

Zebrafish Pou5f1-dependent transcriptional networks in temporal control of early development

Daria Onichtchouk^{1,*}, Florian Geier^{2,6}, Bozena Polok¹, Daniel M Messerschmidt^{3,7}, Rebecca Mössner¹, Björn Wendik¹, Sungmin Song¹, Verdon Taylor³, Jens Timmer^{2,4,5} and Wolfgang Driever^{1,5,*}

¹ Developmental Biology, Institute Biology I, Faculty of Biology, University of Freiburg, Freiburg, Germany, ² Zentrum für Biosystemanalyse ZBSA, University of Freiburg, Freiburg, Germany, ³ Department of Molecular Embryology, Max-Planck-Institute of Immunobiology, Freiburg, Germany, ⁴ Department of Physics, University of Freiburg, Freiburg, Germany and ⁵ Freiburg Institute for Advanced Studies, University of Freiburg, Freiburg, Germany
⁶ Present address: Department of Biosystems Science and Engineering (D-BSSE), ETH Zürich, Mattenstrasse 26, CH-4058 Basel, Switzerland
⁷ Present address: Laboratory of Mammalian Development, Institute of Medical Biology, 8A Biomedical Grove, Singapore 138648, Singapore
* Corresponding authors. D Onichtchouk or W Driever, Developmental Biology, Institute Biology I, Faculty of Biology, University of Freiburg, Hauptstrasse 1, D-79104 Freiburg, Germany. Tel.: +49 761 203 2595; Fax: +49 761 203 2597; E-mail: Daria.Onichtchouk@biologie.uni-freiburg.de or Tel.: +49 761 203 2587; Fax: +49 761 203 2597; E-mail: driever@biologie.uni-freiburg.de

Received 1.10.09; accepted 18.1.10

The transcription factor POU5f1/OCT4 controls pluripotency in mammalian ES cells, but little is known about its functions in the early embryo. We used time-resolved transcriptome analysis of zebrafish *pou5f1* MZspg mutant embryos to identify genes regulated by Pou5f1. Comparison to mammalian systems defines evolutionary conserved Pou5f1 targets. Time-series data reveal many Pou5f1 targets with delayed or advanced onset of expression. We identify two Pou5f1-dependent mechanisms controlling developmental timing. First, several Pou5f1 targets are transcriptional repressors, mediating repression of differentiation genes in distinct embryonic compartments. We analyze *her3* gene regulation as example for a repressor in the neural anlagen. Second, the dynamics of SoxB1 group gene expression and Pou5f1-dependent regulation of *her3* and *foxD3* uncovers differential requirements for SoxB1 activity to control temporal dynamics of activation, and spatial distribution of targets in the embryo. We establish a mathematical model of the early Pou5f1 and SoxB1 gene network to demonstrate regulatory characteristics important for developmental timing. The temporospatial structure of the zebrafish Pou5f1 target networks may explain aspects of the evolution of the mammalian stem cell networks.

Molecular Systems Biology 6: 354; published online 9 March 2010; doi:10.1038/msb.2010.9

Subject Categories: simulation and data analysis; development

Keywords: developmental timing; mathematical modeling; Oct4; transcriptional networks

This is an open-access article distributed under the terms of the Creative Commons Attribution Licence, which permits distribution and reproduction in any medium, provided the original author and source are credited. This licence does not permit commercial exploitation or the creation of derivative works without specific permission.

Introduction

The transcription factor POU5f1/OCT4 controls pluripotency of mouse embryonic inner cell mass cells (Nichols *et al*, 1998), and of mouse and human ES cell lines (Boiani and Scholer, 2005). Although POU5f1/OCT4-dependent pluripotency transcriptional circuits and many transcriptional targets have been characterized (Boyer *et al*, 2005; Loh *et al*, 2006), little is known about the mechanisms by which POU5f1/OCT4 controls early developmental events. In ES cells, POU5f1/OCT4 cooperates with SOXB1 class transcription factors (Masui *et al*, 2007), Nanog (Chambers *et al*, 2003; Mitsui *et al*, 2003), and KLF4 (Jiang *et al*, 2008), to regulate target genes. Together, these core components of the network maintain their own expression and suppress differentiation.

In addition, POU5f1/OCT4 is involved in developmental decisions, including trophectoderm segregation (Strumpf *et al*,

2005) and primordial germ cell survival (Kehler *et al*, 2004). Mouse POU5f1/OCT4 is expressed in epiblast-derived structures from gastrulation to the 16-somite stage (Downs, 2008), suggesting other functions in development. However, detailed understanding of POU5f1/OCT4 functions during mammalian blastocyst and gastrula development as well as studies of the temporal changes in the POU5f1/OCT4-regulated networks are precluded by the early lineage defects in *Pou5f1/Oct4* mutant mice (Nichols *et al*, 1998). Similarly, investigation of potential roles of POU5f1/OCT4 in differentiating ES cells is hampered by critical requirements for POU5f1/OCT4 to suppress the first lineage-specification event—trophectoderm differentiation (Niwa *et al*, 2000, 2002).

Pou5f1 gene homologues have been identified in birds (*cPouV*; Laval *et al*, 2007), *Xenopus* (*XIPou91*, *XIPou25*, and *XIPou60*; Morrison and Brickman, 2006), axolotl (*Axoct4*; Bachvarova *et al*, 2004), and zebrafish (*pou2*; Takeda *et al*, 1994). According to the current view, a common ancestor of

jawed vertebrates had a single PouV class gene, syntenic to zebrafish *pou5f1/pou2* (Niwa *et al*, 2008; Frankenberg *et al*, 2009). This *pou5f1/pou2*-type gene was duplicated to give rise to *pou5f1/pou2* (in *Xenopus* and chick) and *Pou5f1/Oct4* (in *Axolotl*, mouse and human). All five sequenced fish species have only a single *pou5f1/pou2*-type gene (Niwa *et al*, 2008; Frankenberg *et al*, 2009), suggesting that the *pou5f1/pou2* gene duplication occurred later in evolution, presumably in the common ancestor of tetrapods. Therefore, zebrafish *pou5f1/pou2* should be considered to be an ortholog of mouse *Pou5f1/Oct4* and all other vertebrate PouV class genes (Koonin, 2005). *pou5f1* genes in vertebrates show broad expression during pregastrulation and gastrulation stages (Belting *et al*, 2001; Burgess *et al*, 2002; Bachvarova *et al*, 2004; Lunde *et al*, 2004; Morrison and Brickman, 2006; Laval *et al*, 2007; Downs, 2008), suggesting that at least in part their function during these stages is conserved. Less well known is that Pou5f1 in fish and mouse is also expressed in the neural plate until midsomitogenesis (Takeda *et al*, 1994; Reim and Brand, 2002; Downs, 2008). In contrast, expression in primordial germ cells is present only in mouse and chick (Kehler *et al*, 2004; Laval *et al*, 2007), but not in zebrafish (Reim and Brand, 2006).

In zebrafish, the zygotic *pou5f1* loss-of-function mutation *spiel ohne grenzen* (*Zspg*) is lethal due to neural plate patterning defects (Belting *et al*, 2001). *pou5f1* mRNA rescue of *Zspg* embryos enables homozygous mutant fish to be established that can generate embryos devoid of maternal Pou5f1, *Mspg* (abbreviated 'M'), in which the zygotes are rescued by expression from the paternal allele; and *MZspg* embryos (abbreviated 'MZ'), which are completely devoid of maternal and zygotic Pou5f1 activity (Lunde *et al*, 2004; Reim *et al*, 2004). MZ embryos have gastrulation abnormalities (Lachnit *et al*, 2008), dorsoventral patterning defects (Reim and Brand, 2006), and do not develop endoderm (Lunde *et al*, 2004; Reim *et al*, 2004). The only direct Pou5f1 transcriptional target characterized in zebrafish so far is *sox17* during endoderm specification (Lunde *et al*, 2004; Reim *et al*, 2004; Chan *et al*, 2009). Interestingly, and in contrast to *Pou5f1/Oct4* mutant mice, which are blocked in development due to loss of inner cell mass, MZ mutant embryos are neither blocked in development nor display a general delay. For example, Nodal-dependent mesendoderm induction proceeds normally as judged by the correct expression of *ntl*, *ndr1*, or *gata5* (Lunde *et al*, 2004; Reim *et al*, 2004). Further, gastrula organizer formation as judged by the onset of *gsc*, *boz*, and *chd* expression is initiated with the same developmental timing as in wild-type (WT) siblings (Reim and Brand, 2006). Even selected later development events, including somitogenesis, proceed at a pace similar to WT (Lunde *et al*, 2004). At the cellular level, the delay in epiboly movement in MZ is a selective delay in deep cell epiboly, while the enveloping layer is less affected (Lachnit *et al*, 2008). Specifically, in contrast to the mammalian embryo, cell cycle and proliferation are normal in MZ during early to midgastrula stages (Lachnit *et al*, 2008). The early synchronous cell cycles in zebrafish are maternally controlled (Kane and Kimmel, 1993), and largely independent of Pou5f1 activity. Therefore, zebrafish present a good model system to identify specific transcriptional targets of Pou5f1 during development without disturbing develop-

ment by the loss of embryonic blastomers (inner cell mass) observed in the mouse *Pou5f1/Oct4* mutant.

To better understand the Pou5f1-regulated transcriptional circuitry in zebrafish, we identified groups of genes activated or repressed by Pou5f1, and analyzed the temporal and spatial expression of these targets during the first 3–8 h of zebrafish development, which correspond to pregastrula and gastrulation stages. A large group of developmental regulators is prematurely expressed in MZ embryos, whereas transcriptional repressors including FoxD3 and Her3 are absent, and the expression of SoxB1 genes is severely reduced. We found that Pou5f1 and SoxB1 proteins share a large set of direct target genes. We characterize *her3* as a novel Pou5f1 target, and demonstrate molecular mechanisms of regulation by Pou5f1 and SoxB1 proteins that can explain the temporal profile of *her3* expression. We developed a model of the regulatory network based on a set of ordinary differential equations to describe the dynamics of Pou5f1–SoxB1 target gene regulation, and provide important insights into regulatory features of the network. Finally, we compare the Pou5f1 targets in zebrafish and mouse, and establish evolutionarily conserved components of Pou5f1 and SoxB1 regulatory subnetworks that are likely critical for the control and timing of vertebrate development.

Results

Pou5f1-dependent changes in the maternal and early zygotic transcriptome

Pou5f1 is expressed maternally and zygotically during early zebrafish development (Takeda *et al*, 1994). Complete loss of Pou5f1 activity in MZ mutant embryos provides an *in vivo* model to study the contribution of Pou5f1 activity to the control of early vertebrate developmental progression and early fate decisions. We investigated transcriptome changes by microarray analysis at 10 distinct time-points during development, from ovaries to late gastrulation. Comparison of expression levels of each probe in WT and MZ genotypes (Figure 1A; Supplementary Table 1) defined 20 stage-specific sets of regulated genes. We grouped the probes independently based on similarity of their temporal expression profiles (Figure 1A and B; Supplementary Table 2; 24 temporal profiles). We then identified those Pou5f1-regulated gene groups that showed strong positive correlations between stage-specific gene sets (SSGS, horizontal axis in Figure 1C and D) and temporal profiles in WT and MZ (TP, vertical axis in Figure 1C and D, respectively). This correlation analysis (see Supplementary information and references therein for details) enabled us to subdivide the large heterogeneous group of Pou5f1 targets into smaller sets of genes with similar temporal behavior (squares in Figure 1C and D). Visualization of gene expression changes throughout developmental time in WT (Figure 1C) versus MZ (Figure 1D) aids comprehension of the temporal changes in the Pou5f1 downstream networks (Supplementary Figure 1).

To assess the biological significance of the high correlation between SSGS and TP for individual groups (squares in Figure 1C and D), or for clusters of neighboring groups (letters A–E in Figure 1C and D), we performed Fisher's exact tests for enrichment in gene ontology (GO) function (Supplementary

