casanova encodes a novel Sox-related protein necessary and sufficient for early endoderm formation in zebrafish

Yutaka Kikuchi,[1,4] Antoine Agathon,[2,4] Jonathan Alexander,[1] Christine Thisse,[2] Steven Waldron,[1] Deborah Yelon,[1,3] Bernard Thisse,[2] and Didier Y.R. Stainier[1,5]

[1]Department of Biochemistry and Biophysics, Programs in Genetics, Human Genetics, and Developmental Biology, University of California, San Francisco, California 94143-0448, USA; [2]Institut de Génétique et Biologie Moléculaire et Cellulaire CNRS/INSERM/ULP, BP 163, 67404 Illkirch cedex, CU de Strasbourg, France

Early endoderm formation in zebrafish requires at least three loci that function downstream of Nodal signaling but upstream of the early endodermal marker *sox17*: *bonnie and clyde (bon)*, *faust (fau)*, and *casanova (cas)*. *cas* mutants show the most severe phenotype as they do not form any gut tissue and lack all *sox17* expression. Activation of the Nodal signaling pathway or overexpression of Bon or Fau/Gata5 fails to restore any *sox17* expression in *cas* mutants, demonstrating that *cas* plays a central role in endoderm formation. Here we show that *cas* encodes a novel member of the Sox family of transcription factors. Initial *cas* expression appears in the dorsal yolk syncytial layer (YSL) in the early blastula, and is independent of Nodal signaling. In contrast, endodermal expression of *cas*, which begins in the late blastula, is regulated by Nodal signaling. Cas is a potent inducer of *sox17* expression in wild-type embryos as well as in *bon* and *fau/gata5* mutants. Cas is also a potent inducer of *sox17* expression in *MZoep* mutants, which cannot respond to Nodal signaling. In addition, ectopic expression of *cas* in presumptive mesodermal cells leads to their transfating into endoderm. Altogether, these data indicate that Cas is the principal transcriptional effector of Nodal signaling during zebrafish endoderm formation.

[*Key Words*: Sox; gut endoderm; Nodal; *bonnie and clyde*; *faust*; *casanova*]

Received February 28, 2001; revised version accepted April 13, 2001.

Vertebrate endoderm formation has recently been the focus of intense investigations. In the late blastula stage zebrafish embryo, endodermal cells form in the four most marginal blastomere tiers (Warga and Nüsslein-Volhard 1999), and involute early during gastrulation to occupy the deep hypoblast, directly overlying the extra-embryonic yolk syncytial layer (YSL; Warga and Kimmel 1990). During gastrulation, the endodermal cells are recognizable by their characteristically large size and flat morphology, distinct from the smaller, rounder mesodermal cells (Warga and Nüsslein-Volhard 1999). Starting at the onset of gastrulation, endodermal cells express *sox17*, a high-mobility group (HMG) domain transcription factor gene (Alexander and Stainier 1999), as well as *foxA2/axial/HNF3β* , a winged helix/forkhead transcription factor gene (Strähle et al. 1993; Schier et al. 1997; Kaestner et al. 2000).

Several zebrafish mutations affect endoderm formation and endodermal *sox17* and *foxA2* expression. The earliest acting mutations, *cyclops* (*cyc*), *squint* (*sqt*), and *oep*, affect Nodal signaling: Cyc and Sqt are themselves Nodal-related proteins (Erter et al. 1998; Feldman et al. 1998; Rebagliati et al. 1998a; Sampath et al. 1998), and Oep is a transmembrane protein essential for Nodal signaling (Zhang et al. 1998; Gritsman et al. 1999). Embryos mutant for both *cyc* and *sqt*, or embryos that lack both maternal and zygotic Oep (*MZoep*), fail to form all endoderm and most mesoderm (Feldman et al. 1998; Gritsman et al. 1999). *casanova* (*cas*), *faust* (*fau*), and *bonnie and clyde* (*bon*) more specifically affect endoderm formation and endodermal *sox17* and *foxA2* expression. The *cas* phenotype is the most dramatic, in that *cas* mutants do not form any gut tissue and fail to express any markers of endoderm differentiation or formation from the onset of gastrulation (Alexander et al. 1999; Alexander and Stainier, 1999).

In contrast to *cas* mutants that completely lack *sox17* expression, *fau* and *bon* mutants exhibit a significant level; at bud stage, *fau* and *bon* mutants contain ~60% and 10% of the wild-type number of *sox17*-expressing endodermal cells, respectively. *fau*, which encodes the

[3]Present address: Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, NY 10016, USA.
[4]These authors contributed equally to this work.
[5]Corresponding author.
E-MAIL didier_stainier@biochem.ucsf.edu; FAX (415) 476-3892.
Article and publication are at www.genesdev.org/cgi/doi/10.1101/gad.892301.

zinc finger transcription factor Gata5 (Reiter et al. 1999), and *bon*, which encodes a Mix-type homeodomain protein (Kikuchi et al. 2000), are both expressed in the marginal domain of the embryo beginning at dome stage, that is, ~80 min before the endodermal cells initiate *sox17* and *foxA2* expression (Alexander et al. 1999; Reiter et al. 1999,2001; Rodaway et al. 1999). The expression of both *bon* and *fau/gata5* is regulated by Nodal signaling (Alexander and Stainier 1999; Rodaway et al. 1999; Reiter et al. 2001). And whereas *bon* expression ceases soon after gastrulation commences (Alexander et al. 1999), *fau/gata5* expression persists and is involved in maintaining endodermal *sox17* and *foxA2* expression during gastrulation (Reiter et al. 2001).

cas also functions downstream of Nodal signaling as indicated by the failure of a constitutively active form of the type-I transforming growth factor β (TGF-β) receptor Taram-a (Renucci et al. 1996), a probable zebrafish ortholog of mammalian Alk4 (Payne et al. 2001) and Nodal receptor, to induce *sox17* expression in *cas* mutants (Alexander and Stainier 1999). Overexpression of *bon* in *cas* mutants also fails to restore endodermal *sox17* expression (Alexander and Stainier 1999), and similar data were obtained when *fau/gata5* was overexpressed in *cas* mutants (Reiter et al. 2001). In addition, mosaic analysis indicates that *cas* acts cell-autonomously in the endodermal progenitors (Alexander et al. 1999).

These and other data have suggested the following pathway for zebrafish endoderm formation: Nodal signaling induces the expression of *bon* and *fau/gata5*; Bon and Fau/Gata5 cooperatively regulate the expression of *sox17* and *foxA2*, but do not regulate each other's expression; Cas is essential cell autonomously for *sox17* expression and appears to function downstream of, or in parallel to, Bon and Fau/Gata5 (Alexander and Stainier 1999; Kikuchi et al. 2000; Reiter et al. 2001). Clearly, many issues remain to be resolved regarding this pathway and the isolation of *cas* should help address some of them.

Studies in *Xenopus* also implicate Nodal-related molecules (Osada and Wright 1999), a Mix-type homeodomain protein, Mixer (Henry and Melton 1998), and Gata5 (Weber et al. 2000) in the regulation of *Xsox17* expression and endoderm formation (Xanthos et al. 2001). XSox17 itself appears to be an important intrinsic regulator of endoderm formation; when overexpressed in animal caps, it activates endodermal gene expression, and overexpression of a dominant interfering variant of itself (*Xsox17-enR*) inhibits endoderm differentiation in both isolated vegetal pole explants and the intact frog embryo (Hudson et al. 1997).

The striking parallels between the pathways regulating endoderm formation in zebrafish and *Xenopus* suggest that similar molecules are likely to function in this process in mammals. Indeed, Nodal signaling has also been implicated in mouse endoderm development (Conlon et al. 1994; Tremblay et al. 2000), and Gata factors are expressed and function in a variety of developing tissues including the endoderm (for review, see Zaret 1999).

Given its central role in zebrafish endoderm formation, we sought to isolate *cas* using a combined approach of positional cloning and candidate gene testing. Here, we show that *cas* encodes a novel member of the Sox family of HMG domain transcription factors. *cas* expression initiates in the dorsal YSL in the early blastula and is later found in the presumptive endodermal progenitors prior to their involution and initiation of *sox17* expression. Cas is a potent inducer of *sox17* expression in wild-type embryos, as well as in *bon*, *fau/gata5*, and *MZoep* mutants. In addition, ectopic expression of *cas* in mesodermal cells leads to their transfating into endodermal cells. These data illustrate the unique potency of Cas as a transcriptional regulator of endoderm formation in zebrafish.

Results

Isolation of cas

We initially mapped cas^{ta56} within 2 cM of the LG7 centromere using half-tetrad analysis (Johnson et al. 1995): 94 wild-type and 88 mutant embryos were generated after blocking the second meiotic division of gynogenotes. Fine mapping revealed that *cas* is tightly linked to the simple sequence repeat markers Z7958 (no recombination in 5405 meiotic events) and Z9869 (no recombination in 550 meiotic events). We used these markers to begin to assemble a contiguous stretch of genomic DNA with yeast artificial chromosome and P1-derived artificial chromosome clones. However, it became clear that this region was underrepresented in the existing large-insert genomic libraries (J. Alexander, S. Waldron, and D.Y.R. Stainier, unpubl.).

In the course of a large-scale, in situ hybridization screen for zebrafish genes expressed in restricted patterns during embryogenesis (C. Thisse and B. Thisse, unpubl.), we isolated *CG569*, a *Sox*-related gene expressed in the YSL and endodermal cells during gastrulation. In light of this expression pattern, we placed *CG569* on the Goodfellow radiation hybrid (RH) panel (Kwok et al. 1999), and found that it mapped in the *cas* region (data not shown), suggesting that *CG569* could correspond to *cas*. The *CG569* genomic locus was characterized and sequenced in the wild-type and cas^{ta56} alleles. In cas^{ta56} we found a T-to-G substitution that introduces a premature translational stop codon at position 170 and leads to a truncation of the protein shortly after the HMG domain (Fig. 1A). During a mosaic F1 screen for cardiac mutations (Alexander et al. 1998), we identified a second mutant allele of *cas*, designated cas^{s4}. The phenotype of cas^{s4} is indistinguishable from that of cas^{ta56}, except that cas^{s4} mutants exhibit some endodermal *foxA2* expression that is not maintained during gastrulation (data not shown). In cas^{s4} we found an A-to-G substitution that introduces an arginine in place of a highly conserved histidine at codon 130 within the HMG domain (Fig. 1A). An HMG domain has a "twisted letter L" or "boomerang" shape, and this histidine (His-65 within the HMG domain) is believed to contribute to the stability of the long arm of the L, which is

important for DNA binding and bending (Werner et al. 1995).

