



Published in final edited form as:

Nature. 2013 November 21; 503(7476): 360–364. doi:10.1038/nature12632.

## Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition

Miler T. Lee<sup>1,\*</sup>, Ashley R. Bonneau<sup>1,\*</sup>, Carter M. Takacs<sup>1</sup>, Ariel A. Bazzini<sup>1</sup>, Kate R. DiVito<sup>1</sup>, Elizabeth S. Fleming<sup>1</sup>, and Antonio J. Giraldez<sup>1,2,‡</sup>

<sup>1</sup>Department of Genetics, Yale University School of Medicine, New Haven, CT 06510

<sup>2</sup>Yale Stem Cell Center, Yale University School of Medicine, New Haven, CT 06520

### Summary

Upon fertilization, maternal factors direct development and trigger zygotic genome activation (ZGA) at the maternal-to-zygotic transition (MZT). In zebrafish, ZGA is required for gastrulation and clearance of maternal mRNAs, which is in part regulated by the conserved microRNA miR-430. However, the factors that activate the zygotic program in vertebrates are unknown. Here, we show that Nanog, Pou5f1 and SoxB1 regulate zygotic gene activation in zebrafish. We identified several hundred genes directly activated by maternal factors, constituting the first wave of zygotic transcription. Ribosome profiling revealed that *nanog*, *sox19b* and *pou5f1* are the most highly translated transcription factors pre-MZT. Combined loss of these factors resulted in developmental arrest prior to gastrulation and a failure to activate >75% of zygotic genes, including miR-430. Our results demonstrate that maternal Nanog, Pou5f1 and SoxB1 are required to initiate the zygotic developmental program and induce clearance of the maternal program by activating miR-430 expression.

In animals, maternal gene products drive early development in a transcriptionally silent embryo, and are responsible for zygotic genome activation (ZGA). ZGA occurs during the maternal-to-zygotic transition (MZT), when developmental control transfers to the embryonic nucleus. This universal transition represents a major reprogramming event that requires (i) chromatin remodeling to provide transcriptional competency, (ii) specific activation of a new transcriptional program and (iii) clearance of the previous transcriptional program. In *Drosophila*, maternal Zelda is required for activating the first zygotic genes through binding of TAGteam cis elements<sup>1,2</sup>. However, the maternal factors that mediate ZGA in vertebrates remain largely unknown<sup>3,4</sup>. In zebrafish, ZGA coincides with the midblastula transition (MBT) ~3 hours post fertilization (hpf), during which genome competency is established through widespread changes in chromatin<sup>5,6</sup> and DNA

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

<sup>‡</sup>To whom correspondence should be addressed [antonio.giraldez@yale.edu](mailto:antonio.giraldez@yale.edu), Tel: 203.785.5423, Fax: 203.785.4415.

\*Equal contribution

Contributions: M.T.L., A.R.B. and A.J.G. designed the project, performed experiments and data analysis. M.T.L., A.R.B., C.M.T. and A.J.G. wrote the manuscript. C.M.T. designed and performed the cycloheximide experiment and contributed to in situ hybridizations. A.B.B. designed and performed ribosome profiling and U1U2 experiments. K.R.D. and E.S.F. assisted with gene validation. Sequencing data are deposited in the Gene Expression Omnibus (GEO) database with accession number GSE47558.

methylation<sup>7,8</sup>. Bivalent chromatin marks are associated with zygotic genes thought to be ‘poised’ for activation<sup>5</sup>. Yet, many loci with active marks appear to be transcriptionally inactive<sup>5</sup>, suggesting that competent genes require induction by additional factors. ZGA is required for epiboly<sup>9</sup>, and the clearance of maternal mRNAs, a process regulated in part by the conserved microRNA (miRNA) miR-430<sup>10-12</sup>. While significant advances have taken place in understanding how vertebrate embryos acquire transcriptional competency and orchestrate the clearance of the maternal program, the factors that control activation of the specific genes during ZGA remain unknown. Here we combine loss-of-function analyses, high-throughput sequencing and ribosome footprinting to identify factors that activate the first wave of zygotic transcription to initiate nuclear control of embryonic development.

## Identifying the first zygotic transcripts

To define factors that mediate transcriptional activation, we first sought to identify the earliest genes transcribed from the zygotic genome. Accurate characterization of the early transcriptome faces two main challenges: (i) zygotic transcription of a gene can be masked by a large maternal contribution, and (ii) poly(A)<sup>+</sup> selection of mRNAs can lead to apparent increases in gene expression, reflecting delayed polyadenylation of maternal mRNAs rather than transcription. We reasoned that maternal mRNAs are spliced during oogenesis, so examining introns from total RNA would allow us to quantify de novo transcription independent of polyadenylation or maternal contribution. We performed Illumina total RNA sequencing on wild type (WT) embryos after the onset of zygotic transcription (4hpf, sphere; and 6hpf, shield) (Fig. 1a) compared to embryos before MBT (2hpf, 64-cell stage) and  $\alpha$ -amanitin treated embryos, which lack zygotic transcription. This analysis identified 608 genes with significant increases in exon or intron expression levels > 5 RPKM (reads per kilobase, per million reads) at sphere stage ( $P < 0.1$ , Benjamini-Hochberg multiple test correction) (Fig. 1b,c, Supplementary Fig. 1a-h). Intron signal identifies an additional 6602 genes with low levels of transcription by 4hpf, and 9330 transcribed genes by 6hpf, expanding the number of zygotically expressed genes previously identified<sup>13,14</sup> (Supplementary Fig. 1i-o). Over 74% of these are genes with maternal contributions (maternal and zygotic genes, M+Z), most of which are only identified by elevated intron signal (Fig. 1b, Supplementary Fig. 1g), reflecting the sensitivity of this method to detect de novo transcription.

Next, we examined which genes are directly triggered by the maternal program in the “first wave” of transcription by 4hpf, versus those activated by zygotic factors. We reasoned that blocking zygotic gene function while leaving maternal factors unaffected would uncouple the first from subsequent waves of zygotic transcription. To this end, we inhibited splicing of zygotic mRNAs using morpholinos complementary to U1 and U2 spliceosomal RNAs (U1U2 MO) (Fig. 1d, Supplementary Fig. 1a-d)<sup>15</sup>. U1U2 MO embryos arrest prior to epiboly (Fig. 1a), despite remaining transcriptionally active. Illumina sequencing revealed an enrichment in intron-exon boundary reads (Fig. 1e) and activation of a subset of zygotic transcripts to levels > 5 RPKM (Methods); these genes constitute the first wave of zygotic transcription (Fig. 1f). To test that these first-wave genes are indeed independent of zygotic factors, we treated embryos with cycloheximide (CHX) prior to MBT (32-cell stage) to selectively block translation of zygotic mRNAs, while allowing translation of maternal

mRNAs. CHX-treated embryos also fail to reach epiboly (Fig. 1a) and have a highly correlated transcriptome profile with U1U2 MO (Pearson's  $R = 0.97$ , Supplementary Fig. 2), confirming first-wave transcription in the absence of zygotic proteins. First-wave genes comprise both embryonic-specific and housekeeping genes ubiquitously expressed in adult tissues (Supplementary Fig. 3a) and are enriched in pattern specification, gastrulation and chromatin modifying functions (Supplementary Fig. 3b). We validated a subset of these genes by RT-PCR, including *klf4b*, *nanog* and *isg15* (Supplementary Fig. 3c-k). Notably, the pri-miR-430 polycistron is highly expressed as part of this first wave ( $>1000$  RPKM) (Fig. 1c, f). Together, these results identify 269 first-wave genes expressed by sphere stage for which maternal factors are sufficient for activation.

