Development 133, 1955-1966 (2006) doi:10.1242/dev.02358

Global analysis of the transcriptional network controlling *Xenopus* endoderm formation

Débora Sinner[1], Pavel Kirilenko[2], Scott Rankin[1], Eric Wei[1], Laura Howard[2], Matthew Kofron[1], Janet Heasman[1], Hugh R. Woodland[2,*] and Aaron M. Zorn[1,*]

A conserved molecular pathway has emerged controlling endoderm formation in *Xenopus* zebrafish and mice. Key genes in this pathway include Nodal ligands and transcription factors of the Mix-like paired homeodomain class, Gata4-6 zinc-finger factors and Sox17 HMG domain proteins. Although a linear epistatic pathway has been proposed, the precise hierarchical relationships between these factors and their downstream targets are largely unresolved. Here, we have used a combination of microarray analysis and loss-of-function experiments to examine the global regulatory network controlling *Xenopus* endoderm formation. We identified over 300 transcripts enriched in the gastrula endoderm, including most of the known endoderm regulators and over a hundred uncharacterized genes. Surprisingly only 10% of the endoderm transcriptome is regulated as predicted by the current linear model. We find that Nodal genes, Mixer and Sox17 have both shared and distinct sets of downstream targets, and that a number of unexpected autoregulatory loops exist between Sox17 and Gata4-6, between Sox17 and Bix1/Bix2/Bix4, and between Sox17 and Xnr4. Furthermore, we find that Mixer does not function primarily via Sox17 as previously proposed. These data provides new insight into the complexity of endoderm formation and will serve as valuable resource for establishing a complete endoderm gene regulatory network.

KEY WORDS: Endoderm, Development, *Xenopus*, Nodal, Sox17, Gata, Mixer, Microarray, Gene regulatory network

INTRODUCTION

Recent work in *Xenopus*, zebrafish and mouse has identified a conserved pathway regulating specification of the embryonic endoderm in vertebrates (Loose and Patient, 2004; Stainier, 2002; Tam et al., 2003; Xanthos et al., 2001). The key zygotic factors in this pathway are the Nodal-related TGFβ signaling ligands, the Mix-like family of homeodomain transcription factors, the Gata4/5/6 zinc-finger transcription factors and the HMG box transcription factor Sox17.

In *Xenopus*, endoderm development is initiated by the maternal T-box transcription factor VegT, which is localized to the presumptive endoderm tissue (Horb and Thomsen, 1997; Stennard et al., 1996; Zhang and King, 1996). VegT is required for endoderm formation and the expression of zygotic factors, including the *Nodal*-related genes *Xnr1*, *Xnr2*, *Xnr4*, *Xnr5* and *Xnr6*, and *Derriere*, *Mix1*, *Mix2*, *Bix1*, *Bix2*, *Bix3*, *Bix4*, *Mixer*, *Gata4*, *Gata5*, *Gata6*, *Sox17α* and *Sox17β* (Xanthos et al., 2001; Zhang et al., 1998). VegT directly activates the transcription of many of these (Casey et al., 1999; Clements and Woodland, 2003; Engleka et al., 2001; Hilton et al., 2003; Tada et al., 1998), but maintenance of their expression and subsequent endoderm formation requires Nodal signaling (Clements et al., 1999; Kofron et al., 1999; Yasuo and Lemaire, 1999). In *Xenopus*, Nodal signaling is necessary and, at high levels, sufficient to induce endoderm development (Agius et al., 2000; Henry et al., 1996; Osada and Wright, 1999) by promoting the expression of the Mix-like, Gata and Sox17 transcription factors, which in turn activate downstream target genes (Afouda et al., 2005; Clements et al., 2003; Hudson et al., 1997; Xanthos et al., 2001).

Ectopic expression of Mixer, Bix1, Bix2, Bix4, Sox17α/β or Gata4-6 can all induce endoderm differentiation in naïve animal cap ectoderm (Casey et al., 1999; Ecochard et al., 1998; Henry and Melton, 1998; Hudson et al., 1997; Tada et al., 1998; Weber et al., 2000) and loss-of-function studies have shown that Mixer, Gata4-6 and Sox17α/β are essential for proper endoderm development (Afouda et al., 2005; Clements et al., 2003; Henry and Melton, 1998; Hudson et al., 1997; Kofron et al., 2004).

Although the precise epistatic relationships between Mixer, Gata4-6 and Sox17 are unresolved, a linear model is commonly proposed where Nodal proteins regulate Mixer and Gata, and these function upstream of Sox17, which in turn activates endoderm target genes (Stainier, 2002; Xanthos et al., 2001). In support of this model, Mixer and Gata5 can induce *Sox17* expression in animal caps and VegT-depleted embryos, but Sox17 cannot induce expression *Mixer* or any of the other *Mix-like* genes (Henry and Melton, 1998; Sinner et al., 2004; Xanthos et al., 2001). Furthermore, a dominant-negative version of Sox17 (Sox17-EnR) has been shown to inhibit Mixer function, but, conversely, a dominant-negative Mixer (Mixer-EnR) cannot inhibit Sox17 function (Henry and Melton, 1998), suggesting that Mixer acts primarily via Sox17.

However, other evidence suggests that endoderm specification is more complex than predicted by the linear model. First, *Sox17* expression precedes *Mixer*, which is principally expressed in equatorial regions of the endoderm (Henry and Melton, 1998), which is inconsistent with Mixer acting primarily via Sox17. Second, studies have suggested that that Sox17α/β and Gata4-6 can regulate the expression of each other (Afouda et al., 2005; Clements et al., 2003; Sinner et al., 2004).

A limitation of many studies to date is that they have relied on only a few early markers, usually *Hnf1β* (Demartis et al., 1994) and *Endodermin* (*Edd*) (Sasai et al., 1996) to assay endoderm

[1]Division of Developmental Biology, Cincinnati Children's Hospital Research Foundation and Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45299, USA. [2]Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK.

*Authors for correspondence (e-mail: h.r.woodland@warwick.ac.uk; aaron.zorn@chmcc.org)

Accepted 13 March 2006

specification and it is unclear if their regulation is indicative of all endoderm genes. In addition, most studies have relied on ectopic overexpression in animal cap ectoderm (Afouda et al., 2005; Clements and Woodland, 2003; Dickinson et al., 2006; Sinner et al., 2004; Taverner et al., 2005), which may lack important co-factors found in the vegetal tissue and it is unclear how accurately animal cap assays reflect endogenous endoderm development.

Here, we have used microarray analysis and functional experiments to better resolve the regulatory network controlling *Xenopus* endoderm formation. We defined a robust set of genes with enriched expression in the gastrula endoderm, containing ~90% of the known endoderm-expressed genes and several hundred uncharacterized sequences. We determined which of these genes were regulated by Nodal signaling, Mixer or Sox17, and found that only ~10% of endoderm genes can be regulated as described by the current linear model of endoderm development. The bulk of endoderm gene regulation appears to be much more complex, with Nodal proteins, Mixer and Sox17 having both shared and distinct sets of target genes. We find that transcriptional repression by Mixer plays a greater role than previously appreciated and that extensive autoregulatory loops exist between Sox17 and Bix1/2/4, between Sox17 and Xnr4, and between Sox17 and Gata4-6. This data challenges the existing models of vertebrate endoderm development and provides an important resource for understanding of the complex gene regulatory network that controls *Xenopus* endoderm development.

MATERIALS AND METHODS

Embryo culture and microinjection

Embryo manipulations and microinjections were preformed as previously described (Zorn et al., 1999b), and embryos were staged according to the normal table of development for *Xenopus laevis* (Nieuwkoop and Faber, 1994). Two-cell stage embryos were vegetally injected with the following doses of antisense morpholino oligos or synthetic RNA: a combination of antisense morpholino oligos to Sox17α1 + Sox17α2 + Sox17β (20 ng each) (Clements et al., 2003); Mixer antisense morpholino oligo (40 ng) (Kofron et al., 2004); *Cerberus-short* RNA (1 ng) (Piccolo et al., 1999); *Mixer* RNA (50-500 pg) (Henry and Melton, 1998); *Sox17β* RNA (10-100 pg; *Xenopus* tropicalis *Sox17β* that is immune to the *laevis Sox17β* morpholino oligo) (D'Souza et al., 2003); and *Gata4*, *Gata5* or *Gata6* RNA (50-100 pg) (Afouda et al., 2005). Rescue experiments were performed with *Gata4*, *Gata5* and *Gata6* with similar results; therefore only the *Gata6* data are shown.

Microarray analysis and data processing

Table 1 lists the different conditions and the number of biological replicates used in the array study. For each biological replicate, ~20 sibling embryos from a single mating or ~50 micro-dissected explants from sibling embryos were used. Total RNA was extracted using Trizol (Invitrogen) and purified on RNAeasy columns (Qiagen). Ten micrograms of total RNA was used for cDNA syntheses and to make labeled RNA probe which was hybridized to Affymetrix *Xenopus* Genechips by the CHRF microarray core facility, using the standard Affymetrix protocol. GeneSpring 7.1 software (Silicon Genetics) was used for data normalization, clustering and filtering. Raw CEL file data from all the samples was pre-normalized using RMA (Robust Multichip Average). The average log intensity of the biological replicates was then normalized to the average log intensity of stage 11 whole embryo. NCBI Unigene cluster nomenclature was used to describe uncharacterized sequences. All of the raw microarray data are available from GEO (series record number GSE4448).

RT-PCR analysis

Real time RT-PCR analysis was performed on an Opticon machine (MJ Research) using Qiagen SYBR green PCR mix as previously described (Sinner et al., 2004). Details of the primer sequences used for the ~60 genes analyzed in this study are available on request. For each new primer pair a melt curve analysis was performed and the PCR product was examined on a gel to ensure that a single fragment of the predicted molecular weight was amplified. The data for each sample was normalized to the expression level of the ubiquitously expressed gene *ornithine decarboxylase* (*ODC*).