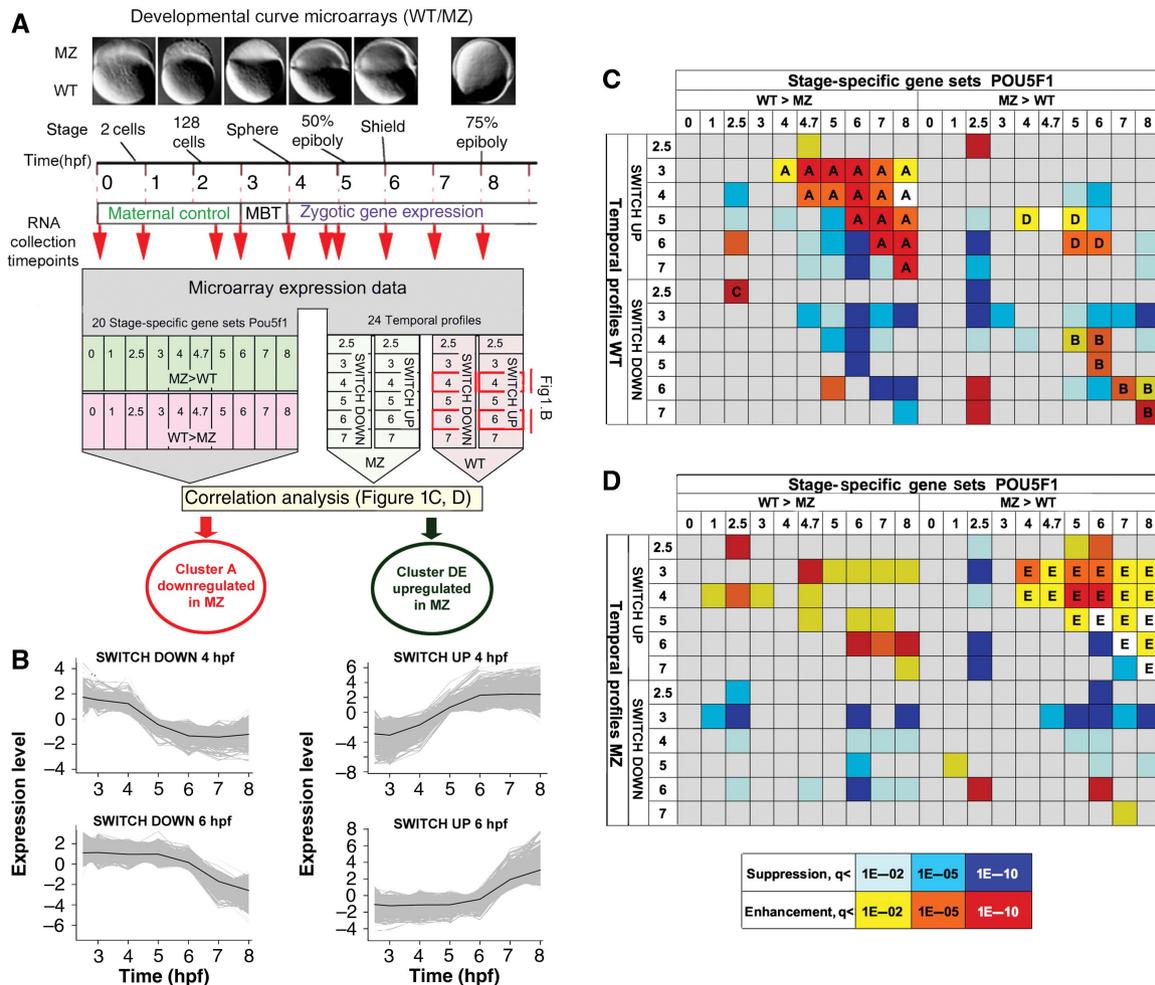


Figure 1 Developmental time series and experimental analysis of Pou5f1-dependent transcriptome changes. **(A)** Scheme of microarray time-series data analysis (0–8 hpf). In all, 20 stage-specific Pou5f1-regulated gene sets were defined based on > 2-fold regulation MZ versus WT (5350 probes at $P < 0.05$). Independently, expression profiles of each gene were analyzed in WT and MZ separately for temporal changes in expression level, as shown in **(B)**, and grouped to 12 temporal profiles in WT and in MZ. **(B)** Examples of temporal profile groups. Reference temporal profiles (black lines), based on maximal switch between high to low (down) or low to high (up) levels were calculated. Probes were grouped based on maximal similarity of temporal expression graphs (gray lines) to one of the reference profiles. **(C, D)** Identification of Pou5f1-dependent target clusters by correlation analysis. Each temporal profile group (left) in WT **(C)** or MZ **(D)** was tested for significant correlations with stage-specific Pou5f1-regulated gene sets (top). Intersection squares with strong correlations in Fisher's exact test are colored (q -value Fisher's exact test, legend at top). The letters A–E indicate clusters showing functional enrichment in GO Fisher exact test.

Table 3) and in specific expression patterns (Supplementary Table 4). Clusters A to E proved to be significantly enriched for specific GO functions in Fisher's exact test. Only Clusters A, D, and E were enriched in genes that function in development and differentiation, and have region-specific expression patterns. We did not specifically consider maternally expressed genes regulated by Pou5f1 (Cluster B, enriched for metabolic functions, and Cluster C, enriched for ribosomal proteins) for further analysis, because genetic evidence indicates that deficiencies in maternal transcripts can be completely rescued by zygotic activity (Lunde *et al*, 2004; Reim *et al*, 2004).

Developmental regulators with loss or delay of expression in MZ (Cluster A)

Cluster A (Supplementary Table 5) contains 1005 probes with loss or delay of expression in MZ, and includes several early

acting transcriptional repressors (*her3*, *klf2b*, *klf4*, *foxD3*, *snail2*), *sox* genes, patterning and differentiation genes, as well as genes encoding diverse signaling pathway components (Figure 2A and B). The reduced expression of *her3*, *foxd3*, *sox2*, *hesx1*, *sox19b*, and *sox11a*, was confirmed by whole-mount *in situ* hybridization (Figure 4E–G) and RT-PCR (Supplementary Figure 6E–G and data not shown). Reduced or lost expression of 165 Cluster A genes was also confirmed by an independent validation microarray experiment using the Affymetrix platform (Supplementary information; Supplementary Table 5).

Premature expression of patterning and differentiation genes in MZ (Cluster DE)

Cluster DE (Supplementary Table 6) contains 844 probes upregulated in MZ, and includes many genes involved in embryonic patterning (*pax6* (Stoykova *et al*, 1996); *gbx2*

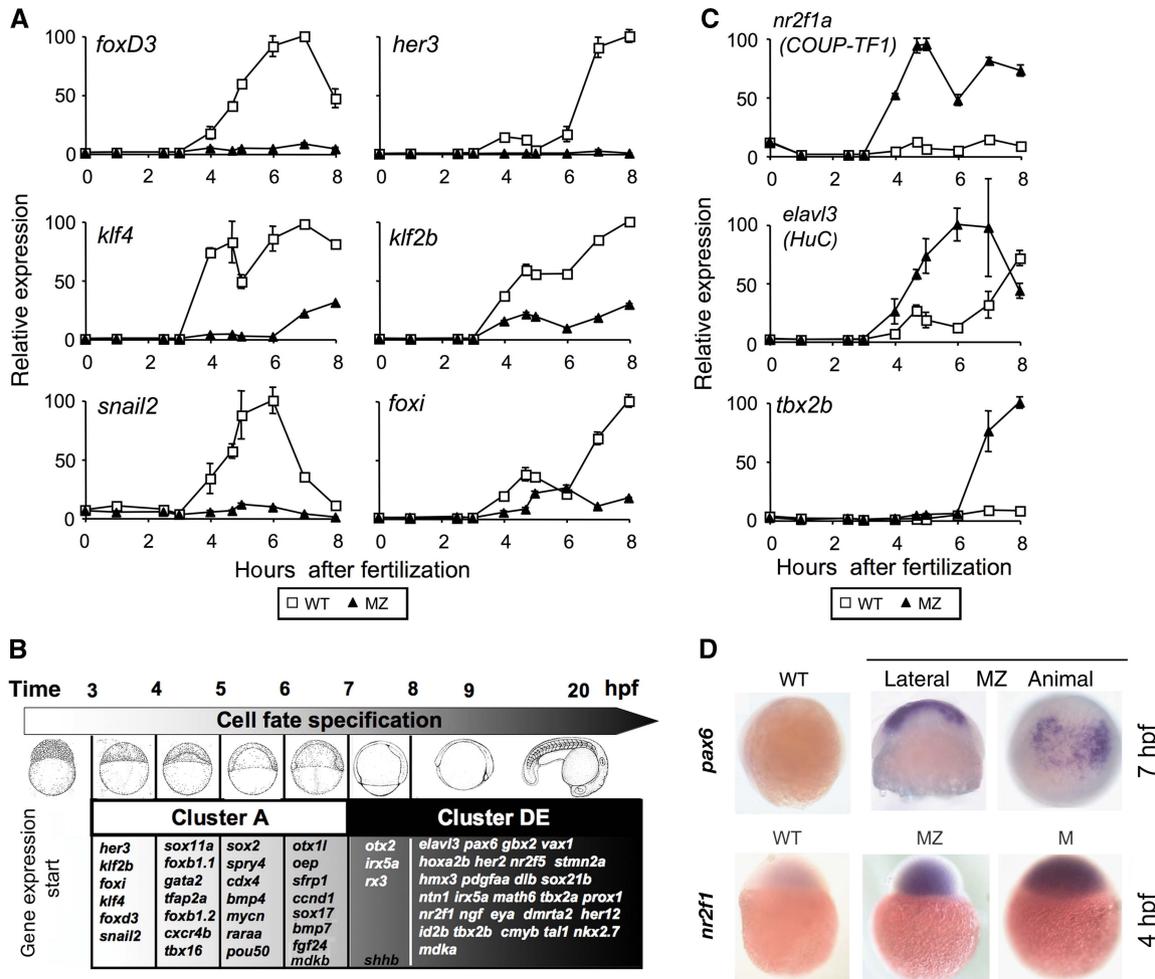


Figure 2 Pou5f1 is required for proper developmental timing of gene expression. **(A)** Microarray expression profiles in WT (white squares) and MZ (black triangles) of Cluster A genes downregulated in MZ. **(B)** Diagram of selected Cluster A (downregulated in MZ, black font) and DE (upregulated in MZ, white font) genes organized by their normal time of activation in WT. In WT, Cluster A genes start to be expressed at 3–8 hpf, whereas in MZ their expression during these stages is reduced or absent. In contrast, in WT most Cluster DE genes start to be expressed at > 8 hpf, and are involved in tissue specification and differentiation. In MZ, expression of Cluster DE genes is prematurely activated at gastrula stages. **(C)** Microarray expression profiles in WT (white squares) and MZ (black triangles) of Cluster DE genes prematurely activated in MZ. **(A, C)** Values were normalized to maximal expression (100). Error bars show s.e.m. of two probes (*foxD3*, *klf2b*, *klf4*, *nr2f1*) or of three biological replicates for one probe. **(D)** Whole-mount *in situ* hybridization, lateral view with animal pole at top, except where animal pole view is indicated. Cluster DE genes with different activation time and spatial pattern in MZ: *pax6a* prematurely appears after 7 hpf in the neuroectodermal region of MZ, *nr2f1* is ubiquitously upregulated already at sphere (4 hpf).

(Rhinn *et al*, 2003); and *rx3* (Stigloher *et al*, 2006)), and differentiation for example of neural tissues (*sox21* (Sandberg *et al*, 2005); COUP-TF family members *nr2f1a*, *nr2f2* and *nr2f5* (Gauchat *et al*, 2004)), or blood progenitors (*tal1* and *lmo2* (Gering *et al*, 2003)). A striking feature of this group is premature initiation of expression in MZ, clearly visible in the microarray-based developmental gene expression curves (Figure 2C; Supplementary Figure 2). Indeed, expression of the majority of Cluster DE genes starts only after 8 hpf in the wild type, but occurs 2–4 h prematurely in MZ (Supplementary Table 7, summarized in Figure 2B). By *in situ* hybridization, we have confirmed premature expression of *nr2f1* (at 4 hpf) and *pax6* (at 7 hpf), in MZ (Figure 2D), several hours before their normal onset of expression. On the basis of the prevailing functions of genes in this group, we classify them as ‘promoters of differentiation’ (PODs). The upregulation of 191

Cluster DE members was confirmed by the independent Affymetrix microarray experiment (Supplementary Methods and Supplementary Table 6).

To evaluate the role of the level of Pou5f1 on the expression dynamics of PODs, we performed transcriptome analysis of M mutants (Supplementary Figure 4; Supplementary Table 11). M mutants are devoid of maternal Pou5f1, and start to express functional *pou5f1* RNA when the zygotic genome is activated at the midblastula transition (MBT, 3 hpf; data not shown). Both Clusters A and DE genes are activated at MBT in M mutants, but expression levels of most Cluster DE genes decline below those of MZ by 5 hpf, and further decline to WT levels (Figure 3A–E). Pou5f1 activating transcriptional repressors, which then in turn repress PODs, may explain this temporal behavior. Among the Pou5f1 targets is the transcriptional repressor *her3*. We performed Her3 overexpression in