## Nanog, SoxB1 & Pou5f1 activate the first wave

Considering the specific, widespread and steep pattern of zygotic gene activation, we hypothesized that the factors that trigger the first wave may include sequence-specific transcriptional regulators highly translated prior to ZGA. We analyzed the translation levels of all maternal mRNAs using ribosome profiling data (Fig. 2a)<sup>16</sup>. We found that Nanog, Sox19b and Pou5f1 (Oct4) are the most highly translated sequence-specific transcription factors (TFs) in the pre-MZT transcriptome (Fig. 2b). Pou5f1, the SoxB1 family (which includes Sox2 and Sox19b) and Nanog are key TFs involved in maintaining pluripotency in embryonic stem cells (ESCs) (reviewed in<sup>17,18</sup>). In zebrafish, Pou5f1 provides temporal control of gene expression<sup>19</sup> and together with SoxB1 regulates dorsal-ventral patterning and neuronal development<sup>18,20-23</sup>, while Nanog is essential for endoderm formation through regulation of zygotic *mtx2*<sup>24</sup>.

To examine the roles of Nanog, Sox19b and Pou5f1 in activating zygotic gene expression, we combined a maternal-zygotic loss-of-function (LOF) Pou5f1 (*MZpou5f1*)<sup>21</sup> with previously published translation blocking morpholinos for Nanog<sup>24</sup> and SoxB1<sup>20</sup> (Methods). Because Sox2, Sox3 and Sox19a have been shown to compensate for Sox19b loss, we used a combination of morpholinos targeting all four *sox* genes<sup>20</sup> (Supplementary Fig. 4a). Simultaneous Nanog LOF in combination with SoxB1 or Pou5f1 resulted in complete developmental arrest prior to gastrulation, with  $>95\%$  of the treated embryos failing to initiate epiboly ( $n=387$  and  $n=52$  respectively) (Fig. 2c, Supplementary Fig. 4b-e). This phenotype resembles that of  $\alpha$ -amanitin injected embryos, suggesting that these factors play a role in activating zygotic genes. We used two different approaches to analyze the activity and specificity of these morpholinos. First, we performed ribosome profiling on WT and Nanog + SoxB1 MO injected embryos pre-MBT<sup>16,25</sup>. Translation efficiency for both Nanog and Sox19b was reduced  $>97\%$  in the morpholino-injected embryos compared to WT (Fig. 2d, Supplementary Fig. 4f), but was largely unaffected for the rest of the transcriptome (Fig. 2e). Second, we co-injected mRNAs encoding *nanog* and *soxB1* with the morpholinos and were able to rescue gastrulation (Fig. 2c, Supplementary Fig. 4c-e). Together, these results show that Nanog, Sox19b and Pou5f1 regulate progression through zygotic development and gastrulation.

Illumina sequencing revealed that combined loss of Nanog, SoxB1 and Pou5f1 results in widespread reduction in first-wave gene expression by 4hpf: 77% for strictly zygotic genes,

50% for M+Z genes. (Fig. 3a, b, Supplementary Fig. 5). By 6hpf, expression loss is systemic, with 86% of strictly zygotic and 79% of M+Z genes failing to be expressed to WT levels (Fig. 3a, b, Supplementary Fig. 5), an effect that was rescued by providing back the cognate mRNAs (Fig. 3c, Supplementary Fig. 5, 6). Comparing the single and double loss-of-function transcriptomes to the triple, we found that regulation is often combinatorial and redundant, with Nanog LOF having the strongest effect and SoxB1 the weakest (Fig. 3d, Supplementary Fig 7a-c). By 6hpf, affected genes include housekeeping genes, general transcription factors (e.g., *gata6*, *otx1*, *irx1b*, *ntla*) and major signaling components in gastrulation, anterior-posterior axis and dorsal-ventral axis specification (e.g., *oep*, *fgf3*, *wnt11*, *chd*, *nog1*, *ndr2*, *bmp2b*) (Supplementary Fig. 7d,e). Together, these results show that Nanog, Pou5f1 and SoxB1 play a fundamental role in activating the first wave, an effect that propagates to subsequent waves resulting in a global impact on zygotic gene expression.

### miR-430 is strongly activated by Nanog

Notably, among the first-wave genes co-regulated by Nanog, Pou5f1 and SoxB1 was miR-430, a miRNA that functions in the clearance of maternal mRNAs in zebrafish and *Xenopus*<sup>10-12</sup>. Northern analysis revealed a strong reduction of mature miR-430 levels in Nanog loss-of-function embryos (Fig. 4a). Although individual loss of SoxB1 or Pou5f1 had no detectable effect on miR-430 expression, when combined with Nanog LOF they reduced miR-430 levels even further, a phenotype that was rescued by co-injecting the respective mRNAs (Fig 4a-c). Nanog MO embryos failed to repress a GFP-reporter of endogenous miR-430 activity<sup>26</sup>, consistent with Nanog's role in activating miR-430 (Supplementary Fig. 8a,b).

To determine whether Nanog specifically binds the miR-430 genomic locus, we analyzed Nanog chromatin immunoprecipitation sequencing (ChIP-Seq) data at high (3.3hpf) and dome stage (4.3hpf)<sup>24</sup>. Consistent with widespread Nanog regulation, 74% of first-wave genes are bound by Nanog, a significant enrichment compared to subsequent-wave genes (Fig. 4d, Supplementary Fig. 9a). miR-430 is expressed from a 17kb genomic region on chromosome 4, which includes 55 repeated miR-430 hairpin sequences. Because this locus is repetitive, it had been excluded from previous analyses; however, the sequences are largely unique relative to the rest of the genome. Reads aligning the miR-430 locus were enriched >16-fold in the Nanog IP compared to whole cell extract (WCE), (Fig. 4e), indicating that strong Nanog binding throughout the locus correlates with strong miR-430 expression at ZGA. When the reads were aligned to the presumptive 5' end of the polycistron, we observed a strong peak of binding in a ~600 nt region between two miR-430 precursors, which contains 3 canonical Nanog binding sites (**CATT[T/G][T/G]CA**)<sup>24,27</sup>.

To determine whether Nanog induces clearance of maternal mRNAs through activation of miR-430, we analyzed the expression of an endogenous miR-430 target, *cd82b*<sup>10</sup>. *cd82b* mRNA is maternally deposited and cleared in WT by 6hpf (Fig. 5a). In contrast, *cd82b* mRNA is stabilized in *MZdicer* mutants or  $\alpha$ -amanitin treated embryos, which lack miR-430 processing and expression respectively. Similar loss of regulation is observed in Nanog+SoxB1 MO, as well as triple LOF embryos, a defect that is rescued by providing the cognate mRNAs (Fig 5b, Supplementary Fig. 8c). To determine the global effect of this

regulation, we examined RNA-Seq levels of maternal mRNAs containing miR-430 target sites. Loss of Nanog alone or in combination with loss of SoxB1 and *MZpou5f1*, resulted in miR-430 target stabilization, identical to *MZdicer*<sup>10,16,26</sup> (Fig. 5c, Supplementary Fig. 8d-f) ( $P < 1E-51$ , two-sided Wilcoxon rank sum test). A significant, but weaker effect was observed in *Pou5f1+SoxB1* LOF embryos ( $P < 1E-25$ ) (Supplementary Fig. 8d). These results show that Nanog together with Pou5f1 and SoxB1 activate miR-430 expression, thus revealing a genetic network that links maternal regulation of zygotic gene expression to zygotic clearance of maternal mRNAs.

## Discussion

Our transcriptome analysis during the maternal-to-zygotic transition provides three major insights. First, maternal factors directly regulate hundreds of mRNAs that constitute the first wave of zygotic transcription. These targets are activated in the absence of zygotic gene function and are enriched for genes that guide early embryonic development. Transcriptional competence coincides with changes in the chromatin and DNA methylation states of the genome<sup>4-8</sup>. Modifications to the epigenetic landscape during the MZT may be sufficient to allow basal levels of transcription; however, we show here that maternal transcription factors play a vital role in shaping transcriptional output.

