A Small Change With a Twist Ending: A Single Residue in EGF-CFC Drives Bilaterian Asymmetry

Marta Truchado-García ⓘ,*,[1,2] Kimberly J. Perry,[3] Florencia Cavodeassi,[1,4] Nathan J. Kenny,[5,6] Jonathan Q. Henry,[3,7] and Cristina Grande*,[1]

[1]Departamento de Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, Madrid, Spain

[2]Department of Molecular and Cell Biology, University of California—Berkeley, 571 Weill Hall, Berkeley, CA 94702

[3]Department of Cell and Developmental Biology, University of Illinois, Urbana, IL 61801

[4]Institute of Medical and Biomedical Education, St George's University of London, Cranmer Terrace, London SW17 0RE, United Kingdom

[5]Natural History Museum, Cromwell Road, London, United Kingdom

[6]Department of Biochemistry (Te Tari Matū Koiora), University of Otago, Dunedin, (Aotearoa) New Zealand

[7]The Marine Biological Laboratory, Woods Hole, MA 02543

*Corresponding authors: E-mails: mtruchado@berkeley.edu; cristina.grande@uam.es.

Associate editor: John Parsch

Abstract

Asymmetries are essential for proper organization and function of organ systems. Genetic studies in bilaterians have shown signaling through the Nodal/Smad2 pathway plays a key, conserved role in the establishment of body asymmetries. Although the main molecular players in the network for the establishment of left-right asymmetry (LRA) have been deeply described in deuterostomes, little is known about the regulation of Nodal signaling in spiralians. Here, we identified orthologs of the *egf-cfc* gene, a master regulator of the Nodal pathway in vertebrates, in several invertebrate species, which includes the first evidence of its presence in non-deuterostomes. Our functional experiments indicate that despite being present, *egf-cfc* does not play a role in the establishment of LRA in gastropods. However, experiments in zebrafish suggest that a single amino acid mutation in the *egf-cfc* gene in at least the common ancestor of chordates was the necessary step to induce a gain of function in LRA regulation. This study shows that the *egf-cfc* gene likely appeared in the ancestors of deuterostomes and "protostomes", before being adopted as a mechanism to regulate the Nodal pathway and the establishment of LRA in some lineages of deuterostomes.

Key words: Nodal, EGF-CFC, cripto, left-right asymmetry, zebrafish, EvoDevo, gene expression pattern, Spiralia, *Crepidula fornicata*.

Article

Introduction

Despite dramatically different body architectures, animals share common signaling pathways and transcriptional networks that regulate their development, a core "genetic toolkit." Here, we examine the toolkit that underlies left-right asymmetry (LRA) in the Spiralia, a group that contains mollusks, annelids, platyhelminthes, and at least eight other phyla, and is the most phyla-rich and arguably morphologically diverse bilaterian clade (Marlétaz et al. 2019).

Most bilaterians exhibit some LRAs in their both external and internal anatomy (e.g., the position of the heart and liver in vertebrates, and the direction of gut torsion in vertebrates and arthropods). Both paired and unpaired organs generate asymmetries, according to their shape, configuration and function. These asymmetries are essential for proper packing, connectivity, and function of these organ systems. Complete reversal of these asymmetries (i.e., *situs inversus*) to form a perfect mirror image is often well tolerated, but partial reversals (e.g., heterotaxies) may severely compromise viability (Lobikin et al. 2012). Within the Spiralia, shell coiling of gastropod snails is perhaps the most obvious morphological asymmetry, but others include the positions of certain sensory organs, various neural connections, the digestive tract, and locomotive cilia (Grande 2010; Martín-Durán et al. 2016). Most snails present a coiled shell (and by extension, body conformation) that has a clear chirality—either dextral (with a clockwise twist, as is more common) or sinistral (with an anticlockwise twist). However, chirality is also evident during embryonic development during the second and third cleavage divisions, and these events define the chirality of the embryo and its future larval/adult body plan (Shibazaki et al. 2004). In gastropods, a dextral third

cleavage results in animals with dextral shells and a sinistral third cleavage results in sinistral shells.

The establishment of LRA has been studied mainly in vertebrate models and in tissue culture, although there are several studies in invertebrates (reviewed in Hamada and Tam 2020). These studies have demonstrated that symmetry breaking starts very early in embryogenesis. The information generated by the early symmetry breaking events is translated into asymmetric deployment of genetic cascades, reinforcing the initial signal, and reliably dictating subsequent asymmetries during organogenesis. At the molecular level, the initial determination of LRA in vertebrates is widely conserved. However, conservation and divergences in these mechanisms among metazoans requires further analysis (Palmquist and Davidson 2017). Signaling through the Nodal/Smad2 pathway plays a role in LRA in diverse deuterostomes (Duboc et al. 2005; Soukup et al. 2015) and some spiralians (Grande and Patel 2009; Grande et al. 2014; Martín-Durán et al. 2016). This pathway controls cellular processes such as differential cell migration, proliferation, adhesion, and asymmetric morphogenesis in both deuterostomes and the Spiralia (Hamada et al. 2002; Duboc et al. 2005; Grande and Patel 2009). Indeed, loss of function (LOF) of components in the pathway can promote *situs inversus* or heterotaxy (Levin et al. 1997; Grande and Patel 2009). However, the Nodal network is absent in Ecdysozoa, which suggests independent evolution of the mechanisms driving LRA in this clade (Schier 2009).

Nodal belongs to the transforming growth factor beta (TGF-β) superfamily of signaling molecules, and signals asymmetrically along a single side of the developing embryo, through a well-conserved cascade (Massagué 2012), to create a phospho-Smad complex, which promotes the activation of target genes, such as *Pitx*. Comparative studies have demonstrated similar expression and regulatory interactions between *nodal* and *Pitx* in diverse deuterostomes (Raya and Belmonte 2004; Yoshida and Saiga 2008; Soukup et al. 2015). In snails, chirality defines the location of *nodal* activation, so that the gene is expressed on the right side of dextral and the left side in sinistral individuals (Grande and Patel 2009; Kuroda et al. 2009). Defective function of the Nodal pathway leads to LRA defects in snails (Grande and Patel 2009). However, little is known about the regulation of Nodal signaling in spiralians.

Unlike in other TGF-β pathways, an epidermal growth factor-cripto/FRL-1/cryptic (EGF-CFC) coreceptor is required for the assembly of the Nodal receptor complex and plays a key role in the regulation of Nodal activity in vertebrates. These signaling factors have diverse functions, from controlling early events during embryonic development to tissue homeostasis. The EGF-CFC protein contains recognition sites for interactions with diverse molecules. For instance, these are able to block signaling of TGF-β ligands such as Activin or enhance the signaling of other ligands such as BMP-4 or Nodal (Minchiotti et al. 2001; Yan et al. 2002; Gray and Vale 2012). Three main domains define

this protein family: (1) the EGF-like domain, involved in ligand recognition (i.e., TGF-βs, BMPs, GDFs, and Tomoregulin), and regulation of the Nodal, MAPK, and PI3K-Akt pathways; (2) the CFC domain, involved in receptor binding; and (3) the GPI domain, involved in membrane anchoring (reviewed in Klauzinska et al. 2014). To date, EGF-CFCs have only been described in deuterostomes, and are considered an evolutionary innovation of the group (reviewed in Shen and Schier 2000; Ravisankar et al. 2011). In mammals, the family includes two genes called *cripto* (Ding et al. 1998) and *cryptic* (Yan et al. 1999). In the zebrafish (*Danio rerio*), one EGF-CFC gene has been described (one-eye-pinhead, *oep*). *Oep* is maternally and zygotically expressed; at later stages, it is expressed in the forebrain, lateral plate, and notochord. Zygotic *oep* mutants show many of the observed phenotypes for *cripto* and *cryptic* mutants in mice, including defects in the establishment of both left-right and anterior-posterior axes (Schier et al. 1997).

The establishment of LRA is one of the most studied functions of EGF-CFC (reviewed in Klauzinska et al. 2014). The interaction between EGF-CFC and Nodal at a molecular level is well described in both mice and humans (Calvanese et al. 2010). A threonine within the EGF-like domain (position T88 in humans) is essential for Nodal function. This residue undergoes O-fucosilation, facilitating the ligand-coreceptor interaction. This O-fucosilation cannot be replaced by any other amino acid, even those capable of being fucosilated (serine) (Schiffer et al. 2001; Shi et al. 2007). Whether this interaction is also necessary for the activation of the Nodal pathway in other organisms remains unknown.

In this manuscript, we have surveyed the presence and/or role in LRA of both Nodal and the EGF-CFC coreceptor in non-deuterostomes, focusing on gastropods, to understand the relationship between these two proteins in non-deuterostomes. We identified orthologs of EGF-CFC in several invertebrate species, the first evidence of its presence in non-deuterostomes. We infer that, despite *egf-cfc* orthologs being present in non-deuterostomes, the Nodal pathway may have acted in an *egf-cfc*-independent manner in the common ancestor of all bilaterians, and this relationship persists in non-deuterostomes. Our experiments in Zebrafish suggest that the role of Nodal in LRA asymmetry is shared between deuterostomes and non-deuterostomes, although the pathway became dependent on the EGF-CFC coreceptor in some lineages of deuterostomes sometime after a single key amino acid change occurred in its protein sequence.

