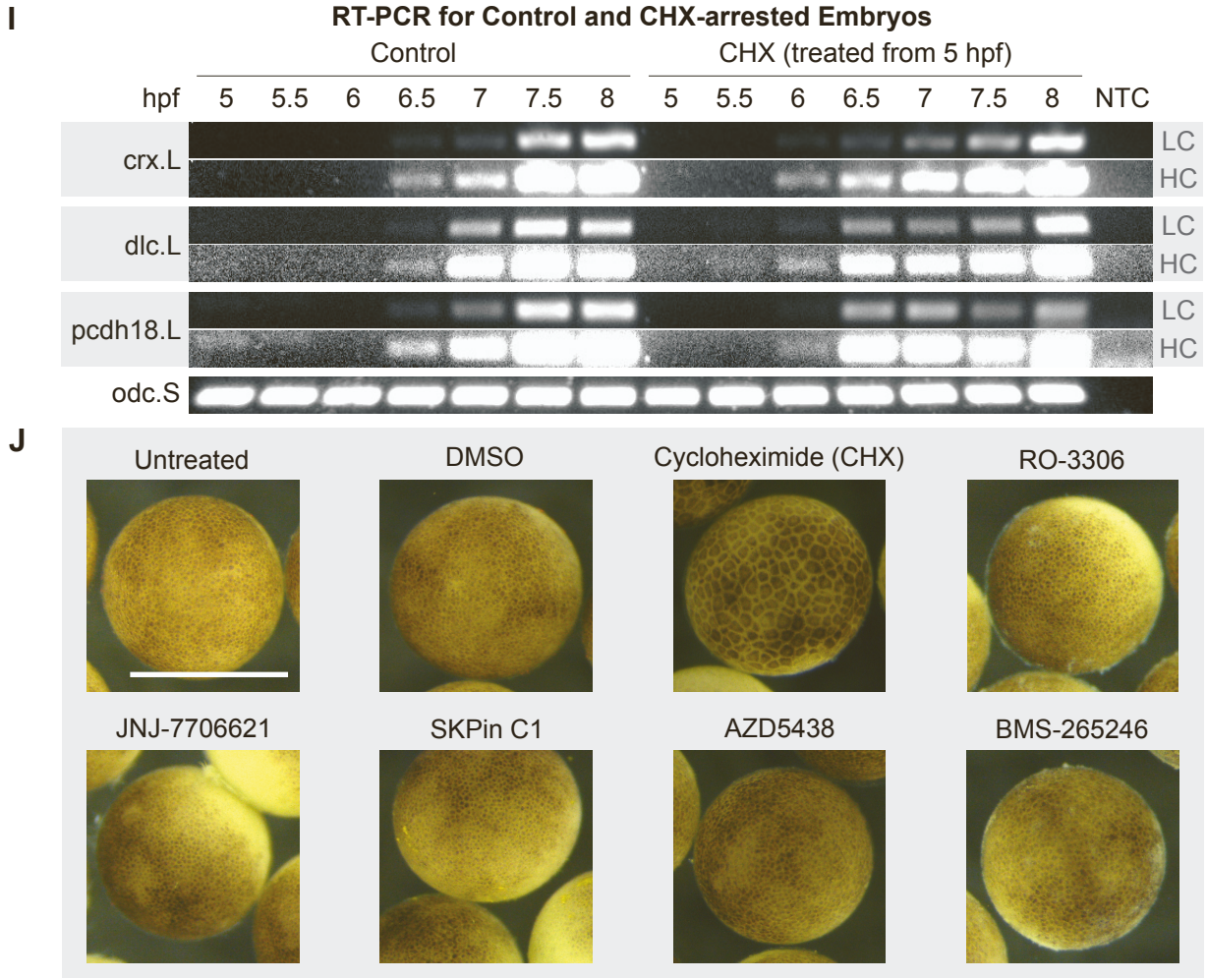


Figure S3. Effect of CHX on nascent zygotic transcription. Related to Figure 3.

(A) DNA amount in nucleus of control embryos or embryos treated with cycloheximide (CHX). DNA was stained with TO-PRO-3 and the nuclear DNA amount was calculated by multiplying the nuclear intensity (subtracted with cytoplasmic background) by the volume of nucleus. Analyzed cell numbers from representative images: N = 66 (Control 5.5 hpf), N = 61 (Control 7.5 hpf), and N = 57 (CHX 5-7.5 hpf). Data are represented with Mean + SD. ns, not significant.

(B-C) Increase in nascent transcription from 5 to 7.5 hpf for control (A) and cell cycle elongated (B) embryos. Using all genes in the genome. Each dot represents individual genes with log2 reads averaged from duplicates quantified by DESeq2. Differentially enriched genes in each group are labelled in red (up) and blue (down), based on a 1.5-fold threshold (dashed lines) and padj <0.05.

(D) Comparison of log2 reads from cell cycle elongated (CHX) versus control embryo at 7.5 hpf, during widespread ZGA onset. Each dot represents each gene in the genome.