Table 1. Samples used in the array study

Name	Description	Number of biological replicates
An	Animal cap (isolated stage 9 cultured to stage 11)	3
Veg	Vegetal region (stage 11)	3
Eq	Equatorial region (stage 11)	2
We	Stage 11 whole embryo	9
Nodal–	Nodal inhibited whole embryo (stage 11) Cerberus-short mRNA (1 ng)	3
Mixer–	Mixer depleted whole embryo (stage 11) Antisense Mixer morpholino oligo (40 ng)	2
Sox17–	Sox17 depleted whole embryo (stage 11) Antisense Sox17 α1/α2/β morpholino oligos (20 ng each)	3
Egg	Unfertilized egg	3
St18	Stage 18 whole embryo	3

DNA constructs and In situ hybridization

Plasmids for validation were generously provided by Professor Naoto Ueno from the NIBB *Xenopus* EST project (Japan) or clones from the NIH *Xenopus* sequencing project were purchased from ATCC. Synthesis of antisense RNA probes and in situ hybridization to bisected gastrula embryos were performed as described (Sive et al., 2000).

RESULTS

Identification of endoderm-enriched transcripts

To better understand the gene regulatory network controlling endoderm formation in *Xenopus*, we first used microarray analysis to identify genes with endoderm-enriched expression in the mid-gastrula (stage 11) embryo. We chose mid-gastrula because cell transplantation studies have shown that the endoderm germ layer is specified by this time (Heasman et al., 1984) and because the known targets of Mixer, Gata4-6 and Sox17 are expressed.

We compared the transcriptional profile of stage 11 whole embryos (We), micro-dissected vegetal (Veg) and equatorial regions (Eq), and animal caps (An) (Table 1). Vegetal regions isolated from stage 11 gastrulae contained mostly endoderm, and small amounts of mesoderm. Equatorial regions isolated from stage 11 gastrulae contained mostly mesoderm but also superficial endoderm. The animal cap tissue isolated from stage 9 embryos and cultured until stage 11, contained ectoderm. For each biological replicate, total RNA was prepared from ~20 whole embryos or ~50 explants from sibling embryos. The RNA was subjected to microarray analysis using the Affymetrix *Xenopus* Genechip and the resulting data were analyzed with GeneSpring software, where the average log intensity of the biological replicates was normalized to the average expression levels in stage 11 whole embryos.

To identify genes with enriched expression in the endoderm, we examined the behavior of the known endoderm genes *Sox17α/β*, *Mixer*, *Bix1-4*, *Gata4-6*, *Hnf1β*, *FoxA1* and *Edd*. From their characteristics, we empirically determined the following parameters for selecting endoderm-enriched transcripts from the ~15,000 sequences on the microarray. After filtering the data to eliminate genes that were not expressed in the gastrula, we selected transcripts

with expression in the vegetal region greater or equal to the expression in the equatorial region. This eliminated many mesoderm-specific genes (e.g. *Xbra*), but retained most genes known to be expressed at the mesoderm-endoderm boundary (e.g. *Eomesodermin*). We then selected genes with threefold or greater expression in the vegetal region than in the animal cap, resulting in a list of 503 sequences that represented 483 genes based on their NCBI Unigene designations (see Table S1 in the supplementary material). This list of 483 genes contained 35 of 40 published genes known to have endoderm-enriched expression (see Table S2 in the supplementary material), providing a strong validation of our approach. Of the five known genes that were not recovered by our selection (*Siamois*, *Hex*, *Xnr4*, *Xnr6* and *FoxA1*) four had low expression levels, just above background, which may explain why they did not behave as predicted. For further analysis, we selected 276 sequences (representing 264 genes) that had statistically significant differences in expression between vegetal and animal cap regions over all biological replicates using Student's *t*-test ($P < 0.05$) (Table S1). Fig. 1A summarizes the transcriptional profile of those 276 sequences and Table S3 in the supplementary material presents the top 60 endoderm-enriched genes.

The predicted molecular function encoded by these 264 endoderm-enriched transcripts, based on NCBI Unigene annotations, Gene Ontology and blast analyses is indicated in Fig. 1B. Over 40% of the genes are uncharacterized, while ~25% encode predicted regulatory proteins (38 transcription factors, 15 secreted ligands/antagonists, nine receptors and eight signal transduction molecules), a number of which have not previously been implicated in endoderm development.

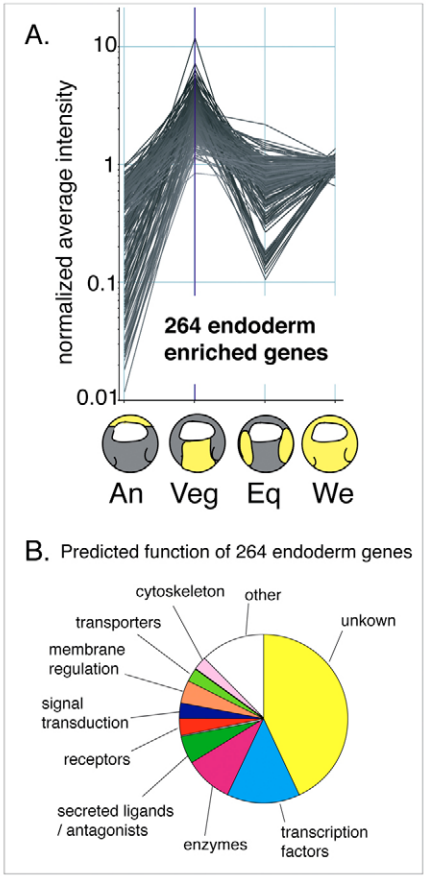

Fig. 1. Endoderm-enriched transcripts identified by microarray analysis. (**A**) The spatial expression profile of 276 endoderm-enriched sequences (264 genes), based on microarray analysis, in different regions of the stage 11 gastrula. The average intensity, normalized to stage 11 whole embryo is shown on a log scale. The diagrams below the graph indicate the different regions of the embryo. An, animal cap ectoderm; Veg, endoderm-enriched vegetal tissue; Eq, mesendoderm-enriched equatorial tissue; We, whole embryo stage 11. (**B**) The pie diagram indicates the predicted functions on the endoderm-enriched genes based on NCBI Unigene annotation, GO ontogeny and blast searches.

Validation of endoderm-enriched transcripts in the *Xenopus* gastrula

We validated the expression of ~25% of the genes not previously known to have endoderm-enriched expression by RT-PCR and in situ hybridization, reasoning that this was a representative sample size. By real time RT-PCR analysis, 51 of 54 (94%) previously uncharacterized genes were expressed at least three times higher in the vegetal region than in the animal cap (Fig. 2A; see Table S4 in the supplementary material). The selection of the 54 genes to validate was largely random, with an emphasis on those genes for which there were full-length cDNAs available in the clone repositories. In situ hybridization to bisected gastrula embryos confirmed that 24 of 35 genes exhibited obvious endoderm enriched expression in the gastrula embryo (Fig. 2B). The remaining 11 transcripts that did not exhibit endoderm restricted expression by in situ, did have enriched endoderm expression by RT-PCR but were either undetectable by in situ or had expression throughout the embryo with slightly higher levels in the vegetal region.

The in situ analysis shows that we identified genes with varying endoderm expression patterns. *Xl.11602*, *Xl.13921*, *Xl.15375*, *Xl.2554*, *Xl.3534* and *Xl.46324* were expressed throughout the endoderm in a pattern similar to *Sox17* (Hudson et al., 1997), while others such as *CXCR4*, *Xl.13033*, *Xl.215*, *Xl.13381*, *Xl.8924*, *Xl.18924*, *FoxA4* and *Xl.5418* had varying expression in the deep endoderm and were enriched at the mesendoderm boundary similar to *Mixer* (Henry and Melton, 1998). *Xl.15758* (*epsin2*) and *Xl.7782* were expressed in the deep endoderm, but not in the superficial layer of the blastopore, reminiscent of *Gata5* (Weber et al., 2000). *Wnt11-R*, *Xl.16040* and *Xl.8924*, which were expressed in the anterior endoderm reminiscent of *Hex* or *Cerberus*, whereas *Pinhead* has ventrolateral expression. Finally one unknown gene *Xl.14891* has an expression pattern that suggests it is expressed in germ plasm. A temporal expression profile of the endoderm-enriched sequences based on microarray analysis of egg, gastrula and stage 18 is available in Fig. S1 in the supplementary material.

These extensive validations indicate that we have identified a robust set of genes with endoderm-enriched expression. The fact that our procedure identified most of the known endoderm regulatory genes suggests that many of the uncharacterized genes may also have important regulatory roles in endoderm development. For our subsequent analyses of endoderm gene regulation, we focused on the 276 sequences that behaved consistently and passed the statistical test, as well as all of the other known endoderm enriched genes, representing in total 301 sequences (see Table S1 in the supplementary material).