Results

Two paralogs for *nodal* and one ortholog for *egf-cfc* are present in *Crepidula fornicata*

Searches in the gastropod *C. fornicata* transcriptome database and subsequent orthology analyses identified two paralog genes for *nodal* (*Cfo-nodal*A and *Cfo-nodal*B; fig. 1A;

Grande et al. 2014). *Cfo-nodalA* gives rise to a 428-amino acid protein, whereas *Cfo-nodalB* generates a 472-amino acid protein. Both proteins display the distinctive seven cysteine residues that exist in the C-terminal region of this family of proteins, and the cleavage site domain, which matches orthologs in other animals (fig. 1B). We observed a similar duplication in another gastropod, *Patella vulgata*, but the phylogenetic analysis performed in Grande et al. (2014) suggests these paralogs may have arisen independently.

Next, we focused on the EGF-CFC family. Comprehensive searches in metazoan databases for the *egf-cfc* orthologs and subsequent orthology analyses were carried out using *cripto* and *cryptic* genes from *Mus musculus*, and *oep* from *D. rerio* as queries. We found orthologous genes for the *egf-cfc* gene in gastropods, annelids, and brachiopods (fig. 1C), whereas exhaustive searches of cnidarian and ctenophore genomes, genomes of bilaterians such as *Xenoturbella*, Acoelomorpha, Ecdysozoa, and Platyhelminthes, and the NCBI nr, TSA and EST databases, retrieved no orthologs as members of this family. This indicates that the last ancestor of ecdysozoans, spiralians, and deuterostomes presented only one EGF-CFC member, lost in the Ecdysozoan lineage, and that the EGF-CFC family is not an evolutionary novelty of deuterostomes. Although only one representative *egf-cfc* gene was found in most bilaterians, our phylogenetic analysis shows the existence of up to three paralogous genes for *egf-cfc* within specific evolutionary lines (*Xenopus laevis*, *Xenopus tropicalis*, *Homo sapiens*, and *P. vulgata*; fig. 1C). The phylogenetic analysis also suggests a duplication in the vertebrate lineage, which gave rise to another clade of proteins in mammals (Cripto).

Sequence analysis showed that all orthologs display a conserved structure that includes an EGF-like domain, a CFC domain and a GPI domain (fig. 1D), generating a 209-amino acid protein in *C. fornicata* that we call Cfo-EGF_CFC. Three critical amino acids in the EGF-like domain of the EFG-CFC coreceptor have been identified in humans for correct interaction with Nodal: glycine (G87), threonine (T88), and phenylalanine (F94) (Minchiotti et al. 2001; Yan et al. 2002; Shi et al. 2007). The substitution of these amino acids leads to a complete LOF of EGF-CFC. Three-dimensional models have confirmed that these three amino acids are exposed on the same side of the protein and form a surface that may serve as an interaction point with Nodal. Glycine (G87) and phenylalanine (F94) are conserved in all the orthologs identified here (fig. 1E). However, threonine (T88) is only conserved in some lineages of deuterostomes, whereas gastropod mollusks have a leucine (L) or arginine (R) at this position, annelids have a methionine (M) or isoleucine (I), and brachiopods have a valine (V) (fig. 1E). Previous works point to threonine (T88) as being critical for Nodal-Cripto interaction, and its substitution by other amino acids leads to a failure on the activation of the pathway (Schiffer et al. 2001; Yan et al. 2002; Shi et al. 2007).

The CFC domain mediates the protein–protein interaction of EGF-CFC with the type-I receptor of the TGF-β pathway (Alk4) (Calvanese et al. 2009), in particular,

positions histidine (H120) and tryptophan (W123) in humans (Foley et al. 2003; Marasco et al. 2006). Our analysis shows that these are conserved in all the identified orthologs, with the exception of tryptophan (W123) for the annelid *Spirobranchus* and the hemichordates, and hence, all of them presumably bind to the type-I receptor (fig. 1E). Last, these proteins can act not only as membrane proteins but also as soluble proteins, if a cleavage on the GPI domain occurs as shown in human cell cultures (Watanabe and Salomon 2010). We have recognized the existence of a GPI domain in all the identified orthologs, so these orthologs may share the ability to act as either membrane or soluble proteins.

Cfo-nodalB Expression Suggests an Involvement in LRA Development

We designed specific riboprobes for *Cfo-nodalA* and *Cfo-nodalB* by selecting regions within the sequences that were divergent enough to avoid a cross-reaction. Whole-mount *in situ* hybridization (WISH) was subsequently carried out at different stages in *C. fornicata* embryos to examine their expression throughout development. *Cfo-nodalA* transcript is clearly present at the mid-gastrula stage (110 h post fertilization; hpf) in every quadrant (A, B, C, and D) and around the blastopore (fig. 2A), which will form the mouth in these animals. At late gastrula stages when organogenesis starts (170 hpf), *Cfo-nodalA* is located along the anterior end and around the stomodeum (future mouth), and in a small posterior region (fig. 2A). In the pre-veliger larva (180 hpf), *Cfo-nodalA* expression is more intense and remains restricted to an area of the mouth in a bilaterally symmetrical manner, as well as in the shell gland (fig. 2A). Earlier, *Cfo-nodalB* is detected when embryos reach 40–50 cells (40 hpf) (fig. 2B). At this stage, expression is detected in a single cell of the C quadrant ($2c^2$). As subsequent divisions take place, *Cfo-nodalB* appears in the $2c^2$ cell progeny ($2c^{21}$ and $2c^{22}$), at 60 hpf (fig. 2B). During gastrulation (110 hpf), the expression remains in the progeny of these cells (fig. 2B). During organogenesis and larval development (from 150 hpf), this expression is asymmetrical, located on the right side of the mouth, the right area of the velum, the rudiment of the foot, and in an ectodermal patch on the right side of the post-trochal region (fig. 2B). This expression becomes more intense and remains in these regions of preveliger stages. Finally, in veliger larvae (200 hpf), expression is restricted to the right side of the visceral mass and the head (fig. 2B). Thus, although *Cfo-nodalA* expression is largely symmetrical, *Cfo-nodalB* shows clear asymmetric expression, concentrated on the right side of the developing embryo (fig. 2A and B).