Second, we observe that Nanog, SoxB1 and Pou5f1, previously implicated in the maintenance of pluripotency, contribute to widespread activation of zygotic genes during the MZT. These maternal factors enhance transcriptional activation of more than 74% of first-wave zygotic genes, and by 6hpf influence expression of >80% genes over all. Simultaneous removal of Nanog with SoxB1 and/or Pou5f1 results in complete block of gastrulation and developmental arrest, similar to global inhibition of zygotic gene expression (Fig. 2c, Supplementary Fig. 9c). Nanog binds 74% of first-wave genes during the early stages of ZGA (Fig. 4d). Additionally, while this manuscript was under review, Pou5f1 and Sox2 were also shown to associate with ~40% of early zygotic genes<sup>28</sup>. However, SoxB1+Pou5f1 LOF is insufficient to block gastrulation and zygotic development<sup>28</sup> (Fig. 2c). This highlights the central role of Nanog, which together with Pou5f1/SoxB1 initiates the zygotic program of development, though it is likely that additional factors cooperate with them to provide genome competency and regulate the timing of ZGA<sup>4</sup>. In mouse, Oct4 and Nanog have been proposed to regulate gene expression at 2-cell stage<sup>29,30</sup> and along with Sox2 are required for specification of the blastocyst lineages<sup>31-33</sup>. In fact, when we analyze early zygotic genes in mouse, we find that they are enriched for Nanog, Oct4 and Sox2 binding in embryonic stem cells (Supplementary Fig. 9b). Conceptually and mechanistically, many parallels exist between the MZT and the cellular reprogramming that occurs in induced pluripotent stem cells (iPSCs)<sup>3,12</sup>. Indeed, reprogramming of terminally differentiated cells was first shown in the context of the early embryo through nuclear transfer<sup>34,35</sup>. The onset of zygotic development can be viewed as a major reprogramming event that occurs upon fusion of two terminally differentiated cells (sperm and oocyte). As shown in ES cells and iPSCs, Pou5f1, Nanog and Sox2 are central players in the induction<sup>36-40</sup> and maintenance<sup>41-43</sup> of pluripotency in vivo and in vitro<sup>17,35</sup>. In these contexts, part of their role is to serve as “pioneering” factors, binding to silent chromatin to facilitate de novo gene expression<sup>44</sup>. This pioneering activity is likely recapitulated during the MZT, where an

endogenous function of Nanog, SoxB1 and Pou5f1 is to mediate activation of the first wave of zygotic genes, establishing a transient pluripotent state.

Third, we show that Nanog together with SoxB1 and Pou5f1 directly regulates miR-430, which is responsible for clearance of maternal mRNAs during the MZT<sup>10-12</sup>, facilitating the transfer of developmental control to the zygotic program (Supplementary Fig. 9c). Members of the conserved miR-430/295/302/372 family of miRNAs stabilize self-renewal fate in ES cells and enhance reprogramming efficiency<sup>45,46</sup>. We hypothesize that in both cases, these miRNAs are ‘clearing the slate’ by accelerating the removal of mRNAs from the previous program, thus facilitating the establishment of new transitional states by reprogramming factors<sup>12</sup>. The dramatic upregulation of miR-430 expression by Nanog, SoxB1 and Pou5f1 provides a central link between the mechanisms that drive zygotic gene activation and the clearance of the previous maternal history.

## Methods

### Zebrafish maintenance

*MZpou5f1*<sup>hi349Tg/hi349Tg</sup><sup>48</sup> were generated as previously described by <sup>21</sup>. Embryos obtained from natural crosses between homozygous *MZpou5f1*<sup>hi349Tg/hi349Tg</sup> mutants were injected with 30pg of *pou5f1* mRNA at the one-cell stage. *MZdicer*<sup>hu896/hu896</sup> fish were generated as described in <sup>26</sup>. Zebrafish wild type embryos were obtained from natural crosses of TU-AB and TLF strains of mixed ages (5-17 months). Selection of mating pairs was random from a pool of 60 males and 60 females allocated for a given day of the month. Fish lines were maintained in accordance with AAALAC research guidelines, under a protocol approved by Yale University IACUC.

### Treatments and mRNA injection

Embryos from all wild type crosses were pooled following collection and distributed equally between experimental conditions. Unless otherwise stated, a minimum of 30 wild type embryos were subjected to each treatment in each experimental replicate. Morpholinos were obtained from Gene Tools and resuspended in nuclease-free water. Unless otherwise stated, one nanoliter of MO solution was injected into dechorionated embryos at the one-cell stage. A combination of two MOs were used to target each gene in a 1:1 ratio as described in <sup>49</sup>, with one SoxB1 morpholino targeting a conserved region of both *sox2* and *sox3*. Nanog and SoxB1 MOs were previously described in <sup>24</sup> and <sup>49</sup> respectively. For individual and combinatorial loss-of function, wild type and *MZpou5f1* embryos were injected with 1ng of each SoxB1 MO (0.125mM) and 5ng of Nanog MO (0.6mM). For inhibition of splicing, one MO (1.25mM) complementary to U1 and two MOs (0.6mM each) complementary to isoforms of U2 spliceosomal RNAs (U1U2) were used <sup>15,50,51</sup>. Divergence of the U2 genes in zebrafish requires the use of two different morpholinos to block activity.

Zebrafish Nanog and SoxB1 capped mRNA was generated by in vitro transcription using mMessage mMachine Sp6 Kit (Ambion) in accordance to the manufacturer’s instructions. For Nanog MO rescue, zebrafish *nanog* was cloned into a pCS2 vector and sense mutations introduced during PCR amplification (indicated in lowercase):

5' ATGGCaGAtTGGAAaATGCCgGTGAGTTAC. SoxB1 rescue constructs were kindly provided by Yusuke Kamachi<sup>49</sup>. To rescue the loss-of-function phenotype, 50pg of Nanog and 20pg of SoxB1 mRNAs were injected either individually or together into morpholino injected embryos at one-cell stage. Triple loss-of-function embryos were additionally injected with 30pg of Pou5f1 mRNA.

Pol II inhibition:  $\alpha$ -amanitin was obtained from Sigma Aldrich and resuspended in nuclease-free water. Dechorionated embryos were injected with 0.2ng of  $\alpha$ -amanitin at one-cell stage<sup>52</sup>.

Translation inhibition: wild type embryos were collected and dechorionated at one-cell stage. To allow for translation of maternal mRNAs, at 32-cell stage, embryos were transferred to media containing Cycloheximide (50 $\mu$ g/mL) (Sigma Aldrich) and incubated at 28°C. Embryos were collected and frozen in liquid nitrogen at sphere and shield stage. Total RNA was extracted using Trizol (Invitrogen) and resuspended in 10 $\mu$ L RNase-free water.

To assay miR-430 activity, a GFP reporter was used as previously described<sup>26</sup>. GFP and dsRed mRNAs were in vitro transcribed using mMessage mMachine Sp6 Kit (Ambion) in accordance to the manufacturer's instructions. Embryos were injected with 150pg of GFP reporter and 100pg of dsRed loading control at the one-celled embryos.

All phenotypes were initially assayed by one experimenter and blindly confirmed and/or imaged by another. Distribution-free statistics were used to determine significance, except for calculating RNA-Seq differential expression (see below).

### In situ hybridization

Template for in situ probes were amplified from shield stage cDNA and a T7-promoter sequence added for in vitro transcription. Primers are listed below. Antisense digoxigenin (DIG) RNA probes were generated by in vitro transcription in 20 $\mu$ L reactions consisting of 100ng purified PCR product (8 $\mu$ L), 2 $\mu$ L DIG RNA labeling mix (Roche), 2 $\mu$ L 10X transcription buffer (Roche), and 2 $\mu$ L T7 RNA polymerase (Roche) in RNase-free water and purified using a QIAGEN RNEasy kit. In situ protocol was followed as detailed in<sup>26</sup>. To reduce variability, the following conditions were combined in the same tube during in situ hybridization and recognized based on their morphology: 1) wild type and  $\alpha$ -amanitin injected embryos and 2) Nanog + SoxB1 MO with and without rescue mRNA. Prior to photo documentation, embryos were cleared using a 2:1 benzyl benzoate:benzyl alcohol solution. Images were obtained using a Zeiss stereo Discovery.V12.