Regulation of endoderm gene expression by Nodal proteins, Mixer and Sox17

We next determined which of the 301 endoderm-enriched sequences were regulated by either Nodal proteins, Sox17 or Mixer. We focused on these three regulators because specific loss-of-function

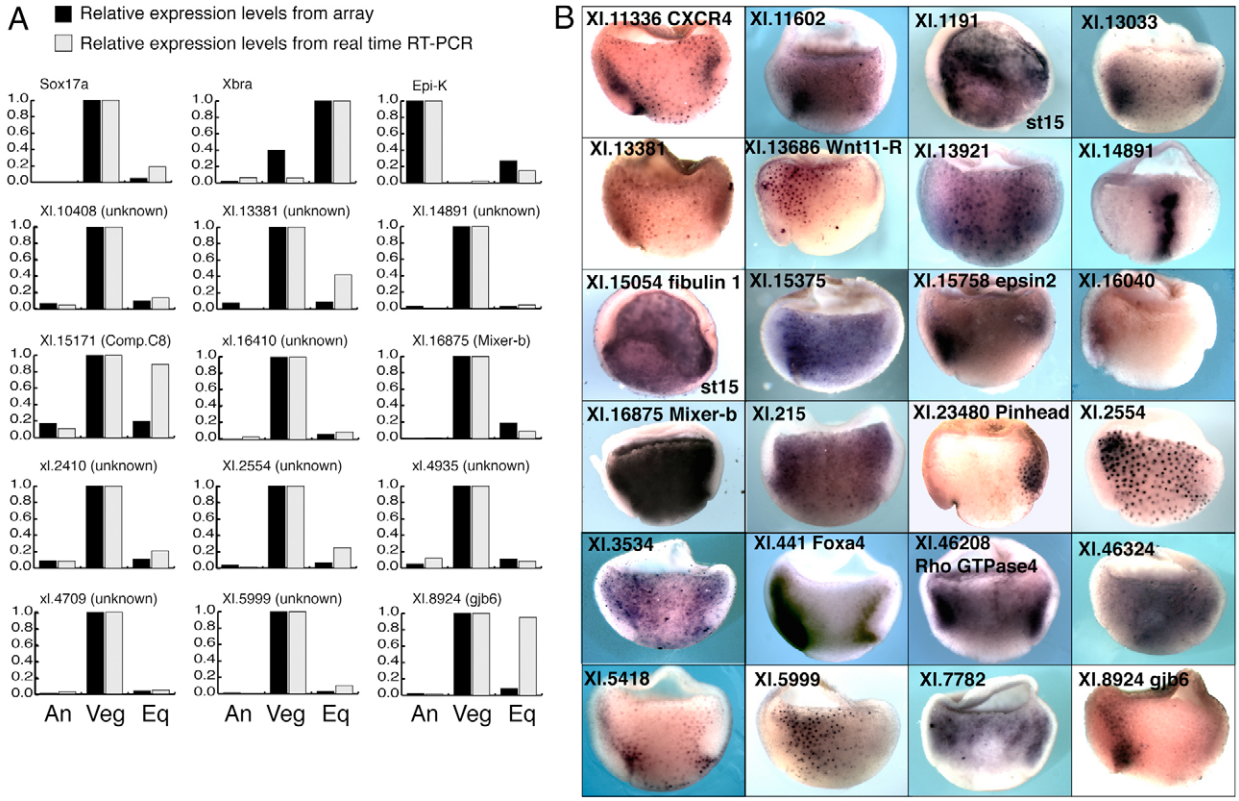

Fig. 2. Validation of endoderm-enriched transcripts. (**A**) At stage 11, RNA was isolated from animal cap ectoderm (An), endoderm-enriched vegetal tissue (Veg) and mesendoderm enriched equatorial tissue (Eq), and assayed by RT-PCR to validate the expression profile of endoderm-enriched transcripts. The black histogram shows relative normalized expression from the array and the grey histograms show the relative expression levels in RT-PCR normalized to the loading control, *ODC*. *Sox17α*, *Xbra* and *Epidermal keratin* (*Epi-K*) are positive controls for the dissections. Fifty-one out of 54 genes were confirmed to have vegetal expression that was three times greater than animal cap expression (see Table S4 in the supplementary material) and 12 representative genes are shown [*Xl.10408, Xl.13381, Xl.14891, Xl.15171*, *Xl.16410, Xl.16875* (*Mixer-b*), *Xl.2410, Xl.2554, Xl.4709, Xl.4935, Xl.5999, Xl.8924* (*gap junction subunit 6*)]. (**B**) In situ hybridization to bisected gastrula with probes to the indicated genes validates their endoderm-enriched expression. *Xl.1191* and *Xl.15054* are not detected until neurula stage (st15).

approaches have been validated for each of them (Agius et al., 2000; Clements et al., 2003; Kofron et al., 2004). Furthermore this allowed us to test the linear model predicting that Nodal genes>Mixer>Sox17>endoderm-target genes. If this model is correct, inhibiting any one of the components should prevent zygotic endodermal gene expression.

We performed a microarray analysis on three types of experimental gastrula stage embryos (Table 1): (1) where nodal signaling was blocked (*Nodal–*) by injection of 1 ng RNA encoding Cerberus-Short (1 ng), a specific antagonist of Nodal proteins (Agius et al., 2000); (2) embryos depleted of Mixer (*Mixer–*) by injection of antisense Mixer morpholino oligos (40 ng) (Kofron et al., 2004); and (3) embryos depleted of Sox17α1/α2/β (*Sox17–*) by injection of three antisense Sox17 morpholino oligos (20 ng each) (Clements et al., 2003). Each of these loss-of-function paradigms has been show to result in specific defects in endoderm development.

Fig. 3 shows the expression profile of the 301 endoderm-enriched transcripts in *Nodal–*, *Mixer–* and *Sox17–* embryos, and a complete list of the average normalized expression levels for each transcript in the different experimental conditions are presented in Table S1 (see supplementary material). The expression profiles in Fig. 3A immediately show that many genes are sensitive to some, but not all, of the experimental conditions. A hierarchal clustering of genes and experimental conditions (Fig. 3B) shows that, as expected, the expression profile of *Nodal–* embryos is more similar to animal cap tissue than to control embryos or vegetal tissue, indicating that both endoderm and mesoderm development was inhibited by *Cerberus-short* RNA injection. Sox17-depeleted embryos had an expression profile more similar to equatorial tissue and *Nodal–* embryos than to control embryos or vegetal tissue, suggesting that mostly endoderm development was compromised, rather than that of mesoderm. Surprisingly, the profile of *Mixer–* embryos was more similar to isolated vegetal tissue than to any other sample. As we will describe in more detail later, this was due to the fact that many mesendoderm genes are upregulated in *Mixer–* embryos.

As an initial validation, we focused on known Nodal, Mixer and Sox17 targets (Agius et al., 2000; Clements et al., 2003; Clements et al., 1999; Henry and Melton, 1998; Hudson et al., 1997; Kofron et al., 2004; Rosa, 1989; Sinner et al., 2004; Xanthos et al., 2001), comparing their expression on the array to that determined by real time RT-PCR (Fig. 4). Although the array tended to under-represent the fold changes observed by RT-PCR, we found that 14/14 known Nodal targets (*Xnr1*, *Xnr2*, *Xnr4*, *Mix1-2*, *Bix1-4*, *Mixer*, *Gata4-6*, *Sox17α/β*, *Edd*, *HNF1β*), 8/10 of the known Mixer targets (*Xnr5*, *Gata5*, *Bix1*, *Bix4*, *Cerberus*, *Sox17α*, *Eomesodermin*, *Edd*, *FGF3*, *eFGF*) and 5/7 of the known Sox17 targets (*Xnr4*, *Gata4-6*, *Foxa1*, *Edd*, *HNF1β*) behaved as expected from published results, providing a strong validation of the array data (Fig. 4; data not shown).

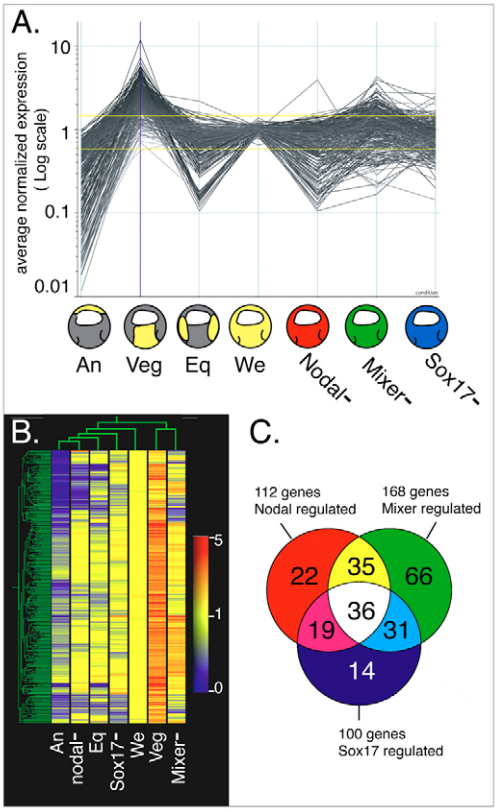

Fig. 3. Regulation of endoderm genes by Nodal proteins, Mixer and Sox17. (**A**) The expression of the 301 endoderm-enriched sequences was determined by microarray analysis at stage 11 in three experimental conditions. (1) Nodal–, embryos where nodal signaling was inhibited by injection of *Cerberus-S* RNA (1 ng/embryo). (2) Mixer–, embryos injected with Mixer antisense morpholino oligos (40 ng/embryo). (3) Sox17–, embryos injected with antisense morpholino oligos to Sox17α1 + Sox17α2 + Sox17β (20 ng each/embryo). The average intensity, normalized to stage 11 whole embryo (We) is shown on a log scale. An, animal cap ectoderm; Veg, endoderm enriched vegetal tissue; Eq, mesendoderm enriched equatorial tissue. The yellow lines indicate the 1.4-fold change threshold from the control. (**B**) Hierarchical clustering indicates which conditions have the most similar expression profiles. Low expression is indicated in blue and high expression in red. A transcript was considered 'regulated' in a given condition if its expression was more than 1.4-fold changed from whole embryo control. (**C**) The Venn diagram shows that based on this 1.4-fold criteria, 112/301 endoderm transcripts are Nodal regulated, 168/301 are Mixer regulated and 100/301 are Sox17 regulated, with a total of 223/301 endoderm transcripts that are regulated by either Nodal proteins Mixer or Sox17. Only 36/301 of transcripts (white) are regulated in a manner consistent with the simple linear model of endoderm development.