(E) Volcano plot showing fold-change in expression comparing cell cycle elongated (CHX) versus control embryos. Labels show top 20 genes most differentially expressed. The numbers of up-regulated (red) and down-regulated (blue) genes by CHX are indicated. The horizontal dashed line indicates 1.3 (a threshold of padj = 0.05) and the two vertical dashed lines indicate -1 and 1, respectively (a threshold of 2-fold change).

(F) Comparison of gene annotation information between down-regulated (top) and up-regulated (bottom) genes in the cell cycle elongated embryos (CHX) compared to control at 7.5 hpf. Showing percentage of genes that are named or unnamed, on chromosomes or unassigned scaffold contigs, or both.

(G) Genome browser view of example genes only weakly induced by cell cycle elongation (CHX) compared to control.

(H) Quantification of cell size in control and CHX-arrested embryos from 5 hpf. Cell size is represented as cell diameter (μm, Left) and cell volume (nl, Right) calculated by assuming a spherical shape of blastomeres. A total of 80-160 cells from at least three embryos were analyzed for each group. Data are represented as mean + SD.

(I) RT-PCR for zygotic genes early activated in CHX-arrested embryos from 5 hpf to 5.5-8 hpf as compared with control embryos. The gene odc.S is used as a loading control. NTC, no-template control. LC, low contrast; HC, high contrast.

(J) Inhibitor incubation from 5-7.5 hpf to characterize cell cycle arrest in blastula stage embryos. The final concentration in media is 0.2 mg/ml for CHX and 100 μM for the indicated Cdk inhibitors (~ 1,000 to 10,000-fold higher than respective Ki). Untreated and DMSO-treated embryos are used as controls. Only CHX is able to permeate *Xenopus* embryos o rapidly induce cell cycle arrest and Cdk inhibitors fail to arrest embryo division. Scale bar, 1 mm.

