

ARTICLE

Received 12 May 2015 | Accepted 10 Nov 2015 | Published 18 Dec 2015

DOI: 10.1038/ncomms10148

OPEN

# Embryonic transcription is controlled by maternally defined chromatin state

Saartje Hontelez<sup>1,\*</sup>, Ila van Kruijsbergen<sup>1,\*</sup>, Georgios Georgiou<sup>1,\*</sup>, Simon J. van Heeringen<sup>1</sup>, Ozren Bogdanovic<sup>2</sup>, Ryan Lister<sup>2,3</sup> & Gert Jan C. Veenstra<sup>1</sup>

Histone-modifying enzymes are required for cell identity and lineage commitment, however little is known about the regulatory origins of the epigenome during embryonic development. Here we generate a comprehensive set of epigenome reference maps, which we use to determine the extent to which maternal factors shape chromatin state in *Xenopus* embryos. Using  $\alpha$ -amanitin to inhibit zygotic transcription, we find that the majority of H3K4me3- and H3K27me3-enriched regions form a maternally defined epigenetic regulatory space with an underlying logic of hypomethylated islands. This maternal regulatory space extends to a substantial proportion of neurula stage-activated promoters. In contrast, p300 recruitment to distal regulatory regions requires embryonic transcription at most loci. The results show that H3K4me3 and H3K27me3 are part of a regulatory space that exerts an extended maternal control well into post-gastrulation development, and highlight the combinatorial action of maternal and zygotic factors through proximal and distal regulatory sequences.

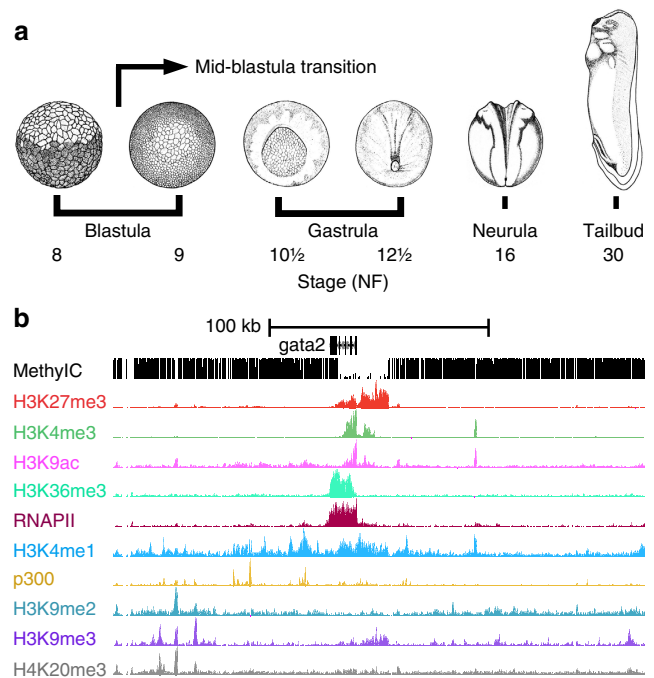
<sup>1</sup>Department of Molecular Developmental Biology, Radboud Institute for Molecular Life Sciences, Faculty of Science, Radboud University, PO Box 9101, 6500 HB Nijmegen, The Netherlands. <sup>2</sup>ARC Center of Excellence in Plant Energy Biology, The University of Western Australia, Perth, Western Australia 6009, Australia. <sup>3</sup>The Harry Perkins Institute of Medical Research, Perth, Western Australia 6009, Australia. \* These authors contributed equally to this work. Correspondence and requests for materials should be addressed to G.J.C.V. (email: g.veenstra@science.ru.nl).

During early embryonic development cells differentiate, acquiring specific transcription and protein expression profiles. Histone modifications can control the activity of genes through regulatory elements in a cell-type-specific manner<sup>1–4</sup>. Recent advances have been made in the annotation of functional genomic elements of mammalian cells, *Drosophila* and *Caenorhabditis* through genome-wide profiling of chromatin marks<sup>5,6</sup>. Immediately after fertilization, the embryonic genome is transcriptionally silent, and zygotic genome activation (ZGA) occurs after a number of mitotic cycles<sup>7</sup>. In *Drosophila* and zebrafish (*Danio rerio*) ZGA starts after 8 and 9 mitotic cycles, respectively, in mammals transcription starts at the two-cell stage<sup>8,9</sup>, whereas in *Xenopus* this happens after the first 12 cleavages at the mid-blastula transition (MBT)<sup>10–12</sup>. Permissive H3K4me3 and repressive H3K27me3 histone modifications emerge during blastula and gastrula stages<sup>13–16</sup>. To date, little is known about the origin and specification of the epigenome in embryonic development of vertebrates, which is essential for understanding physiological cell lineage commitment and differentiation.

To explore the developmental origins of epigenetic regulation we have generated epigenome reference maps during early development of *Xenopus tropicalis* embryos and assessed the need for embryonic transcription in their acquisition. We find a hierarchical appearance of histone modifications, with a priority for promoter marks which are deposited hours before transcription activation on regions with hypomethylated DNA. Surprisingly, the promoter H3K4me3 and the Polycomb H3K27me3 modifications are largely maternally defined (MaD), providing maternal epigenetic control of gene activation that extends well into neurula and tailbud stages. By contrast, p300 recruitment to distal regulatory elements is largely under the control of zygotic factors. Moreover, this maternal-proximal and zygotic-distal dichotomy of gene regulatory sequences also differentiates between early and late Wnt signalling target genes, suggesting that different levels of permissiveness are involved in temporal target gene selection.

## Results

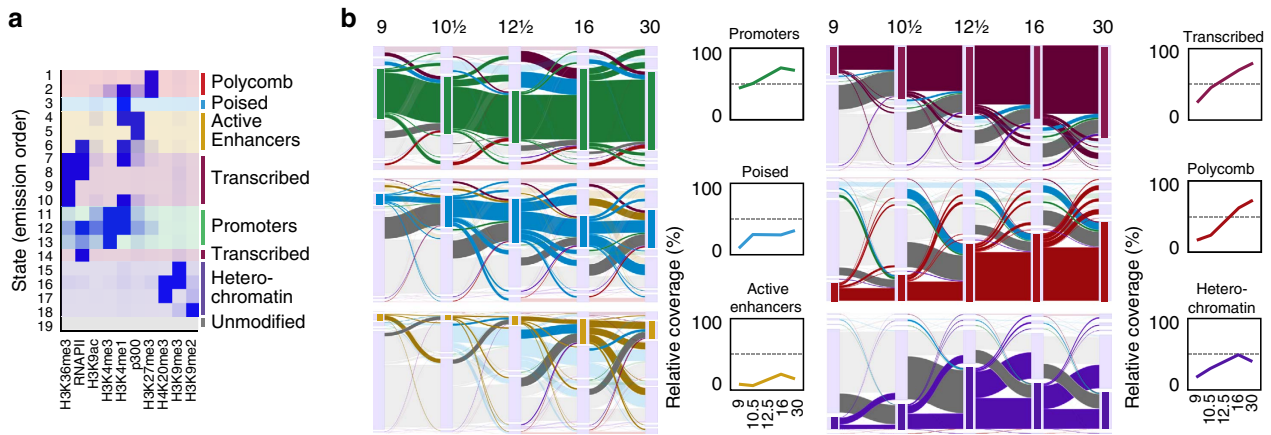
**Progressive specification of chromatin state.** We have performed chromatin immunoprecipitation (ChIP) sequencing of eight histone modifications, RNA polymerase II (RNAPII) and the enhancer protein p300 at five stages of development: blastula (st. 9), gastrula (st. 10.5, 12.5), neurula (st. 16) and tailbud (st. 30). These experiments allow identification of enhancers (H3K4me1, p300)<sup>17–20</sup>, promoters (H3K4me3, H3K9ac)<sup>14,21–23</sup>, transcribed regions (H3K36me3, RNAPII)<sup>22</sup> and repressed and heterochromatic domains (H3K27me3, H3K9me2, H3K9me3 and H4K20me3)<sup>1,14,24,25</sup>. In addition we generated pre-MBT (st. 8) maps for three histone modifications (H3K4me3, H3K9ac and H3K27me3) and single-base resolution DNA methylome maps using whole-genome bisulfite sequencing of blastula and gastrula (st. 9 and 10.5) embryos (Fig. 1; Supplementary Fig. 1). Our data set consists of 2.7 billion aligned sequence reads representing the most comprehensive set of epigenome reference maps of vertebrate embryos to date. Using a Hidden Markov Model approach<sup>26</sup> we have identified 19 chromatin states based on co-occurring ChIP signals (Fig. 2a). This analysis identifies combinations of ChIP signals at specific genomic sequences without distinguishing between overlapping histone modifications that result from regional or cell-type specificity and co-occurrence in the same cells<sup>14</sup>. Seven main groups were recognized, namely (i) Polycomb (H3K27me3, deposited by Polycomb Repressive Complex 2 (PRC2)), (ii) poised enhancers, (iii) p300-bound enhancers, (iv) transcribed regions, (v)