Cfo-nodalB expression in the right side of the embryo is closely followed by asymmetric expression of *Pitx*, a known target gene of Nodal in gastropods and other bilaterians (Hamada et al. 2002; Grande and Patel 2009). *Pitx* is asymmetrically expressed in *C. fornicata* immediately after *Cfo-nodalB* activation (fig. 2C), which is consistent with previous data in other gastropods (Grande and Patel 2009).
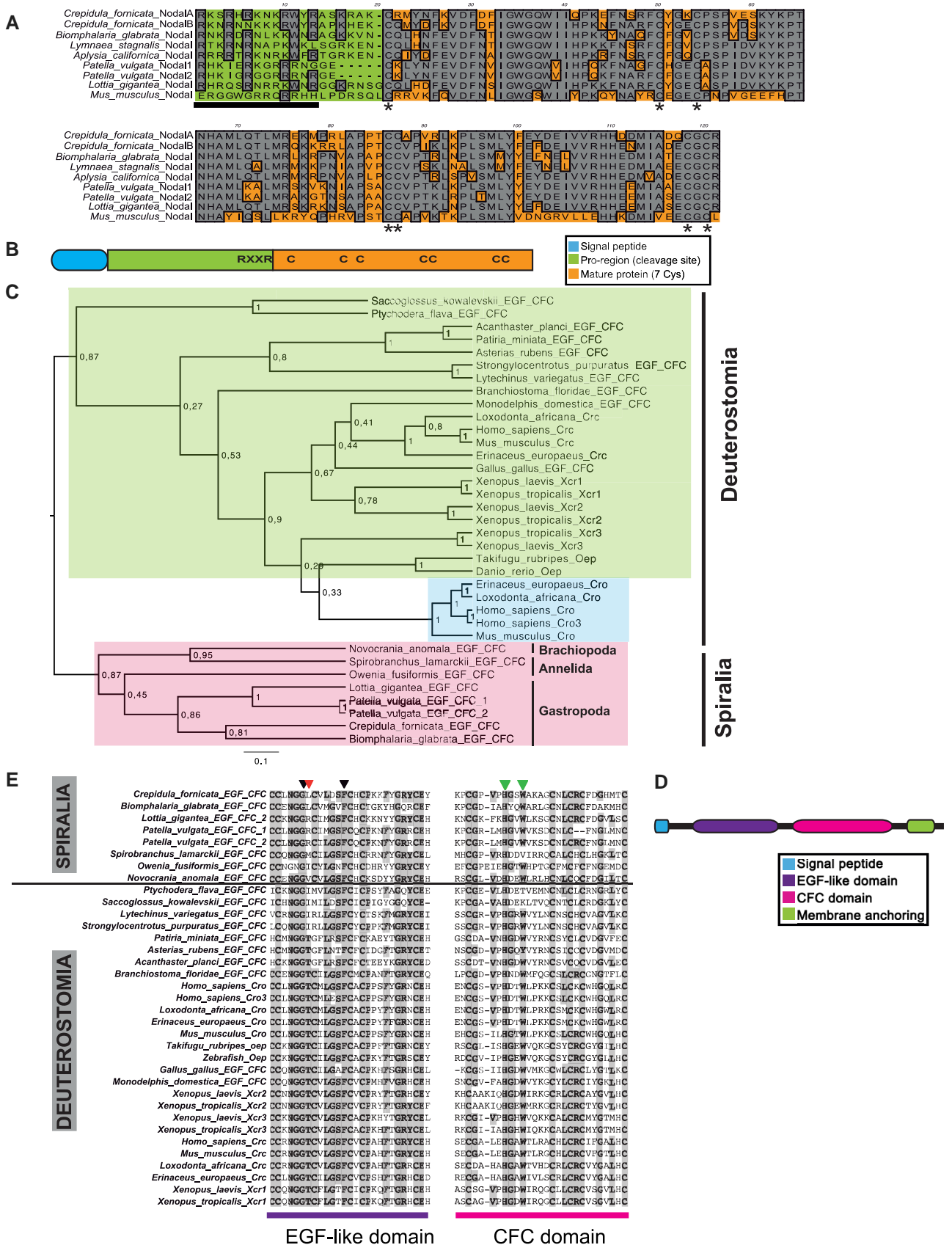
FIG. 1. Two paralogs for Nodal are present in *Crepidula fornicata*, and EFG-CFC orthologs are present in non-deuterostomes. (A) Alignment of several ortholog proteins for Nodal in different gastropod species (*C. fornicata*; *Biomphalaria glabrata*; *Lymnaea stagnalis*; *Aplysia californica*; *Patella vulgata*; *Lottia gigantea*) and mammalian *Mus musculus*. Gray indicates conservation for one given position; asterisks point the seven conserved cysteines that characterize the family; bar delimits the regions were cleavage sites are located. (B) Generic structure of Nodal protein. C, cysteines; (RXXR), cleavage site. (C)
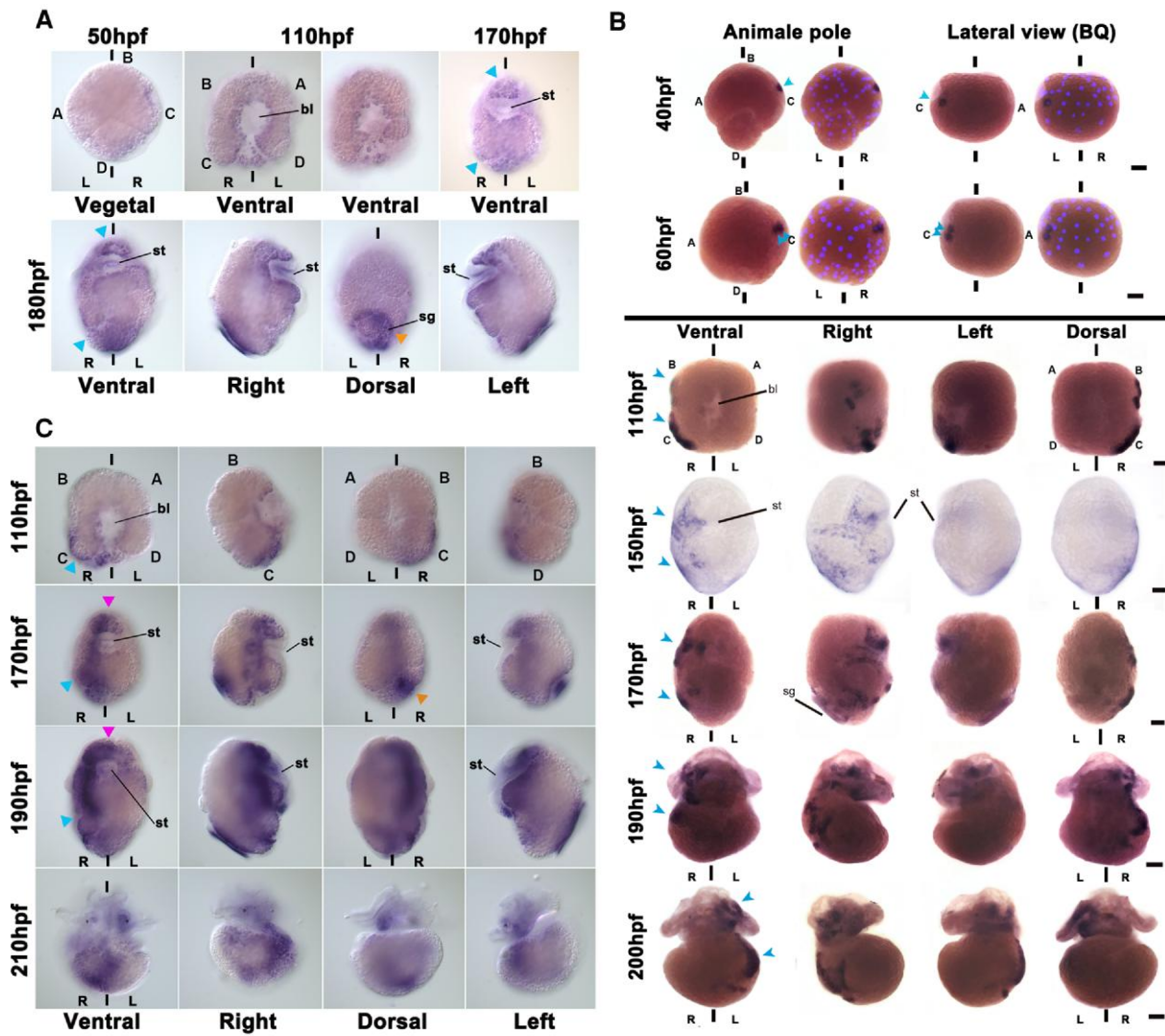
Fig. 2. *Cfo-nodalB* expression suggests involvement in LRA development. (*A–C*) Spatio-temporal localization of (*A*) *Cfo-nodalA*, (*B*) *Cfo-nodalB*, and (*C*) *Pitx* mRNA by WISH during the development of *Crepidula fornicata*. Several views of the same embryo for each stage are indicated on the figure. bl, blastopore; L, left; R, right; sg, shell gland; st, stomodeum. A, B, C, and D indicate the quadrants; hpf, hours post fertilization; bilateral axis is represented by the vertical bars. Scale corresponds to 30 μm. (*A*) Note that the expression is symmetrical in all the stages (blue arrowheads); orange arrowhead, shell gland. In (*B*), superpositions of brightfield images with Hoechst staining have been included for early stages, to identify the cells. Note that the expression is asymmetrical in all the stages (blue arrowheads). (*C*) Pink arrowhead, symmetrical signals; blue arrowhead, asymmetrical signals; orange arrowhead, shell gland.

Pitx expression is asymmetrical on the right side as previously shown in Henry et al. (2010). This expression on the right side is evident at mid-gastrula stages (fig. 2C),

and colocalizes to the same regions of *Cfo-nodalB* (fig. 2B and C). *Pitx* shows two expression sets, one asymmetrical on the right side, and another symmetrical, as is the case

EFG-CFC proteins inter-relationships. Bayesian phylogenetic tree of EFG-CFC inferred proteins from spiralian species, along with previously described sequences of other bilaterians (supplementary S1, Supplementary Material online). The WAG model (Whelan and Goldman 2001) was selected as the best-fit model of protein evolution using ProtTest (Abascal et al. 2005). Tree shown is the result of Bayesian analysis, run for 25,000,000 generations, and analyses run until the SD of split frequencies was below 0.01, with the first 25% of sampled trees discarded as "burn-in." Posterior probabilities can be seen at each branch. Scale bar represents substitutions per site at given distances. Shading shows major protein lineages. (*D*) Generic structure of EGF-CFC protein. (*E*) Alignment of the EGF-like and CFC domains of several ortholog proteins for Cripto, Cryptic, and EGF-CFC in different bilaterians, from both Spiralia and Deuterostomia clades. Shading indicates conservation for one given position. Arrowheads indicate the relevant positions in the ligand-receptor-coreceptor binding; Gly and Phe (black arrows) are conserved, whereas Thr, described as necessary for Nodal signaling, is only present in chordates and some lineages of echinoderms, and is variable in spiralians (red arrow); His and Trp (green arrows) are conserved in all orthologs with the exception for *Spirobranchus* and hemichordates.

Table 1. Results Obtained after mRNA Injection in *C. fornicata* zygotes, of *nodalB* and *egf-cfc* Constructs.

				Observed phenotype	
		mRNA & dose	Total embryos	Wild type	Symmetrical
Nodal	Overexpression	Uninjected control	26	100%	0%
	(mid-dose)	Cfo_NodalB_wt (450 pg)	20	45% (9)	55% (11)
	Overexpression	Uninjected control	30	100%	0%
	(higher dose)	Cfo_NodalB_wt (600 pg)	25	8% (2)	92% (23)
	Defective cleavage site construct	Uninjected control	48	100%	0%
		Cfo_NodalB_ΔCS	74	100%	0%
			Total embryos	Wild type	Radialized
EGF-CFC	Overexpression	Uninjected control	70	100%	0%
		Cfo_egf-cfc_wt	84	100%	0%
	EGF-like domain deletion	Uninjected control	34	100%	0%
		Cfo_egf-cfc_ΔEGF (600 pg)	62	52 (83.9%)	10 (16.1%)
		Cfo_egf-cfc_ΔEGF (800 pg)	61	10 (16.4%) mortality 45.4%	20 (38.2%)
	L56A substitution	Uninjected control	34	100%	0%
		Cfo_egf-cfc_L56A (600 pg)	72	100%	0%

NOTE.—Overexpression of wild-type *nodal*B presented bilateral symmetry phenotype, in a dose-dependent manner. However, a construct lacking the processing region of the pre-protein (Cfo_NodalB_ΔCS) did not show any apparent phenotype, when compared with the control. No phenotype was also observed in embryos injected with either *egf-cfc* wild type or L56A substitution (Cfo_egf-cfc_L56A), whereas radialized larvae appeared when embryos were injected with a defective form lacking the EGF-like domain (Cfo_egf-cfc_ΔEGF), which binds to the ligands. This last construct also increased mortality when a higher concentration was used.