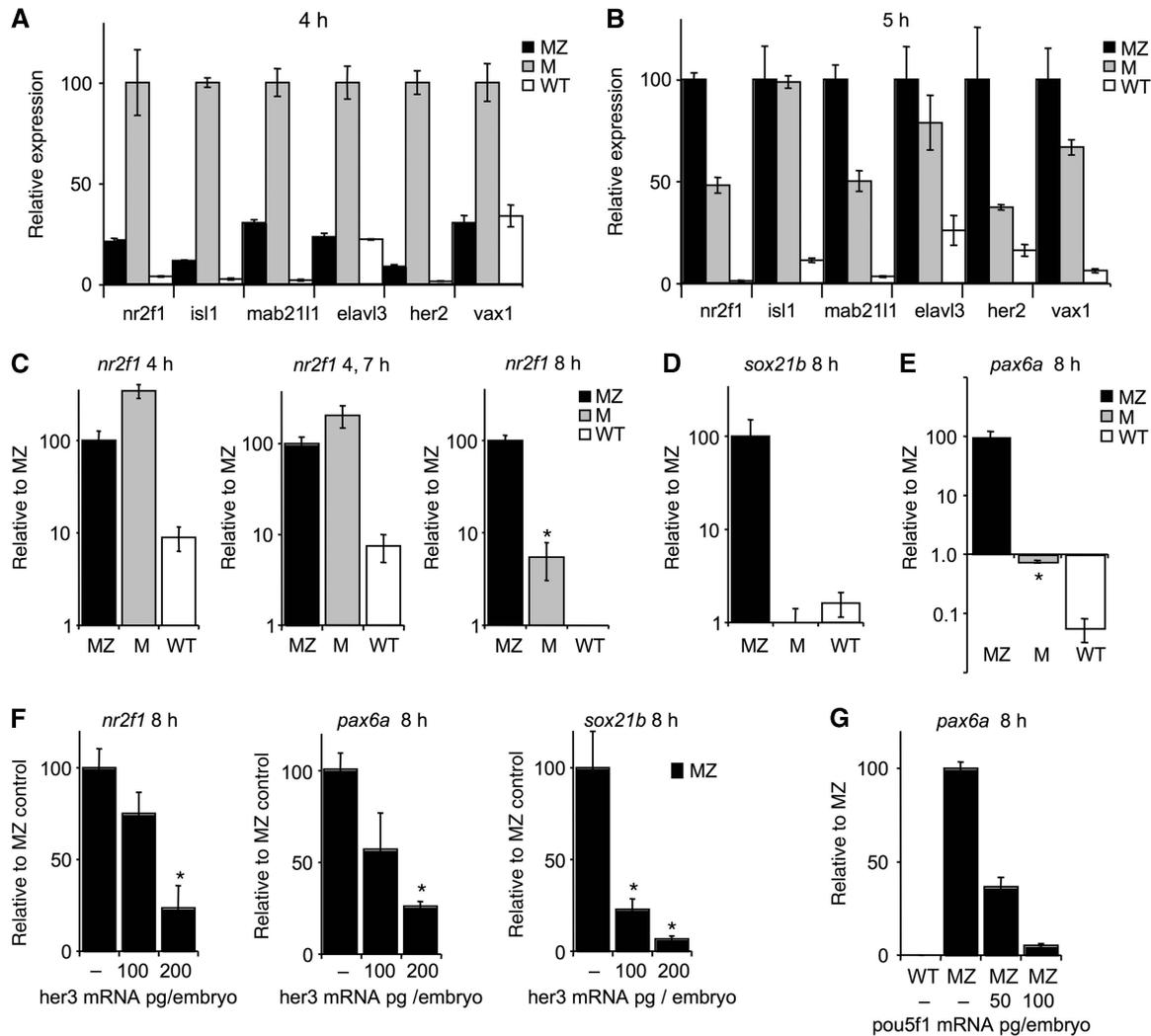


Figure 3 Pou5f1 indirectly represses PODs. (A–E) Temporally delayed repression of PODs in M mutant. (A) Expression of PODs in WT, M, and MZ at 4 hpf. (B) Expression of PODs in WT, M, and MZ at 5 hpf. Relative expression of PODs is shown as percent of maximum (reference to M at 4 hpf but to MZ at 5 hpf, to better reveal changing ratio M to MZ), data are from microarrays. Error bars show standard error of the mean (s.e.m.). (C–G) Expression analysis of PODs (real-time QPCR). RNA was collected from M, MZ, or WT embryos at times indicated. (C) *nr2f1* is an early POD with onset of expression already at 3 hpf in MZ. In M, expression levels of *nr2f1* show gradual decline at 4, 4.7, and 8 hpf. (D) Expression of *sox21b*, a neural POD with late switch time point in MZ (7 hpf, Supplementary Figure 2) was indistinguishable in M and WT at 8 hpf. (E) *pax6a*, neural POD that switches up at 6 hpf in MZ (Supplementary Figure 2), expression levels at 8 hpf. (F) *her3* mRNA injection reduces *nr2f1*, *pax6a*, and *sox21b* expression in MZ. RNA isolated at 8 hpf. (G) *pou5f1* mRNA injection reduces *pax6a* expression in MZ. Averages of three independent experiments are shown, error bars—s.e.m. of three biological replicates. Star indicates significant difference from MZ control ($P < 0.05$, Student's *t*-test).

MZ embryos to address if it could be involved in the repression of *nr2f1*, *pax6*, or *sox21b*, and found that injection of *her3* mRNA significantly reduced the levels of all three PODs (Figure 3F), however, less efficiently than *pou5f1* injection (Figure 3G). Considering that *her3* is expressed soon after MBT (Hans et al, 2004), this suggests that Her3 may be involved in the process of repressing PODs during gastrulation, acting partially redundantly with other Pou5f1-dependent repressors of differentiation (RODs).

Direct and indirect targets of Pou5f1

To determine whether repression of PODs by Pou5f1 is likely to be direct or indirect, we performed a series of Pou5f1 overexpression experiments in which we inhibited translation

of zygotically expressed mRNAs with cycloheximide (CHX; Figure 4A, Supplementary Table 8). We are aware that our CHX-based analysis of zebrafish Pou5f1 targets still needs verification of direct interactions by chromatin immunoprecipitation techniques. Targets induced by Pou5f1 re-expression in MZ in the presence of CHX (termed ‘direct targets’) significantly overlap with Cluster A genes (Figure 4B). In contrast, we found no significant overlap of genes down-regulated by Pou5f1 re-expression in MZ in the presence of CHX with Cluster DE (Supplementary Figure 3; Supplementary Table 9). We analyzed the genomic regulatory regions of genes up- or downregulated in the CHX experiments, as well as those of Clusters A and DE, for enrichment in Pou5f1 consensus DNA-binding sites (Kel et al, 2006). We found an eightfold enrichment ($P < 1e-6$) of Pou5f1-binding sites in the direct

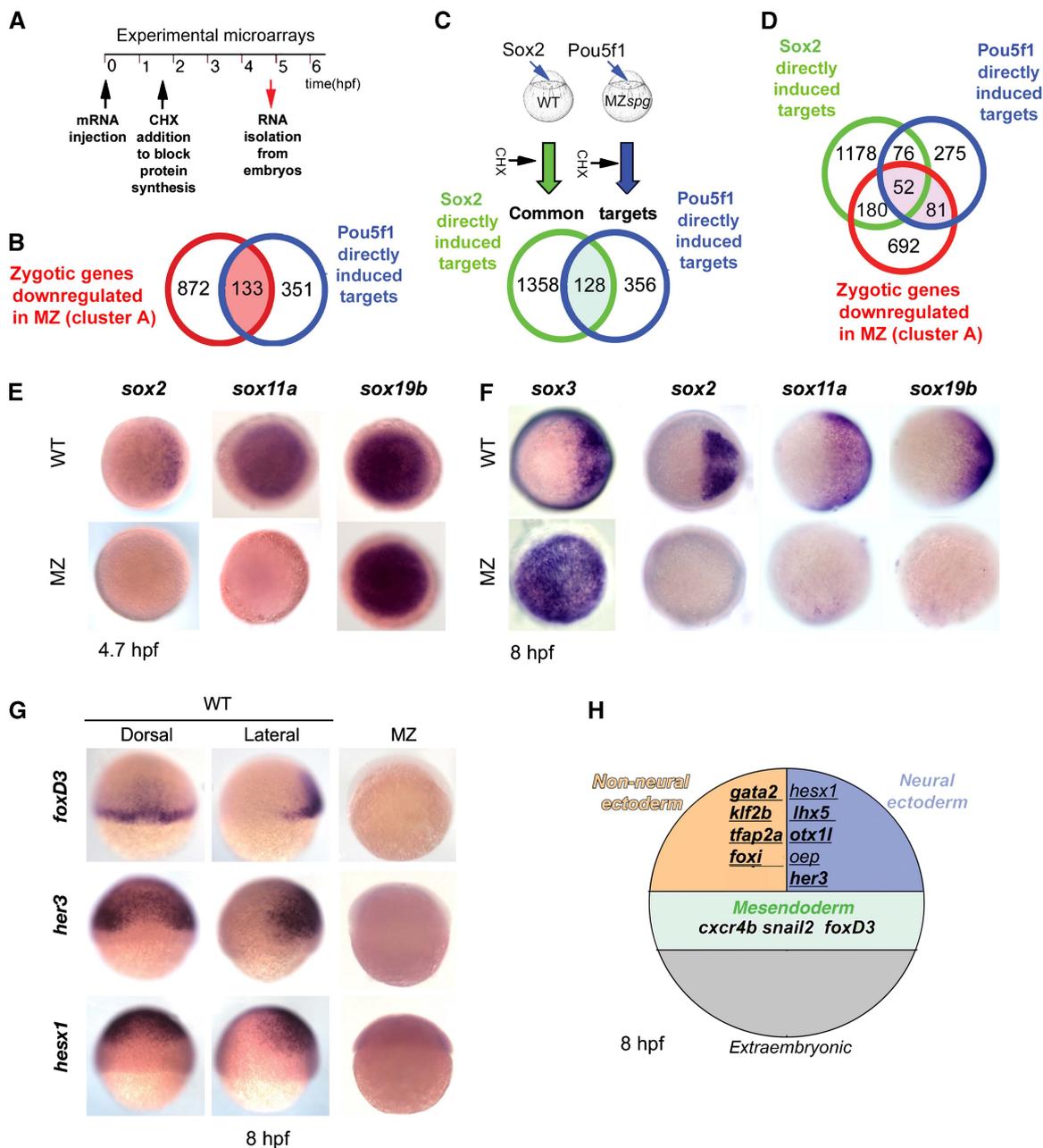


Figure 4 Direct Pou5f1 and Pou5f1/Sox transcriptional targets have spatially restricted expression domains. **(A)** Design of overexpression experiments analyzed by microarray. CHX was added at the 64-cell stage to allow for translation of injected mRNA, but to block translation of the earliest zygotic transcripts. mRNA for analysis was isolated from 4.7 hpf old embryos. **(B)** Overlap of target sets directly activated by injection of Pou5f1 (+ CHX) at 3–4.7 hpf and zygotically regulated clusters A 3–8 hpf. Pink-shaded area highlights 133 developmentally regulated direct targets of Pou5f1. **(C)** Overlap of target sets activated at 3–4.7 hpf by overexpression of Pou5f1 and Sox2 in the presence of CHX, analyzed by microarray. The set of shared direct targets Pou5f1 and Sox2 is shaded green. **(D)** Overlap of target sets directly activated by Pou5f1 (+ CHX) or Sox2 (+ CHX) with Cluster A (3–8 hpf). Pink shading shows Sox-dependent (52) and Sox-independent (81) probe sets of Pou5f1 developmentally regulated Cluster A genes. **(E–G)** Whole-mount *in situ* hybridization, probes and genotypes as indicated. **(E, F)** SoxB1 and SoxC genes require Pou5f1 for proper expression; animal pole views, dorsal at right. **(E)** The expression of *sox2* and *sox11a* is impaired in MZ already by 4.7 hpf; **(F)** *sox2*, *sox19b*, and *sox11a* are strongly downregulated, and *sox3* expression is mislocalized at 8 hpf. **(G)** Pou5f1 direct targets occupy different territories in the embryo: at 8 hpf in WT *foxD3* is expressed in the margin and paraxial mesoderm, whereas the target genes *her3* and *hesx1*, which are regulated by both Pou5f1 and Sox, occupy posterior and anterior neuroectoderm, respectively; no detectable expression in MZ. **(H)** Schematic drawing of zebrafish embryo at 75% epiboly (8 hpf), indicating the spatial segregation of expression domains of Pou5f1 targets. Genes induced by Pou5f1 directly (CHX experiments) are shown in bold; genes induced by Sox2 directly (CHX experiments) are underlined.

targets of Cluster A, in comparison to the background of 6655 genes from the array (Supplementary Table 10). Groups of genes downregulated by Pou5f1 showed no enrichment. Taken

together, these data indicate that Pou5f1 in zebrafish acts primarily as a transcriptional activator, and may represses zygotic genes only indirectly.

Pou5f1 and SoxB1 proteins co-regulate a large target set

Many Pou5f1 target genes in mammalian ES cells seem to be controlled by Oct-Sox enhancers (Boyer *et al*, 2005). We investigated the relative roles of Pou5f1 and SoxB1 proteins by performing Sox2 overexpression in combination with CHX treatment, and compared Sox2 direct targets with those of Pou5f1. The analysis revealed a high positive correlation of genes directly activated by Pou5f1 and Sox2 (128 probes in Figure 4C; Supplementary Figure 3E; Supplementary Table 12), suggesting that Pou-Sox cooperation may regulate many of the identified targets. Comparison of the results from the CHX-treated embryos with Cluster A genes identified 81 probes directly regulated by Pou5f1 alone and 52 probes regulated in combination with Sox2 (Figure 4D). Pou5f1-direct but SoxB1-independent targets include the transcription factors FoxD3, Klf4, Snail2, and Sox11a. In contrast, Her3, Klf2a, Klf2b, Gata2, Foxi, Sox2, and Sox3 expression directly depend on both Pou5f1 and Sox2 (Supplementary Table S12).

SoxB1 proteins act as transcriptional activators and have been reported to have the same DNA-binding specificity in overexpression assays (Okuda *et al*, 2006). In zebrafish, 4 SoxB1 genes are expressed during the first 8 h of development: *sox2*, *sox3*, *sox19a*, and *sox19b*. *sox19b* is present in the zebrafish egg, others are only expressed zygotically (Okuda *et al*, 2006). According to our data, *sox2* is directly regulated by both Pou5f1 and Sox2, *sox19b* is indirectly regulated by Pou5f1 and directly by Sox2; *sox11a* is directly induced in MZ by Pou5f1 in the presence of CHX (Supplementary Table 12 and data not shown). Most of SoxB1 activity, and that of the SoxC gene *sox11a* is significantly reduced in MZ embryos (Figure 4E and F).

Interestingly, tissue-specific expression of Pou5f1 targets correlates with their regulation by Sox2. Out of 12 Pou5f1 transcriptional targets with characterized expression patterns in neural or non-neural ectoderm, 10 are targets of Sox2. In contrast, out of nine mesodermal targets only three are also targets of Sox2 (Supplementary Table 13). SoxB1 group genes are expressed in the whole-ectodermal region until 6 hpf and become restricted to neuroectoderm at 8 hpf (Figure 4E and F; Okuda *et al*, 2006; Dee *et al*, 2007). Our data indicate that SoxB1 activity is sufficient for activation of *her3*, *foxi1*, *klf2b*, *tfap2a*, and *lhx5*. In contrast, *foxD3* and *snail2* are regulated by Pou5f1 independent of zygotic Sox activity, which correlates with their expression in the mesendoderm (Figure 4G, Supplementary Table 13). The spatial distribution of selected Pou5f1 direct targets is summarized in Figure 4H. Taken together, these data suggest that the Pou5f1 transcriptional network is spatially segregated into Sox-dependent (ectodermal) and Sox-independent (mesodermal) subnetworks.

Monophasic and biphasic expression profiles of Pou5f1 target genes

Most Sox-independent direct Pou5f1 targets in WT reach maximal expression levels soon after MBT (*foxD3*, *klf4*, *snail2* in Figure 2A; Supplementary Figure 5). In contrast, Sox2- and Pou5f1-dependent genes tend to have biphasic or delayed

expression, and reach maximum levels at 6–7 hpf (*her3*, *klf2b*, *foxi* in Figure 2A; Supplementary Figure 5). We chose *foxD3* and *her3* as examples for Sox-dependent and -independent targets to study further. MZ embryos injected with *pou5f1* RNA expressed *foxD3* to the level of WT by 8 hpf, whereas *her3* in these embryos reached <10% of WT expression levels (Supplementary Figure 6E and F). In CHX-treated MZ embryos, Pou5f1 and Pou5f1-VP16 (Supplementary Figure 6D and F) were able to activate *her3*; however, the level of *her3* message was about two orders of magnitude below the level in WT embryos. These results suggested that, although the *her3* promoter seems to be directly activated by Pou5f1, additional Pou5f1-dependent indirect input is needed to reach full expression levels.

her3 expression critically depends on a conserved Oct/Sox enhancer

To investigate the mechanism of direct *her3* regulation by Pou5f1, we made a luciferase expression construct Her3_{luc} with 2.2 kb of the *her3* upstream promoter region (Hans *et al*, 2004). We identified a conserved Sox2–Pou5f1 composite ‘SP’ site 11 bp upstream of the TATA-box with high similarity to the Oct4 Position Weight Matrix by Loh *et al* (2006) (Figure 5A). Mutated Her3_{S^m}_{luc} and Her3_{P^m}_{luc} constructs showed reduced activity in WT luciferase assays compared with Her3_{luc} (Figure 5C), and could not be activated by overexpression of Pou5f1 (Figure 5D; Supplementary Figure 7A).