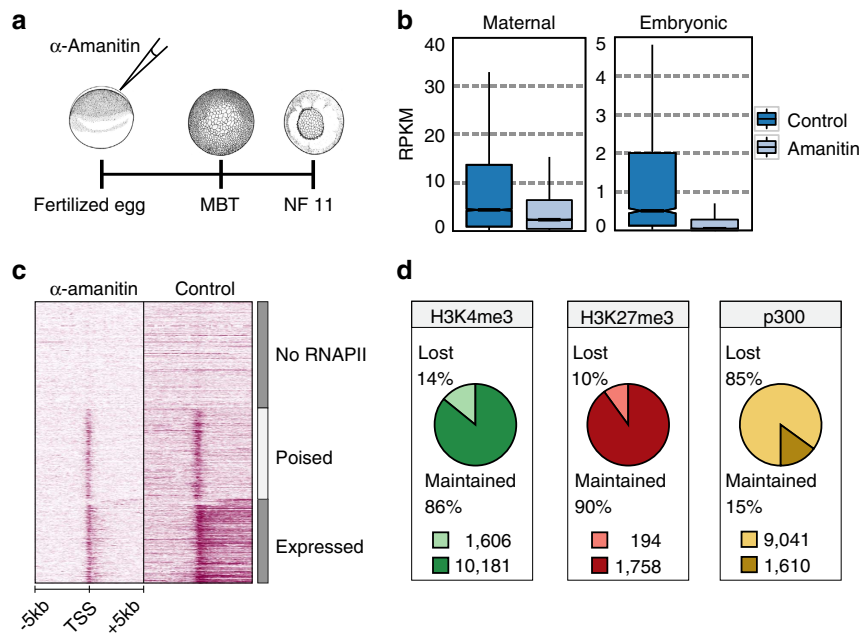
**Figure 1 | Reference epigenome maps of *Xenopus tropicalis* development.** (a) Genome-wide profiles were generated for stages 8 and 9 (blastula, before and after MBT), 10.5 and 12.5 (gastrula), 16 (neurula) and 30 (tailbud). Adapted from Tan, M.H. et al. *Genome Res.* **23**, 201–216 (2013), under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at <http://creativecommons.org/licenses/by/3.0/>. (b) *Gata2* locus with late gastrula (stage 10.5) methyl-C-seq, ChIP-seq enrichment of histone modifications, RNAPII and p300 (cf. Supplementary Figs 1 and 2).

promoters, (vi) heterochromatin and (vii) unmodified regions (Fig. 2a; Supplementary Fig. 2). Alluvial plots of state coverage per stage show that all states increase in coverage during development, except for the unmodified state (Fig. 2b; Supplementary Fig. 2a). Unmodified regions decrease in coverage during development, however, even at tailbud stage 67% of the total epigenome remains naive for the modifications and bound proteins in our data set (Supplementary Fig. 2b). Promoter coverage remains relatively constant during development from blastula to tailbud stages, in contrast to the Polycomb state which increases in coverage during gastrulation. P300-bound enhancers are highly dynamic during development (Fig. 2b). Global enrichment levels of modified regions show similar dynamics, and reveal a priority for promoter marking at or before the blastula stage, followed by enhancer activation and heterochromatic repression during late blastula and gastrulation stages (Supplementary Fig. 3a,b). A detailed time course between fertilization and early gastrulation shows that both H3K4me3 and H3K9ac emerge hours before the start of embryonic transcription (Supplementary Fig. 3c). We and others have previously reported that H3K4me3 is acquired during blastula stages<sup>14</sup>. Indeed, H3K4me3 and H3K9ac levels increase strongly before the MBT, well before embryonic transcription starts. This however raises the question to what extent histone modifications are regulated by maternal or embryonic factors.

**Maternal and zygotic epigenetic regulation.** To determine the maternal and zygotic contributions to chromatin state, we used  $\alpha$ -amanitin to block embryonic transcription (Fig. 3a).



**Figure 2 | Chromatin state dynamics.** (a) Emission states (same for all developmental stages) of the hidden Markov model, identifying the 19 most prevalent combinations of histone modifications and bound proteins. From top to bottom: Polycomb (red), Poised enhancers and promoters (blue), Active Enhancers (gold), Transcribed (dark magenta), Promoter (green), Heterochromatin (purple) and unmodified (grey). (b) Alluvial plots of chromatin state coverage during development. Each plot shows the transitions (to and from the highlighted group of chromatin states) across developmental stages (stages 9–30). The height represents the base pair coverage of the chromatin state relative to the modified genome. The ‘modified genome’ has a chromatin state other than unmodified in any of the stages 9–30. From top to bottom left: promoters (green), poised (blue), p300-bound enhancers (gold). From top to bottom right: transcribed (dark magenta), Polycomb (red) and heterochromatin (purple). Line plots: Chromatin state coverage per stage as a percentage of the modified genome.



**Figure 3 | Developmental acquisition of chromatin states.** (a) Inhibition of embryonic transcription with  $\alpha$ -amanitin, adapted from Tan, M.H. *et al.* *Genome Res.* **23**, 201–216 (2013), under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at <http://creativecommons.org/licenses/by/3.0/>. (b) RNAPII on the TSS of genes in control and  $\alpha$ -amanitin-injected embryos (stage 11). (c) Box plots showing RNA expression levels (RPKM) of maternal and embryonic transcribed genes in control and  $\alpha$ -amanitin-injected embryos (stage 11). Box: 25th (bottom), 50th (internal band), 75th (top) percentiles. Whiskers:  $1.5 \times$  interquartile range of the lower and upper quartiles, respectively. (d) ChIP-seq on chromatin of  $\alpha$ -amanitin-injected and control embryos reveals maternal and zygotic origins of H3K4me3, H3K27me3 or p300 binding. Data from two biological replicates, see Supplementary fig. 4.

$\alpha$ -Amanitin blocks the translocation of RNA polymerase II (RNAPII) on DNA, thereby preventing transcript elongation<sup>27</sup>. It is therefore expected that injection of  $\alpha$ -amanitin into embryos will stall RNAPII, immobilizing it on DNA after its recruitment to pre-initiation complexes. Indeed, both RNAPII elongation and embryonic transcription were effectively blocked in  $\alpha$ -amanitin-injected embryos (Fig. 3b,c; Supplementary Fig. 4a). New transcription is necessary for gastrulation<sup>11,28,29</sup>, but  $\alpha$ -

amanitin-injected embryos survive to the equivalent of stage 11 control embryos. ChIP sequencing of replicates of  $\alpha$ -amanitin-injected and control embryos (stage 11) revealed that the majority of H3K4me3 (86%) and H3K27me3 (90%) regions are consistently modified with these modifications independently of embryonic transcription (Fig. 3d; Supplementary Fig. 4b,c). This is especially surprising given the temporal hierarchy of H3K27me3 and H3K4me3, and the relatively late acquisition