Table 2. Results Obtained After RHAct-A Treatment on *C. fornicata* Embryos.

					Observed phenotype	
Window	Concentration (µM)	In	Out	Total embryos	Wild type	Symmetrical
Cleavage (7 hpf)	Control: 0	7 h	48 h	65	65	0
	10			56	56	0
Early blastula (32 hpf)	Control: 0	32 h	80 h	65	65	0
	10			51	10	41 (80.4%)
Blastula (48 hpf)	Control: 0	48 h	96 h	70	70	0
	5			45	41	4 (8.8%)
	10			43	18	25 (60%)
Gastrula (61 hpf)	Control- 0	61 h	109 h	30	0	0
	10			32	12	20 (62.5%)

NOTE.—Each experiment represents a different treatment at 7, 32, 48, and 61 h post fertilization (hpf). All treatments lasted 48 h, and carried out at 10 µM RHAct-A, except on 48 hpf-embryos for which an extra 5 µM treatment was also done. The resulting embryos presented bilateral symmetry phenotype, in a time- and dose-dependent manner.

in other gastropods and vertebrates, suggesting the presence of independent enhancer areas that regulate its expression (Christiaen et al. 2005; Grande and Patel 2009).

Cfo-nodalB regulates the establishment of LRA and Activin inhibits the pathway in *C. fornicata*

To further investigate the function of *Cfo-nodalB*, we produced wild-type and defective forms for expression in *C. fornicata*. Overexpression experiments carried out through microinjection of an *in vitro* synthesized mRNA for the *Cfo-nodalB* paralog into *C. fornicata* zygotes produced a high percentage of affected individuals (table 1). We observed a consistent phenotype, including loss of LRA in the body plan. This included a symmetrical shell and arrangement of the internal organs (i.e., no hindgut rotation with the lack of torsion; fig. 3B, table 1). The scoring of the embryos with a symmetrical phenotype was dose dependent and directly related to the amount of RNA injected

into the embryos. At the highest injection concentration, 92% of embryos were symmetrical (table 1).

Next, we investigated whether the removal of the putative cleavage sites within the *Cfo-nodalB* precursor would inhibit activation of the *Cfo-nodalB* pathway. The cleavage and dimerization of the ligands are necessary for these to bind to the corresponding receptors (Shen 2007). Such changes are effective in loss of function experiments, as their presence leads to competition with wild-type forms that block the activation of functional dimers (Joseph and Melton 1998). This potentially dominant negative form was designed to disrupt Nodal signaling at the ligand-receptor level. However, when this modified mRNA was injected in *C. fornicata* zygotes, no phenotypes different from wild type were observed in resulting larvae at any concentration (table 1). The utility of this approach may be questioned by studies performed in *Xenopus*, which showed that even non-processed peptides can generate some signaling (Eimon and Harland 2002). However, our result suggests that our design may result in a non-functional
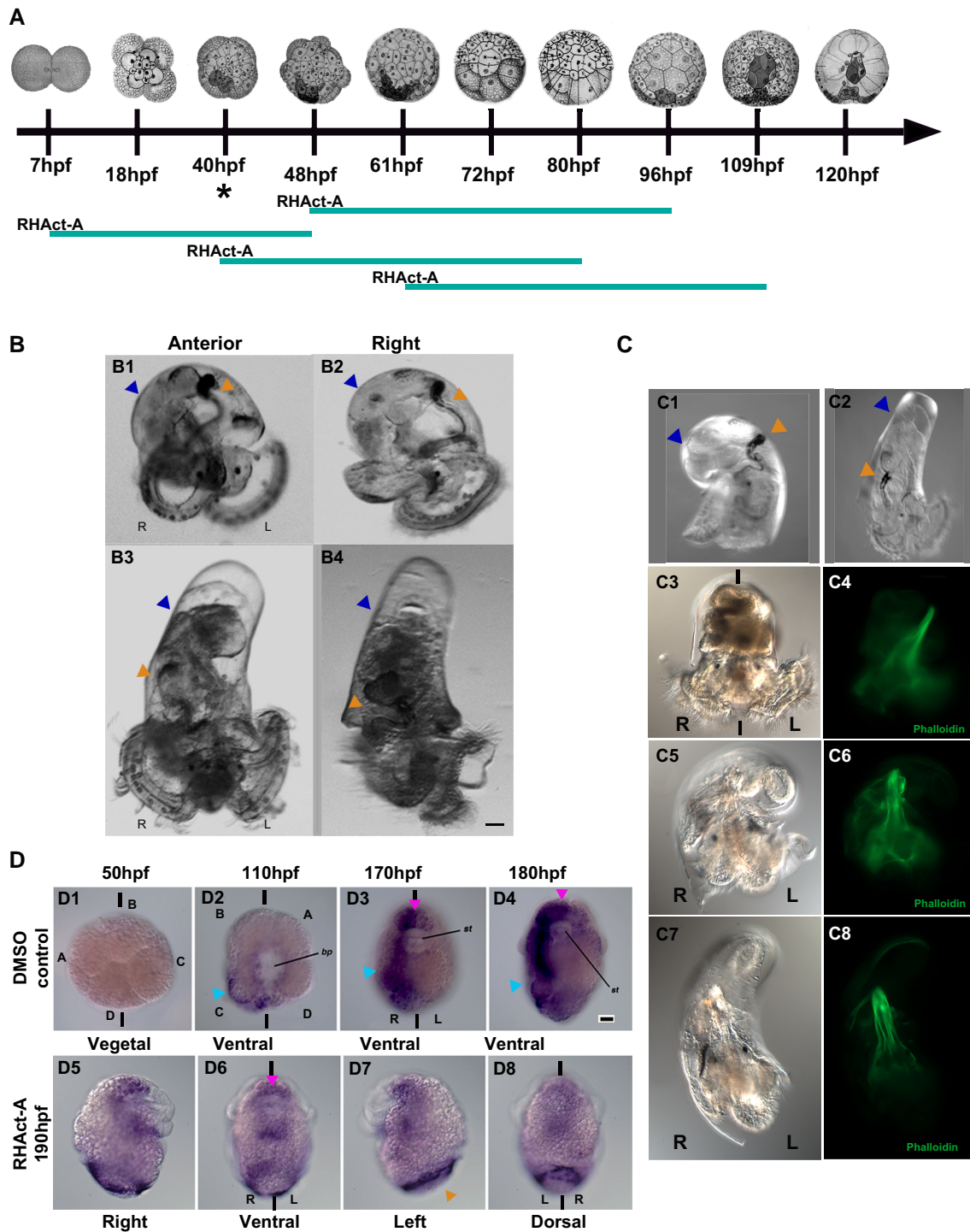
FIG. 3. *Cfo-nodalB* regulates the establishment of LRA, whereas Activin inhibits the Nodal pathway in *Crepidula fornicata*. (A) Diagram of the windows used on RHAct-A treatment. The teal bar represents the time in which the embryos were cultured in RHAct-A, and the asterisk indicates when *nodalB* expression first starts; hpf, hours post fertilization. (B) Overexpression of *Cfo-nodalB*, in a 2-week-old hatched veliger. (B1) Anterior and (B2) right views of a control veliger; (B3) anterior and (B4) right views of a veliger after *Cfo-nodalB* injection. (C) RHAct-A treatment, in a 2-week-old hatched veliger. Right view of a (C1) control veliger and (C2) a treated veliger. (C3–C8) Three individuals with different grades of coiling after treatment. Brightfield of (C3) coiled DMSO control, (C5) hemitorsioned, and (C7) straight larvae after RHAct-A. (C4/C6/C8): note the change of the retractor muscle of those same larvae (C3, C5, C7) shown by phalloidin stain. In both *B* and *C*, note the bilateral symmetry phenotype in *B3*, *B4* and *C2*, *C7*. L, left; R, right; shell coiling, blue arrowhead; gut coiling, orange arrowhead. (D) Spatio-temporal localization of *Pitx* mRNA by WISH, in RHAct-A treated *C. fornicata* embryos. (D1–D4) Control embryos in vegetal/ventral view in four stages during development. *D1*, non-stained blastula (50 hpf); *D2*, mid-gastrula (110 hpf); *D3*, organogenesis (170 hpf); *D4*, preveliger larva (180 hpf). Note the right-side staining (cyan arrowhead) and the symmetrical staining (pink arrowhead) on the anterior region and stomodeum (*D2–D4*). (D5–D8) Views of a preveliger, post-RHAct-A exposure. Note that the symmetrical expression remains (pink arrowhead), whereas the asymmetrical expression is lost on the right side; shell gland, orange arrowhead. st, stomodeum; bl, blastopore; L, left; R, right. A, B, C and D indicate the quadrants; bilateral axis is represented by the vertical bars. Scale bars correspond to 30 µm.
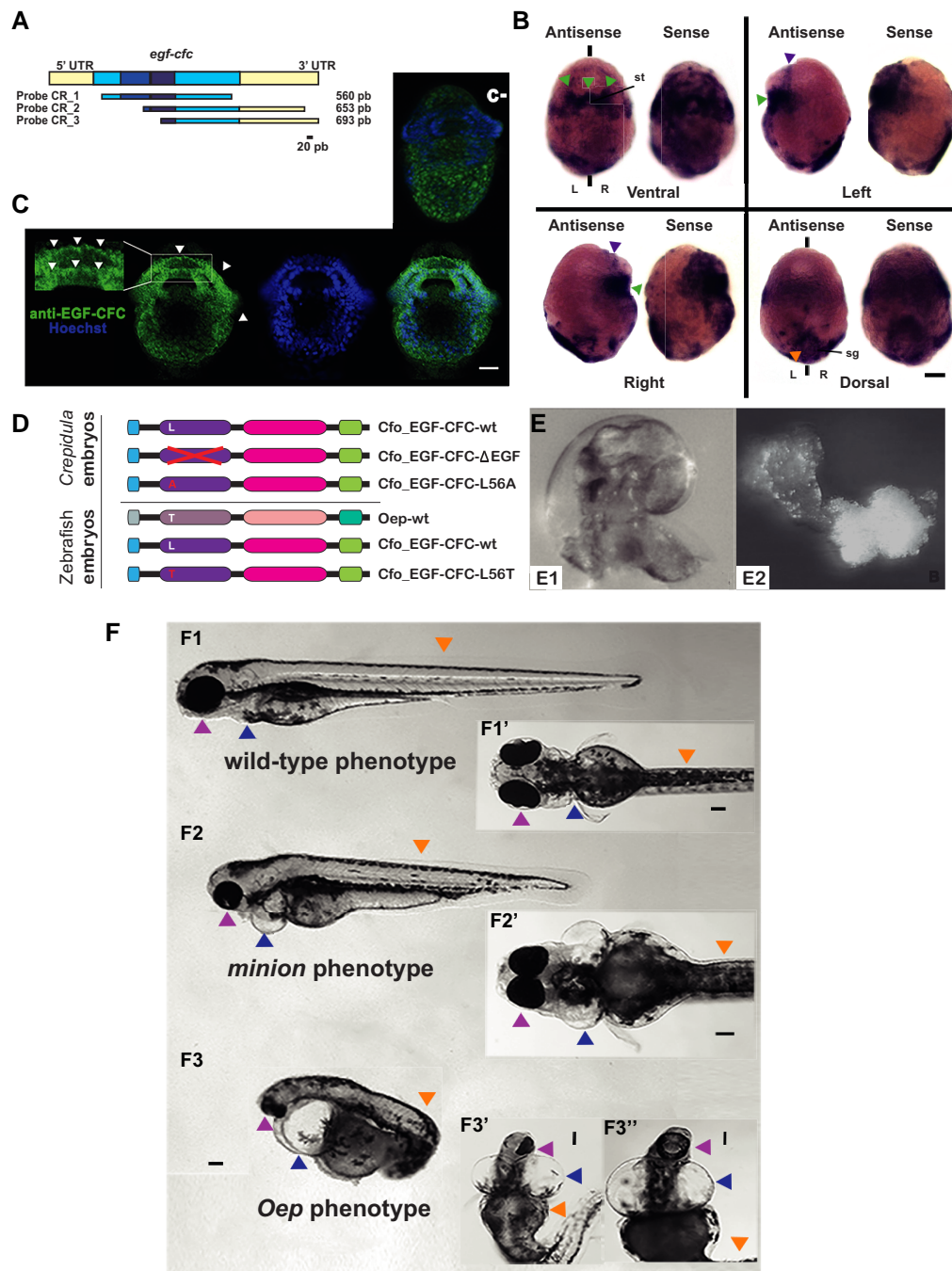
Fig. 4. Cfo-EGF_CFC is not involved in the establishment of LRA in *Crepidula fornicata*, but a single substitution is able to the rescue of Zebrafish Oep mutants. (*A*) Cfo-*egf-cfc* probe cloning designs. Regions in yellow represent the 5′- and 3′-UTRs, and blue corresponds to the coding region. Note that the two darkest fragments correspond to the EGF-like and CFC domains. Probe 1 contains most part of the coding region, whereas probes 2 and 3 contain different lengths of coding region and 3′-UTR. The bar indicates the scale. (*B*) Localization of *Cfo-egf_cfc* transcripts by WISH in a *C. fornicata* preveliger larva with probe 1. The left embryo on each view was incubated with the antisense probe, and the right one corresponds to the sense probe. Green arrowheads point to the symmetrical signal at the stomodeum; purple arrowhead, symmetrical anterior signals; orange, apparent symmetrical shell gland signals. st, stomodeum; sg, shell gland; L, left; R, right; bilateral axis is represented by the vertical bars. Scale corresponds to 30 μm. (*C*) Ventral view of an immunofluorescent staining for EGF-CFC on a *C. fornicata* preveliger larva. Green channel shows the staining from a specific antibody designed against this protein, which signal appears to be symmetrical and consistent with the result of WISH (fig. 4*B*). Magnification of the stomodeum reveals this protein in the cell membranes (white arrows). The blue channel shows nuclei after Hoechst staining; followed by the merge of both channels. An equivalent larva is provided from the negative control reaction (c−); note the lack of staining at the stomodeum. Scale corresponds to 30 μm. (*D*) mRNA constructs injected in *Crepidula* and Zebrafish embryos. (*E1*) Left side of a wild-type *C. fornicata* juvenile; (*E2*) radialized individual resulting from EGF-CFC LOF, after using the construct lacking the EGF-like domain. (*F*) Zebrafish individuals resulting from the rescue experiment on *oep* mutants, after injecting zygotes with substitution L56T on the EGF-CFC construct. (*F1*) Left and (*F1′*) ventral views of a wild-type phenotype; (*F2*) left and (*F2′*) ventral views of a minion phenotype; (*F3*) left and (*F3′* and *F3″*) ventral views of a classic *oep* phenotype. Arrowheads point to the structures affected the phenotype: Pink, lens and head; blue, heart; orange, anterior-posterior axis. Scale bars correspond to 100 μm.