From these comparisons and RT-PCR validations of more than 50 additional genes (see Table S1 in the supplementary material), we found that changes in expression levels of less than 1.4-fold from control whole embryo on the array (Fig. 4; yellow bar) were often not reproducible by RT-PCR. By contrast, changes in expression levels of greater than 1.4-fold up or down from control whole embryo levels represented, in most instances, robust changes that were reproducibly validated, indicating an accurate identification of the expression of a gene and its regulation by the experimental condition. This greater than 1.4-fold cut-off exhibited the highest validation rate, without excluding moderate changes that we knew were real by RT-PCR and published reports. With this criterion, ~74% (n=53; Table S1 in the supplementary material) of the time the array data correctly predicted the behavior of a gene in all three experimental conditions. In ~21% of the cases, the array prediction was partially validated in that the trend in expression change was correct, but the threshold of more than a 1.4-fold change was not met in one or more conditions. Only 5% of the time did the array predict a change that was contradictory to the RT-PCR validation, indicating that the array data were a very good predictor of a the regulation of a gene.

Based on the criterion of more than a 1.4-fold change, we found that 223 of the 301 endoderm enriched genes were regulated by either Nodal signaling, Mixer or Sox17 (Fig. 3C), with 112 Nodal-regulated, 168 Mixer-regulated and 100 Sox17-regulated genes, respectively. Of the 78 genes that were not regulated by Nodal proteins, Mixer or Sox17, 67 had high maternal expression, including germ plasm genes *Dazl* and *Deadsouth* (see Table S1 in the supplementary material). Surprisingly, a Venn analysis indicated that only 36/223 transcripts were similarly regulated by Nodal, Mixer and Sox17, including *HNF1β* and *Edd*, two of the early endoderm markers used to establish the current linear model of endoderm development. This suggests that the transcriptional network controlling endoderm development is more complex than predicted by the current model.

Epistatic relationships between Nodal proteins, Mix-like, Gata4-6 and Sox17α/β

The array analysis revealed a number of previously unappreciated relationships between *Xnrs*, *Mixer*, *Bix1-4*, *Gata4-6* and *Sox17α/β*, and their downstream targets. For example, we found that expression of *Xnr4*, *Mix2*, *Bix1*, *Bix2*, *Bix4* and *Gata* genes were all downregulated in *Sox17*– embryos (Fig. 4), which was unexpected as they were previously thought to act upstream of Sox17. To test these observations, we performed a series of loss-of-function and rescue experiments, comparing the ability of Sox17β, Mixer or Gata6, to rescue gene expression in *Nodal*– embryos with their ability to rescue gene expression in *Sox17*– embryos (Fig. 5A). According to the linear model, all should rescue Nodal inhibition, but only injection of *XtSox17β* RNA (*Xenopus tropicalis* Sox17β mRNA lacking the sequence targeted by the Sox17 morpholino) should rescue gene expression in embryos where endogenous Sox17 protein has been depleted.

Sox17 is involved in multiple autoregulatory loops

Our data, in conjunction with published reports, suggests three major feedback loops: one between Sox17 and Gata4-6; a second unexpected autoregulatory loop between Sox17 and Bix1/Bix2/Bix4; and a third between Sox17 and the Nodal ligand Xnr4.

Sox17 and Gata4-6

In animal caps, Gata4-6 and Sox17α/β are known to induce each others expression, and Gata4-6 are required for full *Sox17α/β* expression levels in the gastrula (Afouda et al., 2005; Clements et al., 2003; Sinner et al., 2004). Here, we show that *Gata4-6* are downregulated in both *Nodal*– and *Sox17*– embryos (Fig. 4) (Clements et al., 2003) and that injection of *XtSox17β* can partially rescue *Gata5-6* expression in both *Nodal*– and *Sox17*– embryos. Similarly, we find that Gata6 can rescue *Sox17α/β* expression in Nodal– embryos (Fig. 5A). Together, these data demonstrate that Sox17 and Gata4-6 autoregulate the expression of one another downstream of Nodal signaling.

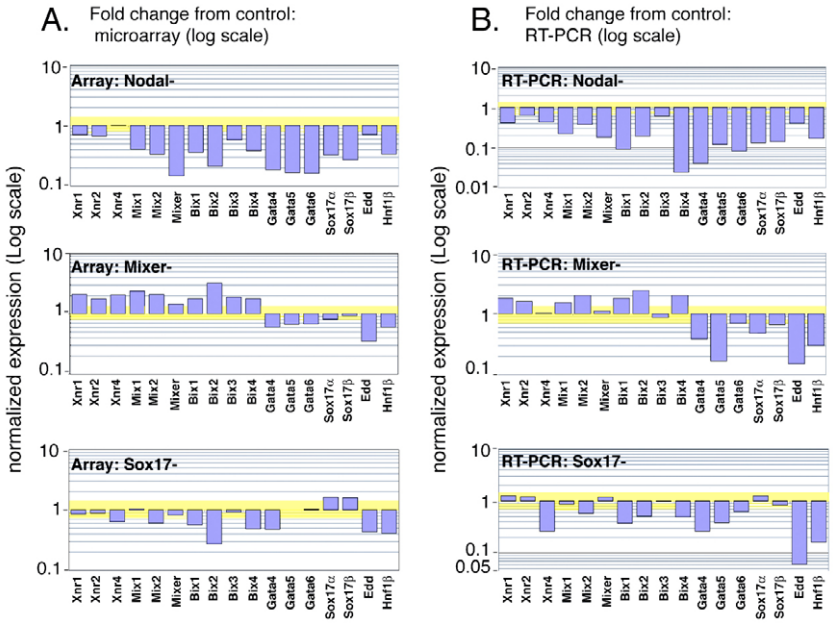

Fig. 4. Expression of known zygotic endoderm regulators in Nodal–, Mixer– and Sox17–embryos. The expression of the known endoderm regulators *Xnr1, Xnr2, Xnr4, Mix1, Mix2, Mixer, Bix1, Bix2, Bix3, Bix4, Gata4, Gata5, Gata6, Sox17α, Sox17β, Edd* and *Hnf1β* from (**A**) the microarray and (**B**) real-time RT-PCR validation was determined from *Nodal*– (Cerberus-S injected 1 ng/embryo), *Mixer*– (Mixer morpholino injected 40 ng/embryo) and *Sox17*– (Sox17α1+α2+β morpholino injected 20 ng each/embryo) embryos at stage 11. Both the array data and RT-PCR is plotted on a log scale. Changes in expression less than 1.4-fold up or down from the control are within the horizontal yellow bar. From these and the validation of more than 40 other genes changes in 1.4-fold were considered robust and were validated in over 75% of the cases.

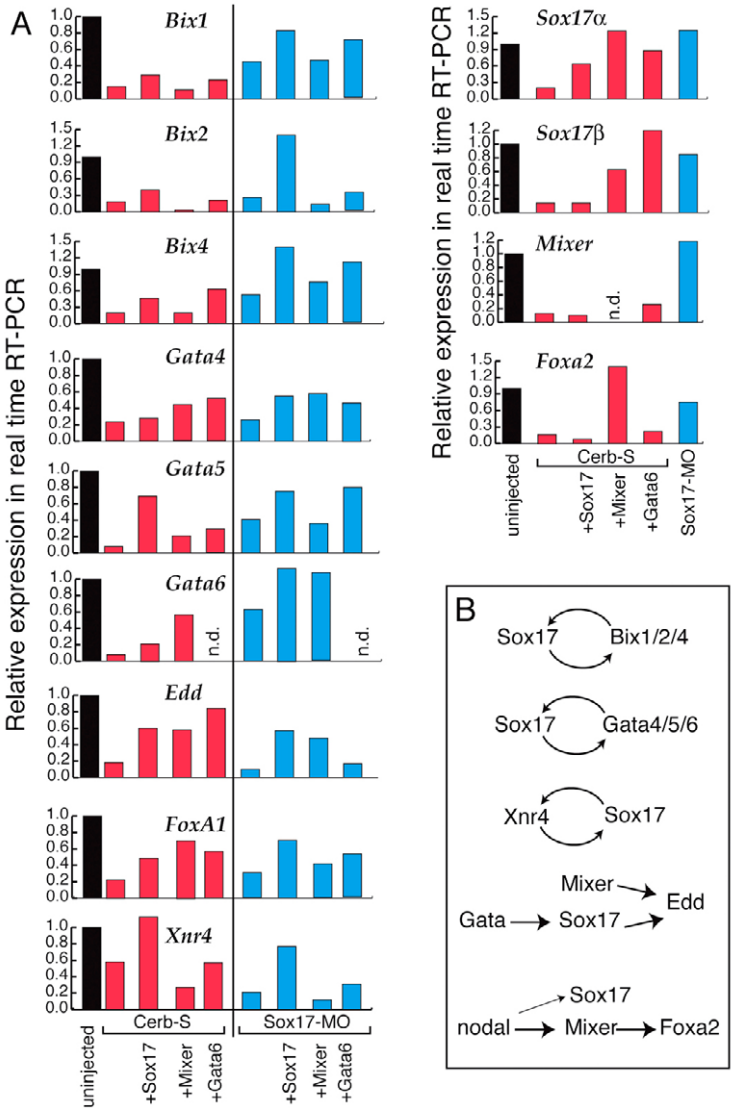

Fig. 5. Functional analysis of hierarchical relationships between zygotic endoderm regulators. (**A**) At the two-cell stage, embryos were injected with either *Cerb-S* RNA to inhibit Nodal signaling (1 ng) or antisense morpholino oligos to Sox17α1+α2+β (Sox17-MO; 20 ng each). Some Cerb-S and Sox17-MO injected embryos were then injected with *XtSox17β* RNA (from *Xenopus tropicalis* and resistant to the Sox17-MOs; 10-100 pg), *Mixer* RNA (50-500 pg) or *Gata6* RNA (50-100 pg). A range of rescue RNA doses was used and the lowest does that gave a reproducible rescue is shown. At stage 11, RNA from the embryos was assayed by real time RT-PCR for expression of *Bix1, Bix2, Bix4, Gata4, Gata5, Gata6, Edd, Foxa1, Xnr4, Sox17α, Sox17β, Mixer* and *Foxa2*. Relative expression normalized to *ODC* is shown and the expression level in control gastrula was set to 1.0. This experiment was repeated three times and a representative example is shown. (**B**) These results, along with previously published reports, support the indicated regulatory relationships, which were previously not described in the existing models of endoderm development. *Foxa2* (Ruiz i Altaba et al., 1993) is not present on the Affymetrix *Xenopus* chip and is distinct from *Pintallavis* (Ruiz i Altaba and Jessell, 1992) or *XFKH1* (Dirksen and Jamrich, 1992), which are *Foxa4a* and *Foxa4b*, respectively (Kaestner et al., 2000).