### Northern analysis

To detect endogenous miR-430, ten wild type and MZ*pou5f1* embryos injected with Nanog MO and SoxB1 MO were collected at 6hpf and flash frozen in liquid nitrogen. Total RNA was extracted using Trizol (Invitrogen) and resuspended in 5 $\mu$ L RNase-free water and 5  $\mu$ L 2X loading buffer (8M urea, 50mM EDTA, 0.2mg/ml xylene cyanol, and 0.2mg/ml bromophenol blue). Northern protocol was followed as detailed in<sup>16</sup>.

## Ribosome Profiling

Fifty wild type embryos injected with one nanoliter of Nanog MO (0.6mM) and SoxB1 MO (0.125mM) and fifty non-injected embryos were collected at 64-cell stage. Embryos were lysed using 800ul of a mammalian cell lysis buffer containing 100ug/ml Cycloheximide as per the manufacturers instruction (ARTseq Ribosome Profiling Kit, RPHMR12126, Epicentre). For nuclease treatment, 3ul of ARTseq Nuclease was used. Ribosome protected fragments were run and 28-29nt fragments were gel purified as previously described in <sup>16</sup> and cloned according to the manufacturers protocol (ARTseq Kit). Illumina libraries were constructed and sequence reads analyzed as in <sup>16</sup>. Subsequent to sequencing, traces of exogenous RNA corresponding to a *nanog* antisense probe, and *ntla* sense and antisense were detected outside the expected size range. Only 28 and 29nt sense sequences were used in the analysis matching the size of the ribosome footprint.

## Reverse Transcription PCR (RT-PCR)

Total RNA from ten embryos was extracted using TRIzol (Invitrogen) at sphere and shield stage for each experimental condition. RNA was treated with TURBO DNase (Ambion) for 30 minutes at 37°C and extracted using phenol chloroform. cDNA was generated by reverse transcription with random hexamers using SuperscriptII (Invitrogen). RT-PCR reactions were carried out at an annealing temperature of 60°C for 35 cycles. Primers are listed below.

## Illumina Sequencing

Total RNA was extracted as above, and strand-specific TruSeq Illumina RNA sequencing libraries were constructed by the Yale Center for Genome Analysis. Prior to sequencing, samples were treated with Epicentre Ribo-Zero Gold kits according to the published protocol, in order to deplete ribosomal RNA. Samples were multiplexed on Illumina HiSeq 2000 / 2500 machines to produce single-end 76 nt reads.

Raw reads were initially filtered by aligning permissively to a ribosomal DNA index using Bowtie v0.12.9 <sup>53</sup> with switches --seedlen 25 -n 3 -k 1 -y -e 10000. Unaligned reads were then aligned to the zebrafish Zv9 (UCSC danRer7) genome sequence using Tophat v2.0.7 <sup>54</sup> with default parameters.

Hybrid gene models were constructed from the union of zebrafish Ensembl r70, RefSeq annotations (downloaded from genome.ucsc.edu on 2/8/2013), and Ensembl RNA-Seq gene models <sup>55</sup>. All overlapping transcript isoforms were merged in order to produce maximal exonic annotations. To quantify exonic expression levels per gene, genome-uniquely aligning reads overlapping  $\geq 10$  nts to the exonic region of a given gene were summed. To quantify intronic expression levels per gene, an annotation mask was first created consisting of repetitive sequences as annotated by RepeatMasker in addition to any region aligned by  $\geq 2$  reads in the  $\alpha$ -amanitin samples; this is to minimize false positive introns due to annotation inconsistencies, under the assumption that the transcriptionally inhibited  $\alpha$ -amanitin transcriptome should contain no intron-containing transcripts. Valid intron-overlapping reads aligned the intronic region uniquely and overlapped no more than 50% to the masked regions. For the purposes of RPKM normalization, we considered intron length to be the number of unmasked nucleotides. We additionally identified reads that mapped to

at most two different genic loci (e.g., two closely related paralogs) and from these calculated “meta gene” expression values. Meta genes were treated as conventional genes for differential expression, but counted as two different genes in subsequent analyses.

The miR-430 locus is internally repetitive; therefore, reads were aligned to miR-430 in a separate step using Bowtie with switches -n 2 -k 1 on the genomic region chr4:27999472-28021845, which spans the presumed mir-430 polycistron. Reads overlapping any of the Ensembl annotated miR-430 hairpins in this region were counted as mir-430 cluster reads. Reads are counted only once, regardless of the number of times they overlap.

### Differential gene expression analysis

Differential expression analysis was performed using the R package DESeq<sup>47</sup> with the parameters fit-type = local and sharingMode = fit-only. For exonic expression comparisons, raw exon-overlapping read counts were assembled for all genes with a raw read count of at least 10 in one or more of the samples. Genes annotated as Ensembl biotypes ‘IG\_C\_pseudogene’, ‘IG\_pseudogene’, ‘IG\_V\_pseudogene’, ‘misc\_RNA’, ‘Mt\_rRNA’, ‘Mt\_tRNA’, ‘non\_coding’, ‘nonsense\_mediated\_decay’, ‘retained\_intron’, ‘rRNA’, ‘sense\_intronic’, ‘sense\_overlapping’, ‘snoRNA’, ‘snRNA’ were excluded. Additionally, all Ensembl miR-430 annotations were excluded, and a meta “miR-430 hairpin” gene added in, based on the quantification described in the previous section. For intronic expression comparisons, since overall counts are lower, variance models for DESeq were calculated using both intronic counts and exonic counts as separate gene entries (i.e., at most 1 intronic count entry and 1 exonic count entry per gene). Differential expression proceeded as normal, except multiple test correction of p values was applied relative only to the intronic counts.

Six sets of differential expression analyses were performed separately: exons and introns for each of (Group 1) WT 64c, WT Sphere, WT Shield,  $\alpha$ -amanitin 4hpf, and  $\alpha$ -amanitin 6hpf, with the two  $\alpha$ -amanitin conditions serving as pseudo replicates for DESeq for variance estimation; (Group 2) Sphere stage WT, -Nanog, -Nanog-Sox, -Pou, -Nanog-Sox-Pou, and 2 biological replicate shield stage WT samples for variance estimation; (Group 3) Shield stage WT, -Nanog, 2 -Nanog-Sox conditions treated as non replicates, -Pou, -Sox-Pou, -Nanog-Sox-Pou, and 2 additional biological replicate shield stage WT samples to parallel Group 2. For Groups 2 and 3, we applied an exonic RPKM  $\geq 1$  and intronic RPKM  $\geq 0.5$  threshold in one or more of the samples.

Zygotic transcription was determined based on significant exon and intron increases in sphere and shield stages relative to  $\alpha$ -amanitin. 64c (pre-MBT) was used as further confirmation when no significant changes in intron level were detected or the gene was intronless (genes with  $<10$  nts of unmasked intron sequence were considered effectively intronless). Increases in either exon signal, intron signal, or both determined positive zygotic transcription. For genes with a maternal contribution, increases in intronic signal due to zygotic transcription can be accompanied by no change or decreases in exonic signal. For genes significantly expressed, zygotic expression contribution is estimated using either intronic RPKM level; or the RPKM difference between the post-MZT condition and the

maximum of 64c and  $\alpha$ -amanitin expression levels. Expression calls are provided in Supplementary Data Table 1.

To define first-wave genes, genes that were detected as transcribed in the U1U2 MO treated embryos above an expression level of 5 RPKM were considered to be first wave, using an estimate for zygotic transcription based on intronic signal for multi-exon genes, or comparison to  $\alpha$ -amanitin and 64c for single-exon genes as described above. Although a cut off of 5 RPKM was used for the main analyses, lower levels of transcription were observed for many genes, indicating weaker degrees of activation. Genes that were not called as transcribed in wild type sphere were removed.