product, given the lack of an overexpression phenotype, or alternatively: even the relatively low endogenous concentration of NodalB is sufficient to get a wild type.

As a third approach to manipulate the Nodal pathway, we treated the embryos with recombinant Activin. Activin is a TGF-β ligand closely related to Nodal and is able to functionally replace it *in vitro* (Namwanje and Brown 2016), but interestingly, it is not present in gastropods (Grande and Patel 2009; Grande et al. 2014). To date, no ortholog of Activin has been identified for Spiralia, but there are several Activin-like paralogs that result from duplication and divergence events of the ancestral gene (Kenny et al. 2014). Some of these paralogs have been analyzed and it has been demonstrated that these are not active during gastropod development (Grande and Patel 2009). Hence, the available evidence suggests the absence of a role in gastropod development for these Activin paralogs. Under these conditions, it is expected that the recombinant human Activin A (RHAct-A) may compete with Nodal for receptors in *C. fornicata* embryos without promoting signal transduction. RHAct-A treatments administered to two-cell stage embryos (7 hpf) to 48 hpf did not show symmetrical larvae. However, embryos treated at 40 hpf (the stage of *Cfo-nodalB* activation in *C. fornicata* according to our WISH results) presented the symmetrical body plan observed in the overexpression experiment (symmetrical shells and loss of torsion in the visceral mass), in a dose-dependent manner in up to 80% of the cases at the highest dose (fig. 3A and C, table 2). Thus, the RHAct-A treatment during this window of time leads to a symmetric body plan, presumably due to RHAct-A outcompeting the endogenous Nodal ligands (supplementary S2, Supplementary Material online). On the other hand, when we shifted the window to 48 and 61 hpf, this percentage decreased to around 60% of cases. The lack of an observable phenotype at 7 hpf treatment can be explained by the fact that *Cfo-nodalB* expression is not present until at least 40 hpf, only 8 h prior to removing the treatment (fig. 2B), and it could take some time for this ligand to reach its physiological concentration.

Analysis of *Pitx* expression revealed that RHAct-A was producing an inhibition phenotype, instead of an overexpression effect (fig. 3D). Indeed, although the symmetrical *Pitx* domain in the anterior region persists in an unaffected manner, the asymmetric domain is lost in those larvae treated with RHAct-A (fig. 3D). Additionally, the retractor muscle (fig. 3C3–C8) is straight in comparison with control larvae (fig. 3C1), suggesting a lack of torsion. Thus, the presence of the RHAct-A protein interferes with LRA in the gastropod *C. fornicata*, when applied after the 40 hpf stage, potentially acting as a dominant-negative inhibitor that interferes with Nodal signaling (supplementary S2, Supplementary Material online). *Cfo-nodalB* expression continues until later stages (200 hpf), colocalizing with the *Pitx* signal (fig. 2B and C). However, the decrease of symmetrical individuals in blastula treatments suggests that the initial activation of the pathway between 48 and 61 hpf may be key for the correct establishment of LRA.