Coexpression of Sox2 and either zebrafish or mouse Pou5f1 led to very strong (20 ×) activation of Her3_{luc}, suggesting cooperative action both in mammalian cell culture (Supplementary Figure 7B) and in zebrafish embryos (Figure 5D). This cooperativity was completely abolished in Her3_{S^m}_{luc} and Her3_{P^m}_{luc} mutant constructs (Figure 5D), suggesting that the interaction between Sox2 and Pou5f1 on the SP-binding site is essential for *her3* promoter activity (Figure 5D). Coexpression of low concentrations of Pou5f1–VP16 strong activator fusion (Lunde *et al*, 2004) with Sox2 in MZ embryos, also resulted in cooperative activation of the Her3_{luc} construct (Supplementary Figure 6H). To test whether Sox2 and Pou5f1 cooperatively bind to the SP site, we performed DNA retardation assays with labeled WT or mutated S^m and P^m oligos (Figure 5B), and either mouse or zebrafish Pou5f1 and mouse Sox2 proteins. Specific complexes with Pou5f1, and Pou5f1/Sox2 were detected when using labeled WT (Supplementary Figure 7C) but not S^m or P^m probes (data not shown). The formation of labeled triple zPou5f1/Sox2/DNA or mPou5f1/Sox2/DNA complexes was effectively prevented with as little as 10-fold excess of unlabeled WT oligo. P^m and S^m unlabeled oligos competed for DNA in these triple complexes with considerably lower efficiency than WT oligo (Figure 5E; Supplementary Figure 7D and E, arrowheads). S^m oligo, which contains a POU-binding half-site, could efficiently compete for DNA in double mPou5f1/DNA or zPou5f1/DNA complexes as efficiently as WT oligo (Figure 5E; Supplementary Figure 7D and E, arrows). Taken together, our data indicate that *her3* promoter activation is likely achieved by cooperative binding of Pou5f1 and SoxB1 proteins to the SP site. Phylogenetic conservation of the SP site and the similar mechanism of binding by mouse and zebrafish Pou5f1 suggest

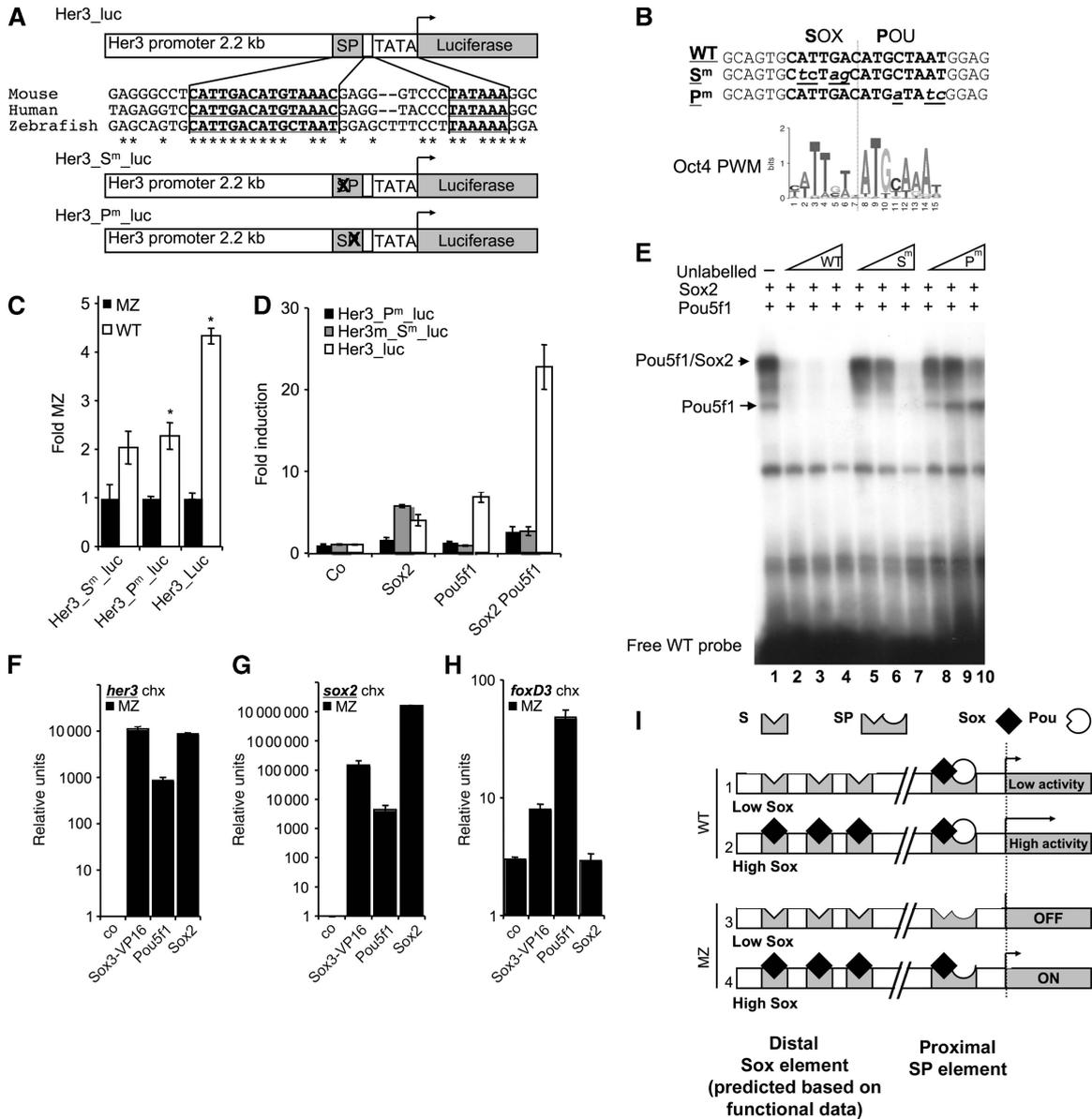


Figure 5 *her3* expression critically depends on Sox–Pou5f1-binding site. Zebrafish Pou5f1 and mouse Sox2 was used in all assays. **(A)** From top to bottom: 2.2 kb *her3* promoter–luciferase constructs with SP–Sox–Pou5f1-binding site. Alignment of the conserved SP-binding site for human, mouse, and zebrafish. Luciferase constructs with mutations in the half-sites of SP. **(B)** Sequence of zebrafish SP-binding site, Sox half-site mutant (S^m) and Pou half-site mutant (P^m). Mutated residues are shown in italics and underlined. Oct4 PWM from Loh *et al* (2006). **(C, D)** Luciferase assays. Zebrafish embryos were injected at 1 cell stage with wild type and mutant reporter constructs and mRNAs as indicated, embryos snap-frozen at 6 hpf, and luciferase activity measured (see Supplementary information for details). **(C)** Her3 promoter activity in WT (white bars) versus MZ (black bars). Asterisks: significant difference of luciferase activity from WT (C) ($P < 0.05$, two-tailed *t*-test), **(D)** Pou5f1 and Sox2 cooperatively induce Her3 promoter construct activity in MZ embryos. **(E)** Gel retardation assay. Labeled WT oligo was incubated with *in vitro* transcribed proteins and unlabeled oligos as indicated at the top. 10 × ; 20 × or 100 × excess of unlabeled WT (lanes 3–5), S^m (6–8) or P^m (9–11) oligo was added. Formation of the Pou5f1/Sox2/DNA complex (arrowhead) can be efficiently prevented by 10 × excess of unlabeled WT, but not mutant oligos. Formation of Pou5f1/DNA complex (arrow) can be efficiently prevented by 10 × excess of unlabeled WT and S^m, but not by P^m mutant oligo. **(F–H)** Quantitative RT–PCR. We expressed Pou5f1, Sox2, or zebrafish Sox3-VP16 fusion in MZ embryos, blocked protein synthesis at the 64-cell stage with CHX, and measured endogenous mRNA at 8 hpf. **(F)** Sox2 and Sox3-VP16 directly activate *her3* to 10-fold higher level than Pou5f1. **(G)** Sox2 or zSox3-VP16 directly activate *sox2* an order of magnitude higher than Pou5f1 activated *sox2*, which is consistent with *sox2* biphasic expression (Supplementary Figure 5). **(H)** In contrast, monophasic *foxD3* is activated by Pou5f1 but not by Sox2. **(I)** Model of biphasic Pou5f1/Sox target activation in WT and MZ embryos. Low and high Sox protein activity are indicated. S—postulated Sox-only binding site, SP—Sox–Pou5f1-binding site, Sox protein—black diamonds, Pou5f1 protein—white notched circle. Regulatory regions of biphasic genes contain both S and SP enhancers. In the presence of Pou5f1 and low concentrations of Sox, SP sites are preferentially occupied, resulting in low transcriptional activity (G1). In the presence of Pou5f1 and high Sox concentrations, S sites are also occupied, resulting in higher transcriptional activity (G2). In the absence of Pou5f1, and at low Sox concentrations, these biphasic genes remain silent (G3), whereas high Sox concentrations alone can induce their transcriptional activity (G4).

that the SP site may have a similar function in the activation of *hes3* (homologue of *her3* in mammals; Hirata *et al*, 2001; Hatakeyama *et al*, 2004).

***her3* responds to high Sox2 levels independent of Pou5f1**

According to our data, *her3* transcription is directly activated by Sox2 overexpression in WT zebrafish embryos. Thus, either (1) in the presence of Pou5f1 and initially low expression of SoxB1 genes, additional Sox2 is required for the full activation of *her3* via SP and other Pou5f1-dependent Sox-binding sites or (2) direct activation by Sox2 occurs via Pou5f1-independent sites in *her3*. Overexpression of Sox2 or Sox3-VP16 in MZ can directly activate *her3* and *sox2* but not *foxD3* (Figure 5F-H), supporting the second mechanism. Taken together, our data favor a mechanistic explanation for the biphasic expression profile of *her3* (Figure 2A) based on independent Pou-Sox and Sox-only regulatory modules (Figure 5I). First, maternal SoxB1 (Sox19b) and Pou5f1 may bind to the SP enhancer and activate a low level of *her3* expression during the period from 3 to 6 hpf (Figure 5I1). In MZ, this activation fails due to the absence of Pou5f1 (Figure 5I3). The second phase of *her3* activation (6–8 hpf) may depend on high SoxB1 activity required to bind Sox-only regulatory modules (Figure 5I2). We hypothesize that the use of Pou-Sox and Sox-only regulatory modules may also explain the biphasic activation of other targets (i.e. *foxi*, *klf2b*, Figure 2A), where activation thresholds to initiate high-level expression may differ for each gene.

The structure of Sox gene control is suited to provide temporal information to the activation of Pou5f1-Sox targets: maternal *sox19b* activity helps to start the network, and then the immediate Pou5f1-dependent target *sox11a* builds up sufficient SoxB1 and SoxC activity to regulate *sox2* to high levels and to activate a larger set of Pou5f1-SoxB1-dependent zygotic targets. Essentially, this enables an early response that is mostly dependent on Pou5f1-Sox19b, and a late response that may be more dependent on additional zygotic SoxB1 group activation. This hypothesis is in agreement with our finding that the majority of the Pou5f1/SoxB1-dependent targets have bimodal or delayed expression, whereas Pou5f1-only regulated targets are activated soon after zygotic transcription start.

A dynamic network model of Pou5f1-dependent temporal control

To summarize and theoretically check the consistency of our current findings on Pou5f1/SoxB1-dependent versus Pou5f1-only regulation, we built a small dynamic network model that links the temporal control of target genes to regulatory principles exerted by Pou5f1 and SoxB1 proteins (Figure 6A; Supplementary information). The model was derived using a phenomenological approach based on binary transcription responses (Veflingstad and Plahte, 2007) to reflect the temporal switching behavior of most genes (Supplementary Figure 1A and B). The model parameters were determined by a fit to the WT and MZ gene expression data (Materials and methods). The optimized model highlights two qualitatively different

temporal expression modes of Pou5f1 downstream targets: monophasic for targets depending only on Pou5f1 (*foxd3*), and biphasic for Pou5f1- and SoxB1-dependent targets (*sox2* and *her3*; Figure 6B). Interestingly, the activation of biphasic targets is dominated by the SoxB1 factors and the timing of expression strongly depends on the SoxB1 activation threshold as deduced from the model parameters. To test whether the model is also able to correctly predict a different genetic condition, we simulated the M mutant, which is lacking maternal Pou5f1 (Figure 6B, blue, dashed curve). The model predicts an overall shift in the developmental program. Most importantly, the *sox2* and *her3* expression is rescued with a delay of about 2 h. The model predictions were checked experimentally by quantitative RT-PCR (Figure 6B, blue dots). Most predictions are in good agreement with the experimental data, for example the delayed rescue of the *sox2* and *her3* expression pattern. Only the *Sox11a* expression pattern differs from the prediction, which points to additional regulatory input not implemented in the model. With respect to the ‘POD’ *nr2f1*, the model correctly predicts the efficient downregulation by zygotic targets of Pou5f1 (Figure 6B).

To understand the quantitative contributions of Pou5f1 and the Sox factors to the temporal regulation of biphasic targets such as *sox2* and *her3* in more detail, we performed a systematic parameter screen (Figure 6C). We varied the Sox activation threshold as well as the relative contributions of Pou5f1 and the Sox factors to the activation of an exemplary biphasic target gene, and calculated the time the target needs to reach its half-maximal expression level. Interestingly, it turns out that the Sox factors control all of the timing aspects of biphasic target gene expression. Thus, the Sox threshold of activation is the major determinant for the start of the second phase of expression. To achieve the temporal control over target gene expression, Sox factors must activate their targets stronger than Pou5f1 alone (see Figure 6C). Otherwise the genes behave monophasically.