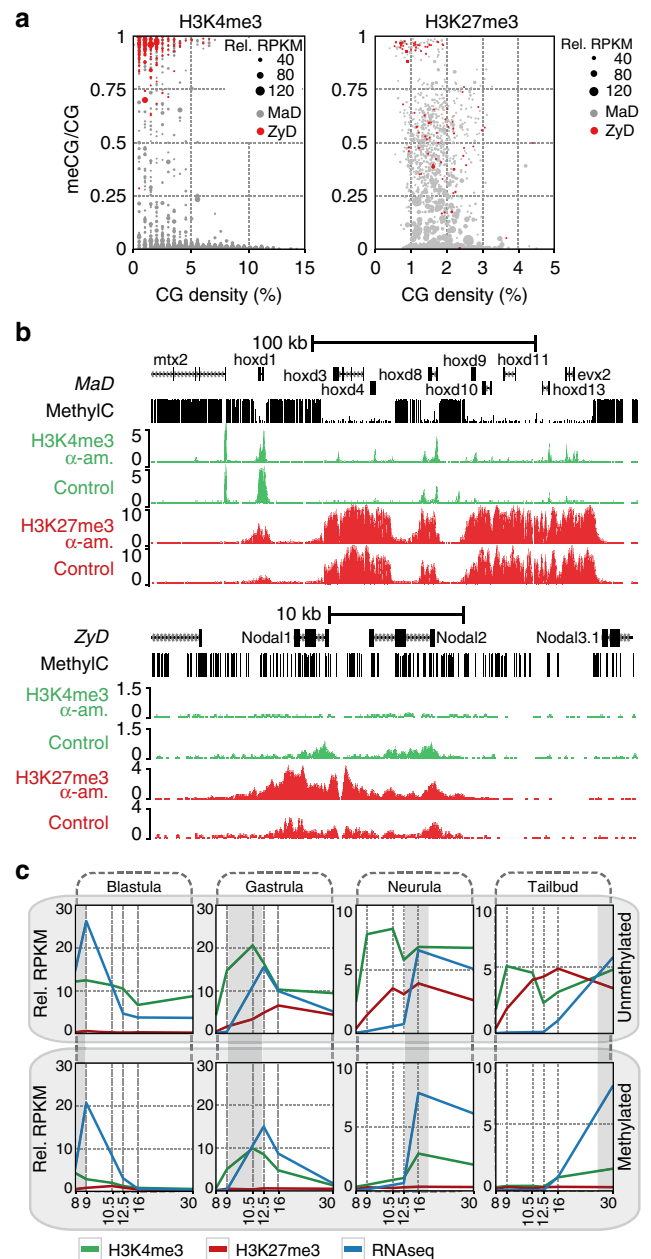
of H3K27me3 (Fig. 2b). By contrast, only 15% of the p300-bound regions recruit p300 independently of active transcription (Fig. 3d). This suggests that the promoter-permissive H3K4me3 mark and the Polycomb-repressive H3K27me3 mark are mostly controlled by maternal factors (maternally defined, MaD), whereas p300 binding to regulatory regions is largely zygotically defined (ZyD). Regions with MaD H3K4me3 and H3K27me3 acquire these modifications more robustly and also earlier during development compared with ZyD regions (Supplementary Fig. 4d). By contrast, ZyD p300-bound regions show more robust p300 recruitment during gastrulation compared with p300 MaD regions. These data show a pervasive maternal influence on the developmental acquisition of key histone modifications.

**DNA methylation logic of maternal control.** Trimethylation of H3K4 and H3K27 has been associated with CpG density and a lack of DNA methylation. The Set1 and related MLL complexes are responsible for H3K4me3 (ref. 10). Set1 is recruited to hypomethylated CpG domains via the Cxxc1 protein (Cfp1)<sup>30–32</sup>. In the absence of H3K4me3, PRC2 binding to hypomethylated CpGs results in H3K27me3 and inhibition of gene activation<sup>13,33</sup>. Using our whole-genome bisulfite sequencing data we determined that MaD H3K4me3 promoters are predominantly hypomethylated (Fig. 4a; Supplementary Fig. 5a; Supplementary Data 1). Conversely, promoters decorated with ZyD H3K4me3 almost exclusively have highly methylated promoters. Demethylation of ZyD promoters was not detected, and methylation levels of MaD and ZyD regions were similar in stage 9 and stage 10.5 (Supplementary Fig. 5a,b). In addition, H3K4me3 often extends asymmetrically from promoters into gene bodies (+1–2 kb from transcription start site (TSS); Supplementary Fig. 5c), likely representing the second and third nucleosomes that are trimethylated via RNAPII-recruited Set1 in actively transcribed genes<sup>34</sup>. Concordantly,  $\alpha$ -amanitin reduces H3K4me3 at downstream positions. Interestingly, we also find poised enhancers that gain H3K4me3 in  $\alpha$ -amanitin-injected embryos and which exhibit intermediate to high levels of DNA methylation (Supplementary Fig. 5d,e).

The majority of promoters with ZyD H3K27me3 shows intermediate to high levels of DNA methylation (Fig. 4a; Supplementary Fig. 5a; Supplementary Data 1). Some of the MaD H3K27me3 regions are methylated, but the highly enriched H3K27me3 domains (larger dots) are almost exclusively both maternally defined and hypomethylated. This is illustrated by the *hoxd* cluster which harbours a large hypomethylated domain with MaD H3K4me3 and H3K27me3 (Fig. 4b). There are also examples of reciprocal changes of H3K4 and H3K27 methylation, for example at the hypermethylated promoters of *nodal1* and *nodal2*.

ZyD p300-bound regions are generally hypermethylated, whereas MaD p300-bound regions show a variable degree of DNA methylation (Supplementary Fig. 5e). However, promoters that overlap with MaD p300 peaks are hypomethylated in 77% of the cases, whereas 96% of the promoters that are associated with ZyD p300 peaks are hypermethylated (Supplementary Fig. 5f), showing that p300-recruiting hypomethylated promoters tend to be under complete maternal control, for both H3K4 methylation and p300 recruitment.

To further explore the relationships between DNA methylation, histone modifications and developmental activation of transcription we determined correlations with different measures of gene activity such as RNA-seq and ChIP-seq of RNAPII and H3K36me3 (Supplementary Fig. 6). We find that H3K36me3 and RNAPII in gene bodies correlate well with each other but less with transcript levels (RNA-seq), presumably due to the effects of RNA stability. A much lower correlation was found between



**Figure 4 | DNA methylation logic of maternally versus zygotically defined H3K4me3 and H3K27me3.** (a) CpG density and methylation at stage 9 of promoters (H3K4me3:  $\pm 100$  bp from TSS; H3K27me3:  $\pm 2.5$  kb from TSS) that contain a zygotically defined (ZyD, lost in  $\alpha$ -amanitin treated embryos, red) or maternally defined (MaD, maintained in  $\alpha$ -amanitin treated embryos, grey) peak for H3K4me3 (left) or H3K27me3 (right) after inhibition of embryonic transcription. The size of the dot indicates the relative RPKM of the histone modification (background corrected). (b) *Hoxd* (MaD) and *nodal1*, -2 (ZyD) loci with stage 9 methylC-seq, H3K4me3 and H3K27me3 in control and  $\alpha$ -amanitin-injected embryos. (c) Developmental profiles of H3K4me3 and H3K27me3 (median background corrected RPKM) at genes without detectable maternal mRNA do correlate with activation for methylated promoters (lower panels) but not for hypomethylated CpG island promoters (upper panels).

either measure of gene activity and the promoter marks H3K4me3 and H3K9ac, especially at early stages. In part this may be caused by time delays of transcriptional activation relative to acquisition of permissive histone modifications<sup>14,15</sup>. It raises



the question to what extent a lack of DNA methylation at promoters, which is associated with MaD H3K4me3, uncouples promoter marking and transcriptional activation. Therefore, we grouped transcribed genes without detectable maternal messenger RNA<sup>35</sup> based on the stage of maximum expression and DNA methylation (Fig. 4c). We find that developmentally activated promoters with hypomethylated CpG islands are trimethylated at H3K4 or H3K27 early on, irrespective of the time of transcriptional activation. By contrast, methylated promoters show a much closer relation between H3K4me3 and gene expression. Although H3K4me3 is known to stabilize the transcription initiation factor Taf3 (a subunit of TFIID) and can also interact with the chromatin remodeller Chd1 (refs 36–38), hypomethylated promoters gain H3K4me3 autonomously with their hypomethylated CpG island status, independent of embryonic transcription.

**ZyD p300-bound domains shape enhancer clusters.** P300 can be recruited by transcription factors that bind to regulatory elements. We therefore modelled transcription factor motif contributions to p300 binding across multiple developmental stages (see Methods). The results predict specific transcription factors to recruit p300 in a stage-specific manner (Fig. 5a). Clustering of MaD and ZyD p300-bound regions with H3K4me3, H3K4me1 and RNAPII data revealed that ZyD p300 is recruited to distal regulatory sequences that lose both p300 and RNAPII binding in the presence of  $\alpha$ -amanitin, whereas MaD p300 binding mostly includes promoter-proximal regions that are H3K4me3-decorated and recruit RNAPII in the presence of  $\alpha$ -amanitin but without elongating (Fig. 5b). Indeed, MaD p300 regions are enriched for promoter-related motifs (Supplementary Fig. 7). Although some ZyD p300-bound regions overlap with annotated transcription start sites (Supplementary Fig. 5f), most of these sequences are decorated with H3K4me1 in the absence of H3K4me3, suggesting they correspond to distal regulatory sequences (Fig. 5b). Both MaD- and ZyD p300-bound regulatory regions recruit embryonically regulated transcription factors such as Otx2, Gsc, Smad2/3, Foxh1, T (Xbra), Vegt and Eomes (Supplementary Fig. 8)<sup>39–41</sup>, suggesting that multiple transcription factors contribute to p300 recruitment.