Sox17 and Bix

In both Cerberus-Short and Sox17-depleted embryos, *Bix1*, *Bix2* and *Bix4* are downregulated, and their expression can be partially rescued by injection of *XtSox17β* RNA (Fig. 5A). This was surprising because, although Bix1, Bix2 and Bix4 have all been shown to induce *Sox17* transcription in animal caps (Casey et al., 1999; Ecochard et al., 1998; Tada et al., 1998), the reverse is not true: Sox17 cannot induce *Bix1-4* transcription in animal cap assays (Sinner et al., 2004). This implies that Sox17 requires additional cofactors to promote Bix gene expression and that these factors are absent from the animal hemisphere. Candidate co-factors include activated Smad2, or possibly Gata4-6 [as Gata6 can partially rescue *Bix1* and *Bix4* expression in *Sox17*– embryos (Fig. 5A)]. Interestingly Sox17-dependent regulation was specific to only a subset of the *Mix-like* gene family and was not observed with *Mix1*, *Bix3* or *Mixer*. Together with published reports, these data suggest that Sox17 and Bix1, Bix2, Bix4 regulate each others expression and that Gatas may also participate in this cross regulatory loop.

Sox17 and Xnr4

The third autoregulatory loop we identified was between Sox17 and Xnr4. We found that *Xnr4* expression, which was thought to act at the top of the zygotic gene hierarchy regulating endoderm development, was dependent on Sox17 (but not Mixer or Gatas). Injection of *XtSox17β* RNA rescued the *Xnr4* expression levels in *Nodal*– and *Sox17*– embryos (Fig. 5A), suggesting that Sox17 may act in part by maintaining Nodal signaling, one of the most upstream components of the endoderm specification pathway. It is intriguing that only *Xnr4* and not any other *Xnrs* are Sox17 dependent, suggesting that Xnr4 may have some unique function.

Mixer does not function primarily via Sox17

Although Mixer rescued *Sox17* expression in *Nodal*– embryos (Fig. 5A), we consistently found that *Sox17* was only moderately downregulated in *Mixer*– embryos (Fig. 4) (Kofron et al., 2004). Furthermore, of the 268 genes regulated by either Sox17 or Mixer, only 67 genes were regulated by both Mixer and Sox17 (Fig. 3). Thus, the modest reduction in *Sox17* levels observed in *Mixer*– embryos cannot account for the Mixer loss-of-function phenotype, indicating that Mixer does not function primarily via Sox17 as commonly cited.

Endoderm target genes are not all coordinately regulated

Contrary to the current model, we found that the early endoderm markers *Edd*, *Hnf1β*, *Foxa1* and *Foxa2* were not all regulated in same way. As expected, the reduction of *Edd*, *Foxa1* and *Hnf1β* expression in *Nodal*– embryos was rescued by injection of *Sox17*, *Mixer* or *Gata6* RNA (Fig. 5A; data not shown). However, in *Sox17*– embryos, only Sox17 or Mixer, but not Gata4-6, rescued *Edd* expression (Fig. 5A; data not shown). These data, along with the fact that *Edd* is downregulated in *Mixer*– embryos (Fig. 4) (Kofron et al., 2004), indicates that Sox17 and Mixer independently contribute to *Edd* expression, and that Gatas regulate *Edd* via Sox17 (Fig. 5B). Finally, *Foxa2*, which can be induced in animal cap experiments by ectopic Sox17 (Sinner et al., 2004), does not require Sox17 for expression and only Mixer (but not Sox17 or Gata6) rescued *Foxa2* expression in *Nodal*– embryos. This suggests that although all three factors, Sox17, Mixer and Gata, participate in *Foxa1* regulation, Mixer is the primary regulator of *Foxa2* expression (Fig. 5B).

These results challenge the existing model of endoderm development and establish a new number of epistatic relationships between the known endoderm regulators. Our data suggests that Sox17 is not the most downstream component of the endoderm specification pathway, as commonly cited, but rather participates in auto regulatory loops with Bix1, Bix2, Bix4, Gata4-6 and Xnr4, all of which were previously considered to be upstream of Sox17. In addition, we find that Mixer does not function primarily via Sox17, as predicted by the linear model, and that different endoderm target genes have varying modes of regulation.

The regulatory network controlling endoderm transcription is complex

Having examined the regulation of the known endodermal genes, we next wanted to determine how endoderm transcription is regulated at a global level. We therefore examined the array data to identify patterns of *Nodal*-, *Mixer*- and *Sox17*-dependent gene expression in all 301 endoderm-enriched transcripts. Based on the criterion of a greater than 1.4-fold change in expression levels relative to controls, we grouped genes into one of three different categories for each of the three experimental condition. Genes with reduced expression in *Nodal*–, *Mixer*– or *Sox17*– embryos were classified as positively regulated (+) by Nodal proteins, Mixer or Sox17, respectively. Genes with increased expression in *Nodal*–, *Mixer*– or *Sox17*– embryos were considered negatively regulated (–) and normally repressed by Nodal proteins, Mixer or Sox17, respectively. Finally, genes exhibiting less than a 1.4-fold change in expression levels in either *Nodal*–, *Mixer*– or *Sox17*– embryos relative to controls were considered to be 'not obviously regulated' (0) by Nodal proteins, Mixer or Sox17, respectively. Based on these criteria, we classified each of the 301 endoderm-enriched genes as positively, negatively or not regulated by Nodal signaling, Mixer and Sox17 (Fig. 6; Table 2; Table S1 in the supplementary material). Overall, Nodal proteins, Mixer or Sox17 regulated 223 of the 301 sequences; of the 78 genes that were not regulated, 67 had significant maternal expression (Fig. 3C, Fig. 6O).

Hypothetically, there are 27 (3^3) possible modes of regulation that a given gene could exhibit (Table 2). Alternatively, if the simplest linear model was correct, then all zygotic endoderm genes would be positively regulated by Nodal proteins, Mixer and Sox17 (+N +M +S). Surprisingly, we found that only 25 out of the 223 sequences (~10%) were regulated in this manner (Fig. 6A; Table 2) and that endoderm gene expression can be classified into 19 different modes of regulation (Fig. 6; Table 2). Table S1 in the supplementary material provides a full list of how each gene was classified and the average normalized expression data for each condition. To validate some of these novel modes of regulation, we performed loss-of-function and rescue experiments, comparing the ability of RNA encoding, Sox17, Mixer or Gata6, to rescue gene expression in *Nodal*– embryos (Fig. 7).

Nodal proteins, Mixer and Sox17 have both shared and distinct downstream targets

First, we confirmed that the genes *Xl.5999* and *Xl.8924* (Fig. 6A) were positively regulated by Nodal signaling, Mixer and Sox17 (Fig. 7A; +N +M +S), similar to *Hnf1β* and *Edd*. Co-injection of *Mixer* RNA produced the best rescue of *Xl.5999* and *Xl.8924* expression in *Nodal*– embryos, while the rescue by Gata and Sox17 was very modest. These results are consistent with the hypothesis that Mixer induces *Sox17* and *Gata* expression (Fig. 5), and then all three of these contribute to *Xl.5999* and *Xl.8924* regulation.

We classified 21 transcripts positively regulated by Nodal proteins and Mixer, but not by Sox17 (Fig. 7B; +N +M 0S); 14 transcripts positively regulated by Nodal and Sox17 but not Mixer (Fig. 7C; +N 0M +S); and 22 transcripts positively regulated by