Using restriction fragment length polymorphisms to analyze linkage, we found that the *ta56* and *s4* mutant alleles of *CG569* segregated with the *cas* phenotype in 240 and 200 meiotic events, respectively (data not shown). Together with the tight genetic linkage between *cas* and Z7958, and the significant physical linkage between Z7958 and *CG569* on the RH map, these data indicate that *CG569* is very closely linked to *cas*. To test further whether the *cas* phenotype is caused by the identified mutations in *CG569*, we injected wild-type *CG569* mRNA (100 pg) into embryos derived from cas^{ta56} heterozygote intercrosses, and genotyped the injected embryos following wholemount in situ hybridization with *sox17*. We found that this mRNA could restore *sox17* expression in cas^{ta56} mutants (Fig. 1B), further suggesting that *CG569* encodes Cas. We also confirmed that the *s4* mutation inactivates Cas function by showing that cas^{s4} mRNA injected into wild-type embryos had no effect on *sox17* expression (data not shown). Considered together, the linkage data, identification of severe molecular lesions, and rescue experiments indicate that *CG569* is *cas*.

Interestingly, we found that *cas* overexpression led to confluent patches of *sox17* expression in both wild-type and *cas* mutant embryos (Figs. 1B, 4B, see below), in contrast to the regularly spaced distribution of *sox17*-expressing endodermal cells observed in uninjected wild-type embryos (Fig. 4A, see below; Alexander and Stainier 1999). These data suggest that *cas* is able to activate *sox17* expression in nonendodermal cells, and possibly direct cells not fated to become endoderm into the endodermal lineage. This hypothesis will be directly tested and the data shown in Figure 6 (below).

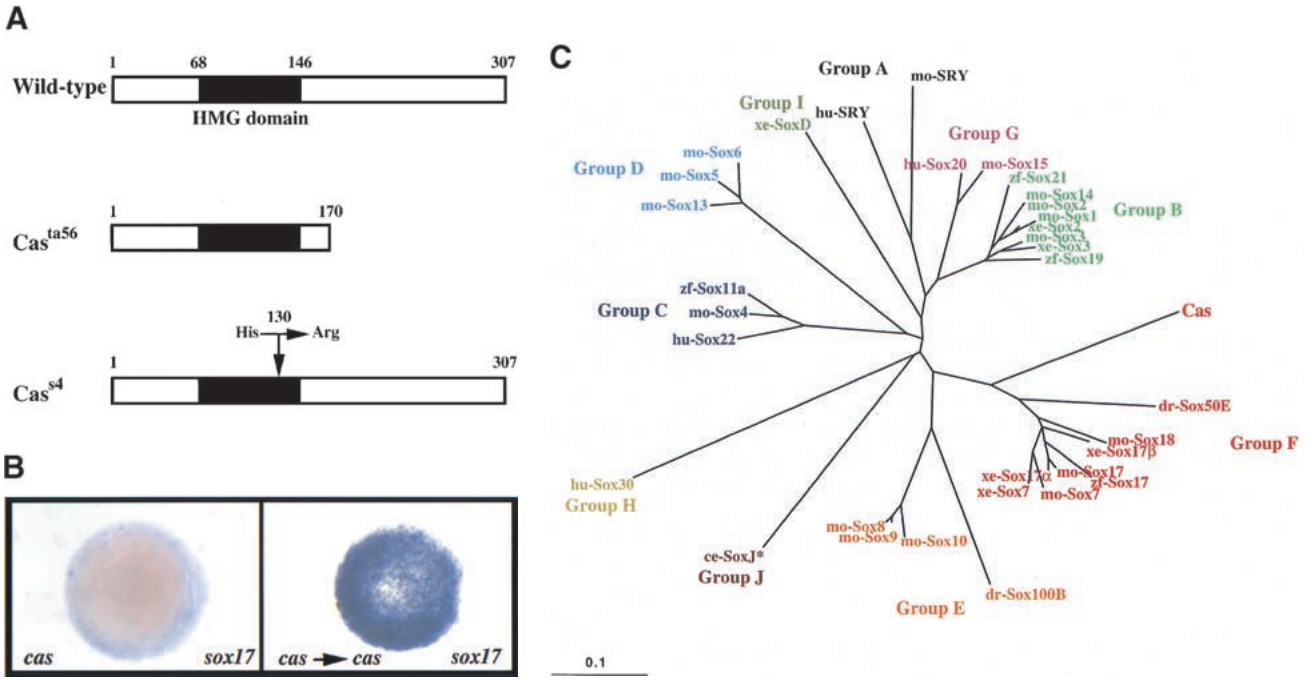
Sequence and clustering analyses indicate that Cas encodes a novel Sox protein of the SoxF group, which includes Sox7, Sox17, and Sox18 (Fig. 1C). Within the HMG domain, Cas shows less sequence identity (<71%) with any of the existing SoxF members than the other SoxF members do with each other (>83%) (Fig. 1D,E). (Contributing to this divergence are substitutions in the RPMNAFMVW signature motif common to the HMG domains of all non-SRY Sox proteins thus far identified [Bowles et al. 2000].) However, Cas shows a higher degree of sequence identity to the SoxF group than to the neighboring SoxE group; for example, Cas and mouse Sox10, a group E member, show only 54% sequence identity in their HMG domains. Thus, considering all currently available data, Cas appears to be a novel member of the SoxF group.

Regulation of cas *expression*

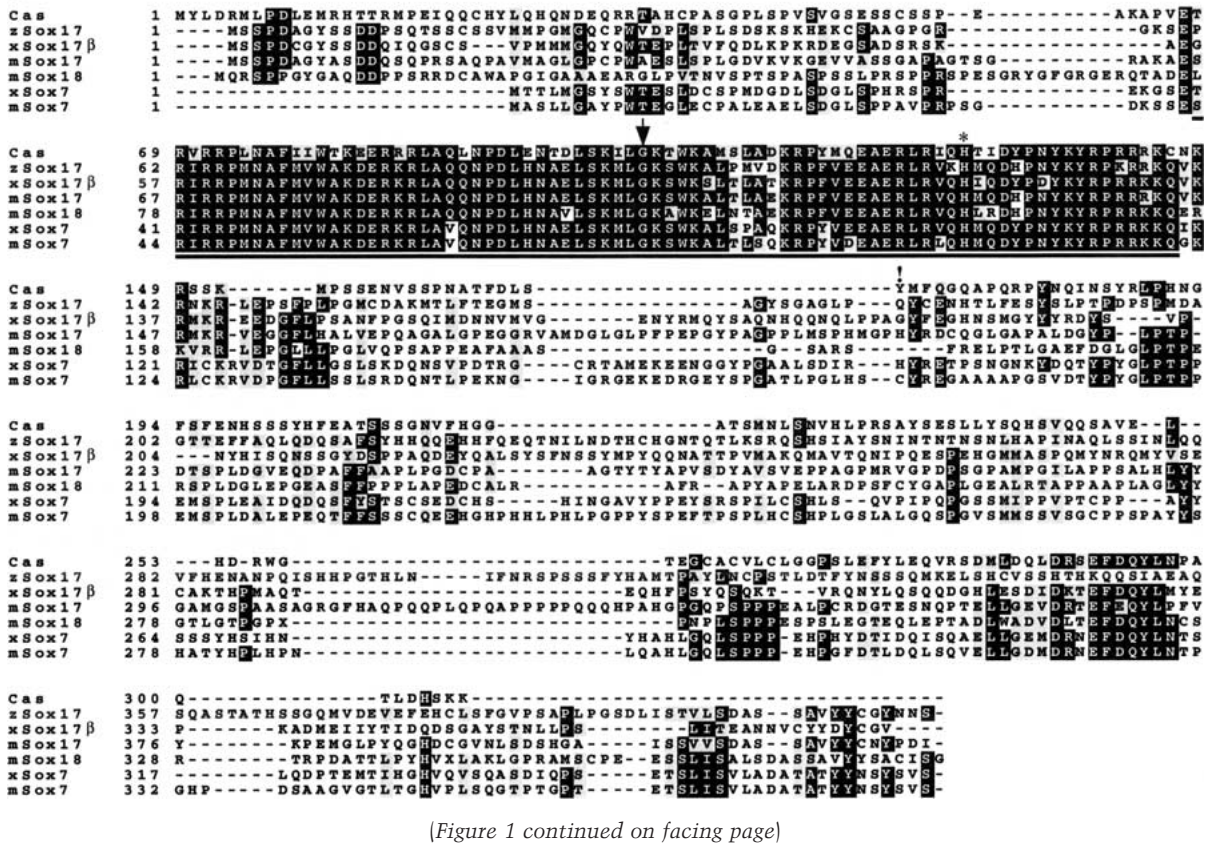
cas expression is first observed at high stage (3.3 h postfertilization [hpf]) in the dorsal YSL (Fig. 2A, arrowhead, B). Embryonic *cas* expression initiates <1 h later in the dorsal marginal blastomeres (Fig. 2C,D), by which time *cas* is expressed at low levels throughout the YSL. The embryonic expression extends progressively around the margin, and by 40% epiboly (5 hpf) *cas* is clearly expressed all around the margin in both the YSL and marginal blastodermal cells (Fig. 2E). A high magnification view clearly shows *cas* expression in both the YSL, distinguished by the large nuclei, and the presumptive endodermal progenitors at the margin (Fig. 2F), whereas adjacent marginal cells, presumably mesodermal progenitors, do not express *cas*. Embryonic *cas* expression precedes *sox17* expression and can be clearly seen during gastrulation, both in the endodermal cells as well as in the forerunner cells (Fig. 2G, fo; I–L), a specialized subset of dorsal cells that also express *sox17* (Alexander and Stainier 1999) and *sqt* (Fig. 2L; Erter et al. 1998; Feldman et al. 1998; Rebagliati et al. 1998a), and later form the lining of Kupffer's vesicle (Cooper and D'Amico 1996). By 90% epiboly, *cas* expression is evident throughout the endoderm (Fig. 2G,H); posterior endodermal cells that express *sox17* but not *foxA2* (Alexander and Stainier 1999) also express *cas*. At this stage *cas* expression is still present in the YSL, but it disappears shortly afterwards from the YSL as well as from embryonic tissues (data not shown).