Large enhancer clusters (ECs) are thought to improve the stability of enhancer–promoter interactions, are associated with genes coding for developmental regulators, and have been implicated in cell differentiation<sup>42–44</sup>. During development the cluster size of p300-bound enhancers grows dynamically by p300 seeding of individual enhancers (Fig. 5c,d, see Methods). Histone modifications and transcript levels of EC-associated genes are developmental stage specific, confirming the association of ECs with developmental genes (Supplementary Fig. 9; Supplementary Data 2). Analysis of the percentage of the total EC regions identified in each stage show that most p300-bound ECs increase in genomic coverage during development by newly gained p300 binding at enhancers (EC clusters 1 and 2), whereas a group of early ECs (EC cluster 3) decrease in coverage as a result of the decreasing number of p300 peaks that contribute to the EC.

We next examined how MaD and ZyD p300-bound regions contribute to p300-bound ECs. Approximately 50% of all ZyD p300-bound enhancers are located in ECs at stage 11. Among MaD p300-bound enhancers this fraction is much reduced (Fig. 5e). Similarly, a much larger fraction of ZyD p300-bound promoters is found in ECs compared with MaD p300-bound promoters. Up to 20% of the developmental ECs that are seeded at stage 9 have a MaD p300 seeding site (Fig. 5f). However, very few ECs can be called based on MaD p300, showing that formation of p300-bound enhancer clusters requires embryonic transcription (Fig. 5g).

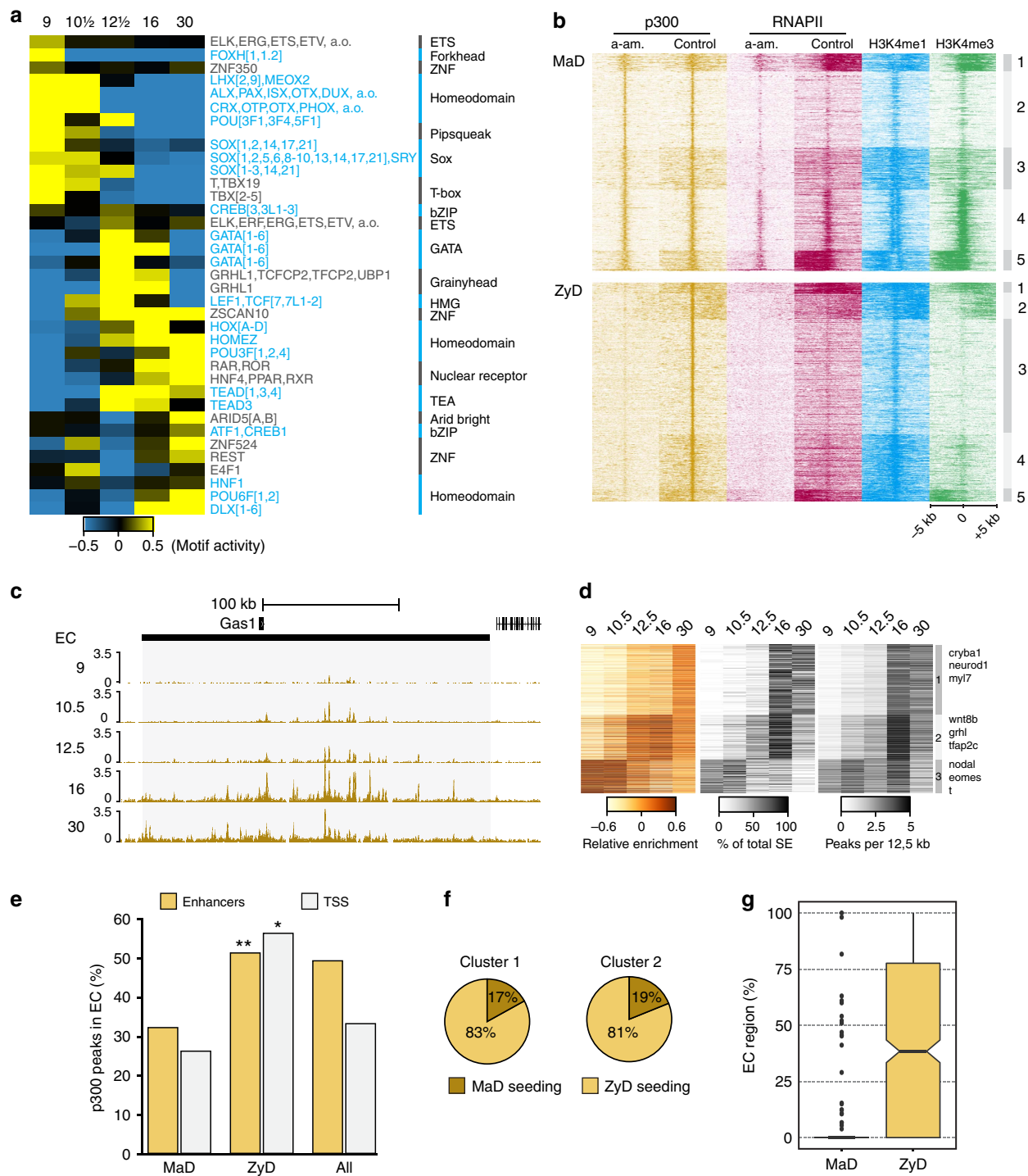
**Extended maternal epigenetic control.** We next examined the extent to which the MaD epigenome is maintained during development. Genes were grouped based on MaD or ZyD trimethylation of H3K4 and H3K27 in the promoter (Supplementary Data 3, see Methods). For p300 we counted the total number of MaD and ZyD peaks in the *cis*-regulatory landscapes of genes (Fig. 6a). Remarkably, MaD H3K4me3-regulated genes represent the majority of all H3K4me3-enriched genes in both early and late developmental stages. Even at neurula and tailbud stages only a small fraction of the H3K4me3-decorated genes are ZyD. Similarly, maternal control of H3K27me3 also extends late into development, albeit to a smaller degree. After gastrulation, the number of MaD H3K27me3 regulated genes slightly decreases, whereas ZyD increases. However, also at neurula stage more than 50% of the Polycomb (PRC2)-regulated genes are under MaD H3K27me3 control. By contrast, p300 in *cis*-regulatory regions of genes is almost exclusively ZyD in all stages (Fig. 6a).

Many genes may maintain MaD H3K4me3 because they are constitutively expressed throughout development. We therefore analysed the regulation of genes that are exclusively embryonically transcribed. We find that 487 of 983 (49.5%) genes which are expressed between blastula and tailbud stages but not expressed in oocytes or before the MBT, feature a MaD H3K4me3 promoter (Supplementary Fig. 10a). Most of the MaD H3K4me3 genes that are modified by PRC2 exhibit MaD H3K27me3. When separating embryonic transcripts based on developmental activation, we find MaD H3K4me3 for 58% of the gastrula genes and up to 74% of the neurula expressed genes (Fig. 6b; Supplementary Fig. 10b). In most cases MaD H3K4me3-regulated genes also have MaD H3K27me3 control. This indicates an important role for the MaD epigenome in the regulation of embryonic transcripts.

To explore the distinctions between expression inside and outside the maternal regulatory space, we analysed Wnt signalling targets. Early Wnt/beta-catenin signalling serves to specify dorsal fates following fertilization, leading to organizer gene expression. This has been shown to depend on Prmt2-mediated promoter poisoning before the MBT<sup>45</sup>. Indeed, we find that seven of eight early Wnt/beta-catenin targets have a hypomethylated island promoter marked with MaD H3K4me3 (Fig. 7a; Supplementary Fig. 10c). Wnt signalling also plays an important role after the MBT, when it ventralises and patterns mesoderm. The majority of these later targets turn out to have a methylated promoter with ZyD H3K4me3. Notably, these ZyD H3K4me3 late Wnt targets are associated with high binding of p300 in their locus; many of the p300 binding events happen at distal regulatory regions. In contrast, MaD H3K4me3 Wnt targets have less p300 binding but are marked with H3K27me3 (Fig. 7a,b). These results illustrate the dichotomy in proximal and distal regulation that is associated with transcriptional activation of maternal and zygotic Wnt target genes, which is paradigmatic of the distinctive maternal and zygotic epigenetic programs that are orchestrated by DNA methylation and exert a long-lasting influence in development (Fig. 8).