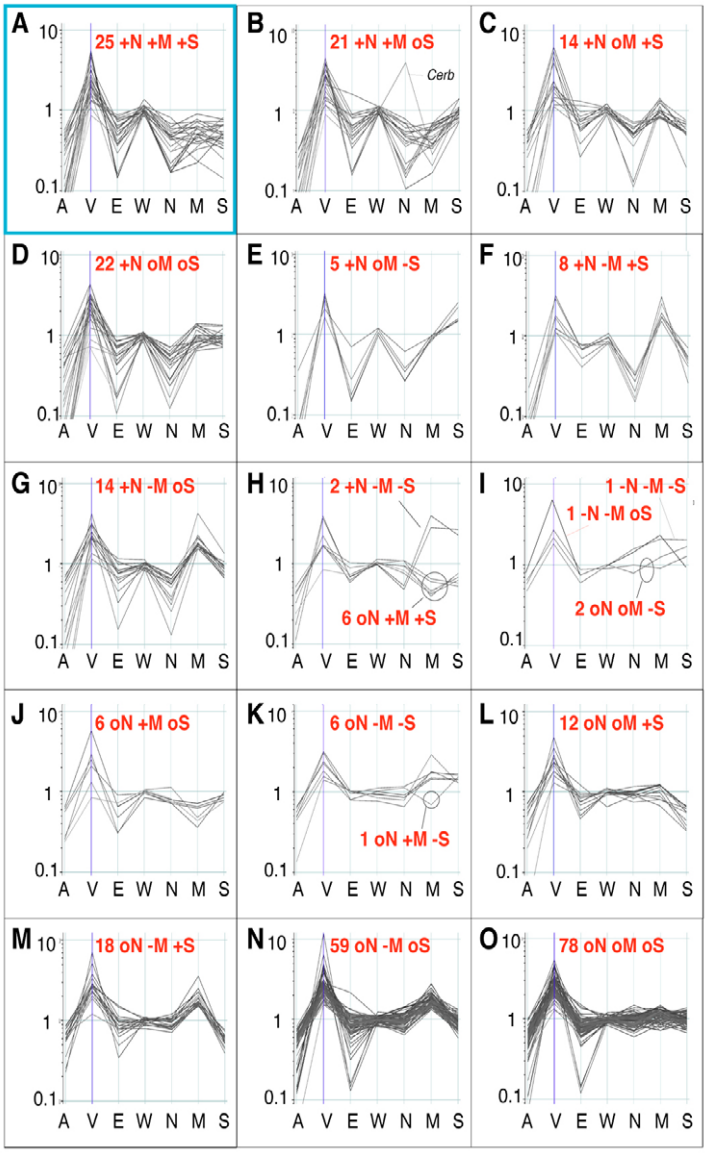

Fig. 6. Novel modes of endoderm gene regulation. (**A-O**) Each of the 301 endoderm-enriched transcripts was placed in a category based on how it was regulated in *Nodal*– (N), *Mixer*– (M) and *Sox17*– (S) embryos, based on the criteria of 1.4-fold change in expression levels relative to stage 11 control embryo. Genes downregulated more than 1.4-fold in *Nodal*–, *Mixer*– or *Sox17*– embryos were classified as positively regulated by Nodal proteins (+N), Mixer (+M) or Sox17 (+S), respectively. Genes upregulated more than 1.4-fold in *Nodal*–, *Mixer*– or *Sox17*– embryos were classified as negatively regulated (–), i.e. normally repressed, by Nodal proteins (–N), Mixer (–M) or Sox17 (–S), respectively. Genes with less than a 1.4-fold change in expression levels relative to controls were considered to be 'not obviously regulated' by Nodal proteins (0N), Mixer (0M) or Sox17 (0S), respectively. The number of genes in each category is indicated in red. The simple linear model of endoderm development predicts only coordinate positive regulation by Nodal proteins, Mixer and Sox17 (A, blue box). The endoderm-enriched transcriptome is complex and can be classified into 19 different categories. A, animal cap; V, vegetal region; E, equatorial region; W, whole embryo.

Nodal proteins, but not regulated by either Mixer or Sox17 alone (Fig. 7D; +N 0M 0S). The regulation of these different groups of genes is more consistent with a model where Nodal proteins, Mixer and Sox17 each have distinct sets of target genes. For example, only co-injection of *Mixer* RNA rescued *Darmin* and *Xl.15089* expression in *Nodal*– embryos, confirming they are positively regulated by Nodal proteins and Mixer, but not by Sox17 or Gata6 (Fig. 7B; +N +M 0S). The rescue experiments also validated genes that were positively regulated by Nodal proteins and Sox17, but not by Mixer, such as *Foxa4* and *Xl.13381* (Fig. 7C; +N 0M +S). In the case of *Xl.13381*, injection of *XtSox17β* RNA alone could not rescue its expression in *Nodal*– embryos, suggesting that other nodal dependent factors are also required for *Xl.13381* transcription. The genes *Xenf* (Nakatani et al., 2000), *Xl.2554* and *Mig30* (Hayata et al., 2002) are examples of the 22 endoderm genes that require Nodal signaling, but are not significantly regulated by Sox17, Mixer or Gata6 (Fig. 6D, Fig. 7D; +N 0M 0S). We hypothesize that these may require the combined action of Sox17, Mixer or Gata6, might be direct Smad2 targets, or might be regulated by some unknown Nodal-dependent factor.

Mixer has a major role in negatively regulating mesendoderm genes

Another unexpected result from the array data was that of the 168 Mixer-regulated transcripts, 108 of them were upregulated in *Mixer*– embryos. An examination of their spatial expression reveals that most of the upregulated genes are highly expressed in the equatorial region (compare low Eq expression in Fig. 6A-E with higher Eq expression in Fig. 6K,M,N), suggesting that that Mixer functions primarily to repress mesendoderm genes, as recently suggested (Kofron et al., 2004). For example the 14 genes classified as +N, –M 0S (Fig. 6G) included the mesendoderm gene *Eomesodermin* and the novel gene *Xl.2967*, which is also enriched in the equatorial region (see Table S3 in the supplementary material). In another example, rescue experiments confirm that *Xl.2967* requires Nodal signaling, probably via Gata proteins and that Mixer represses *Xl.2967* expression (Fig. 7E).

Nodal-independent regulation?

Finally, the array data indicate there are a number of different categories, comprising ~100 genes that were regulated by Sox17 and/or Mixer, but their expression was not

Table 2. Predicted regulation of endoderm-enriched transcripts by Nodal proteins, Mixer and Sox17

Number of sequences on array	Number of genes	Nodal regulation	Mixer regulation	Sox17 regulation	Representative genes
25	23	+	+	+	*Hnf1b, Edd, Gata4, Xl.5999, Xl.8924*
21	18	+	+	0	*Frzb1, Darmin, Ceberus*, Xl.15089*
0	0	+	+	–	
14	13	+	0	+	*Foxa4/pintalavis, PAPC, Xl.13381*
22	22	+	0	0	*Xenf, Hex, Mig30, Xl.2554*
5	3	+	0	–	*Sox17a, Sox17b, ClEBPa*
8	7	+	–	+	*Bix1, Bix2, Bix4, Mix2, Gsc, Otx2*
14	13	+	–	0	*Mix1, Bix3, Eomes, Xnr1, Xnr2*
2	1	+	–	–	*Xnr3*
6	6	0	+	+	*Tbx6-like, Xl.15054*
6	6	0	+	0	*Vex-1, Xl.11188*
1	1	0	+	–	*Wnt11*
12	12	0	0	+	*Otx1, Wnt8, fatvg*
78	75	0	0	0	*Vg1, Xpat, Deadsouth, Xoo1*
2	2	0	0	–	*Xl.5556*
18	18	0	–	+	*Chk1, BMP2, XPTB, Xl.4709*
59	59	0	–	0	*Derriere, Germes, Dazl, Xl.14891*
6	6	0	–	–	*Hermes, Oct-60*
0	0	–	+	+	
0	0	–	+	0	
0	0	–	+	–	
0	0	–	0	+	
0	0	–	0	0	
0	0	–	0	–	
0	0	–	–	+	
1	1	–	–	0	*Dead end*
1	1	–	–	–	*Xl.12017*

+, positively regulated during normal development, expression level >1.4-fold DOWN relative to control; –, negatively regulated during normal development, expression more than 1.4-fold UP relative to control; 0, no obvious regulation, change in expression level less than 1.4-fold relative to control.
*Based on RT-PCR validation because the Nodal– sample had an artificially high signal on the array owing to hybridization of injected Cerb-S RNA.
Where there was more than one sequence on the array for a gene, the average expression level was used.

significantly altered by blocking Nodal signaling (Fig. 6H-N). This suggests that either a significant proportion of endoderm formation is independent of Nodal signaling, and/or that Nodal proteins regulate the expression of both activators and repressors, and the loss of both results in little overall changes in gene expression.

For example, *Xl.4709* is one of the 18 genes with expression unchanged in *Nodal*– embryos, upregulated in *Mixer*– embryo and downregulated in *Sox17*– embryos (Fig. 6M; 0N–M+S). Injection of *Mixer* RNA or the Sox17-MO repressed *Xl.4709* levels, while injection of the Mixer-MO resulted in over expression of *Xl.4709* (Fig. 7F). In a second example, *Xl.1489* was upregulated by depletion of Mixer or injection of *Gata6* RNA (Fig. 7G). We hypothesize that in *Nodal*– embryos both activators (such as Sox17) and repressors (perhaps Mixer) would be missing, resulting in little change in gene expression. However, in the absence of repression by Mixer, activation by Sox17 predominates; while in the absence of Sox17, repression by Mixer predominates and the gene is downregulated.

Sox17 negatively regulates Wnt/β-catenin pathways components

Of the 100 Sox17-regulated sequences we observed, 17 were upregulated in *Sox17*– embryos and thus normally repressed by Sox17 activity (Fig. 6E,H,I,K). Interestingly, at least two of these are components or targets of the Wnt/β-catenin pathway: *Wnt11* and *Xnr3* (McKendry et al., 1997; Tao et al., 2005). This is consistent with reports that Sox17 can antagonize β-catenin/TCF transcriptional activity in vitro (Zorn et al., 1999a) and suggests that in the embryo Sox17 may also restrict Wnt-responsive transcription. In the case of *Xnr3*, our rescue experiments indicate that it is also repressed by Mixer, but positively regulated by Nodal signaling, perhaps via Gata proteins (Fig. 7H).

In summary, we find that the linear model of endoderm formation does not accurately describe the bulk of endoderm gene expression, which is much more complex than previously appreciated. Importantly, this work provides a complete documentation of how each of the 301 endoderm-enriched transcripts were regulated by Nodal signaling, Mixer and Sox17 (see Table S1 in the supplementary material), providing a comprehensive resource for examining the gene regulator network controlling *Xenopus* endoderm formation.

DISCUSSION

Using a combination of microarray analysis and extensive validations, we identified ~300 genes with endoderm-enriched expression, including over a hundred genes uncharacterized in any species. As our strategy identified most of the genes known to control endoderm formation, it is likely that many of these unknown genes may also have important regulatory functions.