Next, we examined whether *cas* expression is regulated by Nodal signaling. Embryos injected with 100 pg of mRNA encoding a constitutively active form of the type I TGFβ receptor Taram-a, designated Taram-a*, showed a massive induction of *cas* (Fig. 2N,O), indicating that Nodal signaling activates *cas* expression. Conversely, injection of 25 pg of mRNA encoding Antivin, a competitive inhibitor of the Activin/Nodal signaling pathway related to mammalian Lefty (Thisse and Thisse 1999; Thisse et al. 2000), led to a complete loss of *cas* expression in marginal blastomeres at late blastula stage, whereas *cas* expression in the YSL appeared to be unaffected (Fig. 2P). Examining *antivin*-injected embryos at 90% epiboly confirmed the loss of *cas* expression in the endodermal and forerunner cells, but not the YSL (Fig. 2Q). Similarly, *MZoep* mutants exhibit no *cas* expression in the embryo, but seemingly normal *cas* expression in the YSL (data not shown). These data demonstrate that the embryonic expression of *cas*, like that of *bon* and *fau/gata5*, is regulated by Nodal signaling, whereas the YSL expression of *cas* is independent of Nodal signaling.

We then used embryonic *cas* expression as a marker of endoderm to examine endoderm formation in *cas*, *bon*, and *fau/gata5* mutants at shield (6 hpf) and 80% epiboly (8.3 hpf) stages. In *cas* mutants, no embryonic *cas* expression is observed, whereas *cas* expression appears to be up-regulated in the YSL (Fig. 3A,D). In *bon* and *fau/gata5* mutants, embryonic *cas* expression behaves essentially like *sox17* expression. *bon* mutants exhibit a profound reduction in the number of *cas*-expressing endodermal cells but normal *cas* expression in the forerunners and YSL (Fig. 3B,E). *fau/gata5* mutants show a smaller, though noticeable, decrease in the number of *cas*-expressing endodermal cells, as well as a decrease in the intensity of *cas* expression in the endodermal cells, whereas *cas* expression appears to be normal in the forerunners and YSL (Fig. 3C,F). These data further support



D



(*Figure 1 continued on facing page*)

the hypothesis that no endoderm forms in *cas* mutants, and also further illustrate the close correspondence between *cas* and *sox17* expression in the endoderm and forerunner cells (Alexander and Stainier 1999; Kikuchi et al. 2000; Reiter et al. 2001).

cas *can induce* cyc *and* sqt *expression*

cas overexpression in wild-type embryos induces widespread *sox17* expression (Fig. 4A,B), and inhibits epiboly and involution (data not shown). These phenotypes are

E

	Identity of HMG domain (%)	
	Cas	**zf-Sox17**
zf-Sox17	**65**	——
xe-Sox17β	**66**	**85**
mo-Sox17	**68**	**91**
mo-Sox18	**65**	**84**
xe-Sox7	**70**	**86**
mo-Sox7	**67**	**84**
mo-Sox10	**54**	**66**
mo-LEF1	**29**	**30**

Figure 1. *cas* encodes a novel member of the Sox family of high-mobility group (HMG) domain transcription factors. (*A*) Schematic representation of wild-type and mutant Cas proteins. The cas^{ta56} mutation leads to a truncation of the protein shortly after the HMG domain, and the cas^{s4} mutation introduces a single amino acid substitution (histidine to arginine) in the third helix of the HMG domain (position 130). (*B*) *sox17* expression in uninjected (left) and *CG569/cas*-injected cas^{ta56} mutant embryos. Animal pole views at shield stage (6 h postfertilization) with dorsal to the right (left embryo). Whereas *sox17* expression is missing in uninjected cas^{ta56} mutants (left), overexpression of *CG569/cas* mRNA always induced *sox17* expression (right). For comparison to wild-type pattern of *sox17* expression, see Fig. 4*A*. (*C*) Unrooted phylogenetic tree for the Sox HMG domain, constructed by the neighbor joining (NJ) method (Saitou and Nei 1987). Branch lengths are representative of the extent of divergence. Ten groups (Groups A–J) are classified by phylogenetic analysis of Sox protein HMG domain sequences (Bowles et al. 2000). The ce-SoxJ* protein (Group J) was identified from databases and previously reported as a Sox protein (Bowles et al. 2000). This phylogenetic analysis indicates that Cas is a divergent member of the SoxF group. (*D*) Sequence comparison of Cas; zebrafish, *Xenopus* and mouse Sox17; mouse Sox18; and *Xenopus* and mouse Sox7. The HMG domain is underlined. Cas shows a relatively high degree of sequence identity in its HMG domain with the members of the SoxF group, but otherwise shows limited sequence conservation. The arrow indicates intron position in *cas*, zebrafish *sox17*, and mouse *Sox17* and *Sox18*, which appears to be the same in all members of the SoxF group but different in most other *Sox* genes (Bowles et al. 2000). The asterisk and exclamation point indicate the position of the cas^{s4} and cas^{ta56} mutations, respectively. (*E*) Percent sequence identity of the Cas and zebrafish Sox17 HMG domains with the HMG domains of the other members of the SoxF group as well as with the HMG domains of mouse Sox10 (SoxE group) and mouse LEF1. Cas shows substantially less sequence identity with the other SoxF group members than does zebrafish Sox17. However, Cas shows even less identity with mouse Sox10 (54%) and mouse LEF1 (29%). Abbreviations: ce, *Caenorhabditis elegans*; dr, *Drosophila melanogaster*; hu, *Homo sapiens*; mo, mouse; xe, *Xenopus laevis*; zf, *Danio rerio*.

very similar to those seen following *taram-a** overexpression (Alexander and Stainier 1999). Although embryonic *cas* expression and function are clearly downstream of Nodal signaling (Fig. 2N–Q; Alexander and Stainier 1999), it seemed possible that *cas* overexpression could act by inducing the expression of *cyc* and/or *sqt*. To investigate this possibility, we overexpressed *cas* (100 pg) in wild-type embryos at the 1–4-cell stage and examined *cyc* and *sqt* expression. We observed a significant increase in *cyc* and *sqt* expression (Fig. 4C–F), indicating that *cas* can activate their expression. However, except for the lack of *sqt* expression in the forerunners, which are missing in *cas* mutants (Alexander et al. 1999), we did not observe overt defects in *cyc* or *sqt* expression in *cas* mutants by wholemount in situ hybridization (data not shown).

Given that *cas* can activate *cyc* and *sqt* expression, it seemed that *cas* could induce *sox17* expression directly or via Nodal induction. To distinguish between these possibilities, we overexpressed *cas* in *MZoep* mutants, which cannot respond to Nodal signaling (Gritsman et al. 1999). We found that *cas* could induce *sox17* expression in *MZoep* mutants (Fig. 4G,H), extending our previous data that Cas functions downstream of Nodal signaling (Alexander and Stainier 1999) by showing that it is sufficient to induce *sox17* expression in the absence of Nodal signaling. In addition, we found that *cas* overexpression in *MZoep* mutants did not induce *cyc* or *sqt* expression (data not shown), indicating that this aspect of Cas function is dependent on Nodal signaling. Interestingly, *cas*-overexpressing *MZoep* mutants exhibited only faint *foxA2* expression and no *bon*, *fau/gata5*, *gsc*, or *ntl* expression (data not shown), suggesting that Cas cannot activate the expression of either *bon* or *gata5*, nor does it induce the expression of mesodermal markers. Rather, Cas can regulate *sox17* expression in a Nodal-independent manner, illustrating its potency as an inducer of endodermal gene expression.

Relationship of Cas, Bon and Fau/Gata5

fau/gata5, *bon*, and *cas* are all induced by Nodal signaling, and each regulates *sox17* expression (Alexander and Stainier 1999; Rodaway et al. 1999; Kikuchi et al. 2000; Reiter et al. 2001). Our previous work had shown that *bon* or *fau/gata5* overexpression in *cas* mutants failed to restore *sox17* expression (Alexander and Stainier 1999; Reiter et al. 2001), indicating that *cas* function is required downstream of *bon* and *fau/gata5*. To test this model further, we injected *cas* into *bon* and *fau/gata5* mutants and examined *sox17* expression. *cas* was able to expand *sox17* expression in *bon* and *fau/gata5* mutants to an extent comparable to that seen in wild-type embryos (Fig. 5A–C), confirming that *cas* is able to function in the absence of *bon* or *fau/gata5*.

Although Cas function clearly acts downstream of Bon and Fau/Gata5, the precise relationship between these proteins remained unclear. It could be that Nodal signaling induces *bon* and *fau/gata5*, which then activate *cas*, which directs endoderm formation. Alternatively, *bon*, *fau/gata5*, and *cas* could each be induced independently by Nodal signaling, and each could direct endoderm formation, with *cas* being uniquely required for *sox17* ex-