## Discussion

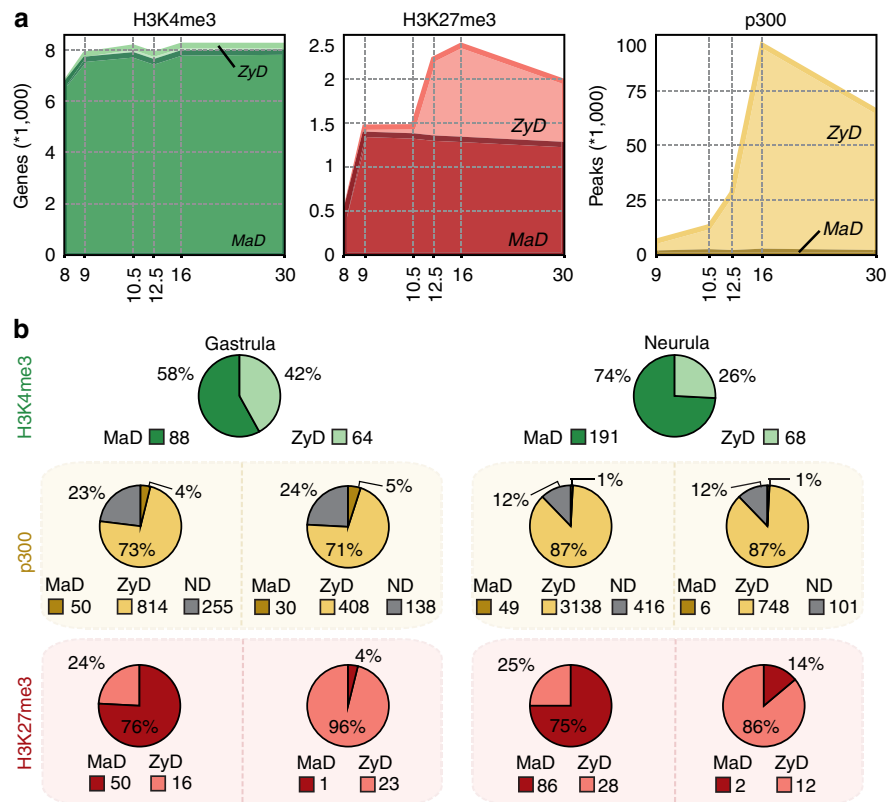
The H3K4me3 modification poises promoters for transcription initiation by stabilizing Taf3/TFIID binding<sup>36,37</sup>. Promoter H3K4 methylation based on an underlying DNA methylation logic driven by maternal factors at the blastula stage sets the stage for a default programme of gene expression. Most constitutively expressed housekeeping genes are within this maternal regulatory space, as well as a subset of developmentally regulated genes. Remarkably, many late expressed genes have hypomethylated promoters and are already poised for activation by H3K4me3



**Figure 5 | Zygotically controlled p300 recruitment shapes enhancer clusters (EC) domains.** (a) Modelled transcription factor motif activity to p300 enrichment (see Methods). Activity reflects modelled contributions in p300 peak RPKM. (b) Heatmaps of MaD (upper panel) and ZyD (lower panel) p300 binding sites in  $\alpha$ -amanitin treated and control embryos. (c) Developmental increase in genomic coverage of the *gas1* EC by acquisition of p300 binding at enhancers. (d) EC dynamics of p300 enrichment (left panel), percentage of total EC region identified in each stage based on stage-dependent p300 binding (middle panel) and number of p300 peaks (per 12.5 kb) in EC. (e) Percentage of zygotic defined (ZyD, lost in  $\alpha$ -amanitin treated embryos) and maternal defined (MaD, maintained in  $\alpha$ -amanitin treated embryos) p300 peaks that map to ECs. Asterisks indicate significance as more or less p300 peaks than expected by chance calculated using cumulative hypergeometric test:  $*P = 6E - 14$ ;  $**P = 5E - 29$  (f) Percentage of ECs that have a MaD or ZyD seeding peak at stage 9. (g) Box plot showing the percentage of the EC region that is defined by MaD or ZyD p300 peaks. Box: 25th (bottom), 50th (internal band), 75th (top) percentiles. Whiskers:  $1.5 \times$  interquartile range of the lower and upper quartiles, respectively. Outliers are indicated with black dots.

during early blastula stages. H3K4me3 is not sufficient for gene transcription and additional embryonic factors are required for activation in many cases. Genes with MaD H3K4me3 generally have fewer p300-bound enhancers associated with them,

suggesting they are regulated by promoter-proximal elements. This further underscores the permissive nature of this regulation, as opposed to zygotically regulated events at both promoters (H3K4me3) and enhancers (recruitment of p300). The



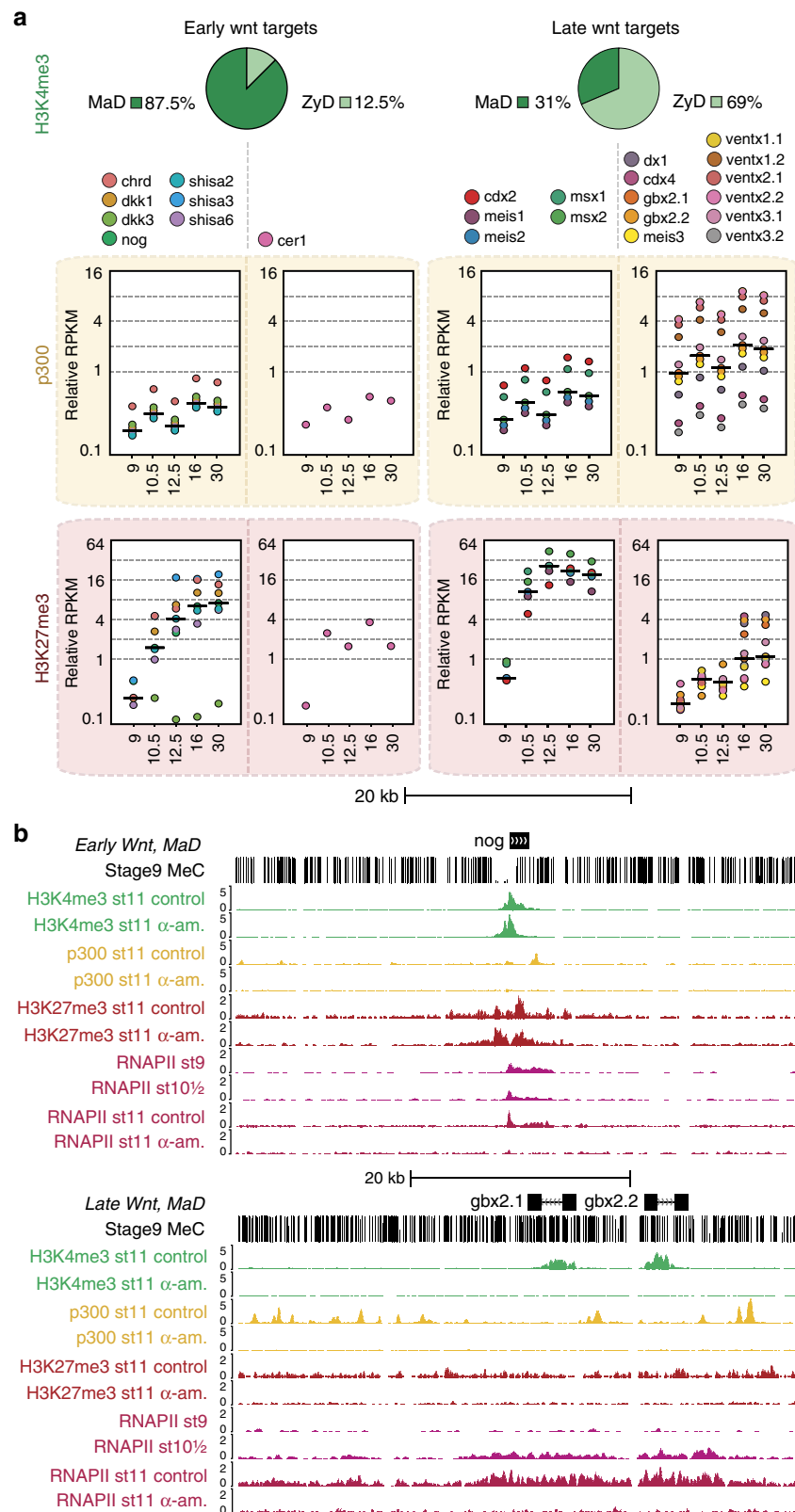
**Figure 6 | Maternal epigenetic control extends beyond gastrulation.** Maternally defined (MaD) peaks emerge at or before stage 11 independent of embryonic transcription. Zygotically defined (ZyD) peaks appear before stage 11 and are lost in  $\alpha$ -amanitin treated embryos, or emerge at or after stage 12. Not determined (ND) peaks are not consistently detected in replicates 1 and 2 and generally have low enrichment values. **(a)** Total number of genes with a MaD or ZyD peak in their promoter (H3K4me3 and H3K27me3), or total number of MaD and ZyD peaks per GREAT region (p300). ND peaks are not shown. **(b)** MaD and ZyD regulation of gastrula and neurula expressed genes. The pie charts show the number genes with a MaD or ZyD peak in their promoter (H3K4me3 and H3K27me3) or the number of MaD, ZyD and ND peaks per *cis*-regulatory region (p300). The H3K27me3 and p300 pie charts represent: Gastrula expressed genes with a MaD (far left) or ZyD (middle left) H3K4me3 peak; neurula expressed genes with a MaD (middle right) or ZyD (far right) H3K4me3 peak.