### Classification of loss of function expression categories

Significant changes in loss of function conditions relative to wild type were determined using either intron or exon signal, depending on the pattern of signal originally used to call the gene as zygotically expressed. For genes with no maternal contribution, decreases in either exon levels relative to wild type are considered to be loss of zygotic expression, while increases in either exon or intron levels are considered to be ectopic increases in zygotic expression. For genes with maternal contribution, we distinguish between two cases: (1) if zygotic transcription was originally detected in wild type only using intronic signal, then loss of zygotic transcription in the loss of function conditions is called only when intronic signal is lost; (2) if zygotic transcription was originally detected in wild type with both exonic and/or intronic signal, then decreases in either intronic levels or exonic levels indicate loss of zygotic expression, with intronic signal taking precedence when the directions of change disagree. For LOF embryos with the *MZpou5f1* genotype, differential expression was additionally performed between uninjected and injected *MZpou5f1* conditions, and expression differences between the injected conditions and wild type were required to be transitively consistent -- e.g., if a gene is called significantly lower in uninjected *MZpou5f1* than wild type, and a gene is significantly lower in injected *MZpou5f1* than uninjected *MZpou5f1*, then the gene must also be considered lower in the injected compared to wild type. To ensure that expression level differences in the *MZpou5f1* background are due to zygotic contributions, in addition to relying on intron signal, we filtered out any genes that were previously reported to be differentially maternally provided in *MZpou5f1*<sup>19</sup>.

### ChIP-Seq analysis

Reanalysis of previously published Nanog ChIP-Seq data (GSE34683) was performed as described in<sup>24</sup>, except using the current version of the zebrafish genome, Zv9. For miR-430 locus alignment, reads were aligned exhaustively to the region chr4:27994413-28019085 (2kb +/- the miR-430 polycistron) using Bowtie with parameters -v 1 --best --strata --all. To estimate read depth and enrichment, reads were normalized by the number of times the read aligned the genome. To focus on the maximally non-redundant region in the locus, reads were preferentially aligned closest to the presumptive 5' boundary of the polycistron (chr4:28000732, corresponding to the 5' end of ENSDARG000000082539).

**Morpholino Oligonucleotide sequence**

sox2-MO2	CTCGGTTTCCATCATGTTATACATT
sox3-MO1	TACATTCTTAAAAGTGGTGCCAAGC
sox3-MO2	GAAGTCAGTCAAAAAGTTCAGAGAGC
sox19a-MO1	GTACATGGCTGCCAACAGAAGTTAG
sox19a-MO2	AAAACGAGAGCGAGCCGTCTGTAAC
sox19b-MO1	GTACATCATGCCACTTCTCGCTTTG
sox19b-MO2	ACGAGCGAGCCTAATCAGGTCAAAC
nanog-MO1	CTGGCATCTTCCAGTCCGCCATTTC
nanog-MO2	AGTCCGCCATTCGCCGTTAGATAA
U1-MO1	GGTATCTCCCCTGCCAGGTAAGTAT
U2-MO1	TGATAAGAACAGATACTACTTGA
U2-MO2	TATCAGATATTAACGATAAGAAC

**In situ Primers**

ntl	TGGAAATACGTGAACGGTGA	*GTACGAACCCGAGGAGTGAA
isg15	AGAAGGGCCAGGTCAAAACT	*CATCACGGCATTGAAAACAC
cebpb	GTATGCAAGCAGCCAGTCAA	*TGACTCGTCGCTGTCCTTG
cldne	TGGTGTCTATGTGCCGAGAG	*CGGCTGGGAGTATTTTCATGT
krt18	ATCACCGCCTAAGAAAGGT	*TCGTACTCCTGCGTCTGATG
foxa3	CTTCAACGATTGCTTCGTCA	*CATCTTCTGCTCGTTGGAC
vent	ACCCAGCAAGTTCTCAGTGG	*TAGCAGCGTGTGAACAGCAT
nnr	CAGAGATGGACAGCGATTCA	*TTCGTTTCCTTCTGGGAGTTT
blf	GTCTCACAAGCGAATCCACA	*GTGTGGGTCTTCTCGTGGTT

**RT-PCR Primers**

nnr	AGCGTTTACAGCGGATCTCA	*AGTGGACGGGGAAATAAACC
isg15	CGAAAGCCTCATTAGCAAC	*GTGCAACTTCATGCCAGACTC
cldne	TGGTGTCTATGTGCCGAGAG	*CGGCTGGGAGTATTTTCATGT
sox11a	CGAAACGGACAGCATGTCTA	GGAGTCGTCATCGTCGCTT
grh13 (1/2)	GAGGAGACCGGATACAAACT	CCAAGTCCACTGTGTTTGT
grh13 (1/3)	GAGGAGACCGGATACAAACT	TTGTAATGCTGCTCTCACG
cldnb	ACTCCCCATGTGGAAAGTCA	GGGGTTGCGTTGTATTTAGC
krt4	GCAACCTCCTCCACTCACTC	AATTGTGGGGTCAATTTCCA
hist1h2aa	CAAAGGCTAAGACTCGCTCCT	TCTGTCTTCTTGGGCAGCAG
tubb4b	AGGTCTGGTCCATTGGTCA	CATCCAGAACGGAATCAACC
klf4b	ACAGTTGTGAATTCCTGGATG	GTTTACATGTGCTCTTCATGTG
vox	GACTGGCTTGTCTCAGAGCTT	GGCCGCTTCACTCTCATAAC
tbx16	AACCTTTACCTTCCCCGAGA	CAAGACTCGGGACTCAAAGC

**qRT-PCR Primers**

blf	CCCTGCTGAGCTTGCATAGT	CCCACACTGAGGACACTTGA
cldne	GGCTTCTTGGGAGCCATTAT	GCGAAAAAGCTGACGATGAT
ctcf	GTTAGCAGAGGCTTGCTTTACTG	GCAGTGAAATTTGCCACA

dact1	AGCCTCGGTTCTTCTTCA	GGAGGATTTGTGCAAGTGGT
dusp1	CTCCAGTAATGTGCGCTTCA	TGGTCGAACTTTTGACCTTCA
ef1a	TGATCTACAAATGCGGTGGA	CAATGGTGATACCACGCTCA
her5	CCAAGCCTTCATGGAGAAA	TAGCTCTGACGTTTGCATGG
mtATP6	CTTTAGCGGCCACAAATGAG	ATGGGGGTTCTTCTGGTAA
mtND5	TTCTTATGCTCAGGGGCAAT	TTAGGGCTCAGGCGTTAAGA
mxtx1	GAAATGCAAGGGTGGA AAAA	ACCCAGTTAGGAGGCATCT
oep	TTCTGGAAAGCCAAAGCAAT	TCATGTCAGTGTGCAGCTTG
pcf11	CCTCGCTGGAAGATCTGACT	CATGTTACAGGCCTCATGTCA
tdp2b	GGAGCCCACCTGCTCTATTA	ACCCTGCCAATTGTGAAGATA

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Y. Kamachi, W. Driever, A.F. Schier, N. Ivanova, and I-H. Park for reagents and fish lines; S. Mane and J. Overton for outstanding sequencing support; E. Zdobnov for hosting genomics data; A. Hubaud, N. Darricarrere and M. Koziol for contribution in the early stages of this project, and all the members of the Giraldez lab for insightful discussions. Supported by NIH grants F32HD071697-02 (M.T.L.), T32GM007499 (A.R.B), F32HD061194-03 (C.M.T.), Pew Fellows Program in Biomedical Sciences (A.A.B), R01GM081602-06, R01GM103789-01, R01HD074078-02, the Pew Scholars Program in the Biomedical Sciences and the Yale Scholars Program (A.J.G.).