Cfo-egf_cfc is Active During Development in *C. fornicata*, Including the Blastopore

We designed several riboprobes for *Cfo-egf_cfc* and performed WISH in *C. fornicata* embryos (fig. 4A and B). *Cfo-egf_cfc* transcripts appear to be maternally provisioned and do not present any asymmetrical expression during development. During early development, expression becomes more intense and is restricted mainly to the region of the blastopore (data not shown). During preveliger stages, expression becomes very intense and is restricted to the stomodeum and part of the velum in a symmetrical pattern (fig. 4B, green arrowhead). Although a discrete signal is also observed in cells on both left and right sides in the posterior region and in the area of the shell gland (fig. 4B, orange arrowhead), the variability between embryos did not allow us to establish a significant pattern. On the other hand, a sense probe for *Cfo-egf_cfc* also revealed expression patterns that showed little difference in localization when compared with those above (fig. 4B). This surprising result could be an artifact; however, several probes designed containing different regions of the mRNA (fig. 4A), including the 3′-UTR domain, showed the same pattern. One possible explanation could be the existence of antisense RNAs, such as natural antisense transcripts (NATs), that may be regulating the *Cfo-egf_cfc* gene product, and that may be encoded by the sense DNA strand (Su et al. 2010; Sun et al. 2017). This hypothesis should be tested in future experiments.

To further analyze *Cfo-egf_cfc* expression, we designed a specific polyclonal antibody in this species to investigate specific areas of translation. Immunofluorescence experiments revealed that although control reactions without the primary antibody do not show any signal (fig. 4C c-), the preveliger larvae present a strong symmetrical signal located only in the membranes of the cells around the blastopore and the velar rudiment (fig. 4C). This spatial distribution is consistent with that observed in WISH results (fig. 4B); thus, confirming this data.

Cfo-EGF_CFC is not Involved in the Establishment of LRA in *C. fornicata*

To investigate whether the EGF-CFC coreceptor is involved in LRA asymmetry outside Deuterostomia, we designed two approaches to interfere with its function in *C. fornicata* (fig. 4D). First, we explored gain of function (GOF) through overexpression of Cfo-EGF_CFC-wt in *C. fornicata* zygotes. No phenotypes different from the controls were observed with this approach (table 1). Second, we explored the loss of function (LOF) through two designs for defective forms, after studying the residues involved in the Nodal pathway. As indicated before, the EGF-like domain is known to be involved in binding the ligand. Initially, we designed a form of Cfo-EGF_CFC devoid of the EGF-like domain (Cfo-EGF_CFC-ΔEGF; fig. 4D). This experiment resulted in an increased mortality rate and those that survived displayed a radialized phenotype, characterized by the absence of dorso-ventral and left-

Table 3. Results Obtained After mRNA Injection of Different *egf-cfc* Constructs, in Zygotes of the *D. rerio oep* Mutant Line.

			Observed phenotype		
	mRNA & dose	Total embryos	Wild type	Oep	Minion
Control experiment (F1 for oep +/− cross)	-	54	41 (75.9%)	13 (24.1%)	0
Rescue with Zf_*oep* gene	Oep-wt (200 pg)	73	65 (89%)	8 (11%)	0
Rescue with *Crepidula* wt	Cfo_egf-cfc_wt (200 pg)	97	76 (78.4%)	21 (21.6%)	0
Rescue with *Crepidula* L56T mutant	Cfo_egf-cfc_L56T (200 pg)	39	33 (84.6%)	4 (10.3%)	2 (5.1%)

NOTE.—The first control experiment provides the baseline proportions of phenotypes within the F1 of crosses between heterozygous *oep* adults; following a classic Mendelian segregation. The second control provides the baseline of rescue using *oep* mRNA, which resulted in 89% of wild-type individuals. When zebrafish zygotes were injected with the wild-type form of *C. fornicata egf-cfc*, no rescue was observed in the resulting specimens. However, the L56T substitution of this same ortholog caused a rescue of nearly 90% of the fish. The minion phenotype is a partial rescue phenotype that strengthens the experimental results.

right axes (fig. 4E2). This result differs from those of Nodal GOF and LOF, suggesting a possible interference with other pathways that this coreceptor is known to regulate. The resulting radialized embryos, which did not undergo gastrulation, had no axes and diminished cell differentiation, are reminiscent of the phenotype obtained when the MAPK pathway is inhibited in *C. fornicata* (Henry and Perry 2008). This, together with the fact that EGF-CFC activates the MAPK pathway in vertebrates (Kannan et al. 1997), suggests that the MAPK pathway may be regulated by Cfo-EGF_CFC in *C. fornicata*. However, more experiments should be performed to test this hypothesis.

As the EGF-CFC coreceptor could be binding different ligands and it is known that specific residues are responsible for each of them, we focused on finding the orthologous amino acid to that located at position 88 (T88) in humans, which is described as critical for Nodal interaction in vertebrates. As described above, this position corresponds to a leucine in *C. fornicata* (L56) (fig. 1E). Hence, we first had to determine whether this amino acid position was as crucial for Cfo-EGF_CFC function in gastropods as T88 was in vertebrates. Finding a leucine in this relevant position made us suspect that the coreceptor may not be able to interact with the ligand Nodal, but we also considered the probability of having other interactions involving this change. Therefore, we generated a mutation at L56 by substituting the leucine with alanine L56A (Cfo-EGF_CFC-L56A; fig. 4D) and injected this mRNA in *C. fornicata* zygotes. Alanine is one of the simplest amino acids from a biochemistry perspective so we wanted to minimize the potential interactions of this residue with some other additional proteins. No phenotype different from a normal wild type was observed in the larvae suggesting that the alanine residue is equivalent to a leucine in *C. fornicata* and that this change has no impact in its LRA (table 1).

L56T Substitution Promotes a Gain of Function in EGF-CFC that can Rescue a Vertebrate Mutant Line for its Ortholog

In order to further investigate whether L56 in *Crepidula* is a true equivalent to position T88 in chordates, we reasoned that we needed to test the function of mutant proteins. The best way to test this hypothesis would be by attempting to rescue an *egf-cfc* LOF mutant by overexpressing the wild-type construct of *C. fornicata*. As we do not have such a mutant in *C. fornicata*, we turned to zebrafish (fig. 4D). An extensively characterized mutant line in the *egf-cfc* gene exists in zebrafish (*oep*). Matings of heterozygous parental individuals display a typical Mendelian segregation (table 3), in which 25% of the embryos show an *oep* phenotype (*oep* −/−; fig. 4F3, F3′, and F3″) and 75% of the embryos present a wild-type phenotype (*oep* +/−; *oep* +/+; fig. 4F1 and F1′). Oep mutants display cyclopia (presence of just one eye due to an undivided forebrain and orbital cavity), anterior-posterior axis deviation with shortening, cardiac hypertrophy and inversion of the heart (Schier et al. 1997). We first confirmed that injection of a wild-type form of zebrafish *oep* mRNA can rescue the mutant phenotype, as previously described. We synthesized the mRNA for *D. rerio oep*, and injected it into zygotes obtained from heterozygous matings (table 3). The results showed an increased number of wild-type phenotypes (89%) in comparison with the expected scoring under normal conditions (75.9%; table 3). Next, we used the two constructs designed for *C. fornicata* Cfo-EGF_CFC-wt and Cfo-EGF_CFC-L56A. Additionally, we created a third mutation, leucine to threonine (Cfo-EGF_CFC-L56T), that attempts to reproduce the form found in deuterostomes (fig. 4D). The injection results in zygotes from the same Zebrafish mating are shown in fig. 4F and table 3. Our findings showed that Cfo-EGF_CFC-wt could not rescue the *oep* phenotype (21.6% mutants; table 3), nor the presumptive defective form Cfo-EGF_CFC-L56A (data not shown). However, Cfo-EGF_CFC-L56T rescued the phenotype (10.3% mutants) in a proportion similar to wild-type zebrafish *oep* (11% mutants; table 3). In some of the cases, a partial rescue that represents an intermediate phenotype was detected and referred to as the "Minion" phenotype, which displays a significant rescue of bilaterality (fig. 4F2 and F2′). The occurrence of the Minion phenotype strengthens our hypothesis of a recovery of function in the Cfo-EGF_CFC-L56T injections, as this showed a noticeable reduction of the cardiac hypertrophy, an apparent complete recovery of the anterior-posterior axis, nearly complete development of the head and the generation of two eyes with partial cyclopia (closer set compared with those in the wild-type embryos). Overall, these results

suggest that even though an EGF_CFC ortholog is present in *C. fornicata*, the product of this gene is probably unable to modulate the Nodal pathway due to a potential inability to bind to Nodal.

Discussion

Nodal has been studied in detail for decades, and research has been extended to non-deuterostomes since its discovery in snails in 2009 (Grande and Patel 2009). The comparative analysis in this study demonstrated conservation of the Nodal pathway and its function among Bilateria. However, the regulatory aspects of the Nodal pathway in gastropods, as well as the effects on its perturbation remain unexplored. Our study provides the first evidence of the presence of an EGF-CFC coreceptor on non-deuterostome representatives, demonstrating that the EGF-CFC family traces back to at least the last common ancestor of Spiralia, Ecdysozoa, and Deuterostomia. Our functional analysis supports the hypothesis that EGF-CFC had an ancestral role unrelated to the Nodal pathway. We propose that a gain of function event driven by a single amino acid change was responsible for facilitating a new role for the coreceptor within the Nodal pathway.

Nodal and EGF-CFC were Present in the Ancestor of Bilateria

As previously reported by Grande et al. (2014), Nodal is present in most bilaterians, and two paralog genes were found for *C. fornicata* (fig. 1A). Studies focusing on the origin and molecular evolution of the EGF-CFC coreceptors identified several members of this family in deuterostomes, which led some authors to conclude that this family originated in the deuterostome ancestor (Ravisankar et al. 2011). The data presented here demonstrate the existence of the EGF-CFC proteins in spiralians (fig. 1C) tracing this coreceptor back to the last ancestor of all bilaterians. Besides *Patella vulgata*, all the non-vertebrate organisms included in the analysis present only one ortholog (fig. 1C). These results can be explained by two gene duplication and divergence events, one in *P. vulgata* and a second in the ancestor of all vertebrates. Subsequently, one of the copies in the vertebrate lineages was lost in some groups such as fish, *Gallus*, or marsupials. Additionally, broader sampling in other groups of metazoans is needed to make real inferences of the evolutionary history of these proteins.