Evolutionary conservation

Mouse orthologs have been identified for 8341 genes in the zebrafish Agilent microarray (conserved genes). Among those genes, 3284 have differential expression between MZ and WT at any time point of the developmental curve. We compared Pou5f1-dependent zebrafish genes to previously identified mouse POU5f1/OCT4 targets. We used data from two microarray studies, based on *Pou5f1*-siRNA-mediated knockdown (Loh *et al*, 2006) (Supplementary Table S17 therein) and tetracycline-inducible loss-of-POU5f1 in ZHBTc4 cells (Sharov *et al*, 2008) (Supplementary Table S4 therein). The Venn diagram (Figure 7A) shows that for 15% of the zebrafish Pou5f1 targets, mouse orthologs were regulated in both microarray experiments. There is an even stronger overlap of 45% between our fish data and the combined set of Oct4 targets from both experiments. Further, 23% of mouse orthologs of Cluster A and 29% of those of Cluster DE were regulated in both mouse experiments (Figure 7B and C). This conservation between fish and mouse target sets may be considered unexpectedly high, given the modest overlap of 9.1% between the POU5f1/OCT4 targets identified for mouse and human (Figure 8A in Loh *et al*, 2006). For example, mouse

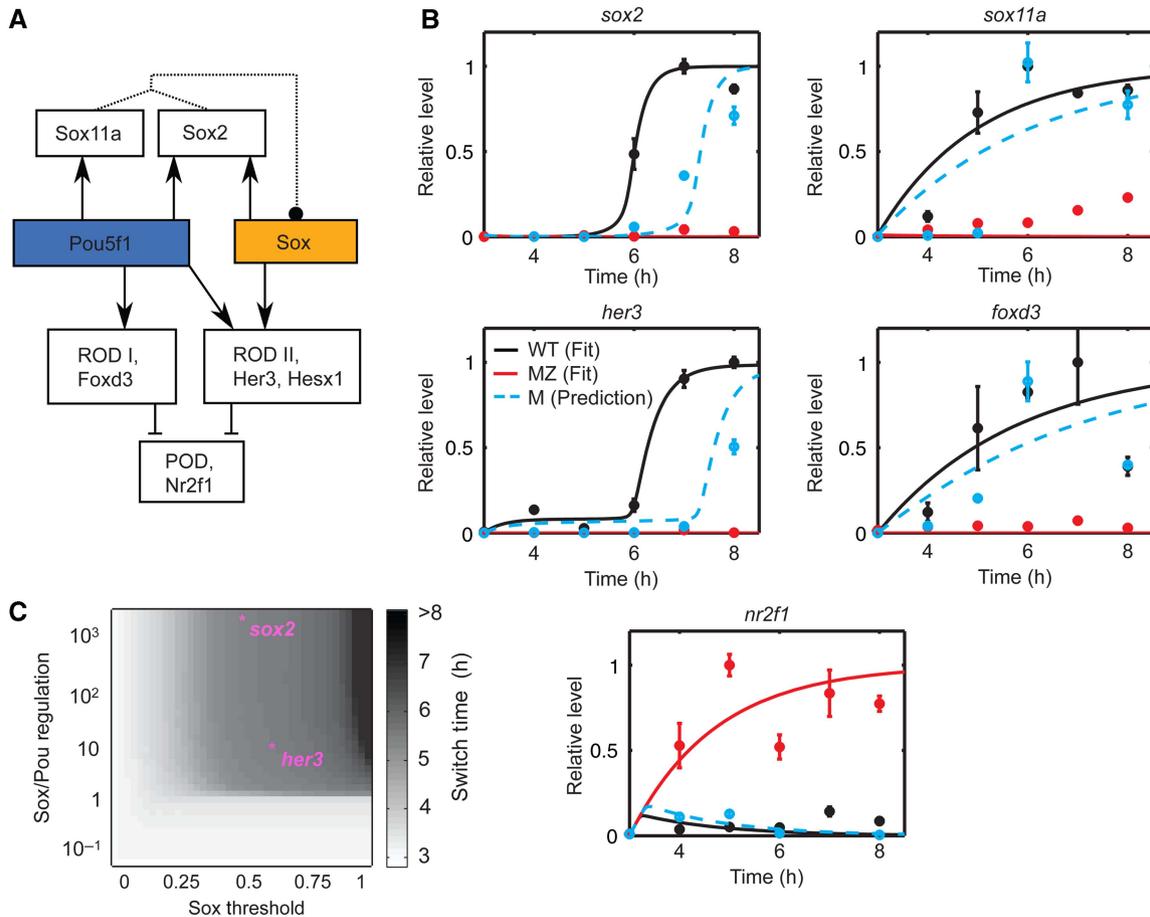


Figure 6 Quantitative mathematical model of Pou5f1 and SoxB1 activity dependent temporal dynamics of target gene expression. **(A)** Interaction chart used to build the model. Arrows indicate direct transcriptional activation. Rounded arrows indicate contributions of individual Sox genes to summarized Sox activity. RODs—repressors of differentiation, PODs—promoters of differentiation, each with exemplary genes listed. **(B)** Simulated time course of selected network components in WT condition (black curve), MZ condition (red curve). Model parameters were determined by a fit to the WT (black circles) and MZ (red circles) expression data (see Supplementary information). The optimized model was used to predict the M condition, that is, non-functional maternal Pou5f1 (blue, dashed curve). The model predictions were subsequently confirmed by additional Q-PCR measurements (blue circles). **(C)** Temporal control of target genes by Pou5f1 and the Sox factors. The onset of the second phase of expression (indicated in gray) was measured by the time when the target gene reaches its half-maximal level and is plotted as a function of the Sox threshold and the relative contribution of the Sox and Pou regulation (see Materials and methods). Note that, to control the timing of target gene expression, the contribution of the Sox factors must exceed the contribution by Pou5f1. The model predicts a smaller contribution of Pou5f1 to the activation of biphasic targets (*her3*, *sox2*) compared with the Sox factors.

hes3 appears as a target on analysis of the primary microarray data (Sharov *et al*, 2008), but was filtered out in the study based on algorithms used to evaluate ChIP and microarray data. To determine whether POU5f1/OCT4 in ES cells regulates *Hes3*, we tested whether *Hes3* expression persists in mouse ES cells when POU5f1/OCT4 expression is turned off. We used the ZHBTc4 ES cell line (Niwa *et al*, 2002) for inducible POU5f1/OCT4 knockdown (Supplementary information; Supplementary Figure 8). *Hes3* expression disappeared shortly after depletion of *Pou5f1/Oct4* mRNA, revealing that *Hes3* expression depends on POU5f1/OCT4. Thus, *her3/Hes3* are indeed evolutionary conserved target of Pou5f1 in zebrafish and mouse. Boyer *et al* (2005) report ChIP binding of POU5f1/OCT4 to the human *HES3* genomic region Chr1:6233720–6234419, suggesting that regulation may also be relevant in human.

To identify a conserved minimal core set of relevant Pou5f1 targets, we compared our Cluster A gene list with the core lists

of mouse POU5f1/OCT4 targets defined by combining microarray and ChIP data (Supplementary Table S17 in Sharov *et al*, 2008). Our analysis identifies a set of 93 genes (Figure 7D; Supplementary Table 14), with many developmental transcription factors, including *sox2*, *foxD3*, *klf2*, *klf4*, and *pou3f1*, and signaling pathway components.

A significant part of the Pou5f1 downstream transcriptional network has been conserved from fish to mammals. Considered together with the similarities in the expression patterns of Pou5f1 during gastrulation stages of all vertebrates studied, similarity of transcriptional targets suggest equivalent Pou5f1 functions during the pregastrulation and gastrulation period of vertebrate embryogenesis. Therefore, we tested whether mouse POU5f1/OCT4 was able to rescue MZ embryos. Injection of mRNA encoding mouse POU5f1/OCT4 into MZ embryos (Figure 8A) was able to restore normal zebrafish development to an extent comparable with zebrafish *pou5f1/pou2* mRNA (Figure 8B and C). In all, 15 out of 115 MZ

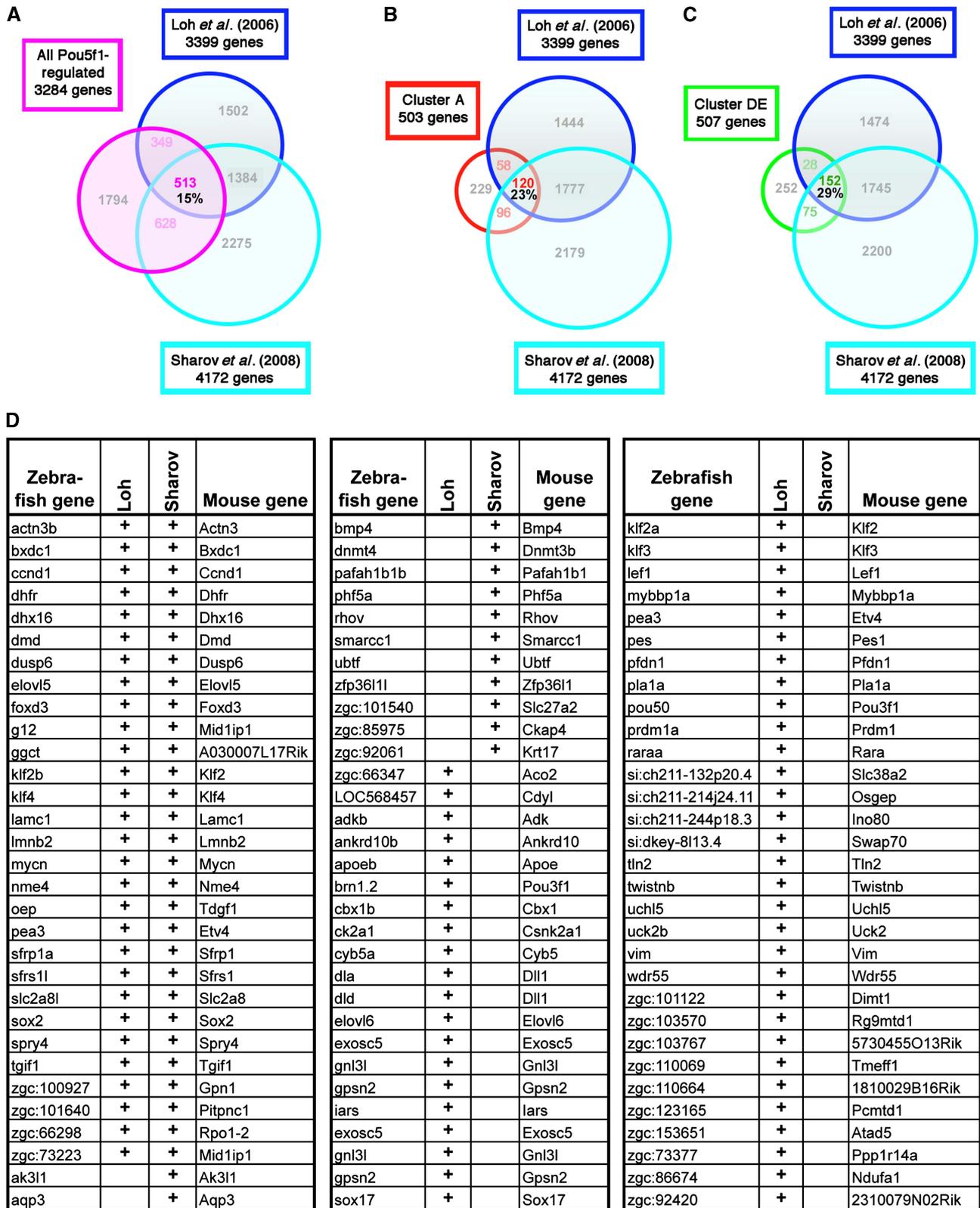


Figure 7 Evolutionary conserved Pou5f1 targets in zebrafish and mouse ES cells. (A–C) We restrict our comparison of Pou5f1 targets to those genes for which orthologs have been identified in zebrafish and mouse. (A) The Venn diagrams show the overlap between zebrafish microarray-based genes list of Pou5f1 targets at 3–8 hpf from this study with two previously reported mouse Pou5f1/Oct4 microarray-based target gene sets (Loh *et al*, 2006; Sharov *et al*, 2008). (B) Overlap of the same mouse gene sets with zebrafish Cluster A, and (C) with Cluster DE. (D) Comparison of our Cluster A to a list of mouse Pou5f1 presumably direct targets selected from microarray and ChIP experiments (Loh *et al*, 2006; Sharov *et al*, 2008), identifies a set of 93 conserved Pou5f1 targets.