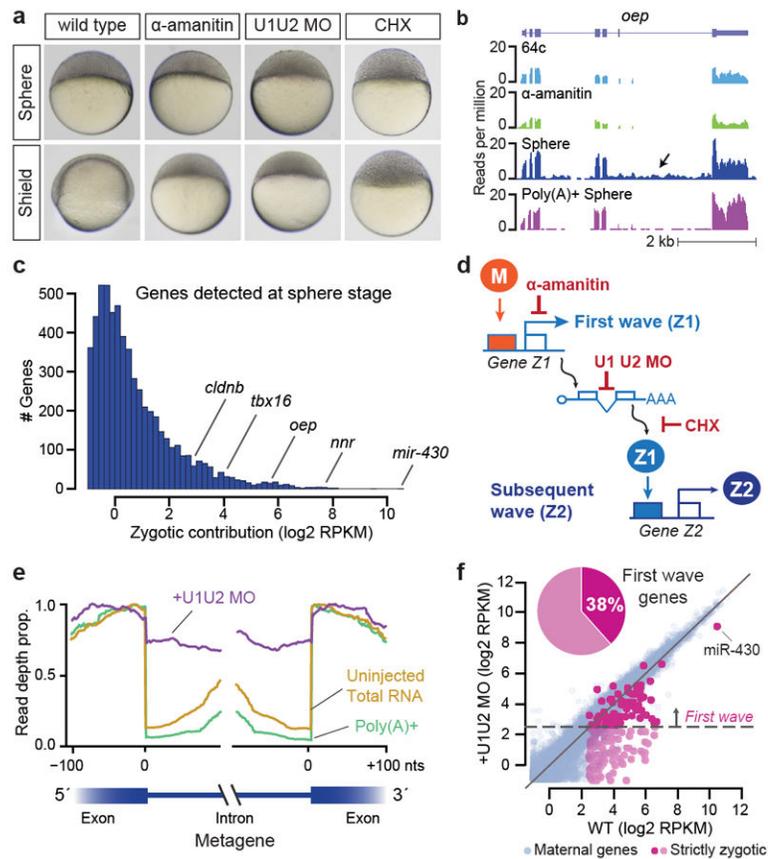
## Bibliography

- ten Bosch JR, Benavides JA, Cline TW. The TAGteam DNA motif controls the timing of *Drosophila* pre-blastoderm transcription. *Development* (Cambridge, England). 2006; 133:1967–1977.10.1242/dev.02373
- Liang HL, et al. The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. *Nature*. 2008; 456:400–403. [PubMed: 18931655]
- Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development* (Cambridge, England). 2009; 136:3033–3042.
- Andersen IS, et al. Epigenetic marking of the zebrafish developmental program. *Curr Top Dev Biol*. 2013; 104:85–112.10.1016/B978-0-12-416027-9.00003-6 [PubMed: 23587239]
- Vastenhouw NL, et al. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature*. 2010; 464:922–926. [PubMed: 20336069]
- Lindeman LC, et al. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. *Developmental cell*. 2011; 21:993–1004.10.1016/j.devcel.2011.10.008 [PubMed: 22137762]
- Potok ME, Nix DA, Parnell TJ, Cairns BR. Reprogramming the Maternal Zebrafish Genome after Fertilization to Match the Paternal Methylation Pattern. *Cell*. 2013; 153:759–772. [PubMed: 23663776]
- Jiang L, et al. Sperm, but Not Oocyte, DNA Methylome Is Inherited by Zebrafish Early Embryos. *Cell*. 2013; 153:773–784. [PubMed: 23663777]
- Kane DA, Kimmel CB. The zebrafish midblastula transition. *Development*. 1993; 119:447–456. [PubMed: 8287796]
- Giraldez AJ, et al. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* (New York, N Y. 2006; 312:75–79.10.1126/science.1122689
- Lund E, Liu M, Hartley RS, Sheets MD, Dahlberg JE. Deadenylation of maternal mRNAs mediated by miR-427 in *Xenopus laevis* embryos. *RNA* (New York, N Y. 2009; 15:2351–2363.

12. Giraldez AJ. microRNAs, the cell's Nepenthe: clearing the past during the maternal-to-zygotic transition and cellular reprogramming. *Current opinion in genetics & development*. 2010; 20:369–375.10.1016/j.gde.2010.04.003 [PubMed: 20452200]
13. Harvey SA, et al. Identification of the zebrafish maternal and paternal transcriptomes. *Development (Cambridge, England)*. 2013; 140:2703–2710.10.1242/dev.095091
14. Aanes H, et al. Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition. *Genome research*. 2011; 21:1328–1338.10.1101/gr.116012.110 [PubMed: 21555364]
15. Kaida D, et al. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*. 2010; 468:664–668.10.1038/nature09479 [PubMed: 20881964]
16. Bazzini AA, Lee MT, Giraldez AJ. Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science*. 2012; 336:233–237.10.1126/science.1215704 [PubMed: 22422859]
17. Chambers I, Tomlinson SR. The transcriptional foundation of pluripotency. *Development (Cambridge, England)*. 2009; 136:2311–2322.10.1242/dev.024398
18. Onichtchouk D. Pou5f1/oct4 in pluripotency control: insights from zebrafish. *Genesis*. 2012; 50:75–85.10.1002/dvg.20800 [PubMed: 21913309]
19. Onichtchouk D, et al. Zebrafish Pou5f1-dependent transcriptional networks in temporal control of early development. *Mol Syst Biol*. 2010; 6:354.10.1038/msb.2010.9 [PubMed: 20212526]
20. Okuda Y, Ogura E, Kondoh H, Kamachi Y. B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. *PLoS genetics*. 2010; 6:e1000936.10.1371/journal.pgen.1000936 [PubMed: 20463883]
21. Lunde K, Belting HG, Driever W. Zebrafish pou5f1/pou2, homolog of mammalian Oct4, functions in the endoderm specification cascade. *Curr Biol*. 2004; 14:48–55. [PubMed: 14711414]
22. Hauptmann G, et al. spiel ohne grenzen/pou2 is required for zebrafish hindbrain segmentation. *Development (Cambridge, England)*. 2002; 129:1645–1655.
23. Belting HG, et al. Pou5f1 contributes to dorsoventral patterning by positive regulation of vox and modulation of fgf8a expression. *Developmental biology*. 2011; 356:323–336.10.1016/j.ydbio.2011.05.660 [PubMed: 21621531]
24. Xu C, et al. Nanog-like regulates endoderm formation through the Mxtx2-Nodal pathway. *Developmental cell*. 2012; 22:625–638.10.1016/j.devcel.2012.01.003 [PubMed: 22421047]
25. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science (New York, N Y)*. 2009; 324:218–223.
26. Giraldez AJ, et al. MicroRNAs regulate brain morphogenesis in zebrafish. *Science (New York, N Y)*. 2005; 308:833–838.10.1126/science.1109020
27. Loh YH, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics*. 2006; 38:431–440.10.1038/ng1760 [PubMed: 16518401]
28. Leichsenring M, Maes J, Mossner R, Driever W, Onichtchouk D. Pou5f1 transcription factor controls zygotic gene activation in vertebrates. *Science (New York, N Y)*. 2013; 341:1005–1009.10.1126/science.1242527
29. Foygel K, et al. A novel and critical role for Oct4 as a regulator of the maternalembryonic transition. *PLoS ONE*. 2008; 3:e4109.10.1371/journal.pone.0004109 [PubMed: 19129941]
30. Tan MH, et al. An Oct4-Sall4-Nanog network controls developmental progression in the pre-implantation mouse embryo. *Molecular systems biology*. 2013; 9:632.10.1038/msb.2012.65 [PubMed: 23295861]
31. Keramari M, et al. Sox2 is essential for formation of trophoderm in the preimplantation embryo. *PLoS ONE*. 2010; 5:e13952.10.1371/journal.pone.0013952 [PubMed: 21103067]
32. Frum T, et al. Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Developmental cell*. 2013; 25:610–622.10.1016/j.devcel.2013.05.004 [PubMed: 23747191]
33. Messerschmidt DM, Kemler R. Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. *Developmental biology*. 2010; 344:129–137.10.1016/j.ydbio.2010.04.020 [PubMed: 20435031]