Nodal and EGF-CFC are Expressed During Development in *C. fornicata*

As observed in previous work in *P. vulgata* (Kenny 2014), in *C. fornicata* only one *nodal* paralog displays asymmetrical expression from early stages, and remains asymmetrically active during development, whereas the second paralog does not (fig. 2A and B). Gene duplication events can lead to paralog sub-functionalization and the emergence of new functions. Further comparative analysis will provide interesting

data in relation to these events. Expression domains of *nodal* in the Bilateria seem to be highly divergent (Hamada et al. 2002; Grande and Patel 2009; Martín-Durán et al. 2016). The early asymmetrical expression of *Cfo-nodalB* is consistent for non-deuterostome clades (Grande and Patel 2009; Martín-Durán et al. 2016), but differs from the asymmetrical establishment of Nodal in deuterostomes. Indeed, in deuterostomes *nodal* expression is initially symmetrical and transitions into an asymmetrical pattern through several regulatory mechanisms (Duboc et al. 2005; Ermakov 2013). It remains unclear what mechanisms activate *nodal* expression directly in an asymmetrical way in spiralians, or whether those activators are shared with deuterostomes, and if so, how they evolved and diverged. Consequently, clear differences appear to exist related to the regulation of *nodal* expression among different bilaterians. On the other hand, we considered that *Cfo-nodalB* could be related to LRA given its gene expression pattern during development (fig. 2B).

The Nodal target *Pitx* is also asymmetrically expressed following *nodal* activation (Hamada et al. 2002; Grande and Patel 2009). Some domains of *Pitx* expression patterns are symmetric in *Crepidula* (fig. 2C), as is the case in other gastropods and vertebrates, suggesting the presence of independent enhancer areas that regulate this expression (Christiaen et al. 2005; Grande and Patel 2009). *Pitx* is the only target described so far in gastropods. Identifying additional targets will provide further insight about body plan evolution.

Our analysis revealed that *egf-cfc* is symmetrically expressed around the blastopore lip in *C. fornicata* (fig. 4B). This structure is a key signaling center during gastrulation in gastropods (Van den Biggelaar et al. 2002). Hence, these results show that the *egf-cfc* gene may have a relevant role during gastrulation in gastropods, as is the case in vertebrates (Ding et al. 1998; Chu et al. 2005). Despite our efforts and region/condition modifications, both the antisense and sense probes showed the same restricted and recurring pattern (fig. 4A and B). These results suggest that there may be an antisense and a sense transcript, both active during *Crepidula* development. NATs are RNA molecules complementary to the sense mRNA, endogenously found and active in eukaryotes and "prokaryotes". NATs have been found to influence embryonic development in some vertebrate species (reviewed in Su et al. 2010; Sun et al. 2017). Up to 72% of mRNAs have an antisense RNA in mouse and human transcriptomes (Beiter et al. 2009). It is plausible that the *egf-cfc* sense transcript in *C. fornicata* may be regulated by an antisense transcript. Nevertheless, more experiments would be necessary to properly validate this hypothesis. To overcome the limitations of our expression analysis, a specific antibody was designed for this work. Immunostainings are consistent with the WISH experiment, given the presence of signal in the membranes of the cells surrounding the blastopore lip in larvae. This signal spreads to the anterior velar rudiment region and along both sides in a symmetrical manner (fig. 4C), and supports the potential relevance of EGF-CFC during gastrulation in gastropods.
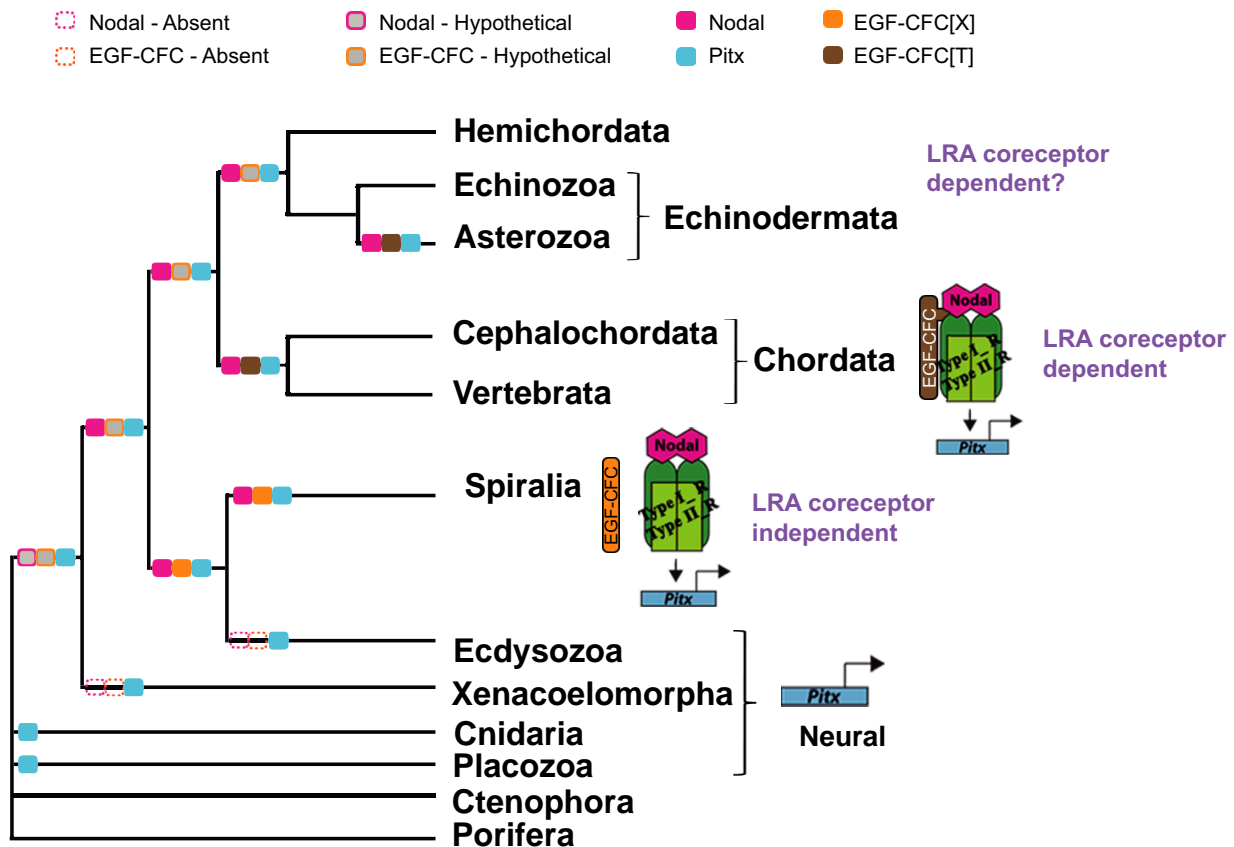
Fig. 5. Proposed model for the evolution of EGF-CFC role in Metazoa. The orthologs for EGF-CFC in spiralian representatives display different amino acids on this relevant position in the EGF-like domain, whereas all representatives of Chordata and some echinoderms display a threonine. The change of amino acid in this specific position would have happened either in the ancestor of Chordata or in the ancestor of all deuterostomes (see supplementary S3, Supplementary Material online). This gave the coreceptor the ability to interact with the ligand, and regulate *Pitx* expression, thereby participating in the establishment of LRA. In Spiralia and its ancestor, the function of EGF-CFC would not be linked to LRA, but to gastrulation and axis determination. The regulation of LRA is proposed to be independent of the coreceptor for Spiralia. Pitx ancestral function would be related to neural specification, and it would remain as the only function of this gene in those clades that lack the ligand and coreceptor. The color filled boxes represent presence of the protein in a given clade. The gray filled boxes represent inferred hypothetical presence of the protein. The dotted line boxes represent the absence of an ortholog. X, any amino acid; T, threonine; LRA, left-right asymmetry.

The Role of Nodal and EGF-CFC in Establishment of LRA

GOF and LOF experiments confirm that Nodal is involved in LRA in *C. fornicata*, as previously reported for deuterostomes (Hamada et al. 2002) and other non-deuterostomes (Grande and Patel 2009). Ectopic activation of the Nodal pathway leads to loss of asymmetric coiling shell growth, and the characteristic twist of the digestive tract (fig. 3B, table 1). This result is consistent with previous reports of loss of bilateral specialization by side identity disruption in vertebrates (Hamada et al. 2002; Duboc et al. 2005), and experiments on differential growth rate of shell asymmetrically produced by the shell gland cells (Grande and Patel 2009; Kurita and Wada 2011). On the other hand, the lack of phenotype for the LOF construct missing the cleavage site questions the functionality of the design (table 1). LOF by RHAct-A confirmed the importance of this pathway for LRA, since *Crepidula* embryos at 30 hpf treated for 2 days led to 80% of symmetric juveniles (fig. 3A and C; table 2). Moreover, these embryos did not show any asymmetrical expression of *Pitx* compared with controls (fig. 3D). These data indicate

that the RHAct-A treatment is an effective approach to block the Nodal pathway in *Crepidula*, and leads to LRA loss (supplementary S2, Supplementary Material online). The symmetric expression domains of *Pitx* remained comparable with those detected in untreated embryos (figs. 2C and 3D), and are consistent with LOF experiments previously reported for other mollusks (Grande and Patel 2009). We hypothesize that a different combination of enhancers may be controlling the expression of *Pitx* in the region symmetrically expressing the gene, or even cross-talk with another pathway could regulate its expression in these regions unaffected by RHAct-A (i.e., Wnt, Notch; Luo 2017). This Pitx-positive region could be related to the possible ancestral role of Pitx, presumably related to neural specification. The appearance of Nodal could have promoted the co-option of the *Pitx* gene for a new function, now involving it in LRA (Grande et al. 2014; fig. 5).