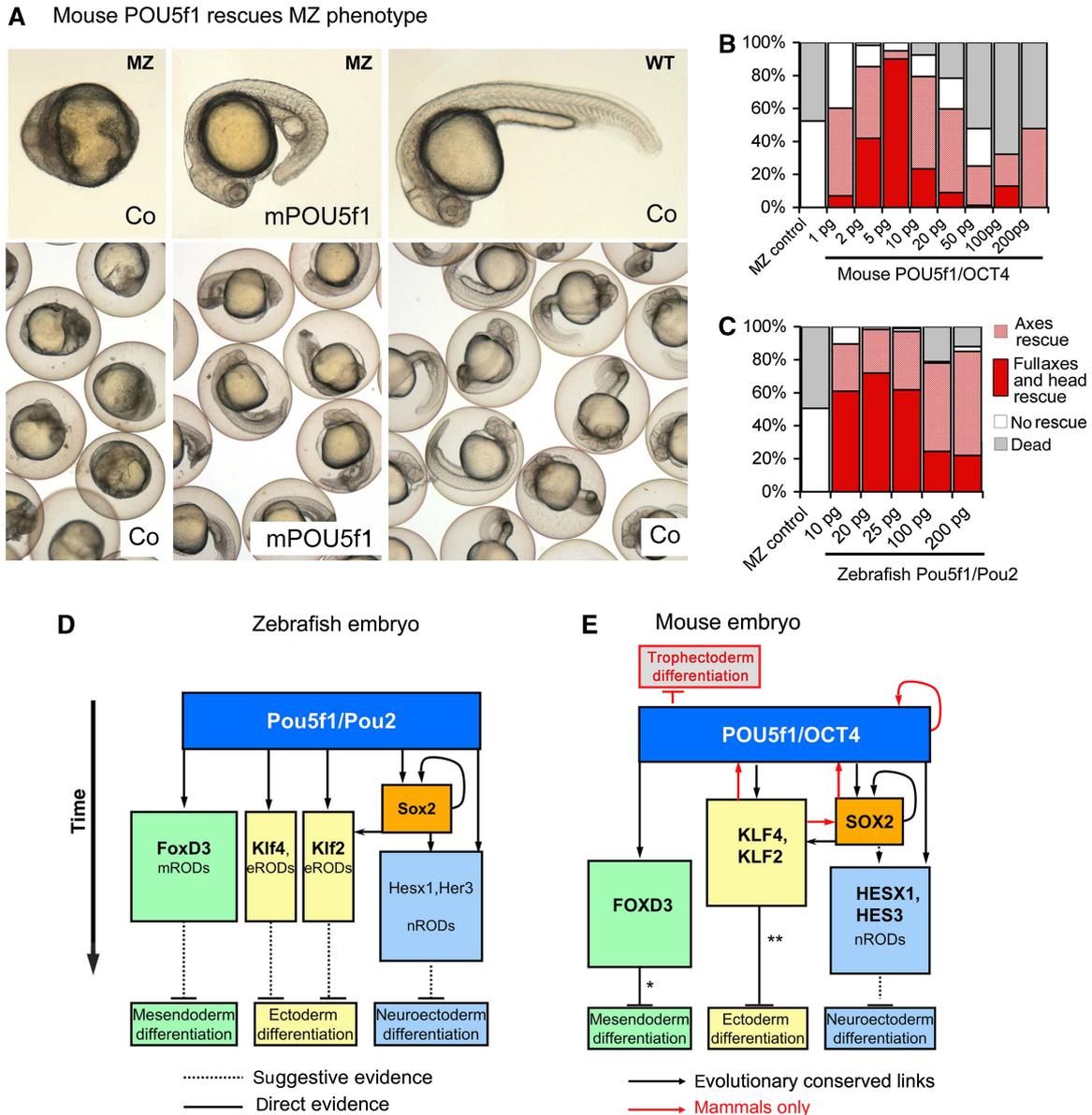


Figure 8 Mouse POU5f1 rescues zebrafish Pou5f1 deficiency. Hypothetical scheme for evolution of the vertebrate downstream Pou5f1 network. **(A)** Mouse POU5f1 rescues zebrafish *pou5f1* deficiency. From left to right: Phenotypes of 24 hpf old non-injected MZ embryos, MZ embryos injected with mouse *Pou5f1* mRNA (2 pg/embryo), and non-injected WT embryos. Top row, individual embryos dechorionated, lateral views, dorsal up. Bottom row: overview of embryo groups from same injection experiment. **(B)** Rescue statistics for different amounts of injected mouse *Pou5f1* mRNA. MZ embryos were injected with the control RNA at 25 pg/embryo or mouse *Pou5f1* mRNA at the amounts indicated, phenotypic classes were counted at 24 hpf and expressed as percent of whole. $N=20-213$ embryos per condition. **(C)** Rescue statistics for different amounts of injected zebrafish *pou5f1* mRNA. MZ embryos were injected with control RNA at 100 pg/embryo or zebrafish *pou5f1* at the amounts indicated. Phenotypic classes were counted at 24 hpf and expressed as percent of whole. $N=102-186$ embryos per condition. (B, C). Phenotypic classes: red solid bars—full rescue of body axes, head, tail, eyes; red-shaded bars—incomplete rescue (body axes and tail form, malformed head, no eyes); white bars—no rescue (no body axis); and gray bars—dead embryos. Note that the rescue capacity of mouse *Pou5f1* at 2–5 pg/embryo (B) is comparable with that of zebrafish *pou5f1* at 10–25 pg/embryo (C). **(D, E)** Hypothesis on Pou5f1 downstream network evolution in vertebrates. **(D)** Pou5f1 downstream transcriptional network in zebrafish embryogenesis from MBT until 8 hpf. Pou5f1 directly activates RODs, which regulate the timing of differentiation by repressing their targets (PODs) in a tissue-specific manner. mRODs, eRODs, nRODs—repressors of mesendodermal, ectodermal, and neural differentiation, respectively. Dotted line—suggestive evidence, solid line—direct evidence from our work. **(E)** A conserved subset of mouse POU5f1/OCT4 transcriptional network (RODs) is used during mouse gastrulation to ensure the timing of lineage-specification events (black arrows), similar to zebrafish. Evolutionary novel, mammalian-specific elements of the network architecture (red arrows) operate at earlier developmental stages, serve the function of inhibiting trophoctoderm differentiation and ensure self-maintenance of ES cell culture. We hypothesize that a subnetwork of POU5f1/OCT4 transcriptional targets, including RODs known from ES cell studies, is also used in mouse development, and acts in a tissue-specific manner during pregastrulation and gastrulation stages. If this is true, loss-of-function of those conserved RODs in ES cells may result in the tissue specificity of differentiation. Indeed, two recent publications support this idea. Asterisks indicate references demonstrating tissue-specific action of RODs (*) FoxD3 (Jiang *et al*, 2008) and (**) Klf4 (Liu and Labosky, 2008) in ES cells. We want to emphasize that the schemes depict only a subset of biological functions of Pou5f1 with an emphasis on germ layer development and temporal control, but do not cover, for example, activities controlling chromatin state or proliferation.

embryos (13%) injected with 2 pg/embryo, and 9 of 19 MZ embryos (47%), injected with 5 pg/embryo of mouse *Pou5f1/Oct4* mRNA, respectively, developed into swimming larvae beyond 6 days postfertilization. This result strongly supports the hypothesis of evolutionary conserved functions of Pou5f1 orthologs during blastula and gastrula stages.

Discussion

We identified changes in Pou5f1 target gene expression both with respect to their expression level and temporal behavior. Several targets directly activated by Pou5f1 encode known RODs, of which we analyzed *her3* in detail. Expression of a second, large group of genes encoding developmental regulators of differentiation normally acting during organogenesis (PODs) is temporally shifted during gastrulation stages in MZ. Our analysis of potential direct transcriptional interactions by suppression of translation of intermediate zygotic Pou5f1 or Sox targets, enabled us to distinguish SoxB1-dependent and -independent subgroups of the Pou5f1 transcriptional network. In summary, Pou5f1, rather than controlling a small number of decisions, instructs the global gene regulatory landscape in the embryo by controlling temporal dynamics of gene expression within diverse developmental modules by Pou5f1-only and Pou5f1-SoxB1-dependent mechanisms.

Pou5f1 directly controls ROD

Pou5f1 overexpression experiments in combination with translational block suggest that Pou5f1 on its own acts directly as a transcriptional activator. Thus, other direct Pou5f1 targets probably mediate repression of PODs by Pou5f1. Several genes directly activated by Pou5f1 have previously been shown to repress differentiation, or to encode transcriptional repressors: *Klf4* and *Klf2b* (Rowland *et al*, 2005), *FoxD3* (Yaklichkin *et al*, 2007), *Her3* (Hans *et al*, 2004), and *Hesx1* (Kazanskaya *et al*, 1997; Quirk and Brown, 2002). We found that *her3* mRNA injection can suppress the premature expression of *nr2f1*, *pax6*, and *sox21b* in MZ embryos. However, we could not directly demonstrate the contribution of *FoxD3* and *Her3/Hesx1* to the repression of premature differentiation in zebrafish, likely because of redundancies involving other ROD controlled by Pou5f1. M embryos are devoid of maternal Pou5f1 expression and have an MZ transcriptional landscape until zygotic transcription starts. During the first hour after zygotic activation of Pou5f1, its activity leads to an upregulation of both RODs and prematurely expressed PODs. However, as RODs appear, expression of PODs starts to decline, compatible with a Pou5f1-indirect mode of repression.

Regulatory relationships of Pou5f1, SoxB1, and their target genes

Our data indicate that Pou5f1 directly controls most SoxB1 and C activity. As Pou5f1 executes many of its functions with a SoxB1 protein as its binding partner (Remenyi *et al*, 2003; Masui *et al*, 2007), we evaluated the role of Sox proteins in the

zebrafish Pou5f1 regulatory network, and found that the SoxB1 group proteins and Pou5f1 coactivate a large target gene set. SoxB1/Pou5f1 and Pou5f1 targets differ in their spatial expression: SoxB1/Pou5f1 targets are predominantly restricted to SoxB1 zygotic expression domains (ectoderm and neural ectoderm), whereas the majority of SoxB1-independent Pou5f1 targets are expressed in mesendoderm, a territory largely free of Sox transcripts.

her3 and *foxd3* regulation reveal SoxB1-dependent and SoxB1-independent Pou5f1 control mechanisms

We identified a Pou5f1–SoxB1 composite regulatory site in the proximal *her3* promoter, and show that this site is required for proper activation of *her3* expression. Pou5f1 and Sox2 mutually enhance the binding of the partner protein to this site, enabling Pou5f1 to activate transcription even at low levels of Sox2. The SoxB1 site in the *her3* proximal element deviates from the canonical SoxB1-binding site, and may thus be strongly bound by SoxB1 only in the presence of Pou5f1. A similar regulation has been previously shown for composite Sox-Pax sites (Kamachi *et al*, 2001). Our data also suggest the existence of distinct SoxB1-dependent elements outside the *her3* proximal promoter, which cause the activation of *her3* by high Sox2 concentration even in the absence of Pou5f1. *Her3* first appears at low levels directly after zygotic transcription starts, and rises to high levels in neural ectoderm from midgastrula onwards. Distinct SoxB1- and Pou5f1/SoxB1-dependent regulatory modules, with different requirements for SoxB1 protein levels, combined with the dynamics of SoxB1 expression, can explain this biphasic activation of *her3*. We hypothesize that other bimodal Pou5f1 targets may be subject to similar control mechanisms. In contrast, monophasic direct Pou5f1 targets, like *foxD3*, are directly activated by Pou5f1, and strictly depend on Pou5f1 in their activation—high levels of Sox2 overexpression alone cannot rescue *foxD3* expression in MZ mutants.

Dynamic model of the Pou5f1–Sox network

We evaluated the quantitative contributions of Pou5f1 and the Sox factors to target gene regulation by mathematical modeling. The model predicts two qualitatively different expression behaviors: monophasic expression of genes activated by Pou5f1 alone and biphasic expression of genes initially activated by Pou5f1 and subsequently by the Sox factors. This biphasic pattern requires a dominant quantitative contribution of the Sox factors to the activation of biphasic targets, as expressed in terms of relative transcription rate. This requirement correlates with the stronger capacity of Sox2 to activate the biphasic targets *her3* and *sox2* in comparison with Pou5f1 (Figure 5F and G). Interestingly, the activation threshold of the Sox factors exerts the main temporal control of biphasic targets. As a consequence, the timing of the expression of Pou5f1 and Sox-dependent targets is relatively independent of the maternal Pou5f1 level (as long as maternal Pou5f1 is above a certain threshold),

while it is very sensitive to zygotic Sox levels. The surprising finding that the zebrafish early regulatory network is relatively independent of total Pou5f1 levels is experimentally validated: M, *Zspg*, as well as MZ embryos rescued by Pou5f1 mRNA injection all develop normally until the end of gastrulation, despite the vastly different Pou5f1 activity levels in these embryos. Thus, in contrast to ES cells, which may be sensitive to POU5f1/OCT4 levels (Niwa *et al*, 2000, 2002), zebrafish embryos are less susceptible to changes in Pou5f1 concentration.

Pou5f1-dependent temporal control of development

Pou5f1 controls a set of transcriptional repressors specific for each major embryonic compartment (Figure 4H), which appears to delay differentiation by repressing differentiation genes until the end of gastrulation. Our findings emphasize two distinct mechanisms by which Pou5f1 contributes to this repression. First, Pou5f1 alone can activate repressor genes (ROD I, e.g. *foxD3*, Figure 6A) immediately after midblastula transition (MBT). Second, Pou5f1 acts together with SoxB1 proteins to activate repressor genes (ROD II, e.g. *her3*) with 2–3 h of delay after MBT. For this second mechanism, it does not appear that Pou5f1 concentration alone determines timing, but that the level of SoxB1 activity is crucial for setting the time point of target gene activation. This system may provide a mechanism to fine-tune the order of Pou5f1/SoxB1 target gene activation based on changes in the affinity of the SoxB1 regulatory elements: lower affinity elements would drive the initiation of expression only at later stages when sufficient SoxB1 activity has accumulated. Both systems together appear to determine a significant portion of the temporal regulatory landscape of the embryo as judged from the changes we observe in timing of target gene activation (Supplementary Figure 1).