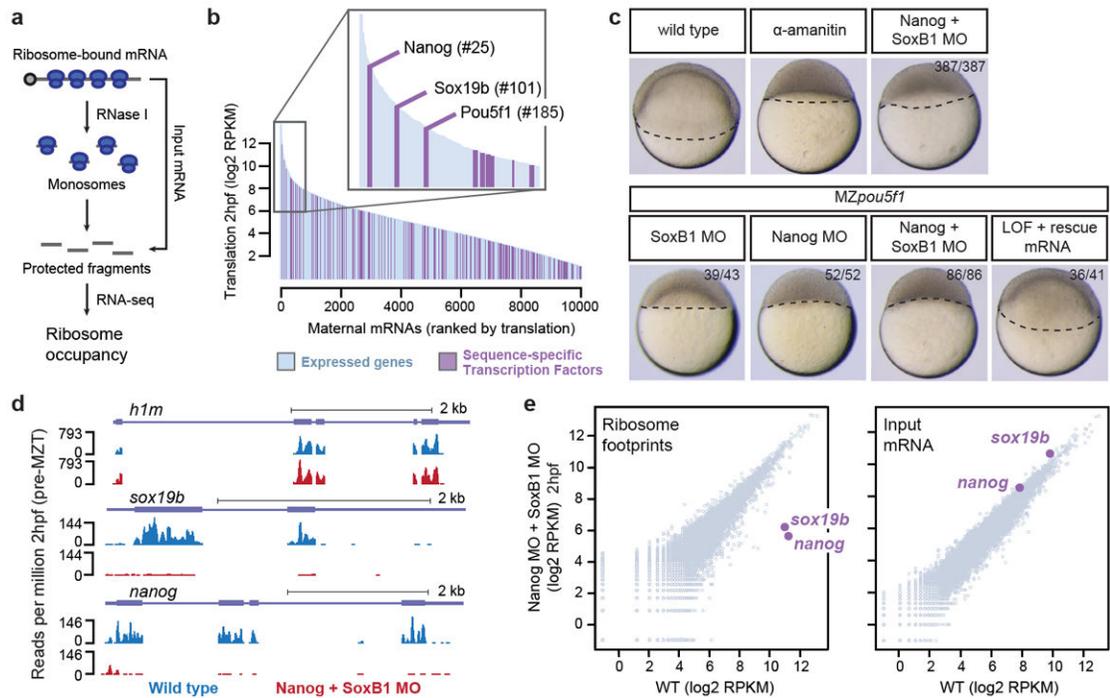
34. Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature*. 1958; 182:64–65. [PubMed: 13566187]
35. Gurdon JB, Melton DA. Nuclear reprogramming in cells. *Science (New York, N Y)*. 2008; 322:1811–1815.10.1126/science.1160810
36. Mitsui K, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*. 2003; 113:631–642. [PubMed: 12787504]
37. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126:663–676. [PubMed: 16904174]
38. Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007; 131:861–872. [PubMed: 18035408]
39. Wernig M, et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*. 2007; 448:318–324.10.1038/nature05944 [PubMed: 17554336]
40. Yu J, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N Y)*. 2007; 318:1917–1920.10.1126/science.1151526
41. Chambers I, et al. Nanog safeguards pluripotency and mediates germline development. *Nature*. 2007; 450:1230–1234.10.1038/nature06403 [PubMed: 18097409]
42. Masui S, et al. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol*. 2007; 9:625–635.10.1038/ncb1589 [PubMed: 17515932]
43. Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet*. 2000; 24:372–376.10.1038/74199 [PubMed: 10742100]
44. Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*. 2012; 151:994–1004.10.1016/j.cell.2012.09.045 [PubMed: 23159369]
45. Judson RL, Babiarz JE, Venere M, Blelloch R. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nature biotechnology*. 2009; 27:459–461.
46. Subramanyam D, et al. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat Biotechnol*. 2011; 29:443–448.10.1038/nbt.1862 [PubMed: 21490602]
47. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome biology*. 2010; 11:R106.10.1186/gb-2010-11-10-r106 [PubMed: 20979621]
48. Amsterdam A, et al. Identification of 315 genes essential for early zebrafish development. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:12792–12797.10.1073/pnas.0403929101 [PubMed: 15256591]
49. Iwafuchi-Doi M, et al. The Pou5f1/Pou3f-dependent but SoxB-independent regulation of conserved enhancer N2 initiates Sox2 expression during epiblast to neural plate stages in vertebrates. *Developmental biology*. 2011; 352:354–366.10.1016/j.ydbio.2010.12.027 [PubMed: 21185279]
50. Matter N, Konig H. Targeted 'knockdown' of spliceosome function in mammalian cells. *Nucleic acids research*. 2005; 33:e41.10.1093/nar/gni041 [PubMed: 15731334]
51. Amsterdam A, et al. Identification of 315 genes essential for early zebrafish development. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:12792–12797.10.1073/pnas.0403929101 [PubMed: 15256591]
52. Iwafuchi-Doi M, et al. The Pou5f1/Pou3f-dependent but SoxB-independent regulation of conserved enhancer N2 initiates Sox2 expression during epiblast to neural plate stages in vertebrates. *Developmental biology*. 2011; 352:354–366.10.1016/j.ydbio.2010.12.027 [PubMed: 21185279]
53. Matter N, Konig H. Targeted 'knockdown' of spliceosome function in mammalian cells. *Nucleic acids research*. 2005; 33:e41.10.1093/nar/gni041 [PubMed: 15731334]
54. Rosel TD, et al. RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation. *The EMBO journal*. 2011; 30:1965–1976.10.1038/emboj.2011.106 [PubMed: 21468032]
55. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dyn*. 1995; 203:253–310. [PubMed: 8589427]

56. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*. 2009; 10:R25.10.1186/gb-2009-10-3-r25 [PubMed: 19261174]
57. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)*. 2009; 25:1105–1111.
58. Collins JE, White S, Searle SM, Stemple DL. Incorporating RNA-seq data into the zebrafish Ensembl genebuild. *Genome Res*. 2012; 22:2067–2078.10.1101/gr.137901.112 [PubMed: 22798491]



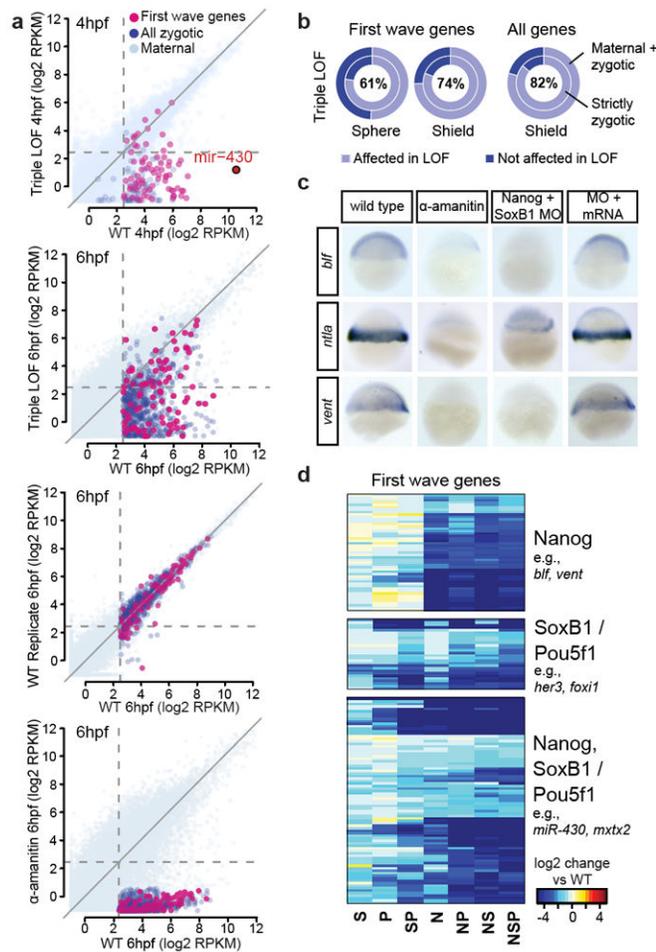
### Figure 1. Characterization of the zygotic transcriptome