Using this robust gene list, we interrogated the existing models of endoderm development determining how global endoderm gene expression was regulated Nodal proteins, Mixer and Sox17. In addition to identifying many novel Nodal, Mixer and Sox17 targets, these experiments indicate that the transcriptional hierarchy controlling endoderm gene expression is much more complicated than previously appreciated, with only 10% of the endoderm transcriptome being regulated as predicted by the linear model commonly cited in the literature. Our analysis classified endoderm gene expression

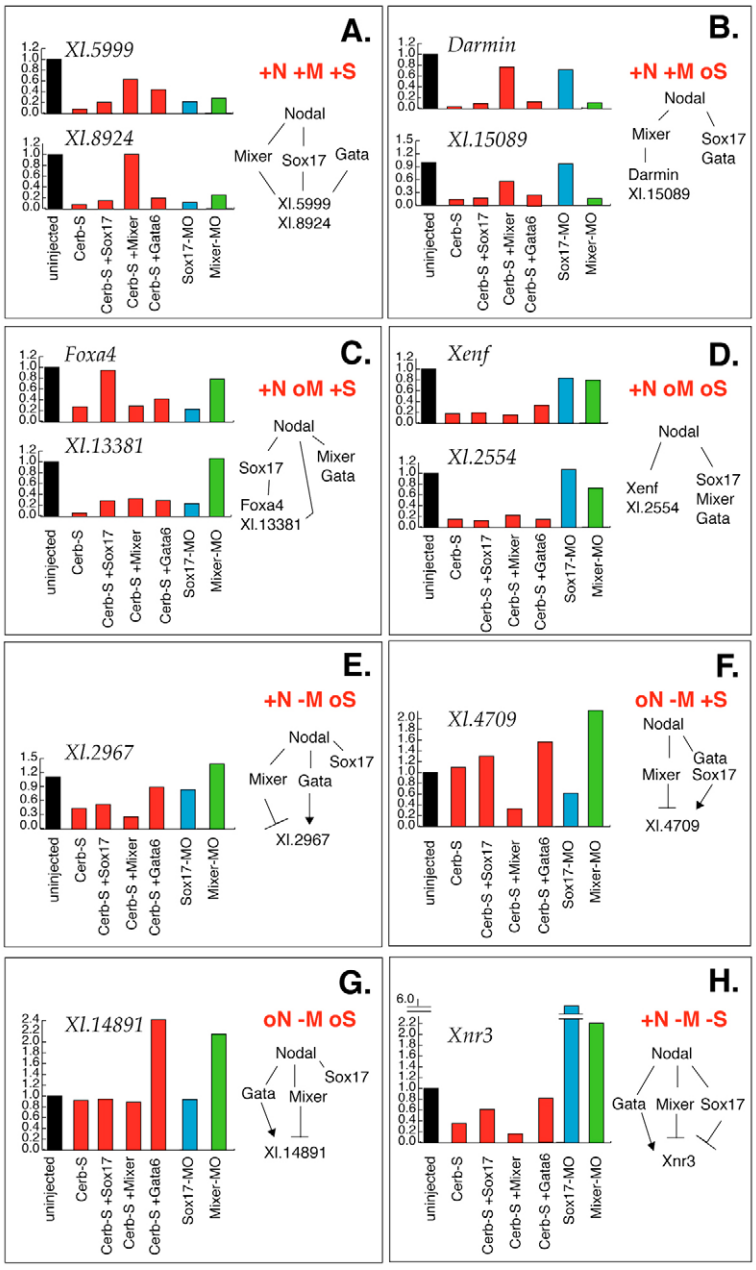

Fig. 7. Testing novel modes of endoderm gene regulation predicted by array analysis. (**A-H**) At the two-cell stage, embryos were injected with either *Cerb-S* RNA to inhibit nodal signaling (1 ng), antisense morpholino oligos to Sox17α1+α2+β (Sox17-MO; 20 ng each) or Mixer antisense morpholino oligos (40 ng/embryos). Some *Cerb-S*-injected embryos were subsequently injected with *XtSox17β* RNA (10-100 pg), *Mixer* RNA (50-500 pg) or *Gata6* RNA (50-100 pg) to rescue target gene expression. A range of rescue RNA doses was used and the lowest does that gave a reproducible rescue is shown. At stage 11, RNA from the resulting embryos was assayed by real-time RT-PCR for the expression of the indicated transcripts. Relative expression normalized to *ODC* is shown and the expression level in control gastrula was set to 1.0. Each experiment was repeated at least twice and a representative experiment is shown. Potential regulatory pathways are shown for each class of regulation (right).

into 19 different categories of regulation with Nodal proteins, Mixer and Sox17 having both shared and distinct sets of target genes.

We validated a number of novel epistatic relationships between the genes known to regulate endoderm formation. We found strong evidence for three auto regulatory loops: one between Sox17 and Bix1, Bix2 and Bix4; a second between Sox17 and Gata4-6; and a third between Sox17 and Xnr4. This was surprising because *Xnr4*, *Mix2*, *Bix1*, *Bix2*, *Bix4* and *Gata4* were all previously thought to act upstream of Sox17, based on animal caps experiments (Afouda et al., 2005; Casey et al., 1999; Clements et al., 1999; Sinner et al., 2004; Tada et al., 1998). These results clearly demonstrate that Sox17 is not the most downstream component of the pathway regulating endoderm formation, as commonly described. We also found that Mixer does not function primarily via Sox17 as commonly cited. Although Mixer probably participates in maintaining *Sox17* expression, much of Mixer function is independent of Sox17, and Mixer (but not Sox17) has a major role in negatively regulating the expression of over a hundred mesendoderm genes.

Two other recent studies have also used microarrays to identify the genes involved in *Xenopus* endoderm development: one by Taverner et al. (Taverner et al., 2005) looking at VegT targets; and another by Dickinson et al. (Dickinson et al., 2006) attempting to identify Mixer and Sox17β target genes. An important distinction between those studies and this one is that they both used overexpression in animal caps to identify downstream targets. By contrast, we defined the endogenous endoderm transcriptome and used loss-of-function approaches to examine gene regulation. Although clearly useful, overexpression animal cap studies have limitations. For example, we found only 30 of the 71 Mixer and Sox17 target genes identified by Dickinson et al. in our primary list of ~500 endoderm-enriched transcripts. When we examine the expression profile of the other 41 genes that were not in our list, only

four were endoderm enriched in our array data and the rest were either enriched in the egg or ectoderm tissue. This indicates that animal cap assays often do not recapitulate endogenous endoderm development, possibly because animal cap cells do not contain all the endogenous co-factors that normally interact with the endoderm transcription factors.

The study we performed here also has limitations. We have not tested cases where two or more factors are required redundantly to regulate gene expression. Furthermore, we focused on genes with endoderm-enriched expression, but clearly there will be genes that are not only expressed in the endoderm that have crucial functions in endoderm development. In addition, gene regulatory networks are known to evolve during developmental time (Bolouri and Davidson, 2003; Loose and Patient, 2004) and so far we have only focused on stage 11. It is likely that Nodal proteins, Mixer and Sox17 may have different functions at different times and in different regions of the embryo (Clements and Woodland, 2003; Yasuo and Lemaire, 1999).

Even with these limitations, we believe that our global analysis adds substantially to the emerging gene regulatory network describing *Xenopus* mesendoderm formation (Loose and Patient, 2004). In addition to identifying much of the endoderm transcriptome, this work provides an essential reference point from which future functional and epistatic experiments can be devised. In the future, it will be important to identify which regulatory events described here are directly controlled at the level of transcription factors binding to promoter elements as opposed to secondary events, which is an essential step establishing a robust gene regulatory network.

Based on the data from this study, along with previously published reports, we propose that a 'core' auto-regulatory network exists between the Nodal proteins, Mix-like, Gata4-6 and Sox17 factors, with the expression of any one component promoting the expression of the other components. This feed-forward system allows for the rapid establishment of an endoderm transcription profile in vegetal cells in the hours between activation of zygotic transcription at the early blastula to the gastrula stage, when endodermal fate is specified. Coupled with the repressive activity of Mixer and Mix1, such a system could also help establish both the endoderm and its boundary with the mesoderm.

We hypothesize that different species could initiate this conserved 'core' zygotic pathway by different means producing a similar outcome. In *Xenopus*, the core pathway is activated by maternal VegT, while in mouse and zebrafish the pathway may be activated at the level of Nodal proteins by some unknown mechanism (Tam et al., 2003). Indeed, a comparison of our data with a transcription profile the mouse gastrula endoderm (Gu et al., 2004) identified a number of common genes, suggesting that the global regulation of endoderm gene expression may be conserved.

We believe this is the first global analysis of the conserved molecular pathway controlling vertebrate endoderm formation during gastrulation. Our data challenge many aspects of existing models of vertebrate endoderm development and provide an important resource for further studies of the complex gene regulatory network controlling *Xenopus* endoderm development.

We are grateful to Drs Ueno, De Robertis and Patient for generously providing cDNA constructs. We thank Jim Wells for discussions and critical comments on the manuscript. Affymetrix hybridizations and bioinformatics support was provided by CHRF core facilities and supported in part by a DDRDC grant from the NIH (DK064403). This work was supported by grants from the Wellcome Trust to H.R.W., and from the NIH to J.H. (HD33002) and A.M.Z. (HD42572).

Supplementary material

Supplementary material for this article is available at
http://dev.biologists.org/cgi/content/full/132/10/1955/DC1

References

Afouda, B. A., Ciau-Uitz, A. and Patient, R. (2005). GATA4, 5 and 6 mediate TGFbeta maintenance of endodermal gene expression in Xenopus embryos. *Development* **132**, 763-774.

Agius, E., Oelgeschlager, M., Wessely, O., Kemp, C. and De Robertis, E. M. (2000). Endodermal Nodal-related signals and mesoderm induction in Xenopus. *Development* **127**, 1173-1183.