Evolution of the Pou5f1–Sox core regulatory network

Systematic analysis of the POU5f1/OCT4-dependent stem cell network in mouse and human ES cells has revealed a surprising complexity of regulated genes and interactions (Boyer *et al*, 2005; Chickarmane *et al*, 2006; Loh *et al*, 2006; Masui *et al*, 2007; Zhou *et al*, 2007; Jiang *et al*, 2008; Kim *et al*, 2008; Ying *et al*, 2008). Although a significant level of mechanistic understanding has been achieved, complexity and logic of the mammalian stem cell regulatory network has been difficult to comprehend in the absence of knowledge on how it may have evolved. The significant overlap between zebrafish and mammalian Pou5f1 targets (Figure 7) together with the ability of mouse POU5f1/OCT4 to functionally replace zebrafish Pou5f1 (Figure 8A–C), suggests that the mammalian network may have evolved from a basal situation similar to what is observed in teleosts. We propose models that emphasize the evolution of Pou5f1-dependent transcriptional networks during development of the zebrafish (Figure 8D) and mammals (Figure 8E). Our representation separates the evolutionary ancient subnetworks downstream of Pou5f1,

which are presumably used for controlling the timing of differentiation during gastrulation in all vertebrates (Figure 8D and E, black arrows). In this conserved subnetwork, Pou5f1 switches on the expression of germlayer-specific RODs to prevent precocious differentiation in mesendoderm, non-neural and neural ectoderm. We hypothesize that some components of these core Pou5f1-downstream subnetworks may have been co-opted for additional evolutionary novel functions during early developmental stages in mammals. This functionality may be based on addition of some novel interacting partners as well as feedback regulatory loops (Figure 8E, red arrows) to the existing interaction network. The model predicts lineage-specific differentiation on knock-out of RODs. Indeed, inducible knockout of FoxD3 in ES cells and embryoid bodies leads to abnormal differentiation towards mesendodermal lineages without interfering with the differentiation towards ectoderm and neuroectoderm (Liu and Labosky, 2008). Klf-knockout induces lineage-specific differentiation towards ectodermal fate with simultaneous downregulation of the self-renewal genes (Jiang *et al*, 2008). The phenotype of both knockouts can thus be explained by evolutionary conservation of tissue-specific functions of FoxD3 and the Klf factors in the vertebrate lineage.

How could changes between an ancient zebrafish-like network and the mammalian network have evolved? The ancient part, including the downstream direct targets Klf, FoxD3, Sox2, and Her3, and the structure of Pou5f1/SoxB1 enhancers, remains a component of the mammalian network, and is conserved between zebrafish and mouse. At the experimental level, this conservation has been confirmed, as overexpression of mouse Pou5f1/Oct4 can completely rescue MZ mutant embryos (Figure 8A–C). However, the degree of this rescue is not reciprocal: zebrafish Pou5f1 replaces mouse POU5f1/OCT4 function in the maintenance of self-renewal in mouse ES cells only to a limited extent (Niwa *et al*, 2008). This suggests that, although POU5f1/OCT4 has acquired additional functions specific for mammalian embryogenesis, it may have also acquired novel interacting partners during evolution to perform these functions. The self-activation loop of *Pou5f1*, characteristic for mammals, is not present in early zebrafish development, as Pou5f1 or SoxB1 cannot efficiently activate *pou5f1* expression in zebrafish in the presence of CHX (Supplementary information; Supplementary Figure 6A–C). In Figure 8E, we have indicated the stem cell maintenance ‘feedback’ components in red, including red arrows back from Sox2 activating *Pou5f1*.

One evolutionarily novel component that stabilizes the system could be Nanog. Nanog in fish is probably not integrated into a pluripotency regulatory circuit (Camp *et al*, 2009). In chick, a *nanog* ortholog is present, and chicken Pou5f1 can rescue *Pou5f1*-deficient mouse ES cells (Laval *et al*, 2007). This new Nanog module may have contributed to the establishment of a mutual autoregulatory loop with *Pou5f1*, and the *SoxB1* genes. Interestingly, *Klf4* and *Klf2* have also been drawn into this regulatory loop during evolution.

In summary, our data indicate that elements of the evolutionarily ancient embryonic Pou5f1-dependent subnetworks that control developmental timing and differentiation have been integrated as modules of Pou5f1 pluripotency

control in ES cells in mammals. As the Pou5f1 downstream regulatory nodes revealed in our zebrafish model are likely conserved across vertebrates, we envision that their knowledge will contribute to the effort of directing differentiation of pluripotent stem cells to defined cell fates.

Materials and methods

Microarray-based transcriptome analysis

For transcriptome analysis, WT embryos of AB × TÜB strain crosses (<http://www.ZFIN.org>) and MZ or M embryos carrying the *m793* allele of the *spg* mutation were used (Belting *et al.*, 2001) ZFIN ID: ZDB-GENE-980526-485, ZDB-GENO-081023-1). Embryos of the genotypes indicated were precisely staged either following *in vitro* fertilization ($t=0$ hpf), or using natural crosses and staging at the four-cell stage ($t=1$ hpf).

RNA preparation for microarrays: 60–100 embryos per sample were snap-frozen in liquid nitrogen, and total RNA was isolated using the RNA Easy kit (Qiagen). Sample quality was assessed in an Agilent Bioanalyzer 2100, using the RNA 6000 nano Assay Kit. Samples were processed by Agilent Technologies Two-Color Microarray-Based Gene Expression Analysis kit, hybridized with Agilent 22 K zebrafish arrays (number 015064 with 21527 features), scanned on an Agilent scanner and processed using the GE-v5_95_Feb07 Agilent protocol. All experiments were performed in triplicate using three independent RNA isolations; except the 1 h time point of the developmental curve, and the WT4, WT5 time points in the M experiment, which were performed in duplicate. Data were normalized across both time curves and genotypes using quantile normalization (Genedata). Independent validation of the Agilent microarray results was performed in three ways: (1) comparison with an Affymetrix Gene Chip microarray experiment, (2) real-time QPCR, and (3) *in situ* hybridization (see Supplementary information).

The primary microarray data from all Agilent and Affymetrix arrays generated in this study have been submitted to GEO (<http://www.ncbi.nlm.nih.gov/geo/>) and are stored under accession series number GSE 17667.

Evaluation of expression time-series data for identification of switch behavior

The estimation of switching times followed the approach of Sahoo *et al.* (2007). The optimal fit of the temporal expression profile to a reference pattern including a single transition (=switch) from high to low or from low to high level was computed using F-statistics. The switch time of a gene refers to the transition point of the best matching reference pattern. The significance level was controlled at an estimated FDR of 10%. The shift in gene expression between WT and MZ was calculated by the difference in estimated switching times between both genotypes. All statistical calculations were performed with MATLAB.

Plasmids used in this study

Zebrafish expression constructs CS2 + Pou5f1 and CS2 + Pou5f1-VP16 have been described (Lunde *et al.*, 2004). Mouse expression constructs CS2 + Oct4, CS2 + mycOct4, and CS2 + Sox2 were kindly provided by A Tomilin. To obtain Her3 expression construct, CS2 + Her3, the Her3 ORF was PCR amplified from cDNA using primers with incorporated restriction enzyme sites for BamHI and XbaI, subcloned into PCRII-Topo vector (Invitrogen), sequenced and subcloned into CS2 + vector (Turner and Weintraub, 1994) via BamHI/XbaI sites. To obtain the Her3-Luc reporter construct, 2.2 kb upstream of Her3 coding sequence was PCR amplified from Her3-Gal4 plasmid kindly provided by S Hans (Hans *et al.*, 2004), using PCR primers with incorporated KpnI/BglII sites, subcloned into PCRII-Topo vector, sequenced and subcloned into pGL4.26 (Promega) using KpnI/BglII sites.

Correlation analysis

We used the implementation of Fisher's exact test (Fisher, 1962) and Gene Ontology Fisher's exact test by Genedata Analyst (Genedata AG, Basel, Switzerland). For details see Supplementary information.

Cycloheximide experiment

Embryos were injected with mRNA or left non-injected, and were treated with protein inhibitor cycloheximide (CHX, Calbiochem), 15 µg/ml in egg water, starting from 1.5 hpf until the embryos were frozen for RNA isolation. In presence of CHX, direct Pou5f1 targets are transcribed, but these mRNAs are not translated, avoiding indirect downstream regulatory effects. As genetic control, we used MZ embryos for all experiments, and re-expressed Pou5f1 by mRNA microinjection at the one cell stage as indicated. CHX was added at the 64-cell stage to allow for translation of injected *pou5f1* mRNA, but to block translation of the earliest zygotic transcripts. Loss of *ntl* expression in CHX embryos was used as control for efficient inhibition of translation.

Calculation of statistical enrichment for Pou5f1-binding sites

We used the ExPlain tool (Kel *et al.*, 2006), to look for statistical enrichment in Pou5f1-predicted binding sites in regulatory regions of selected genes in comparison with a background set. As background set for all cases, we used the inclusive set of 6655 gene regions, which are non-redundant ENSEMBL genes in the Agilent microarray. Putative regulatory regions (defined as 15 kb upstream and 5 kb downstream from transcribed sequence) for all genes were obtained via the UCSC Table browser at <http://genome.ucsc.edu/cgi-bin/hgTables>. Foreground gene sets were the subsets of this background set. First, we ran the Match program on our POU_CHX_UP set (Supplementary Table 8, 247 regulatory regions) using the vertebrate non-redundant database of Positional Weight Matrices (PWMs). We found that the PWMs for Oct4, (V\$OCT4_02 and V\$OCT4_01 (Loh *et al.*, 2006)) are enriched over the background set with high statistical significance. Then we used the more rigorous algorithm F-Match to test all of our gene sets for enrichment in these two matrices (see Supplementary Table 10).

Mathematical modeling of gene-regulatory network

The regulatory relationships between the major components of the early pluripotency network were modeled using a phenomenological approach based on steep transcriptional responses (Veflingstad and Plahte, 2007). Within this framework, the regulatory potential of a transcription factor is characterized by two parameters: (1) the threshold level (τ) for the activation/repression of target gene expression and (2) the rate of target gene expression (α). Gene expression is modeled using the Hill function

$$s(Z, \tau) = Z^n / (\tau^n + Z^n)$$

as the central building block where we choose $n=10$ to achieve steep transcriptional responses. The activation of a target gene X by a single transcription factor Z is given as $\alpha_{X,Z}s(Z, \tau_{X,Z})$, whereas repression is given by $\alpha_{X,Z}(1-s(Z, \tau_{X,Z}))$. Let $\bar{\alpha}_X$ be the maximal expression rate of gene X . Starting from zero expression, the maximal level gene X can attain is $\bar{\alpha}_X/\lambda_X$, where λ_X is its degradation rate, which is related to the half-life ω_X as $\lambda_X = \ln(2)/\omega_X$. As we were only interested in the temporal shape of gene expression and not interested in absolute expression levels, we rescaled each gene by its maximal expression level. As a result, the new dimensionless scale is $X' = X/(\bar{\alpha}_X/\lambda_X)$ and all expression levels are confined between 0 and 1. In addition, the thresholds and expression rates also become relative quantities: $\tau'_{X,Z} = \tau_{X,Z}/(\bar{\alpha}_X/\lambda_X)$ and $\alpha'_{X,Z} = \alpha_{X,Z}/\bar{\alpha}_X$. For example a threshold of $\tau'_{X,Z} = 0.01$ indicates a highly efficient transcription factor Z that is already effective at 1% of its maximal level. Similarly, a Z -induced transcription rate of $\alpha'_{X,Z} = 0.3$ means a transcriptional activation by

30% of the maximal transcription rate of X . In the following, we drop the primes for clarity.

The model includes Pou5f1, targets activated by Pou5f1 (*foxD3*, *sox11a*), targets activated by Pou5f1 and zygotic Sox activity (*her3*, *sox2*), and indirectly repressed targets (PODs, *nr2f1* listed as example). We considered two types of RODs: those regulated by Pou5f1 alone (example *foxd3*) and those regulated by Pou5f1 and Sox (example *her3*). The model is described by the following set of ordinary differential equations.

$$\partial_t[\text{Sox2}] = \lambda_{\text{Sox2}}(v_{\text{Sox2}}[\text{Pou}] + (1 - v_{\text{Sox2}})s([\text{Sox}], \tau_{\text{Sox2, Sox}}) - [\text{Sox2}])$$

$$\partial_t[\text{Sox11a}] = \lambda_{\text{Sox11a}}([\text{Pou}] - [\text{Sox11a}])$$

$$\partial_t[\text{Her3}] = \lambda_{\text{Her3}}(v_{\text{Her3}}[\text{Pou}] + (1 - v_{\text{Her3}})s([\text{Sox}], \tau_{\text{Her3, Sox}}) - [\text{Her3}])$$

$$\partial_t[\text{FoxD3}] = \lambda_{\text{FoxD3}}([\text{Pou}] - [\text{FoxD3}])$$

$$\partial_t[\text{Nr2f1}] = \lambda_{\text{Nr2f1}}(1 - s([\text{FoxD3}], \tau_{\text{Nr2f1, FoxD3}}) - [\text{Nr2f1}])$$

In addition, the following definitions hold.

$$[\text{Pou}] = \begin{cases} 1 & \text{WT} \\ 0 & \text{MZ} \\ 1 - \exp(-\lambda_{\text{Pou}}t) & \text{M} \end{cases}$$

$$[\text{Sox}] = ([\text{Sox2}] + [\text{Sox11a}])/2$$

Note that, Pou5f1 is assumed to be constant in WT conditions, absent in MZ conditions and replenished with a rate $\lambda_{\text{Pou}} = \ln(2)/\omega_{\text{Pou}}$ in M conditions. The two transcription rates for Sox2 and Her3 are defined as $v_{\text{Sox2}} = 1/(1 + \alpha_{\text{Sox2, Sox}}/\alpha_{\text{Sox2, Pou}})$ and $v_{\text{Her3}} = 1/(1 + \alpha_{\text{Her3, Sox}}/\alpha_{\text{Her3, Pou}})$, respectively. Note that only the relative fraction of the transcription rates induced by the Sox and Pou factors are important for our modeling framework.

Parameter estimation

We rescaled each temporal profile of a gene by its maximal expression level in WT and MZ condition to obtain values between 0 and 1. We fit the rescaled system of equations to the rescaled data assuming that the maximal level a gene can attain within our model framework is equal to the observed maximal level. The minimized functional was

$$G(\vec{p}) = \sum_{j=1}^{\text{NVar}} \sum_{k=1}^{\text{NMes}} (y_j^{\text{WT}}(t_k, \vec{p}) - D_j^{\text{WT}}[t_k])^2 + \sum_{j=1}^{\text{NVar}} \sum_{k=1}^{\text{NMes}} (y_j^{\text{MZ}}(t_k, \vec{p}) - D_j^{\text{MZ}}[t_k])^2$$

Here, \vec{p} is a vector of all model parameters, $y_j^{\text{WT}}(t_k, \vec{p})$ is the solution of variable j at the k th time point in WT conditions and $D_j^{\text{WT}}[t_k]$ is the mean value of the corresponding measured data point. The same holds for the MZ condition. The median and interquartile range (IQR; difference between the 75th and the 25th percentiles) of the parameter estimates was determined by a bootstrap method. Mean values of each time point were calculated by sampling with replacement from the raw data. Model parameters were estimated from the resulting time series and the procedure was repeated for 200 times giving the median and IQR given in the Table below.