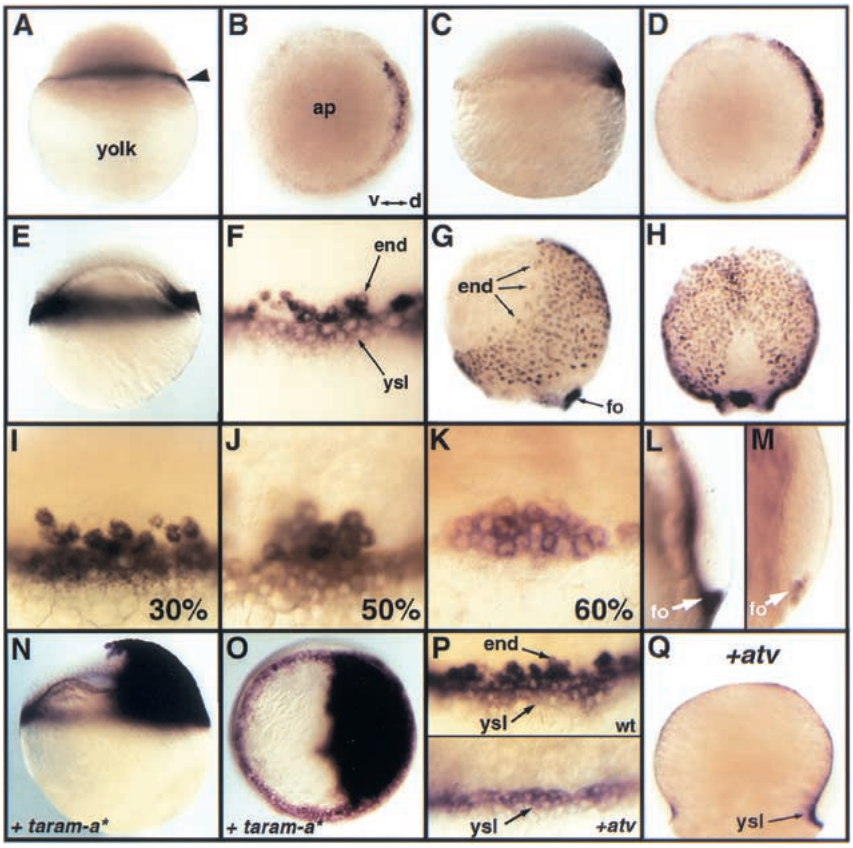


Figure 2. *cas* expression in zebrafish embryos at high (3.3 h postfertilization [hpf]; *A,B*), sphere (4 hpf; *C,D*), 30% epiboly (4.7 hpf; *I*), 40% epiboly (5 hpf; *E,F,N,O,P*), 50% epiboly (5.3 hpf; *J*), 60% epiboly (7 hpf; *K,L*), and 90% epiboly (9 hpf; *G,H,Q*) stages. *sqt* expression at 60% epiboly (7 hpf; *M*). *A, C, E, F, G, L, M, N, P,* and *Q* show lateral views with dorsal to the right. *B, D,* and *O* show animal pole views. *H, I, J,* and *K* show dorsal views with anterior to the top. (*A,B*) *cas* expression is first observed at the high stage in the dorsal YSL. (*C,D*) By sphere stage *cas* expression is detected throughout the marginal (yolk syncytial layer) YSL and in dorsal blastomeres. (*E,F*) At 40% epiboly *cas* is expressed in a subset of blastomeres all around the margin (see high magnification in *F*), and is maintained in the marginal YSL. (*G,H*) As endodermal cells gastrulate, they continue to express *cas*, and a group of noninvoluting cells, the forerunner cells (fo), also begins to express *cas*. At 90% epiboly, expression is seen in all endodermal and forerunner cells and resembles *sox17* expression. (*I*) *cas* expression in presumptive forerunner cells. (*J,K*) By 60% epiboly, forerunners, which continue to express *cas*, form a compact group of cells that do not involute and stay at the vegetal tip of the axial mesoderm throughout gastrulation. (*L,M*) Forerunners show strong *cas* expression (*L*) as well as *sqt* (*M*) expression throughout gastrulation. (*N,O*) Injection of 100 pg of *taram-a** mRNA into one cell at the two-cell stage led to a high level of *cas* expression in the injected half of the embryo. (*P,Q*) Injection of a low dose of *antivin* (*atv*) mRNA (25 pg) at the one-cell stage led to the disappearance of embryonic *cas* expression, whereas the YSL expression appeared unaffected. High magnification views at 40% epiboly are shown in *P*. The optical cross section in *Q* shows *cas* expression present in the YSL, but absent in the endoderm of *atv* injected embryos at 90% epiboly. ap, animal pole; d, dorsal; end, endoderm; fo, forerunner cells; v, ventral; wt, wild type; ysl, yolk syncytial layer.

pression. To distinguish between these models, we tested whether *bon* or *fau/gata5* could induce *cas* expression in cells that do not normally express *cas*. Wild-type embryos were injected with *bon*, *fau/gata5*, or both at the 1–4-cell stage and examined for *cas* expression at 40% epiboly. Although endodermal *cas* expression at the margin was somewhat increased, presumably reflecting increased endoderm formation, no *cas* expression was observed outside the margin (Fig. 5D–F; data not shown), showing that *bon* or *fau/gata5*, alone or in combination, is not sufficient to activate *cas* expression in nonmarginal cells.

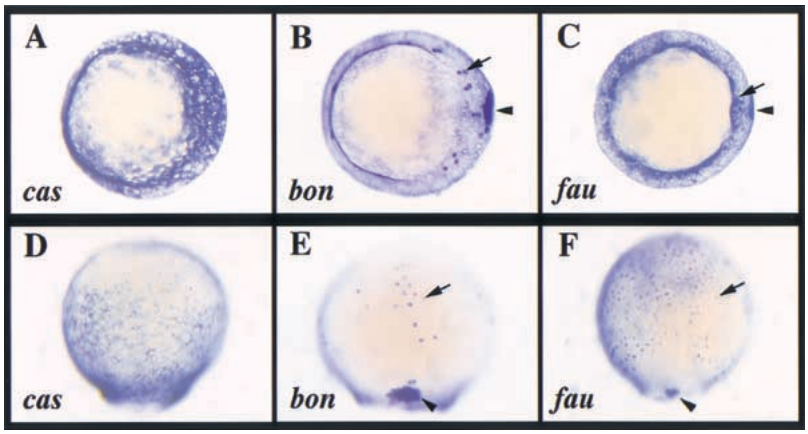


Figure 3. *cas* expression in *cas, bon,* and *fau/gata5* mutant embryos. Animal pole views with dorsal to the right at shield stage (6 h postfertilization [hpf]). (*A–C*), and dorsal views with anterior to the top at 80% epiboly (8.3 hpf; *D–F*). (*A,D*) In *cas* mutants, no *cas* expression is observed in the embryo proper, whereas *cas* appears to be up-regulated in the yolk syncytial layer (YSL). (*B,E*) In *bon* mutants, *cas* expression is observed in a few endodermal cells (arrows), the forerunner cells (arrowheads), and the YSL. (*C,F*) In *fau/gata5* mutants, *cas* expression in the endodermal cells appears weaker (arrows), whereas *cas* expression in the forerunner cells (arrowheads) and the YSL appears normal.

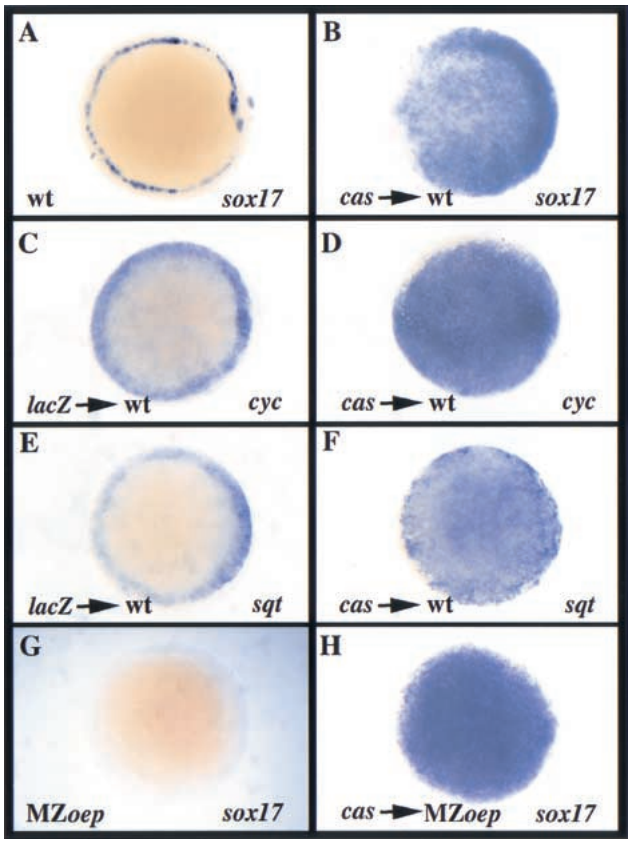


Figure 4. *cas* overexpression induces *cyc* and *sqt* expression in wild-type embryos, and induces *sox17* expression in *MZoep* mutants. (*A,B,G,H*) Animal pole views at shield stage (6 h postfertilization [hpf]) with dorsal to the right (*A*). (*C–F*) Animal pole views at 50% epiboly (5.3 hpf) with dorsal to the right (*C,E*). (*A,B*) A characteristic salt-and-pepper distribution of *sox17*-expressing endodermal cells is observed throughout the margin in uninjected wild-type embryos (*A*), whereas injection of *cas* mRNA at the 1–4-cell stage led to confluent patches of *sox17* expression throughout the embryo (*B*). (*C–F*) Injection of *cas* mRNA at the 1–4-cell stage induced ectopic *cyc* and *sqt* expression in wild-type embryos (*D,F*). (*G,H*) Whereas no *sox17* expression is observed in uninjected *MZoep* mutants (*G*), injection of *cas* mRNA at the 1–4-cell stage induced strong *sox17* expression (*H*). wt, wild type.-

Ectopic cas *expression leads to the transfating of mesodermal cells into endoderm*

cas overexpression leads to widespread *sox17* expression, suggesting that *cas* is able to activate *sox17* expression in nonendodermal cells and possibly direct them to become endoderm. To test this possibility, we ectopically expressed *cas* and *GFP* (100 pg each) in single marginal blastomeres at the 16-cell stage and determined the position of the injected cells at shield stage (6 hpf). For example, Figure 6A shows a clone resulting from a dorsal injection. First we examined *sox17* expression in the injected embryos at shield stage and found that, in all cases, strong *sox17* expression was induced around the labeled clone (Fig. 6B). Next, we examined the fate of dorsal, dorsolateral, lateral, and ventral clones. Examination at 24 hpf showed that the descendants of injected dorsal cells, which would normally populate prechordal plate and notochord, had colonized the entire extent of the gut tube (Fig. 6C). The conversion of dorsal axial mesoderm into endoderm led this embryo (shown in Fig. 6C) to be cyclopic and exhibit U-shaped somites (Fig. 6D,E). Of 106 embryos injected dorsally, cyclopia was seen in 44 (41.5%) and U-shaped somites in 68 (64.1%). We further characterized the loss of anterior axial mesoderm in dorsally injected embryos by examining the expression of *gsc*, a prechordal plate marker, and *hgg1*, a hatching gland marker (Thisse et al. 1994), at 50% and 80% epiboly, respectively: of 74 embryos injected dorsally, we found a complete absence of *gsc* expression in 4 (5.4%), and a strong reduction, such as that seen in Figure 6G, in 54 (73%); of 80 embryos injected dorsally, we found a complete absence of *hgg1* in 16 (20%), and a strong reduction, such as seen in Figure 6I, in 56 (70%). To examine the axial mesoderm in dorsally injected embryos, we analyzed *shh* expression at 24 hpf, which, at this stage, is present in the notochord, floor plate, and part of the diencephalon (Krauss et al. 1993). Of 76 embryos injected dorsally, we found a complete absence of *shh* expression in 4 (5.2%), and patchy expression, such as that seen in Figure 6K, in 68 (89.5%).