H3K27me3 modification is gradually acquired between blastula and gastrula stages on spatially regulated genes, repressing lineage-specific genes in other lineages<sup>13,14</sup>. The acquisition of this modification in the absence of transcription indicates that it is uncoupled from the inductive events of the early embryo, suggesting a default maternal response to a lack of transcriptional activation. The results indicate that maternal factors set permissions and time-dependent constraints on a subset of genes with reduced CpG methylation at their promoter. These permissions and constraints are likely to channel embryonic cell fates into a limited number of directions by controlling hierarchical developmental progression by master regulators. Previously we observed that DNA methylation does not lead to transcriptional repression in early embryos, whereas it does in oocytes and late embryos<sup>46</sup>. The observations described here suggest a new role of DNA methylation in defining a maternal-embryonic programme of gene expression. In zebrafish, the maternal methylome is reprogrammed between fertilization and ZGA, to match the paternal methylome. This also occurs in maternal haploid fish, and appears to align with CG content<sup>47,48</sup>, suggesting an intrinsic maternal mechanism that sets the stage for the MaD epigenome.

Gene expression outside maternal regulatory space could be mediated by p300-associated enhancers, most of which require new transcription for recruitment of p300. Promoter and enhancer activation in the ZyD regulatory space likely involves binding of specific factors. Indeed, we find that both MaD- and

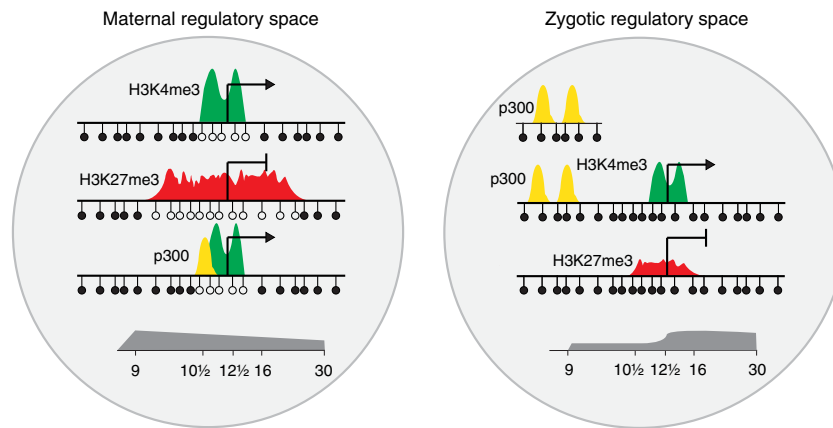
ZyD p300-bound regulatory regions recruit embryonically regulated transcription factors. Enhancers often contain binding sites for many different proteins, which can play different roles in opening up chromatin, recruitment of co-activators and establishing looping interactions with promoters. Future experiments will shed light on the maternal-zygotic hierarchy and the regulatory transitions underlying these events and the roles of maternal and zygotic pioneer factors. We find that ZyD p300-bound enhancers shape enhancer clusters. These form dense hubs of regulatory activity, and EC p300 binding is generally correlated with the expression of the associated genes. The work reported here suggests that recruitment of p300 to 'seeding' enhancers precedes establishing cluster-wide activity of the local enhancer landscape. Future work will also need to address to which extent seeding causes relaxation and opening of the local chromatin and activity of neighbouring enhancers.

Key proteins of the molecular machinery involved in DNA methylation (Dnmt3a, Tet2), H3K4me3 (Mll1-4, Kdm5b/c), H3K27me3 (Ezh2, Eed, Kdm6a/b) and enhancer histone acetylation (p300) are not only highly conserved between species but also frequently mutated in cancer<sup>49-51</sup>. Moreover cancer-specific hypermethylated regions tend to correspond to Polycomb-regulated loci in embryonic stem cells and DNA methylation may restrict H3K27 methylation globally<sup>52,53</sup>. In addition, the sequence signatures of hypomethylated regions that acquire H3K4me3 or H3K27me3 are conserved between fish, frogs and humans<sup>13</sup>. These observations suggest



**Figure 7 | Maternal and zygotic regulatory space separates early and late Wnt target genes. (a)** The number of genes with MaD or ZyD H3K4me3 (pie charts) and relative RPKM (dot plots, horizontal line: median) of p300 in *cis*-regulatory regions of genes and H3K27me3 on promoters ( $\pm 2.5$  kb from TSS) at different developmental stages that have maternally or zygotically defined H3K4me3 at the promoter. Early targets *sia1* and *sia2* are not included, these genes lose H3K4me3 after stage 9 and cannot be assigned to MaD or ZyD space based on our stage 11  $\alpha$ -amanitin data. H3K4me3 on these genes is acquired at stage 8, before embryonic transcription. **(b)** Browser views of the early Wnt target *nog* (*noggin*) and the late Wnt targets *gbx2.1* and *gbx2.2* with ChIP-seq enrichment of H3K4me3, p300 and RNAPII on control and  $\alpha$ -amanitin-injected embryos and RNAPII on stages 9 and 10.5.





**Figure 8 | Model of maternal and zygotic regulatory space.** This shows the segregation of maternal regulatory space, which contains hypomethylated promoters that are mainly controlled by maternal factors, and zygotic regulatory space, which includes methylated promoters and enhancers that are under zygotic control. Most p300-bound enhancers are in zygotic space, however, they can regulate promoters in both maternal and zygotic space, crossing the regulatory space border. This may contribute to varying degrees of permissiveness to transcriptional activation. Maternal regulatory space extends well into neurula and tailbud stages and includes many embryonic genes which are activated at specific stages of development. Zygotic regulatory space requires zygotic transcription, is established from the mid-blastula stage onwards but increases in relative contribution during development.

that the molecular mechanisms that orchestrate the maternal and zygotic regulatory space are conserved. One key difference between mammals and non-mammalian vertebrates is the specification of extra-embryonic lineages between zygotic genome activation and the blastocyst stage in mammals<sup>10</sup>, so it is likely that the way this plays out for specific genes differs between species. In summary, our results provide an unprecedented view of the far reach of maternal factors in zygotic life through chromatin state. The dichotomy of maternal promoter-based and embryonic enhancer regulation demarcates an epigenetic maternal-to-zygotic transition that is maternal permissive to the expression of some embryonic genes and restrictive to others. This highlights the combinatorial interplay of maternal and zygotic factors through distinct mechanisms.

## Methods

**Animal procedures.** *X. tropicalis* embryos were obtained by *in vitro* fertilization, dejellied in 3% cysteine and collected at the indicated stage. Fertilized eggs were injected with 2.3 nl of 2.67 ng  $\mu\text{l}^{-1}$   $\alpha$ -amanitin and developed until the control embryos reached mid-gastrulation (stage 11). Animal use was conducted under the DEC permission (Dutch Animal Experimentation Committee) RU-DEC 2012–116 and 2014–122 to G.J.C.V.

**ChIP sequencing and RNA sequencing.** Chromatin for ChIP was prepared as previously described<sup>24,55</sup>, with minor modifications. Antibody was incubated with chromatin overnight, followed by incubation with Dynabeads Protein G for 1 h. The following antibodies were used: anti-H3K4me1 (Abcam ab8895, 1  $\mu\text{g}$  per 15 embryo equivalents (Eeq)), anti-H3K4me3 (Abcam ab8580, 1  $\mu\text{g}$  per 15 Eeq), anti-H3K9ac (Upstate/Millipore 06-942, 1  $\mu\text{g}$  per 15 Eeq), anti-H3K36me3 (Abcam ab9050, 1  $\mu\text{g}$  per 15 Eeq), anti-H3K27me3 (Upstate/Millipore 07-449, 1  $\mu\text{g}$  per 15 Eeq), anti-H3K9me2 (Diagenode C15410060, 1  $\mu\text{g}$  per 15 Eeq), anti-H3K9me3 (Abcam ab8898, 2  $\mu\text{g}$  per 15 Eeq), anti-H4K20me3 (Abcam ab9053, 2  $\mu\text{g}$  per 15 Eeq), anti-p300 (Santa Cruz sc-585, 1  $\mu\text{g}$  per 15 Eeq) and anti-RNAPII (Diagenode C15200004, 1  $\mu\text{g}$  per 15 Eeq). For all ChIP-seq samples of the epigenome reference maps and RNAPII ChIP-seq samples of the  $\alpha$ -amanitin experiments three biological replicates of different chromatin isolations of 45 embryos were pooled. Two biological replicates for H3K4me3 ( $\alpha$ -amanitin injected: 90 and 56 Eeq; control: 45 and 67 Eeq), H3K27me3 ( $\alpha$ -amanitin injected: 90 and 180 Eeq; control: 45 and 202 Eeq) and p300 ( $\alpha$ -amanitin injected: 112 and 56 Eeq; control: 112 and 67 Eeq) ChIP-seq samples of the  $\alpha$ -amanitin experiments were generated. For RNA-seq samples of the  $\alpha$ -amanitin experiments RNA from five embryos from one biological replicate was isolated and depleted of ribosomal RNA as previously described<sup>35</sup>. Samples were subjected to a qPCR quality check pre- and post preparation. Libraries were prepared with the Kapa Hyper Prep kit (Kapa Biosystems), and sequencing was done on the Illumina HiSeq2000 platform. Reads were mapped to the reference *X. tropicalis* genome JGI7.1, using STAR (RNA-seq) or BWA (ChIP-seq) allowing one mismatch.