**a.** Embryos showing the effects of  $\alpha$ -amanitin, U1U2 morpholino (U1U2 MO) and cycloheximide (CHX). **b.** Sequencing read density across *oep*. Intronic signal increases with zygotic expression in total RNA. **c.** Expression histogram of zygotic genes. **d.** Maternal (M) but not zygotic factors (Z1) can activate transcription upon splice or translation inhibition. **e.** Metagene of read density across exon-intron boundaries in first-wave genes. U1U2 MO shows enriched intron signal (purple). **f.** Biplot comparing expression in wild type and U1U2 MO. Points above 5 RPKM in U1U2 MO are considered first-wave genes.



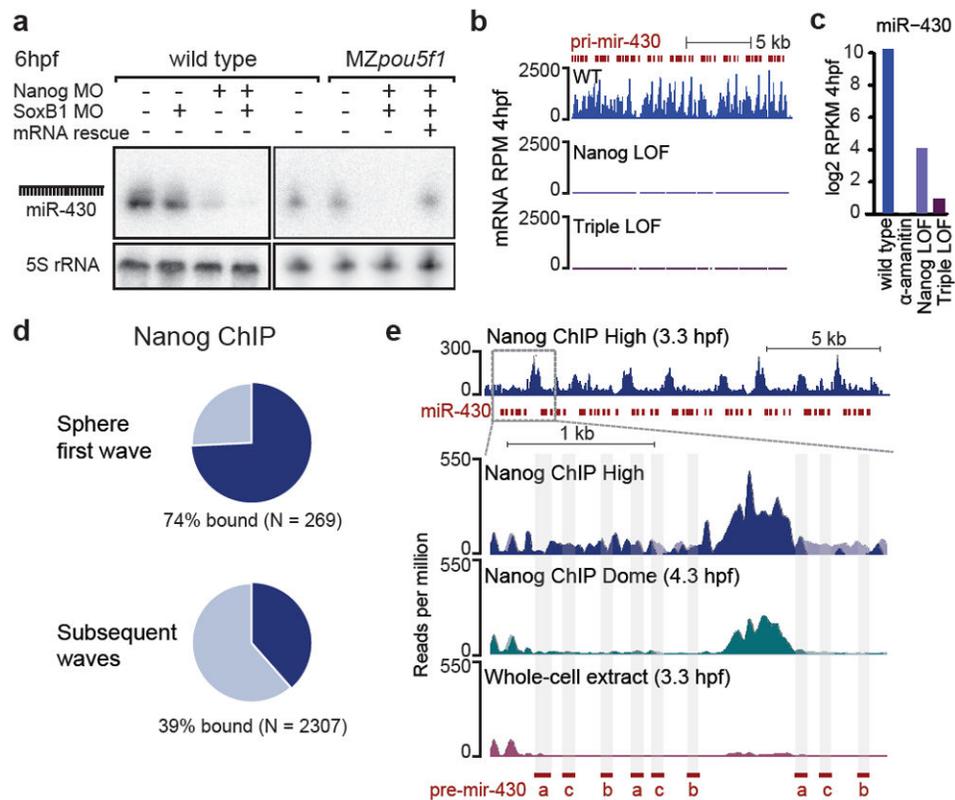
**Figure 2. Identification of Nanog, Sox19b and Pou5f1 as zygotic gene regulators**

**a.** Schematic illustrating ribosome profiling. **b.** Rank plot showing translation levels pre-MZT. Sequence-specific transcription factors are highlighted. **c.** Embryos with combined loss of Nanog+SoxB1, Nanog+Pou5f1 or triple LOF arrest similar to α-amanitin and are rescued with mRNA injection. **d.** Ribosome footprints for *h1m*, *sox19b* and *nanog* in wild type and Nanog MO + SoxB1 MO. *sox19b* and *nanog* are highly depleted in the MO conditions. **e.** Biplots comparing wild type and MO ribosome footprints and input mRNA.



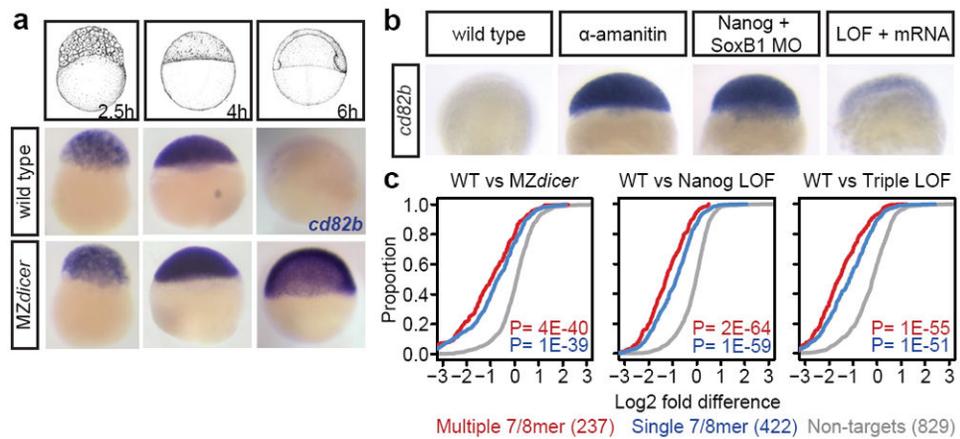
**Figure 3. Transcriptome-wide effects of loss of Nanog, SoxB1 and Pou5f1**

**a.** Biplots showing widespread gene expression loss in the triple LOF at 6hpf. **b.** Donut plots showing global effects of LOF. Percentages show the combined effect for strictly zygotic and maternal+zygotic (M+Z) gene groups. **c.** In situ hybridization shows expression defects in LOF embryos, which is rescued by mRNA injection. **d.** Heatmap showing first-wave zygotic genes in single and combined LOF conditions (N, Nanog MO; S, SoxB1 MO; P, MZ*pou5f1*). Patterns shown are regulation by Nanog predominantly; SoxB1/Pou5f1; or Nanog in combination with SoxB1/Pou5f1.



**Figure 4. miR-430 expression is regulated by Nanog**

**a.** Northern blot shows miR-430 is severely reduced in Nanog LOF and nearly undetectable in the triple LOF **b.** RNA-Seq read levels of the pri-mir-430 polycistron in wild type and LOF. **c.** Bar plot of total miR-430 aligning reads. **d.** First-wave genes are highly bound by Nanog. **e.** Nanog binding across the miR-430 region (top panel) and a zoomed region where reads are preferentially aligned to the 5' end. Binding profiles show a strong peak between two precursors. pre-miR-430a, b and c are marked in red.



### Figure 5. miR-430 activity is abrogated by Nanog LOF

**a.** In situ showing degradation of miR-430 target *cd82b* at 6hpf in wild type, compared to stabilization in *MZdicer* (lacking miR-430 activity). **b.** *cd82b* is stabilized in the Nanog-*SoxB1* LOF embryo, indicating loss of miR-430 activity. The effect is rescued with injection of *nanog* and *soxB1* mRNA. **c.** Cumulative plots showing stabilized expression of miR-430 targets in *MZdicer* and LOF embryos, compared to wild type. P values are for two-sided Wilcoxon rank sum tests comparing each miR-430 target group to non-targets.