Embryos injected dorsolaterally with *cas* and *GFP* also exhibited a decrease in notochordal tissue, as evidenced by loss of *shh* expression and kinking of the embryos (data not shown). Laterally injected embryos showed a unilateral loss of somitic muscle with a concomitant enlargement of *foxA2* expression on the *cas*-injected side (data not shown). Embryos injected ventrally with *cas* and *GFP* exhibited an accumulation of GFP-expressing cells in the tail, which displayed an abnormal shape characterized by a large, empty ventral blister, instead of forming ventral mesenchyme (Fig. 6L). To analyze the fate of the *cas*-expressing ventral cells, which would normally give rise to blood, we examined the expression of the *carbonic anhydrase* gene (C. Thisse and B. Thisse, unpubl.), which is specifically expressed in blood cells and the otic vesicle (Fig. 6M). Of 81 embryos injected ventrally, we observed a total absence of *carbonic anhydrase* expression in blood tissue in 18 (22%), and a partial reduction, such as that seen in Figure 6N, in 48 (59.3%), whereas the otic vesicle expression was unaffected.

Figure 6O–R show our interpretation of these data by representing the fate maps of a noninjected embryo (Fig. 6O), dorsally injected embryos (a smaller clone in Fig. 6P and a larger clone in Fig. 6Q), and a ventrally injected embryo (Fig. 6R). *cas* misexpression in a clone of cells located centrally within the embryonic shield at the onset of gastrulation resulted in the transfating of the prechordal plate progenitors into endoderm, as revealed partly by the loss of *gsc* and *hgg1* expression (Fig. 6G,I,P). *cas* misexpression in a larger dorsal clone affected both anterior and posterior axial mesoderm, resulting in embryos that lack both prechordal and notochordal tissues (Fig. 6K,Q). *cas* misexpression in ventral clones resulted in embryos that lacked nearly all blood cells, but whose

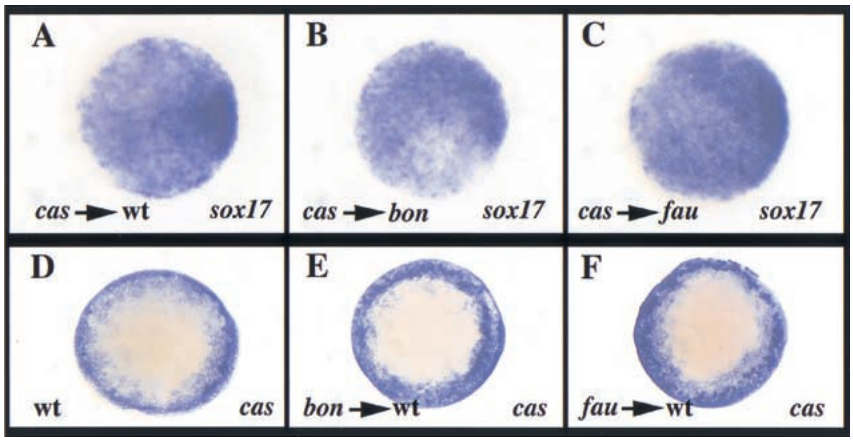


Figure 5. *cas* overexpression induces ectopic *sox17* expression in *bon* and *fau/gata5* mutant embryos, whereas *bon* and *fau/gata5* cannot induce ectopic *cas* expression in wild-type embryos. Animal pole views at 6 h postfertilization (hpf; *A–C*) and 40% epiboly stage (5 hpf; *D–F*). (*A–C*) Injection of *cas* mRNA at the 1–4-cell stage expanded *sox17* expression to a comparable extent in wild-type, *bon*, and *fau/gata5* mutant embryos. (*D–F*) Injection of *bon* or *fau/gata5* mRNA in wild-type embryos was not able to induce ectopic *cas* expression away from the margin.

other mesodermal derivatives appeared unaffected (Fig. 6N,R).

cas misexpression at the animal pole at the 256-cell stage (Fig. 6S,T), a territory normally fated to become epidermis, telencephalon, and eye, led to ectopic *sox17* expression (Fig. 6W), but did not affect the morphology of the injected embryos (Fig. 6U). Injections performed at earlier developmental stages (32–128-cell stage) gave essentially the same results. However, due to the extensive spreading and mixing of cells derived from these earlier stage blastomeres, some of them reached the margin during the blastula stage and became incorporated in the endoderm. Localized injections in single animal pole blastomeres at the 256-cell stage gave rise to smaller clones of cells that remained in the epiblast; at 24 hpf, labeled cells appeared in various ectodermal locations, sometimes within differentiated structures such as the neural tube, and often on top of the yolk (Fig. 6V). Nevertheless, disappearance of ectodermal derivatives, indicative of transfating, was never observed. One possible explanation for these observations is the finding that *cas* misexpression in the animal pole does not induce ectopic *sqt* expression in the embryo proper, but only in a few cells of the enveloping layer (evl; Fig. 6X), whereas *cas* misexpression in the margin induces strong *sqt* expression in embryonic cells (Fig. 6Y). In conditions in which Nodal activity is inhibited by Antivin, injection of *cas* in a marginal blastomere at the 16-cell stage resulted in the induction of *sqt* only in evl cells (Fig. 6Z), demonstrating that the ability of *cas* to induce *sqt* expression in the embryo proper is dependent on Nodal signaling. The finding that *sqt* is not induced in the embryo proper by *cas* misexpression at the animal pole may similarly be explained by the fact that this region has previously been shown to be a ground state for the Activin/Nodal signaling pathway (Thisse et al. 2000).

Discussion

cas *encodes a novel member of the SoxF group of HMG domain transcription factors*

We have provided genetic and molecular evidence that *cas* encodes a novel member of the SoxF group of transcription factors. Sox17, whose *Xenopus* orthologs can activate endodermal gene expression in animal caps (Hudson et al. 1997), also belongs to this group, but several points argue against Cas and XSox17 being equivalent. First, sequence analyses suggest that *cas* is a novel SoxF member and not a recently duplicated version of zebrafish *sox17*. Second, *cas* but not *Xsox17* can restore *sox17* expression in *cas* mutants (J. Alexander and D.Y.R. Stainier, unpubl.), indicating a clear functional difference between them.

In both zebrafish and *Xenopus*, Gata5 and a Mix-type protein (Bon in zebrafish and Mixer in *Xenopus*) have been shown to regulate *sox17* expression (Henry and Melton 1998; Alexander and Stainier 1999; Kikuchi et al. 2000; Reiter et al. 2001; Xanthos et al. 2001). And based on the temporal characteristics of the *Xsox17* expression pattern, an additional upstream regulator probably initiates its expression in *Xenopus* (Xanthos et al. 2001). It is thus possible that, much like Cas regulates *sox17* expression in zebrafish, another member of the SoxF group regulates *Xsox17* expression in *Xenopus*.

The conservation seen in the pathways responsible for endoderm formation in zebrafish and *Xenopus* suggest that similar factors may be involved in this process in mammals. Although mouse *Sox18* does not appear to be expressed in the developing endoderm (Pennisi et al. 2000), it will be interesting to investigate the early expression of *Sox7* and *Sox17* in the mouse embryo, as well as any possible roles they may play in endoderm formation.

cas *may regulate* sox17 *directly*

cas is initially expressed in the YSL and later appears in marginal blastomeres. The embryonic expression of *cas* is clearly regulated by Nodal signaling, as shown by its absence in *antivin*-injected embryos and *MZoep* mutants. Unlike previously identified genes, *cas* expression in late blastula-stage embryos may be restricted to the endodermal progenitors; *cas* is expressed in a subset of *fau/gata5*-expressing cells, appearing more spottily distributed throughout the margin (Fig. 2F,P). *sox17* expression appears to be activated in all *cas*-expressing cells,