**MethylC-seq.** Genomic DNA from *Xenopus* embryos stages 9 and 10.5 was obtained as described before<sup>56</sup>. MethylC-seq library generation was performed as described previously<sup>57</sup>. Library amplification was performed with KAPA HiFi HotStart Uracil + DNA polymerase (Kapa Biosystems, Woburn, MA, USA), using six cycles of amplification. Single-read MethylC-seq libraries were processed and aligned as described previously<sup>58</sup>.

**Quantitative PCR.** PCR reactions were performed on a CFX96 Touch Real-Time PCR Detection System (BioRad) using iQ Custom SYBR Green Supermix (BioRad). We performed RNA expression PCR (RT-qPCR (quantitative PCR)) and ChIP-qPCR for H3K4me3 and H3K9ac on promoters of *odc1*, *ee1a10*, *rnf146*, *tor1a*, *zic1*, *cdc14b*, *eomes*, *xrcc1*, *drosha*, *gdf3*, *t*, *tbx2*, *fastkd3*, *gs17* (see Supplementary Methods for primer sequences). ChIP-qPCR enrichment over background was calculated using the average of 5 negative loci.

**Detection of enriched regions.** We used MACS2 (ref. 59) with standard settings and a  $q$ -value of 0.05. Fragment size was determined using phantom-peakqualtools<sup>60</sup>. Broad settings (--BROAD) were used for H3K4me1, H3K36me3, H3K27me3, H3K9me2, H3K9me3, H4K20me3 and RNAPII. Broad and narrow peaks were merged for H3K4me3. For H3K9ac narrow peaks were used. For p300 broad peaks were used in the ChomHMM analysis, narrow p300 peaks were used for super-enhancer and MaD versus ZyD analyses. All peaks were called relative to an input control track. Peaks that showed at least 75% overlap with 1 kb regions that have more than 65 input reads, and peaks that have a ChIP-seq RPKM higher than the 95 percentile of random background regions are excluded from further analysis. Only scaffolds 1–10 (the chromosome-sized scaffolds) were included in the analysis. Relative RPKM was calculated by dividing the ChIP-seq RPKM of the peaks by the ChIP-seq RPKM of the 95 percentile of random background regions.

We used MANorm<sup>61</sup> to determine differentially enriched regions in  $\alpha$ -amanitin and control embryos. We used merged peak sets of replicate 1, replicate 2 and stage 10.5 to avoid bias caused by peak calling. Lost, gained and unchanged peaks per biological replicate were determined using the following parameters: lost peaks have  $M$ -values  $> 1$  and a  $-\log$  base 10( $P$  value)  $> 5$  (for H3K27me3) or 1.3 (for H3K4me3 and p300) and have a relative RPKM (background corrected)  $> 1$  in stage 11 control (no cut-off was used for st.11 control of H3K27me3 rep.1), stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3); *increased peaks* have  $M$ -values smaller than -1 and a  $-\log$  base 10( $P$  value)  $> 5$  (H3K27me3) or 1.3 (H3K4me3 and p300) and have a rel. RPKM greater than 1 in stage 11  $\alpha$ -amanitin, stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3); *unchanged peaks* are neither gained nor lost and have a rel. RPKM  $> 1$  in stage 11 control (no cut-off was used for st.11 control of H3K27me3 rep.1), stage 11  $\alpha$ -amanitin, stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3). Maintained peaks are peaks that are not lost and have a rel. RPKM  $> 1$  in stage 11 control (no cut-off was used for st.11 control of H3K27me3 rep.1), stage 11  $\alpha$ -amanitin, stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3). Common lost, gained, unbiased and maintained peaks are present in both replicates. All other peaks are considered not defined (ND). Replicate-specific peaks were only used for Supplementary Fig. 4b, for all other figures the common peaks were used.

DNA methylation levels in Supplementary Fig. 4d was calculated using previously published Bio-CAP data<sup>62</sup>. Bio-CAP RPKM levels of stage 11–12 were calculated for H3K4me3, H3K27me3 and p300 peaks, and corrected for input

values. For Fig. 4c genes were considered 'hypomethylated' if the Bio-CAP/Input ratio on the promoter ( $\pm 1$  kb from TSS) was  $> 1$ .

RNA expression analysis was performed as previously published<sup>35</sup>. Embryonic transcripts were separated based on the clustering of maximum expression levels per stage in Fig. 3d of Paranjpe *et al.*<sup>35</sup> (cluster 1 = blastula, cluster 5 = gastrula, clusters 3 and 4 = neurula, clusters 2 and 6 = tailbud).

Enhancer clusters were called as previously described<sup>43</sup>. Enhancer Clusters are called per stage and merged to determine the total Enhancer Cluster region. Percentage of the EC region is calculated relative to the total Enhancer Cluster region.

**MaD and ZyD classification.** MaD peaks emerge at or before stage 11 and are also acquired in  $\alpha$ -amanitin treated embryos in both replicates. Zygotically defined (ZyD) peaks appear at or before stage 11 and are lost in  $\alpha$ -amanitin treated embryos in both replicates, or emerge after stage 11. To classify MaD and ZyD H3K4me3 genes we ran MANorm on promoters ( $\pm 250$  bp from TSS) only, using similar restrictions as described in 'Detection of enriched regions'. MaD H3K4me3 genes have a maintained promoter in both replicates, ZyD H3K4me3 genes have a lost promoter H3K4me3 peak in both  $\alpha$ -amanitin replicates, or a peak that emerges after stage 11. MaD H3K27me3 genes have at least one MaD peak in the vicinity of their promoter ( $\pm 2.5$  kb from TSS). ZyD H3K27me3 genes have at least one ZyD peak in their promoter and lack a MaD peak. ND peaks or genes do not meet the criteria for neither MaD nor ZyD. For p300 the total number of ZyD and MaD peaks was counted in GREAT<sup>63</sup> regions of genes.

**ChomHMM analysis.** Chromatin states were discovered and characterized using ChromHMM v1.10 (ref. 26), an implementation of a hidden Markov model. As input we used the enriched regions from ten tracks (H3K27me3, H3K36me3, H3K4me1, H3K4me3, H3K9ac, H3K9me2, H3K9me3, H4K20me3, p300 and RNAPII) across five developmental stages. We trained and ran the model with a range of states, and determined the 19 emission states model as the optimal number of states that could sufficiently capture the biological variation in co-occurrence of chromatin marks. We subsequently classified the states into seven main groups based on the presence and absence of specific chromatin marks.

The segmentation files of the seven main groups per stage were binned in 200 base pairs intervals. An  $m \times n$  matrix was created, where  $m$  corresponds to the 200 base pair intervals and  $n$  to the developmental stages (9–30). Each element  $a(i,j)$  represents the chromatin state of interval  $i$  at stage  $j$ . For each chromatin group occurrences were counted per stage  $n$ . The changes between stage  $n$  and  $n+1$  were plotted using Sankey diagrams (<https://github.com/tamc/Sankey>), a flow diagram closely related to alluvial diagrams.

**Motif analyses.** For the prediction of motif contribution to p300 recruitment (Fig. 5a) we have implemented the ISMARA method developed by Balwierz *et al.*<sup>64</sup> This method uses motif activity response analysis to determine the transcription factors that drive the observed changes in chromatin state across samples. As input we used the number of known motifs found per p300 binding site and the RPKM of the p300 peaks per developmental stage. The model infers the unknown motif activities from the equation in which the changes in signal levels are explained with the number of binding sites and the unknown motif activities. Motifs that showed a z-score activity that was  $> 13$  are shown in Fig. 5a. Enriched motifs (Supplementary Fig. 7) were detected with gimme diff, a tool from the GimmeMotifs package<sup>65</sup>. The vertebrate motifs used in this script were obtained from CISBP (<http://cisbp.cccb.utoronto.ca/>)<sup>66</sup> and clustered using gimme cluster from GimmeMotifs. The motifs are available at <http://dx.doi.org/10.6084/m9.figshare.1555851> (ref. 67).