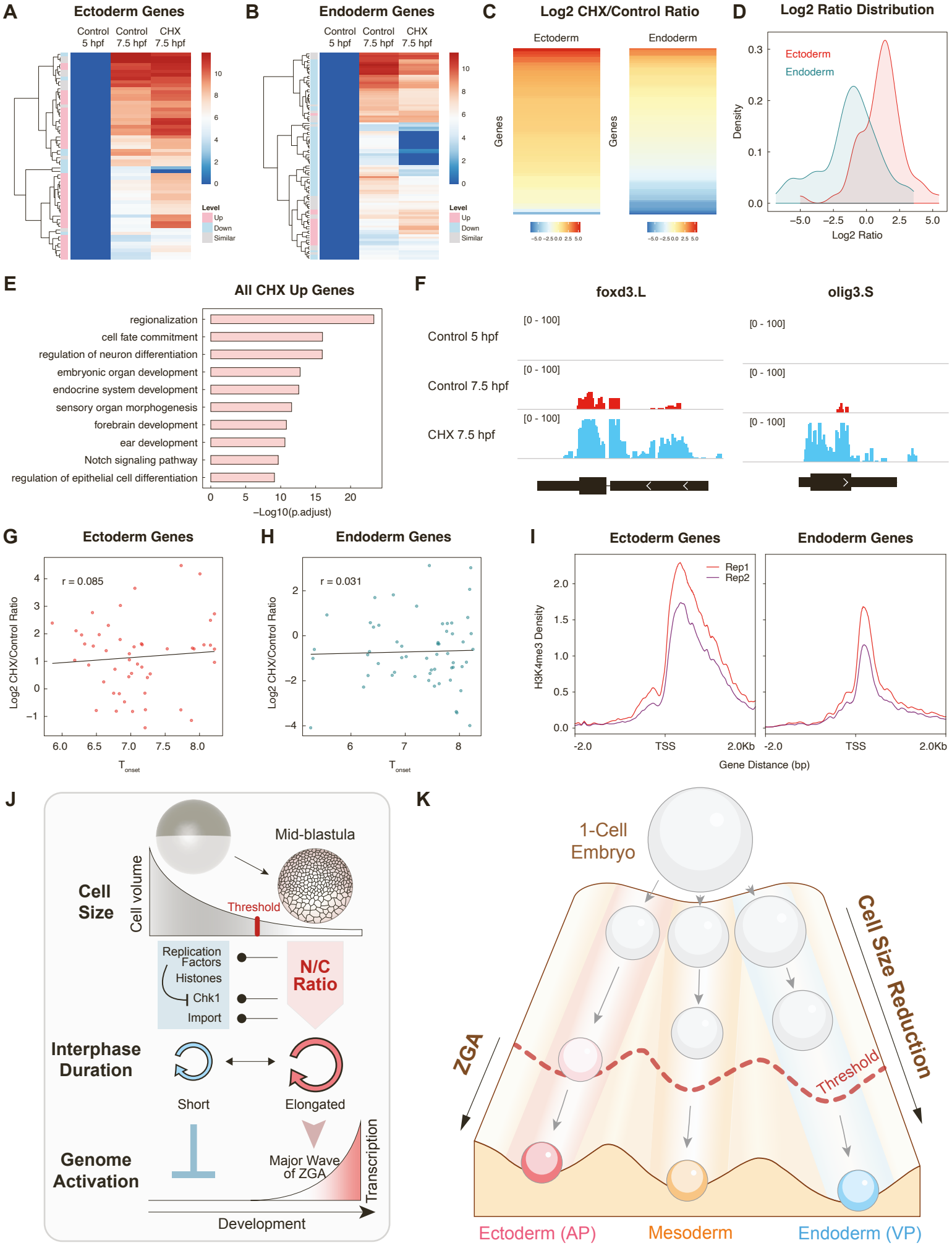

A
Ectoderm Genes
Control 5 hpf
Control 7.5 hpf
CHX 7.5 hpf
Level
Up
Down
Similar
B
Endoderm Genes
C
Log2 CHX/Control Ratio
Ectoderm
Endoderm
Genes
D
Log2 Ratio Distribution
Density
Log2 Ratio
E
All CHX Up Genes
regionalization
cell fate commitment
regulation of neuron differentiation
embryonic organ development
endocrine system development
sensory organ morphogenesis
forebrain development
ear development
Notch signaling pathway
regulation of epithelial cell differentiation
–Log10(p.adjust)
F
foxd3.L
olig3.S
Control 5 hpf
Control 7.5 hpf
CHX 7.5 hpf
[0 - 100]
G
Ectoderm Genes
r = 0.085
Log2 CHX/Control Ratio
T onset
H
Endoderm Genes
r = 0.031
I
Ectoderm Genes
Endoderm Genes
Rep1
Rep2
H3K4me3 Density
TSS
2.0Kb
Gene Distance (bp)
J
Cell Size
Cell volume
Mid-blastula
Threshold
Replication Factors
Histones
Chk1
Import
N/C Ratio
Interphase Duration
Short
Elongated
Genome Activation
Major Wave of ZGA
Transcription
Development
K
1-Cell Embryo
ZGA
Cell Size Reduction
Threshold
Ectoderm (AP)
Mesoderm
Endoderm (VP)

Figure S4. Premature cell cycle elongation over-induces the ectoderm germ layer program. Related to Figure 4.

(A and B) Heatmaps for log2 reads of the ectoderm (A) and endoderm (B) genes with or without CHX treatment. Data shown are selected for genes that are induced in control embryos from 5-7.5 hpf. The up-regulated (red), down-regulated (blue) and not significantly changed (grey) genes are categorized based on a 1.5-fold difference threshold.

(C) Heatmaps for log2 ratio of gene induction in CHX/Control for ectoderm (left) and endoderm (right) genes at 7.5 hpf.

(D) Histogram for log2 ratio of gene induction in CHX/Control for ectoderm (red) and endoderm (blue) genes at 7.5 hpf. Ectoderm is largely over-induced by cell cycle elongation; endoderm is largely under-induced.

(E) Top 10 enriched GO terms for genes that are significantly over-induced by cell cycle elongation (CHX).

(F) Genome browser view of example ectoderm genes over-induced in CHX-treated embryos.

(G and H) No correlation between activation onset time and log2 ratio of gene induction in CHX/Control for ectoderm (G) and endoderm genes (H) as in A and B.

(I) Profile plots for ChIP-seq peaks of H3K4me3 for ectoderm and endoderm genes, respectively, in embryos at stage 10.5. The regions of transcription start site (TSS) +/- 2kb from two replicates are shown (data source: Session et al., *Nature* 2016). The level of H3K4me3 in the indicated regions is higher in ectoderm genes than endoderm genes.

(J) Mechanistic regulation of ZGA via cell size, N/C ratio and cell cycle elongation in *Xenopus* early embryo. During *Xenopus* early development, a fertilized egg undergoes rapid reductive cell divisions without cell growth until mid-blastula stages, resulting in progressive reduction in blastomere cell size. Because the DNA amount remains constant in each cell, the N/C ratio increases. When the cell size or N/C ratio reaches a critical threshold, blastomeres elongate their cell cycle via Chk1, potentially via limiting replication factors and fork stalling, or reduced histone dependent repression of Chk1; RNAPII occupancy may also impact fork stalling. Early rapid divisions block zygotic transcript accumulation because interphase duration is too short for transcriptional accumulation. Once the cell cycle is elongated, a majority of nascent transcripts quickly build up, providing the basis for widespread gene expression during the major wave of ZGA.

(K) A Waddington model for germ layer initiation whose timing is linked to regulated ZGA: cell size reduction down the landscape determines the rate at which ZGA initiates and subsequent germ layer specification. The fertilized egg undergoes asymmetric divisions generating small cells at the animal pole (AP), the presumptive ectoderm, and large cells at the vegetal pole (VP), the presumptive endoderm. Smaller AP cells reach a cell size threshold faster than the large VP cells. AP cells are the presumptive ectoderm and thus early ZGA leads to earlier expression of the ectoderm program. VP cells are the presumptive endoderm and delayed ZGA onset delays expression of the endoderm program. Note: the timing of initiation of the mesoderm program is inferred.