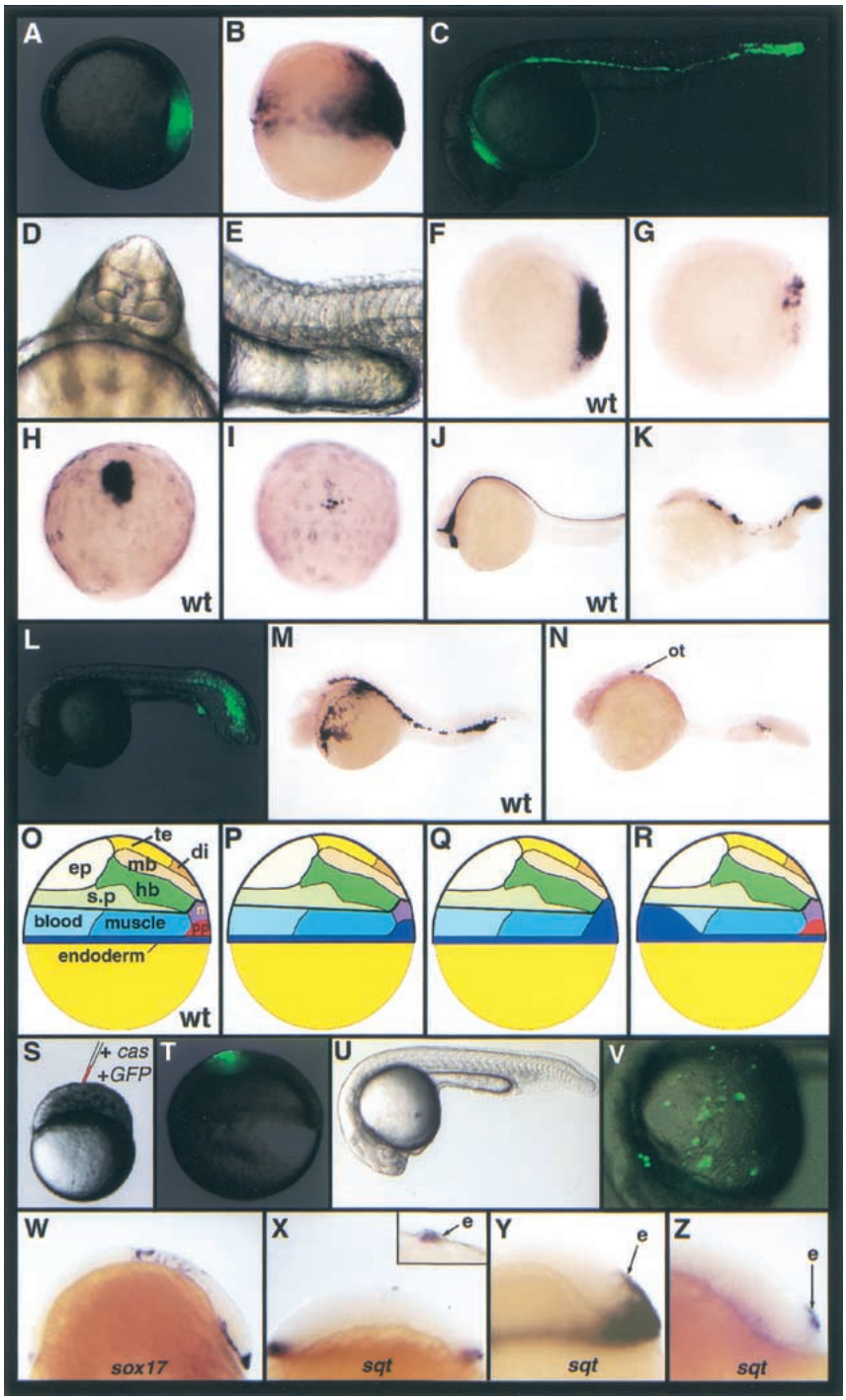


Figure 6. Misexpression of *cas* results in the transfating of presumptive mesodermal cells into endoderm. GFP fluorescence at shield stage (6 hours postfertilization [hpf]; *A*; animal pole view; *T*; lateral view), and at 24 hpf (*C,L,V*; lateral views). Lateral view with dorsal to the right of *sox17* expression at shield stage (*B,W*). *gsc* expression at 50% epiboly (5.3 hpf; *F,G*; animal pole views, dorsal to the right). *hgg1* expression at 80% epiboly (8.3 hpf; *H,I*; dorsal views, anterior to the top). *shh* expression at 24 hpf (*J,K*; lateral views). *carbonic anhydrase* expression at 24 hpf (*M,N*; lateral views). Fate map, as adapted from Thisse et al. (2000) of uninjected (*O*) and *cas*-injected embryos (*P–R*) at the onset of gastrulation (5.3 hpf). *sqt* expression at 30% epiboly (4.7 hpf; *X*) and 40% epiboly (5 hpf; *Y,Z*); lateral views, dorsal to the right. (*A*) Coinjection of 100 pg of *cas* mRNA and 100 pg of *GFP* mRNA into a marginal blastomere at the 16-cell stage gives rise to marginal clones of cells which were screened at the onset of gastrulation to determine their localization along the dorso–ventral axis. A dorsal clone is shown here. (*B*) Embryo with a labeled dorsal clone was examined for *sox17* expression, which was strongly induced dorsally. (*C*) Embryo shown in *A* at 24 hpf. GFP-expressing cells accumulate in the entire endoderm, from the level of the pharyngeal arches anteriorly to the anal opening in the posterior trunk region. More caudally, GFP-expressing cells are observed underlying the notochord in the anterior part of the tail, and in the notochord and tail bud more posteriorly. (*D*) High-magnification view of the head of the embryo shown in *C*, revealing a cyclopic phenotype. (*E*) The same embryo exhibits U-shaped somites, indicative of a reduction in axial mesoderm. (*F,G*) *gsc* expression in a dorsally injected embryo (*G*) is severely reduced as compared to wild type (*F*). (*H,I*) Expression of *hgg1*, a prechordal plate marker, in a dorsally-injected embryo (*I*) is nearly abolished as compared to wild type (*H*). (*J,K*) *shh* expression in a dorsally-injected embryos (*K*) is severely reduced as compared to wild type (*J*). (*L*) Embryo with a labeled ventral clone at 24 hpf. GFP-expressing cells accumulate in the tail, which displays an abnormal shape characterized by a large, empty ventral blister instead of ventral mesenchyme. (*M,N*) The blood expression of *carbonic anhydrase* has almost completely disappeared in a ventrally-injected embryo, whereas the otic vesicle (ot) expression is unaffected. (*O*) Fate map of a wild-type embryo at the onset of gastrulation (di, diencephalon; ep, epidermis; hb, hindbrain; mb, midbrain; n, notochord; pp, prechordal plate; sp, spinal chord; te, telencephalon). (*P*) Misexpression of *cas* in a clone of cells located in the central part of the embryonic shield results in the transfating of the prechordal plate into endoderm, resulting in the loss of *gsc* (*G*) and *hgg1* (*I*) expression. (*Q*) A larger dorsal clone affects both anterior and posterior axial mesoderm, resulting in embryos that lack both prechordal and notochordal mesoderm, as shown in *K*. (*R*) Ventral clones result in embryos that lack almost all blood cells but appear unaffected in the other mesodermal derivatives, as shown in *N*. (*S,T*) An embryo injected with *cas* and *GFP* mRNA in an animal pole blastomere at the 256-cell stage (*S*) is shown at shield stage in *T*. (*U,V*) At 24 hpf, embryos injected at the animal pole have a wild-type morphology (*U*) and the GFP-expressing cells do not populate the endoderm but are often found on top of the yolk (*V*). (*W,X*) Injection of *cas* at the animal pole induces *sox17* (*W*) but not *sqt* (*X*) expression in the embryo proper. In such embryos, *sqt* expression is only induced in a few cells of the enveloping layer (evl cells; e), an extraembryonic structure. (*Y,Z*) In contrast, marginal *cas* injection induces strong *sqt* expression in embryonic cells around the margin (*Y*) and this induction is blocked by Atv overexpression (*Z*; this embryo was injected with 25 pg of *atv* at the two-cell stage and with *cas* at the 16-cell stage. In such embryos, *sqt* expression is only induced in a few evl cells). wt, wild type.-

except the YSL, starting ~80 min after the onset of *cas* expression. *cas* overexpression can induce *sox17* expression in *MZoep* mutants as well as in *antivin*-injected embryos (C. Thisse and B. Thisse, unpubl.), indicating that *cas* can regulate *sox17* expression independently of Nodal signaling. The regulation of endodermal *sox17* expression by Cas is cell autonomous as shown by cell transplantation experiments (Alexander et al. 1999), as well as by the finding that *cas* overexpression in the YSL, in contrast to *cas* overexpression throughout the embryo, does not up-regulate *sox17* expression (Y. Kikuchi and D.Y.R. Stainier, unpubl.). In addition, HMG domain-binding sites are present upstream of *sox17* (Y. Kikuchi and D.Y.R. Stainier, unpubl.), suggesting that the regulation of *sox17* by Cas may be direct.

Potential roles for cas *in the YSL*

cas is first expressed in the dorsal YSL at high stage, which coincides with the initial expression of *sqt* (Feldman et al. 1998; Rebagliati et al. 1998a; Sampath et al. 1998). *cas* and *sqt* expression subsequently spread around the margin and by 30% epiboly *cas*, *cyc*, and *sqt* expression are evident all around the margin (Erter et al. 1998; Rebagliati et al. 1998b). This coexpression in the margin and the regulation of embryonic *cas* expression by Nodal signaling led us to further examine their relationship; although *cas* expression in the YSL is clearly independent of Nodal signaling, interestingly, overexpression of *cas* throughout the embryo led to widespread *cyc* and *sqt* expression. (These data are in contrast to the inability of *bon*, which is also initially expressed in the dorsal YSL [Alexander et al. 1999], to induce *cyc* and *sqt* expression [Kikuchi et al. 2000].) Apart from the forerunner cell phenotype, *cyc* and *sqt* expression appear normal in *cas* mutants, leaving the physiological role of *cas* in *cyc* and *sqt* expression unclear. It may be that defects in *cyc* and *sqt* expression in *cas* mutants are too subtle to detect by in situ, that other regulators of *cyc* and *sqt* expression are able to compensate for the lack of Cas function, or that Cas normally regulates *sqt* expression only in the forerunners.

It is interesting to note that *cas* mutants, like *cyc* mutants, exhibit left–right asymmetry defects (Liang et al. 2000); gene expression that is left-sided in the anterior lateral plate mesoderm (LPM) and diencephalon of wild-type embryos becomes randomized in *cas* and *cyc* mutants (Rebagliati et al. 1998b; Sampath et al. 1998; Liang et al. 2000). In thinking about the basis for the laterality defects in *cas* mutants, it is useful to consider *bon* mutants, in which, despite a nearly complete absence of endoderm, left–right asymmetry appears normal (J. Alexander and D.Y.R. Stainier, unpubl.). Another difference between *cas* and *bon* mutants is the formation of forerunner cells, which is defective in *cas* mutants (Alexander et al. 1999) but apparently normal in *bon* mutants (Kikuchi et al. 2000). Thus, the laterality defects observed in *cas* mutants are unlikely to be due to the absence of endoderm but, rather, a consequence of either the lack of forerunner cells or the lack of *cas* function in the YSL. Given the physical proximity of the YSL to the diencephalon during gastrulation (Woo and Fraser 1995) and to the anterior LPM during mid- to late-segmentation stages (Stainier et al. 1993), it is possible that *cas* function in the YSL influences left–right asymmetry. Alternatively, given the fact that both *cas* and *sqt* are expressed in the forerunner cells, it is possible that these cells are involved in setting up left–right asymmetry. The endocytic properties of forerunner cells (Cooper and D'Amico 1996) could in fact be a key part of this function and lead to differential signaling on the left and right sides of the embryo. It will be interesting to investigate further whether either or both of these tissues are involved in setting up left–right asymmetry as well as whether *cas* regulates left–right asymmetry by regulating *cyc* and/or *sqt* expression or through some other mechanism.