Table of model parameters

Parameter	Allowed range	Fitted value [median (IQR)]
$\tau_{\text{Sox2, Sox}}$	0–1	0.49 (0.02)
$\tau_{\text{Her3, Sox}}$	0–1	0.64 (0.01)
$\tau_{\text{Nr2f1, FoxD3}}$	0–1	0.08 (<0.001)
$\alpha_{\text{Sox2, Sox}}/\alpha_{\text{Sox2, Pou}}$	0.001–1000	899.83 (257.3)
$\alpha_{\text{Her3, Sox}}/\alpha_{\text{Her3, Pou}}$	0.001–1000	11.32 (0.09)
$\lambda_{\text{Sox2}} [\text{h}^{-1}]$	$\ln(2)/5 - \ln(2)/0.1$	4.2 (1.13)
$\lambda_{\text{Sox11a}} [\text{h}^{-1}]$	$\ln(2)/5 - \ln(2)/0.1$	0.5 (<0.001)
$\lambda_{\text{Her3}} [\text{h}^{-1}]$	$\ln(2)/5 - \ln(2)/0.1$	2.74 (0.14)
$\lambda_{\text{FoxD3}} [\text{h}^{-1}]$	$\ln(2)/5 - \ln(2)/0.1$	0.36 (<0.001)
$\lambda_{\text{Nr2f1}} [\text{h}^{-1}]$	$\ln(2)/5 - \ln(2)/0.1$	0.58 (<0.001)

IQR, interquartile range.

The prediction for M conditions shown in Figure 6B, blue-dashed curve was calculated with a Pou5f1 half-life of 2 h.

Calculation of expression timing for biphasic targets

We extended the model by an additional Pou5f1/Sox target gene to quantify the timing of expression for an exemplary biphasic target gene X in dependence of the Sox threshold ($\tau_{X, \text{Sox}}$) and the Pou5f1/Sox regulation ($v_X = 1/(1 + \alpha_{X, \text{Sox}}/\alpha_{X, \text{Pou}})$). The ODE of the exemplary biphasic target reads

$$\partial_t[X] = \lambda_X(v_X[\text{Pou}] + (1 - v_X)s([\text{Sox}], \tau_{X, \text{Sox}}) - [X])$$

The decay rate λ_X was set to the estimated value for *Her3*. Note that the decay rate influences the timing of expression, however, in this context we are only interested in the impact of the two other parameters v_X and $\tau_{X, \text{Sox}}$. We characterized the timing of expression of X by the time needed to reach the half-maximal expression level, starting from zero expression level at MBT. The resulting values are indicated as gray levels in Figure 6C. The Y axis depicts the relative expression rate $\alpha_{X, \text{Sox}}/\alpha_{X, \text{Pou}}$. All numerical analysis was performed with MATLAB. All MATLAB scripts are available upon request.

For further details on statistical analysis and standard techniques (RT-PCR, whole-mount *in situ* hybridization, luciferase assays, point mutagenesis, and gel retardation assays), see Supplementary information.

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (<http://www.nature.com/msb>).

Acknowledgements

We thank Dr Kurz and M Klein for excellent microarray services in the ZBSA genomics core facility, and Drs Li and Rensing of ZBSA/FRISYS for establishing the GENEDATA infrastructure. We thank S Hans, A Tomilin, H Kondoh, the Zebrafish International Resource Center and the zebrafish community for plasmids, and H Niwa for ZTBHc4 cells. We thank the fish facility, especially S Götter, for fish care; A Fuchs and AA Sharov for scientific discussion, and K Lunde, G Pyrowolakis, and A Filippi for discussion of the manuscript. This work was supported by the BMBF FORSYS center FRISYS (WD), the DFG Collaborative Research Center grant SFB592-A3 (WD and DO), and by the Excellence Initiative of the German Federal and State Governments (bioss and FRIAS).

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Bachvarova RF, Masi T, Drum M, Parker N, Mason K, Patient R, Johnson AD (2004) Gene expression in the axolotl germ line: Axdazl, Axvh, Axoct-4, and Axkit. *Dev Dyn* **231**: 871–880
- Belting H-G, Hauptmann G, Meyer D, Abdelilah-Seyfried S, Chitnis A, Eschbach C, Söll I, Thisse C, Thisse B, Artinger KB, Lunde K, Driever W (2001) spiel ohne grenzen/pou2 is required during establishment of the zebrafish midbrain-hindbrain boundary organizer. *Development* **128**: 4165–4176
- Boiani M, Scholer HR (2005) Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol* **6**: 872–884
- Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**: 947–956

- Burgess S, Reim G, Chen W, Hopkins N, Brand M (2002) The zebrafish *spiel-ohne-grenzen* (*spg*) gene encodes the POU domain protein Pou2 related to mammalian Oct4 and is essential for formation of the midbrain and hindbrain, and for pre-gastrula morphogenesis. *Development* **129**: 905–916
- Camp EM, Sanchez-Sanchez AV, Garcia-Espana A, Desalle R, Odqvist L, O'Connor JE, Mullor JL (2009) Nanog regulates proliferation during early fish development. *Stem Cells* **27**: 2081–2091
- Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**: 643–655
- Chan TM, Chao CH, Wang HD, Yu YJ, Yuh CH (2009) Functional analysis of the evolutionarily conserved cis-regulatory elements on the *sox17* gene in zebrafish. *Dev Biol* **326**: 456–470
- Chickarmane V, Troein C, Nuber UA, Sauro HM, Peterson C (2006) Transcriptional dynamics of the embryonic stem cell switch. *PLoS Comput Biol* **2**: e123
- Dee CT, Gibson A, Rengifo A, Sun SK, Patient RK, Scotting PJ (2007) A change in response to Bmp signalling precedes ectodermal fate choice. *Int J Dev Biol* **51**: 79–84
- Downs KM (2008) Systematic localization of Oct-3/4 to the gastrulating mouse conceptus suggests manifold roles in mammalian development. *Dev Dyn* **237**: 464–475
- Fisher RA (1962) Confidence limits for a cross-production ratio. *Aust J Stat* **4**: 41
- Frankenberg S, Pask A, Renfree MB (2009) The evolution of class V POU domain transcription factors in vertebrates and their characterisation in a marsupial. *Dev Biol* **19**: 19
- Gauchat D, Escriva H, Miljkovic-Licina M, Chera S, Langlois MC, Begue A, Laudet V, Galliot B (2004) The orphan COUP-TF nuclear receptors are markers for neurogenesis from cnidarians to vertebrates. *Dev Biol* **275**: 104–123
- Gering M, Yamada Y, Rabbitts TH, Patient RK (2003) Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. *Development* **130**: 6187–6199
- Hans S, Scheer N, Riedl I, v Weizsacker E, Blader P, Campos-Ortega JA (2004) *her3*, a zebrafish member of the hairy-E(spl) family, is repressed by Notch signalling. *Development* **131**: 2957–2969
- Hatakeyama J, Bessho Y, Katoh K, Ookawara S, Fujioka M, Guillemot F, Kageyama R (2004) Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. *Development* **131**: 5539–5550
- Hirata H, Tomita K, Bessho Y, Kageyama R (2001) Hes1 and Hes3 regulate maintenance of the isthmic organizer and development of the mid/hindbrain. *EMBO J* **20**: 4454–4466
- Jiang J, Chan YS, Loh YH, Cai J, Tong GQ, Lim CA, Robson P, Zhong S, Ng HH (2008) A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* **10**: 353–360
- Kamachi Y, Uchikawa M, Tanouchi A, Sekido R, Kondoh H (2001) Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev* **15**: 1272–1286
- Kane DA, Kimmel CB (1993) The zebrafish midblastula transition. *Development* **119**: 447–456
- Kazanskaya OV, Severtzova EA, Barth KA, Ermakova GV, Lukyanov SA, Benyumov AO, Pannese M, Boncinelli E, Wilson SW, Zaraisky AG (1997) *Anf*: a novel class of vertebrate homeobox genes expressed at the anterior end of the main embryonic axis. *Gene* **200**: 25–34
- Kehler J, Tolkunova E, Koschorz B, Pesce M, Gentile L, Boiani M, Lomeli H, Nagy A, McLaughlin KJ, Scholer HR, Tomilin A (2004) Oct4 is required for primordial germ cell survival. *EMBO Rep* **5**: 1078–1083
- Kel A, Voss N, Jauregui R, Kel-Margoulis O, Wingender E (2006) Beyond microarrays: finding key transcription factors controlling signal transduction pathways. *BMC Bioinformatics* **7**(Suppl 2): S13
- Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**: 1049–1061
- Koonin EV (2005) Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet* **39**: 309–338
- Lachnit M, Kur E, Driever W (2008) Alterations of the cytoskeleton in all three embryonic lineages contribute to the epiboly defect of Pou5f1/Oct4 deficient MZspg zebrafish embryos. *Dev Biol* **315**: 1–17
- Lavial F, Acloque H, Bertocchini F, Macleod DJ, Boast S, Bachelard E, Montillet G, Thenot S, Sang HM, Stern CD, Samarut J, Pain B (2007) The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. *Development* **134**: 3549–3563
- Liu Y, Labosky PA (2008) Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. *Stem Cells* **26**: 2475–2484
- Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD, Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL et al (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**: 431–440
- Lunde K, Belting HG, Driever W (2004) Zebrafish *pou5f1/pou2*, homolog of mammalian Oct4, functions in the endoderm specification cascade. *Curr Biol* **14**: 48–55
- Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, Ko MS, Niwa H (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9**: 625–635
- Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**: 631–642
- Morrison GM, Brickman JM (2006) Conserved roles for Oct4 homologues in maintaining multipotency during early vertebrate development. *Development* **133**: 2011–2022
- Nichols J, Zevnik B, Anastasiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**: 379–391
- Niwa H, Masui S, Chambers I, Smith AG, Miyazaki J (2002) Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. *Mol Cell Biol* **22**: 1526–1536
- Niwa H, Miyazaki J, Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**: 372–376
- Niwa H, Sekita Y, Tsend-Ayush E, Grützner F (2008) Platypus Pou5f1 reveals the first steps in the evolution of trophectoderm differentiation and pluripotency in mammals. *Evol Dev* **10**: q671–682
- Okuda Y, Yoda H, Uchikawa M, Furutani-Seiki M, Takeda H, Kondoh H, Kamachi Y (2006) Comparative genomic and expression analysis of group B1 sox genes in zebrafish indicates their diversification during vertebrate evolution. *Dev Dyn* **235**: 811–825
- Quirk J, Brown P (2002) Hex31 homeodomain protein represses transcription as a monomer and antagonises transactivation of specific sites as a homodimer. *J Mol Endocrinol* **28**: 193–205
- Reim G, Brand M (2002) *Spiel-ohne-grenzen/pou2* mediates regional competence to respond to Fgf8 during zebrafish early neural development. *Development* **129**: 917–933
- Reim G, Brand M (2006) Maternal control of vertebrate dorsoventral axis formation and epiboly by the POU domain protein Spg/Pou2/Oct4. *Development* **133**: 2757–2770
- Reim G, Mizoguchi T, Stainier DY, Kikuchi Y, Brand M (2004) The POU domain protein *spg* (*pou2/Oct4*) is essential for endoderm formation in cooperation with the HMG domain protein *casanova*. *Dev Cell* **6**: 91–101
- Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR, Wilmanns M (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev* **17**: 2048–2059

- Rhinn M, Lun K, Amores A, Yan YL, Postlethwait JH, Brand M (2003) Cloning, expression and relationship of zebrafish *gbx1* and *gbx2* genes to Fgf signaling. *Mech Dev* **120**: 919–936
- Rowland BD, Bernards R, Peeper DS (2005) The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. *Nat Cell Biol* **7**: 1074–1082
- Sahoo D, Dill DL, Tibishirani R, Plevritis SK (2007) Extracting binary signals from microarray time-course data. *Nucleic Acids Res* **35**: 3705–3712
- Sandberg M, Kallstrom M, Muhr J (2005) Sox21 promotes the progression of vertebrate neurogenesis. *Nat Neurosci* **8**: 995–1001
- Sharov AA, Masui S, Sharova LV, Piao Y, Aiba K, Matoba R, Xin L, Niwa H, Ko MS (2008) Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. *BMC Genomics* **9**: 269
- Stigloher C, Ninkovic J, Laplante M, Geling A, Tannhauser B, Topp S, Kikuta H, Becker TS, Houart C, Bally-Cuif L (2006) Segregation of telencephalic and eye-field identities inside the zebrafish forebrain territory is controlled by Rx3. *Development* **133**: 2925–2935
- Stoykova A, Fritsch R, Walther C, Gruss P (1996) Forebrain patterning defects in Small eye mutant mice. *Development* **122**: 3453–3465
- Strumpf D, Mao CA, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, Rossant J (2005) Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**: 2093–2102
- Takeda H, Matsuzaki T, Oki T, Miyagawa T, Amanuma H (1994) A novel POU domain gene, zebrafish *pou2*: expression and roles of two alternatively spliced twin products in early development. *Genes Dev* **8**: 45–59
- Turner DL, Weintraub H (1994) Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev* **8**: 1434–1447
- Veflingstad SR, Plahte E (2007) Analysis of gene regulatory network models with graded and binary transcriptional responses. *Biosystems* **90**: 323–339
- Yaklichkin S, Steiner AB, Lu Q, Kessler DS (2007) FoxD3 and Grg4 physically interact to repress transcription and induce mesoderm in *Xenopus*. *J Biol Chem* **282**: 2548–2557
- Ying QL, Wray J, Nichols J, Battle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008) The ground state of embryonic stem cell self-renewal. *Nature* **453**: 519–523
- Zhou Q, Chipperfield H, Melton DA, Wong WH (2007) A gene regulatory network in mouse embryonic stem cells. *Proc Natl Acad Sci USA* **104**: 16438–16443



Molecular Systems Biology is an open-access journal published by *European Molecular Biology Organization* and *Nature Publishing Group*.

This article is licensed under a Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Licence.