Bolouri, H. and Davidson, E. H. (2003). Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. *Proc. Natl. Acad. Sci. USA* **100**, 9371-9376.

Casey, E. S., Tada, M., Fairclough, L., Wylie, C. C., Heasman, J. and Smith, J. C. (1999). Bix4 is activated directly by VegT and mediates endoderm formation in Xenopus development. *Development* **126**, 4193-4200.

Clements, D. and Woodland, H. R. (2003). VegT induces endoderm by a self-limiting mechanism and by changing the competence of cells to respond to TGF-beta signals. *Dev. Biol.* **258**, 454-463.

Clements, D., Friday, R. V. and Woodland, H. R. (1999). Mode of action of VegT in mesoderm and endoderm formation. *Development* **126**, 4903-4911.

Clements, D., Cameleyre, I. and Woodland, H. R. (2003). Redundant early and overlapping larval roles of Xsox17 subgroup genes in Xenopus endoderm development. *Mech. Dev.* **120**, 337-348.

D'Souza, A., Lee, M., Taverner, N., Mason, J., Carruthers, S., Smith, J. C., Amaya, E., Papalopulu, N. and Zorn, A. M. (2003). Molecular components of the endoderm specification pathway in Xenopus tropicalis. *Dev. Dyn.* **226**, 118-127.

Demartis, A., Maffei, M., Vignali, R., Barsacchi, G. and De Simone, V. (1994). Cloning and developmental expression of LFB3/HNF1 beta transcription factor in Xenopus laevis. *Mech. Dev.* **47**, 19-28.

Dickinson, K., Leonard, J. and Baker, J. C. (2006). Genomic profiling of Mixer and Sox17beta targets during Xenopus endoderm development. *Dev. Dyn.* **235**, 368-381.

Dirksen, M. L. and Jamrich, M. (1992). A novel, activin-inducible, blastopore lip-specific gene of Xenopus laevis contains a fork head DNA-binding domain. *Genes Dev.* **6**, 599-608.

Ecochard, V., Cayrol, C., Rey, S., Foulquier, F., Caillol, D., Lemaire, P. and Duprat, A. M. (1998). A novel Xenopus mix-like gene milk involved in the control of the endomesodermal fates. *Development* **125**, 2577-2585.

Engleka, M. J., Craig, E. J. and Kessler, D. S. (2001). VegT activation of Sox17 at the midblastula transition alters the response to nodal signals in the vegetal endoderm domain. *Dev. Biol.* **237**, 159-172.

Gu, G., Wells, J. M., Dombkowski, D., Preffer, F., Aronow, B. and Melton, D. A. (2004). Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* **131**, 165-179.

Hayata, T., Tanegashima, K., Takahashi, S., Sogame, A. and Asashima, M. (2002). Overexpression of the secreted factor Mig30 expressed in the Spemann organizer impairs morphogenetic movements during Xenopus gastrulation. *Mech. Dev.* **112**, 37-51.

Heasman, J., Wyile, C. C., Hausen, P. and Smith, J. C. (1984). Fates and states of determination of single vegetal pole blastomeres of X. laevis. *Cell* **37**, 185-194.

Henry, G. L. and Melton, D. A. (1998). Mixer, a homeobox gene required for endoderm development. *Science* **281**, 91-96.

Henry, G. L., Brivanlou, I. H., Kessler, D. S., Hemmati-Brivanlou, A. and Melton, D. A. (1996). TGF-beta signals and a pre-pattern in Xenopus laevis endoderm development. *Development* **122**, 1007-1015.

Hilton, E., Rex, M. and Old, R. (2003). VegT activation of the early zygotic gene Xnr5 requires lifting of Tcf-mediated repression in the Xenopus blastula. *Mech. Dev.* **120**, 1127-1138.

Horb, M. E. and Thomsen, G. H. (1997). A vegetally localized T-box transcription factor in Xenopus eggs specifies mesoderm and endoderm and is essential for embryonic mesoderm formation. *Development* **124**, 1689-1698.

Hudson, C., Clements, D., Friday, R. V., Stott, D. and Woodland, H. R. (1997). Xsox17alpha and -beta mediate endoderm formation in Xenopus. *Cell* **91**, 397-405.

Kaestner, K. H., Knochel, W. and Martinez, D. E. (2000). Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* **14**, 142-146.

Kofron, M., Demel, T., Xanthos, J., Lohr, J., Sun, B., Sive, H., Osada, S., Wright, C., Wylie, C. and Heasman, J. (1999). Mesoderm induction in Xenopus is a zygotic event regulated by maternal VegT via TGFbeta growth factors. *Development* **126**, 5759-5770.

Kofron, M., Wylie, C. and Heasman, J. (2004). The role of Mixer in patterning the early Xenopus embryo. *Development* **131**, 2431-2441.

Loose, M. and Patient, R. (2004). A genetic regulatory network for Xenopus mesendoderm formation. *Dev. Biol.* **271**, 467-478.

McKendry, R., Hsu, S. C., Harland, R. M. and Grosschedl, R. (1997). LEF-1/TCF

proteins mediate wnt-inducible transcription from the Xenopus nodal-related 3 promoter. *Dev. Biol.* **192**, 420-431.
Nakatani, J., Mizuseki, K., Tsuda, H., Nakanishi, S. and Sasai, Y. (2000). Xenopus Xenf: an early endodermal nuclear factor that is regulated in a pathway distinct from Sox17 and Mix-related gene pathways. *Mech. Dev.* **91**, 81-89.
Nieuwkoop, P. D. and Faber, J. (1994). *Normal Table of Xenopus Laevis (Daudin): A Systematical and Chronological Survey of the Development from the Fertilized Egg till the end of Metamorphosis*. New York: Garland Publishing.
Osada, S. I. and Wright, C. V. (1999). Xenopus nodal-related signaling is essential for mesendodermal patterning during early embryogenesis. *Development* **126**, 3229-3240.
Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T. and De Robertis, E. M. (1999). The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* **397**, 707-710.
Rosa, F. M. (1989). Mix.1, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of Xenopus embryos. *Cell* **57**, 965-974.
Ruiz i Altaba, A. and Jessell, T. M. (1992). Pintallavis, a gene expressed in the organizer and midline cells of frog embryos: involvement in the development of the neural axis. *Development* **116**, 81-93.
Ruiz i Altaba, A., Prezioso, V. R., Darnell, J. E. and Jessell, T. M. (1993). Sequential expression of HNF-3 beta and HNF-3 alpha by embryonic organizing centers: the dorsal lip/node, notochord and floor plate. *Mech. Dev.* **44**, 91-108.
Sasai, Y., Lu, B., Piccolo, S. and De Robertis, E. M. (1996). Endoderm induction by the organizer-secreted factors chordin and noggin in Xenopus animal caps. *EMBO J.* **15**, 4547-4555.
Sinner, D., Rankin, S., Lee, M. and Zorn, A. M. (2004). Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes. *Development* **131**, 3069-3080.
Sive, H. L., Grainger, R. M. and Harland, R. M. (2000). *Early Development of Xenopus Laevis: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
Stainier, D. Y. (2002). A glimpse into the molecular entrails of endoderm formation. *Genes Dev.* **16**, 893-907.
Stennard, F., Carnac, G., and Gurdon, J. B. (1996). The Xenopus T box gene Antipodean encodes a vegetally localized maternal mRNA that can trigger mesoderm formation. *Development* **122**, 4179-4188.
Tada, M., Casey, E. S., Fairclough, L. and Smith, J. C. (1998). Bix1, a direct target of Xenopus T-box genes, causes formation of ventral mesoderm and endoderm. *Development* **125**, 3997-4006.
Tam, P. P., Kanai-Azuma, M. and Kanai, Y. (2003). Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. *Curr. Opin. Genet. Dev.* **13**, 393-400.
Tao, Q., Yokota, C., Puck, H., Kofron, M., Birsoy, B., Yan, D., Asashima, M., Wylie, C. C., Lin, X. and Heasman, J. (2005). Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in Xenopus embryos. *Cell* **120**, 857-871.
Taverner, N. V., Kofron, M., Shin, Y., Kabitschke, C., Gilchrist, M. J., Wylie, C., Cho, K. W., Heasman, J. and Smith, J. C. (2005). Microarray-based identification of VegT targets in Xenopus. *Mech. Dev.* **122**, 333-354.
Weber, H., Symes, C. E., Walmsley, M. E., Rodaway, A. R. and Patient, R. K. (2000). A role for GATA5 in Xenopus endoderm specification. *Development* **127**, 4345-4360.
Xanthos, J. B., Kofron, M., Wylie, C. and Heasman, J. (2001). Maternal VegT is the initiator of a molecular network specifying endoderm in Xenopus laevis. *Development* **128**, 167-180.
Yasuo, H. and Lemaire, P. (1999). A two-step model for the fate determination of presumptive endodermal blastomeres in Xenopus embryos. *Curr. Biol.* **9**, 869-879.
Zhang, J. and King, M. L. (1996). Xenopus VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesoderm patterning. *Development* **122**, 4119-4129.
Zhang, J., Houston, D. W., King, M. L., Payne, C., Wylie, C. and Heasman, J. (1998). The role of maternal VegT in establishing the primary germ layers in Xenopus embryos. *Cell* **94**, 515-524.
Zorn, A. M., Barish, G. D., Williams, B. O., Lavender, P., Klymkowsky, M. W. and Varmus, H. E. (1999a). Regulation of Wnt signaling by Sox proteins: XSox17 alpha/beta and XSox3 physically interact with beta-catenin. *Mol. Cell* **4**, 487-498.
Zorn, A. M., Butler, K. and Gurdon, J. B. (1999b). Anterior endomesoderm specification in Xenopus by Wnt/beta-catenin and TGF-beta signalling pathways. *Dev. Biol.* **209**, 282-297.