The finding that *cas* is expressed in the YSL also suggests an alternative hypothesis for the mesodermal defects observed in *cas* mutants. These defects include the failure of the bilateral myocardial populations to migrate and fuse at the midline (cardia bifida), as well as similar defects in the vasculature, blood, and pronephros (Alexander et al. 1999). We originally proposed that the absence of endoderm may underlie these defects (Alexander et al. 1999). It may be, however, that defects in the YSL, which is also closely apposed to several mesodermal tissues in mid- to late-segmentation stages, may cause the mesodermal defects seen in *cas* mutants, or that the YSL and endoderm both contribute to the migration of mesodermal tissues. Tissue specific loss-of-function analyses will be required to further investigate the basis for these mesodermal defects.

cas, bon *and* fau/gata5

cas, bon and *fau/gata5* encode transcription factors that directly or indirectly regulate *sox17* expression. These three genes exhibit a similar pattern of expression; all initiate expression at approximately the same time in the YSL and the marginal zone of the embryo. In the margin *bon* expression is the broadest (Alexander at al. 1999) and *cas* the most restricted. We have previously shown that, prior to gastrulation, *bon* and *fau/gata5* expression appear normal in *cas*, *bon* and *fau/gata5* mutants (Alexander et al. 1999; Reiter et al. 2001), suggesting that Cas function downstream of, or in parallel to, Bon and Fau/Gata5. These data were confirmed by the finding that *cas* overexpression in *MZoep* mutants does not induce *bon* or *fau/gata5* expression.

Neither *bon* nor *fau/gata5* overexpression restores *sox17* expression in *cas* mutants (Alexander and Stainier 1999; Reiter et al. 2001), further indicating that *cas* functions downstream of, or in parallel to, *bon* and *fau/gata5*. We confirmed these data by showing that *cas* overexpression in *bon* and *fau/gata5* mutants led to widespread *sox17* expression. To further examine whether Cas acts downstream of, or in parallel to, Bon and Fau/Gata5, we overexpressed *bon* and *fau/gata5* in wild-type embryos and examined *cas* expression. We found that *bon* or *fau/*

gata5, alone or in combination, were not able to induce *cas* expression in nonmarginal cells. These data are consistent with the observations that when *bon* and *fau/gata5* were overexpressed in wild-type embryos, epiboly proceeded normally and no ectopic *sox17* expression was induced (Reiter et al. 2001), in contrast to what was observed in *cas*-injected embryos. Furthermore, coinjection of *bon* and *fau/gata5* in *cyc;sqt* double mutants was not able to restore significant *sox17* expression (Y. Kikuchi and D.Y.R. Stainier, unpubl.), in contrast to what was observed in *cas*-injected *MZoep* mutants. All these data are consistent with at least two different models: *cas*, *bon*, and *fau/gata5* may be regulated independently by Nodal signaling and function in parallel; alternatively, *bon* and *fau/gata5* may act upstream of *cas*, with a third factor, X, being restricted to the marginal zone and necessary for *cas* expression. Some combination of these two models is also possible, and molecular dissection of the *cas* promoter may help to further investigate this question.

cas *is sufficient to convert mesoderm into endoderm*

Ectopic expression of *cas* in presumptive mesodermal cells leads to their transfating into endoderm. *cas*-expressing dorsal mesodermal cells lose *gsc* and *hgg1* expression (Fig. 6G,I), and colonize the entire length of the gut tube (Fig. 6C). Such embryos display cyclopia due to the loss of prechordal plate, and U-shaped somites due to the loss of notochord (Fig. 6D,E). It is interesting to note that *cas*-expressing cells, unlike *taram-a**-expressing cells (Peyriéras et al. 1998), never contribute to the hatching gland. We therefore suggest that this organ is not of endodermal origin as previously suggested (Peyriéras et al. 1998), but, rather, it most likely represents a mesodermal derivative whose formation, similar to that of the endoderm, requires high level, and/or sustained, Nodal signaling.

Ventral mesodermal cells, normally fated to give rise to blood as well as to other tissues, lose their ability to form blood when expressing *cas*. However, *cas* expression is not able to redirect their migration, so these cells end up colonizing the posterior gut tube and the tail region (Fig. 6L). Importantly, despite the ability of *cas* to up-regulate *cyc* and *sqt* expression when expressed throughout the embryo, ectopic expression in 1 of 16 blastomeres does not lead to general defects, such as broadened or duplicated axes, indicative of ectopic *sqt* expression (Erter et al. 1998). Rather, the effects on embryonic development appear to be limited to the *cas*-expressing cells. Altogether, these data indicate that *cas* is able to initiate an endoderm-specific program in cells not fated to become endoderm. However, this ability appears to be restricted to the mesoderm; when *cas* is overexpressed in animal pole blastomeres at the 256-cell stage, these cells do not transfate from ectoderm into endoderm. In summary, only when overexpressed near the margin is *cas* able to transfate cells into endoderm. It will be important to identify the factors responsible for this spatial restriction. In this context, it is interesting to reiterate that *cas* is unable to induce in *MZoep* mutants strong expression of endodermal markers other than *sox17*, such as *foxA2* and *fau/gata5*, indicating that other Nodal-regulated components are essential for the full endodermal program. *cas* misexpression in the animal pole, unlike that in the margin, does not lead to ectopic *nodal* gene expression, offering one possible explanation for the inability of these cells to transfate into endoderm.

Conclusions

Phenotypic analyses have shown that *cas* is the central regulator of *sox17* expression and endoderm formation in zebrafish (Alexander et al. 1999; Alexander and Stainier 1999; Reiter et al. 2001). The isolation of *cas*, analysis of its expression pattern, and gain-of-function experiments lead us to propose the following model: *cas* expression in the YSL initiates independently of Nodal signaling. At late blastula stages, Nodal signaling induces *bon*, *fau/gata5*, and *cas* expression in marginal blastomeres. Bon, Fau/Gata5, and Cas directly or indirectly regulate *sox17* expression in the forming endodermal cells, which also continue to express *fau/gata5* and *cas* during gastrulation. Several important issues remain to be addressed, in particular the differential regulation of *bon*, *fau/gata5*, and *cas* expression in the margin. These genes may become activated at different thresholds of Nodal signaling, and/or incorporate inputs from other marginally localized factors acting autonomously or nonautonomously. Resolution of this issue, along with the identification and analyses of additional regulators of endoderm formation, should further our understanding of the molecular events that direct individual marginal blastomeres toward an endodermal fate in vertebrate embryos.

Materials and methods

Zebrafish strains

Adult zebrafish and embryos were maintained and staged as described (Westerfield 1995). Heterozygous carriers were identified by random intercrosses and intercrossed to obtain mutant embryos. Adult *MZoep*^{m134} fish were a generous gift of Will Talbot (Stanford University). The following mutant alleles were used: *bon*^{s9} (Kikuchi et al. 2000), *fau*^{tm236a} (Chen et al. 1996; Reiter et al. 1999), *cas*^{ta56} (Chen et al. 1996), and *cas*^{s4}.

Isolation of the CG569 gene

A full-length clone was isolated from a gastrula cDNA library (gift of Thierry Lepage) inserted between *Xho*I and *Eco*RI of pBluescript IISK^{+}. The GenBank accession number for *cas* is AF362749.

Wholemount in situ hybridization

For *CG569/cas* probe synthesis, *Not*I was used to linearize the plasmid and T7 RNA polymerase to generate antisense RNA. Other riboprobes were prepared according to published instructions. In situ hybridizations shown in Figure 2 and Figure 6 were performed as described by Thisse and Thisse (http://www-

igbmc.u-strasbg.fr/zf_info/zbook/chapt9/9.82.html); the others as described in Alexander et al. (1998).

Genetic mapping and genomic sequencing of cas^{ta56} and cas^{s4}

CG569 was mapped on the Goodfellow zebrafish radiation hybrid panel (Research Genetics) using the following PCR primers: 5′-CCAGATTGCTGCTGATTTTG-3′ and 5′-TCAATACAAA CTGGTCAAGTTAACAAC-3′.

To extract genomic DNA from single embryos, 30 hpf embryos were digested overnight in 10mM Tris-HCl pH7.5, 1mM EDTA, 0.1% SDS, 100mM NaCl, 20 µg/mL proteinase K at 55°C. Mutant genomic DNA was purified from ten 30 hpf embryos using genomic DNA extraction kits (Scotlab Bioscience). To sequence the genomic locus, 4 primer pairs were designed based on *CG569* sequence data. The PCR fragments were cloned into the pGEM-T vector (Promega). We performed three independent PCR reactions for each primer pair, picked two independent clones from each PCR reaction, and sequenced them using an ABI 377 DNA sequencer.

Linkage analysis

CG569 genomic fragments were amplified with PCR primers (5′-CTGGCAGATAAGCGTCCGTACATGC-3′ and 5′-CTGA GTAAGCAGACCGTGGCAGATGC-3′). PCR conditions were 94°C for 5 min (1 cycle), then 94°C for 30 sec, 63°C for 30 sec, and 72°C for 1 min (40 cycles), followed by 72°C for 5 min (1 cycle). Amplified products were digested with *Bfa*I for cas^{ta56} or *Rsa*I for cas^{s4}, and resolved on 2% agarose gels. Following in situ hybridization, embryos were genotyped as described above.

mRNA injections

To introduce the *s4*-type point mutation into a *cas* injection vector, we performed site-directed mutagenesis using QuickChange™ Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instruction using the primer pair (5′-GCAGAAAGACTACGGATCCAGCG*TACCATTGACTATC CTAACTAC-3′ and 5′ -GTAGTTAGGATAGTCAATGGTA C*GCTGGATCCGTAGTCTTTCTGC-3′; * indicates the site of the mutation) and pCS2$^+$-*cas* template. The coding region of cas^{s4} was verified by sequencing both strands. Capped mRNA encoding *cas* or cas^{s4} was transcribed in vitro from pCS2$^+$-*cas* or pCS2 $^+$-cas^{s4}, using the SP6 mMESSAGE mMACHINE system (Ambion). Capped *antivin, taram-a*, bon, fau/gata5, lacZ,* and *gfp* mRNA were prepared according to published instructions (Peyriéras et al. 1998; Thisse et al. 1999; Kikuchi et al. 2000; Reiter et al. 2000). Embryos were injected with the following mRNAs: *antivin* (25 pg), *taram-a** (100 pg), *bon* (100 pg), *cas* (100 pg), *fau/gata5* (60 pg), *lacZ* (100 pg). Embryos injected at the 1–4-cell stage with *cas* mRNA fail to undergo epiboly and do not exhibit a clear dorsal axis; they were fixed for in situ at 6 hpf, which is equivalent to shield stage. For the transfating experiments, 100 pg each of *cas* and *gfp* mRNA were coinjected into a single marginal blastomere of 16-cell stage embryos, or with a 120 ng/µL solution of *cas* and *GFP* mRNA in an animal pole blastomere at the 256-cell stage.