**Generation of plots and heatmaps.** All heatmaps were generated using fluff (<http://simonvh.github.com/fluff/>)<sup>15</sup> or gplots (<http://cran.r-project.org/web/packages/gplots/index.html>). For all heatmap clustering, the Euclidean distance metric was used. Other plots were generated using ggplot2 (<http://ggplot2.org/>).

## References

- Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
- Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
- Lee, J. S., Smith, E. & Shilatifard, A. The language of histone crosstalk. *Cell* **142**, 682–685 (2010).
- Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
- Gerstein, M. B. *et al.* Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* **330**, 1775–1787 (2010).
- Roy, S. *et al.* Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **330**, 1787–1797 (2010).
- Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033–3042 (2009).
- Kane, D. A. & Kimmel, C. B. The zebrafish midblastula transition. *Development* **119**, 447–456 (1993).
- O'Farrell, P. H., Stumpff, J. & Su, T. T. Embryonic cleavage cycles: how is a mouse like a fly? *Curr. Biol.* **14**, R35–R45 (2004).
- Paranjpe, S. S. & Veenstra, G. J. Establishing pluripotency in early development. *Biochim. Biophys. Acta* **1849**, 626–636 (2015).
- Newport, J. & Kirschner, M. A major developmental transition in early *Xenopus* embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* **30**, 675–686 (1982).
- Newport, J. & Kirschner, M. A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell* **30**, 687–696 (1982).
- van Heeringen, S. J. *et al.* Principles of nucleation of H3K27 methylation during embryonic development. *Genome Res.* **24**, 401–410 (2014).
- Akkers, R. C. *et al.* A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev. Cell* **17**, 425–434 (2009).
- Vastenhouw, N. L. *et al.* Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464**, 922–926 (2010).
- Lindeman, L. C. *et al.* Pre-patterning of developmental gene expression by modified histones before zygotic genome activation. *Dev. Cell* **21**, 993–1004 (2011).
- Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
- Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
- Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
- Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318 (2007).
- Roh, T. Y., Wei, G., Farrell, C. M. & Zhao, K. Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome Res.* **17**, 74–81 (2007).
- Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
- Santos-Rosa, H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407–411 (2002).
- Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- Schotta, G. *et al.* A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18**, 1251–1262 (2004).
- Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216 (2012).
- Chafin, D. R., Guo, H. & Price, D. H. Action of alpha-amanitin during pyrophosphorolysis and elongation by RNA polymerase II. *J. Biol. Chem.* **270**, 19114–19119 (1995).
- Sible, J. C., Anderson, J. A., Lewellyn, A. L. & Maller, J. L. Zygotic transcription is required to block a maternal program of apoptosis in *Xenopus* embryos. *Dev. Biol.* **189**, 335–346 (1997).
- Skirkanich, J., Luxardi, G., Yang, J., Kodjabachian, L. & Klein, P. S. An essential role for transcription before the MBT in *Xenopus laevis*. *Dev. Biol.* **357**, 478–491 (2011).
- Clouaire, T. *et al.* Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714–1728 (2012).
- Clouaire, T., Webb, S. & Bird, A. Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. *Genome Biol.* **15**, 451 (2014).
- Thomson, J. P. *et al.* CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**, 1082–1086 (2010).
- van Kruijsbergen, I., Hontelez, S. & Veenstra, G. J. Recruiting polycomb to chromatin. *Int. J. Biochem. Cell Biol.* **67**, 177–187 (2015).
- Ng, H. H., Robert, F., Young, R. A. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* **11**, 709–719 (2003).
- Paranjpe, S. S., Jacobi, U. G., van Heeringen, S. J. & Veenstra, G. J. A genome-wide survey of maternal and embryonic transcripts during *Xenopus tropicalis* development. *BMC Genomics* **14**, 762 (2013).
- Lauberth, S. M. *et al.* H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).
- Vermeulen, M. *et al.* Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).
- Sims, 3rd R. J. *et al.* Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* **28**, 665–676 (2007).

39. Gentsch, G. E. *et al.* In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell Rep.* **4**, 1185–1196 (2013).
40. Yasuoka, Y. *et al.* Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification. *Nat. Commun.* **5**, 4322 (2014).
41. Chiu, W. T. *et al.* Genome-wide view of TGFbeta/Foxh1 regulation of the early mesendoderm program. *Development* **141**, 4537–4547 (2014).
42. Pott, S. & Lieb, J. D. What are super-enhancers? *Nat. Genet.* **47**, 8–12 (2015).
43. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
44. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
45. Blythe, S. A., Cha, S. W., Tadjuidje, E., Heasman, J. & Klein, P. S. beta-Catenin primes organizer gene expression by recruiting a histone H3 arginine 8 methyltransferase, Prmt2. *Dev. Cell* **19**, 220–231 (2010).
46. Bogdanovic, O. *et al.* Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. *Genome Res.* **21**, 1313–1327 (2011).
47. Potok, M. E., Nix, D. A., Parnell, T. J. & Cairns, B. R. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153**, 759–772 (2013).
48. Jiang, L. *et al.* Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* **153**, 773–784 (2013).
49. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).
50. Huether, R. *et al.* The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat. Commun.* **5**, 3630 (2014).
51. Wu, G. *et al.* The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. *Nat. Genet.* **46**, 444–450 (2014).
52. Simmer, F. *et al.* Comparative genome-wide DNA methylation analysis of colorectal tumor and matched normal tissues. *Epigenetics* **7**, 1355–1367 (2012).
53. Brinkman, A. B. *et al.* Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128–1138 (2012).
54. Jallow, Z., Jacobi, U. G., Weeks, D. L., Dawid, I. B. & Veenstra, G. J. Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in *Xenopus*. *Proc. Natl Acad. Sci. USA* **101**, 13525–13530 (2004).
55. Akkers, R. C., Jacobi, U. G. & Veenstra, G. J. Chromatin immunoprecipitation analysis of *Xenopus* embryos. *Methods Mol. Biol.* **917**, 279–292 (2012).
56. Bogdanovic, O., Fernandez-Minan, A., Tena, J. J., de la Calle-Mustienes, E. & Gomez-Skarmeta, J. L. The developmental epigenomics toolbox: ChIP-seq and MethylCap-seq profiling of early zebrafish embryos. *Methods* **62**, 207–215 (2013).
57. Lister, R. *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).
58. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
59. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
60. Kharchenko, P. V., Tolstourkov, M. Y. & Park, P. J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.* **26**, 1351–1359 (2008).
61. Shao, Z., Zhang, Y., Yuan, G. C., Orkin, S. H. & Waxman, D. J. MAAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol.* **13**, R16 (2012).
62. Long, H. K. *et al.* Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *eLife* **2**, e00348 (2013).
63. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
64. Balwiercz, P. J. *et al.* ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome Res.* **24**, 869–884 (2014).
65. van Heeringen, S. J. & Veenstra, G. J. GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. *Bioinformatics* **27**, 270–271 (2011).
66. Weirauch, M. T. *et al.* Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).
67. van Heeringen, S. J. *et al.* Vertebrate motif clusters v3.0. *figshare* <http://dx.doi.org/10.6084/m9.figshare.1555851> (2015).

## Acknowledgements

This work is supported by the US National Institutes of Health (NICHD, grant R01HD069344). Part of this work was carried out on the Dutch national e-infrastructure with the support of SURF Foundation. S.J.v.H. is supported by the Netherlands Organization for Scientific research (NWO-ALW, grant 863.12.002). O.B. is supported by an Australian Research Council Discovery Early Career Researcher Award - DECRA (DE140101962).

## Author contributions

ChIP-seq data production was done by S.H. and I.v.K. Bisulfite sequencing was done by O.B. and R.L. G.G. performed ChromHMM analysis. All other analyses were done by S.H. with support from G.G. and S.v.H. G.J.C.V. and S.H. designed the study and wrote the paper. S.H., I.v.K. and G.G. contributed equally to the study. All authors discussed the results and commented on the manuscript.

## Additional information

**Accession codes:** The data generated for this work have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE67974. Visualization tracks are available at the authors' web site (<http://www.ncmls.nl/gertjanveenstra>).

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Hontelez, S. *et al.* Embryonic transcription is controlled by maternally defined chromatin state. *Nat. Commun.* **6**:10148 doi: 10.1038/ncomms10148 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>