Acknowledgments

We thank Sarah Howard and Alex Navarro for outstanding technical assistance; Will Talbot for *MZoep* fish; Peter Koopman for expert help with sequence and clustering analyses; Don Kane for unpublished information about the LG7 centromere; Hiroyuki Takeda and Takuya Sakaguchi for open discussions about unpublished data; Janet Heasman for discussions about endoderm formation; Leonard D'Amico for discussions about forerunner cells; Jeremy Reiter, Le Trinh, Elke Ober, and Heather Verkade for critical comments on the manuscript. Y.K. was funded by a long-term fellowship from the Human Frontier Science Program and is now an American Heart Association Postdoctoral Fellow. A.A is funded by the ministère de l'Education Nationale et de la Recherche. J.A. was supported by an American Heart Association Predoctoral Fellowship and D.Y. by an Amgen fellowship of the Life Sciences Research Foundation and a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund. This work was supported in part by funds from the Institut National de la Santé et de la Recherche Médicale, the Centre National de la Recherche Scientifique, the Hôpital Universitaire de Strasbourg, the Association pour la Recherche sur le Cancer and the Ligue Nationale Contre le Cancer to B.T. and C.T., and from the American Heart Association, the Packard Foundation, the Life and Health Insurance Medical Research Fund and the NIH (DK58181) to D.Y.R.S.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Alexander, J. and Stainier, D.Y.R. 1999. A molecular pathway leading to endoderm formation in zebrafish. *Curr. Biol.* **9:** 1147–1157.

Alexander, J., Stainier, D.Y.R., and Yelon, D. 1998. Screening mosaic F1 females for mutations affecting zebrafish heart induction and patterning. *Dev. Genet.* **22:** 288–299.

Alexander, J., Rothenberg, M., Henry, G.L., and Stainier, D.Y.R. 1999. *casanova* plays an early and essential role in endoderm formation in zebrafish. *Dev. Biol.* **215:** 343–357.

Bowles, J., Schepers, G., and Koopman, P. 2000. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev. Biol.* **227:** 239–255.

Chen, J.N., Haffter, P., Odenthal, J., Vogelsang, E., Brand, M., van Eeden, F.J.M., Furutani-Seiki, M., Granato, M., Hammerschmidt, M., Heisenberg, C.P., et al. 1996. Mutations affecting the cardiovascular system and other internal organs in zebrafish. *Development* **123:** 293–302.

Conlon, F.L., Lyons, K.M., Takaesu, N., Barth, K.S., Kispert, A., Herrmann, B., and Robertson, E.J. 1994. A primary requirement for *nodal* in the formation and maintenance of the primitive streak in the mouse. *Development* **120:** 1919–1928.

Cooper, M.S. and D'Amico, L.A. 1996. A cluster of noninvoluting endocytic cells at the margin of the zebrafish blastoderm marks the site of embryonic shield formation. *Dev. Biol.* **180:** 184–198.

Erter, C.E., Solnica-Krezel, L., and Wright, C.V. 1998. Zebrafish *nodal-related 2* encodes an early mesendodermal inducer signaling from the extraembryonic yolk syncytial layer. *Dev. Biol.* **204:** 361–372.

Feldman, B., Gates, M.A., Egan, E.S., Dougan, S.T., Rennebeck, G., Sirotkin, H.I., Schier, A.F., and Talbot, W.S. 1998. Zebrafish organizer development and germ-layer formation require nodal-related signals. *Nature* **395:** 181–185.

Gritsman, K., Zhang, J., Cheng, S., Heckscher, E., Talbot, W.S.,

and Schier, A.F. 1999. The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. *Cell* **97:** 121–132.

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

Hudson, C., Clements, D., Friday, R.V., Stott, D., and Woodland, H.R. 1997. Xsox17α and -β mediate endoderm formation in *Xenopus*. *Cell* **91:** 397–405.

Johnson, S.L., Africa, D., Horne, S., and Postlethwait, J.H. 1995. Half-tetrad analysis in zebrafish: Mapping the *ros* mutation and the centromere of linkage group I. *Genetics* **139:** 1727–1735.

Kaestner, K.H., Knöchel, W., and Martinez, D.E. 2000. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes & Dev.* **14:** 142–146.

Kikuchi, Y., Trinh, L.A., Reiter, J.F., Alexander, J., Yelon, D., and Stainier, D.Y.R. 2000. The zebrafish *bonnie and clyde* gene encodes a Mix family homeodomain protein that regulates the generation of endodermal precursors. *Genes & Dev.* **14:** 1279–1289.

Krauss, S., Concordet, J.P., and Ingham, P.W. 1993. A functionally conserved homolog of the *Drosophila* segment polarity gene *hh* is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* **75:** 1431–1444.

Kwok, C., Critcher, R., and Schmitt, K. 1999. Construction and characterization of zebrafish whole genome radiation hybrids. In *Methods in Cell Biology. The Zebrafish: Genetics and genomics* (eds. H.W. Detrich, M. Westerfield, and L.I. Zon), pp. 287–302. Academic Press, San Diego, CA.

Liang, J.O., Etheridge, A., Hantsoo, L., Rubinstein, A.L., Nowak, S.J., Izpisua Belmonte, J.C., and Halpern, M.E. 2000. Asymmetric nodal signaling in the zebrafish diencephalon positions the pineal organ. *Development* **127:** 5101–5112.

Osada, S.I. and Wright, C.V. 1999. *Xenopus nodal*-related signaling is essential for mesendodermal patterning during early embryogenesis. *Development* **126:** 3229–3240.

Payne, T.L., Postlethwait, J.H., and Yelick, P.C. 2001. Functional characterization and genetic mapping of *alk8*. *Mech. Dev.* **100:** 275–289.

Pennisi, D., Gardner, J., Chambers, D., Hosking, B., Peters, J., Muscat, G., Abbott, C., and Koopman, P. 2000. Mutations in *Sox18* underlie cardiovascular and hair follicle defects in ragged mice. *Nat. Genet.* **24:** 434–437.

Peyriéras, N., Strähle, U., and Rosa, F. 1998. Conversion of zebrafish blastomeres to an endodermal fate by TGF-β- related signalling. *Curr. Biol.* **8:** 783–786.

Rebagliati, M.R., Toyama, R., Fricke, C., Haffter, P., and Dawid, I.B. 1998a. Zebrafish nodal-related genes are implicated in axial patterning and establishing left-right asymmetry. *Dev. Biol.* **199:** 261–272.

Rebagliati, M.R., Toyama, R., Haffter, P., and Dawid, I.B. 1998b. *cyclops* encodes a nodal-related factor involved in midline signaling. *Proc. Natl. Acad. Sci.* **95:** 9932–9937.

Reiter, J.F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N., and Stainier, D.Y.R. 1999. Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes & Dev.* **13:** 2983–2995.

Reiter, J.F., Kikuchi, Y., and Stainier, D.Y.R. 2001. Multiple roles for Gata5 in zebrafish endoderm formation. *Development* **128:** 125–135.

Renucci, A., Lemarchandel, V., and Rosa, F. 1996. An activated form of type I serine/threonine kinase receptor TARAM-A reveals a specific signalling pathway involved in fish head-organiser formation. *Development* **122:** 3735–3743.

Rodaway, A., Takeda, H., Koshida, S., Broadbent, J., Price, B., Smith, J.C., Patient, R., and Holder, N. 1999. Induction of the mesendoderm in the zebrafish germ ring by yolk cell-derived TGF-β family signals and discrimination of mesoderm and endoderm by FGF. *Development* **126:** 3067–3078.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.

Sampath, K., Rubinstein, A.L., Cheng, A.M., Liang, J.O., Fekany, K., Solnica-Krezel, L., Korzh, V., Halpern, M.E., and Wright, C.V. 1998. Induction of the zebrafish ventral brain and floorplate requires cyclops/nodal signalling. *Nature* **395:** 185–189.

Schier, A.F., Neuhauss, S.C., Helde, K.A., Talbot, W.S., and Driever, W. 1997. The *one-eyed pinhead* gene functions in mesoderm and endoderm formation in zebrafish and interacts with *no tail*. *Development* **124:** 327–342.

Stainier, D.Y., Lee, R.K., and Fishman, M.C. 1993. Cardiovascular development in the zebrafish. I. Myocardial fate map and heart tube formation. *Development* **119:** 31–40.

Strähle, U., Blader, P., Henrique, D., and Ingham, P.W. 1993. *Axial*, a zebrafish gene expressed along the developing body axis, shows altered expression in *cyclops* mutant embryos. *Genes & Dev.* **7:** 1436–1446.

Thisse, C., Thisse, B., Halpern, M.E., and Postlethwait, J.H. 1994. *goosecoid* expression in neurectoderm and mesendoderm is disrupted in zebrafish *cyclops* gastrulas. *Dev. Biol.* **164:** 420–429.

Thisse, C. and Thisse, B. 1999. Antivin, a novel and divergent member of the TGFβ superfamily, negatively regulates mesoderm induction. *Development* **126:** 229–240.

Thisse, B., Wright, C.V., and Thisse, C. 2000. Activin- and Nodal-related factors control antero-posterior patterning of the zebrafish embryo. *Nature* **403:** 425–428.

Tremblay, K.D., Hoodless, P.A., Bikoff, E.K., and Robertson, E.J. 2000. Formation of the definitive endoderm in mouse is a Smad2-dependent process. *Development* **127:** 3079–3090.

Warga, R.M. and Kimmel, C.B. 1990. Cell movements during epiboly and gastrulation in zebrafish. *Development* **108:** 569–580.

Warga, R.M. and Nüsslein-Volhard, C. 1999. Origin and development of the zebrafish endoderm. *Development* **126:** 827–838.

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.

Werner, M.H., Huth, J.R., Gronenborn, A.M., and Clore, G.M. 1995. Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell* **81:** 705–714.

Westerfield, M. 1995. The Zebrafish book: *A guide for the laboratory use of zebrafish.* University of Oregon Press, Eugene, OR.

Woo, K. and Fraser, S.E. 1995. Order and coherence in the fate map of the zebrafish nervous system. *Development* **121:** 2595–2609.

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.

Zaret, K. 1999. Developmental competence of the gut endoderm: Genetic potentiation by GATA and HNF3/fork head proteins. *Dev. Biol.* **209:** 1–10.

Zhang, J., Talbot, W.S., and Schier, A.F. 1998. Positional cloning identifies zebrafish *one-eyed pinhead* as a permissive EGF-related ligand required during gastrulation. *Cell* **92:** 241–251.