Our results suggest that EGF-CFC, described as a master regulator of the pathway in vertebrates (Kelber et al. 2008), does not have a role in LRA, nor does it regulate Nodal signaling in *C. fornicata*. However, the data regarding its

expression and protein location together with LOF experiments performed through EGF-CFC_Cfo_ΔEGF (fig. 4E, table 1), suggests a conserved role in gastrulation and body axis formation. The key amino acid for Nodal pathways is leucine in *C. fornicata* (fig. 1E). However, the mutation at this position did not produce any phenotype (table 1), suggesting that the ancestral regulation of Nodal signaling was independent of the coreceptor, and an EGF-CFC-dependent pathway would have evolved only in deuterostomes (fig. 5). Indeed, this is supported by previous evidence that Nodal has Cripto-independent activity in vertebrates, and the proposal of the existence of two pathways for this morphogen (a Cripto-dependent and a Cripto-independent pathway; Brennan et al. 2001; Liguori et al. 2008).

The Connection Between Nodal and EGF-CFC Emerges in Deuterostomes

We explored the ability of the *C. fornicata* ortholog to rescue the zebrafish mutant line for EGF-CFC, *oep* (fig. 4D). We reasoned that this approach would allow us to determine the conservation in the function of EGF-CFC across phyla. Overexpression of the wild-type form of *C. fornicata* EGF-CFC does not rescue the *oep* phenotype. Remarkably, changing a single amino acid (L56T) in *C. fornicata* EGF-CFC is necessary and sufficient to restore EGF-CFC activity and rescue *oep* mutants (fig. 4F, table 3). We propose that, although Nodal and EGF-CFC were present in the last ancestor of deuterostomes, spiralians, and ecdysozoans, their functional connection occurred at some point during the evolution of deuterostomes. Prior to that EGF-CFC may have had an ancestral role unrelated to Nodal, but important for gastrulation and body axis formation. A close revision of the identity of the amino acid in position T88 (humans) in several deuterostomes showed two equally parsimonious alternative hypotheses (supplementary S3, Supplementary Material online) on the origin of the mutation. Either threonine occurred in the common ancestor of all deuterostomes and was subsequently modified in the stem lineages that lead to some echinoderms and hemichordates (changing it to an isoleucine), or the change to a threonine occurred in the common ancestor of chordates and the presence of a threonine in some echinoderms is the result of convergent evolution. In any case, the presence of threonine appears to represent a change that occurred in the evolutionary history of deuterostomes and correlates with a novel function for this vital coreceptor and a new gene regulatory network for the Nodal ligand.

Materials and Methods

Experimental Model and Subject Details

All animal experiments were performed in accordance with governmental and institutional guidelines. *Crepidula fornicata*—Adults of *C. fornicata* were harvested from local waters near Woods Hole, MA by the MRC at the Marine Biological Laboratory. Embryos were collected and reared as previously described (Henry et al. 2010; Perry et al. 2015; see also https://www.life.illinois.edu/henry/tools.html). Embryos and larvae were fixed as described in Perry et al. (2015). *Danio rerio*—AB and tupl wild-type zebrafish strains and the *oep*[tz57] mutant line were maintained and bred according to standard procedures. Embryos were harvested and grown in system water at 28°C in a humid incubator. All experiments conform to the guidelines from the European Community Directive and the British and Spanish legislation for the experimental use of animals.

Sequence Analysis and Orthology Establishment

Nodal candidates for *C. fornicata* were identified previously in the laboratory (Grande et al. 2014). In order to identify the candidate EGF-CFC proteins, we searched for representatives in other metazoan clades in different databases in on NCBI (Altschul et al. 1990) and JGI databases (supplementary S1, Supplementary Material online), and we used this information in subsequent searches in different transcriptomic database. EGF-CFC orthologs for *C. fornicata* and *Biomphalaria glabrata* were retrieved from RNA-seq data sets (Perry et al. 2015; Kenny et al. 2016). Full orthology analysis methodology is described in supplementary S4, Supplementary Material online.

Cloning, Probe Synthesis, Whole-mount *in situ* Hybridization

High quality total RNA was extracted and purified as previously described (Kenny et al. 2016). cDNA was generated by using SMARTer polymerase chain reaction (PCR) cDNA Synthesis kit (Clontech). Gene-specific primers were designed for each gene of interest (supplementary S5, Supplementary Material online). PCR products were cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA). Linearized template DNA (amplified from plasmid with T7/SP6 primers) was used to synthesize RNA probes with either T7 or SP6 RNA polymerase (Life Technologies), and DIG-labeling mix (Roche, Indianapolis, IN, USA). WISH protocol was performed, as previously described (Perry et al. 2015).

mRNA Templates, Transcription, and Microinjection

*nodal*2 and *egf-cfc* paralogs of *C. fornicata* were amplified using Phusion High-Fidelity DNA polymerase (New England, Biolabs). The PCR product was digested with *Eco*RI/*Sal*I (*nodal*2) or *Bam*HI/*Xho*I (*egf-cfc*), and inserted into a previously digested pCS2+. Full construct details and microinjection methodology is described in supplementary S6, Supplementary Material online.

Antibody Production and Antibody Staining

To generate Cfo_EGF-CFC-specific antibodies, a peptide of this protein (HZ-Ahx-LSEKPCGPVPHGSWAKA-amide) was synthetized and inoculated in rabbits, by 21st Century Biochemicals, Inc. Resulting polyclonal antibodies were subsequently purified by the same company. Whole specimens were permeabilized on PBT (0.2% X-Triton100

in 1× Phosphate Buffered Saline (PBS)), and blocked in 2% Bovine serum albumin (BSA) in PBT. Sample was incubated with the primary antibody (anti-Cfo_EGF-CFC 1:10), overnight at 4°C. Several washes in PBT were performed prior incubation with the secondary antibody (goat anti-IgG rabbit 488, Molecular Probes, A-11008), at room temperature for 2 h. Nuclei were stained with Hoechst 1:1000 (Molecular Probes, H3570) for 30 min in PBT, at room temperature and darkness. Last, specimens were washed several times in PBT and final in 1× PBS. Samples were stored in 80% glycerol-PBS for imaging.

Recombinant Human Activin Treatment in *C. fornicata*

Crepidula fornicata embryos at three different stages (two-cell stage [7–8 h post fertilization; hpf], morula [32 hpf], early gastrula [48 hpf; 45% epiboly], and late blastula [61 hpf; 65% epiboly]) were treated with 5 or 10 µM human recombinant Activin A protein, or AFSW (Artificial Filtered Sea Water) with 0.1% DMSO, as a control. Embryos were cultured as previously described (Henry et al. 2010) for 48 h, after which point the protein was washed out by replacement with 100% AFSW, repeated thrice.

Microscopy

Slides were prepared as previously described in Perry et al. (2015). Brightfield imaging was carried out on a Zeiss Axioskop2 plus (Carl Zeiss, Inc., Munich, Germany). Images were captured with a CoolSnap FX color camera (Roper Scientific) running Metavue 7.10 software (Universal Imaging). Multifocal stacks were processed using Helicon Focus stacking software 6.7.1 (Helicon Soft Ltd., Kharkov, Ukraine). Confocal imaging was carried out on a LSM710 AxioImager M2 (Zeiss), running Zen 2010 software (Zeiss). Snail larvae were anesthetized with magnesium chloride, as previously described (Perry et al. 2015) and imaged with a DFC350Fx camera docked to a M205FA scope (Leica Microsystems). *Danio rerio* embryos were anesthetized in 0.01% tricaine. ImageJ v2.0.0-rc-56/1.51h and Adobe Photoshop CS5 extended v12.1 ×64 were used for processing the images.

Acknowledgments

Supplementary material

Supplementary data are available at *Molecular Biology and Evolution* online.

References

Abascal F, Zardoya R, Posada D. 2005. Prottest: selection of best-fit models of protein evolution. *Bioinformatics* **21**(9):2104–2105.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* **215**(3):403–410.

Beiter T, Reich E, Williams RW, Simon P. 2009. Antisense transcription: a critical look in both directions. *Cell Mol Life Sci.* **66**(1):94–112.

Brennan J, Lu CC, Norris DP, Rodriguez TA, Beddington RS, Robertson EJ. 2001. Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411**(6840):965–969.

Calvanese L, Marasco D, Doti N, Saporito A, D'Auria G, Paolillo L, Ruvo M, Falcigno L. 2010. Structural investigations on the Nodal-Cripto binding: a theoretical and experimental approach. *Biopolymers* **93**(11):1011–1021.

Calvanese L, Saporito A, Oliva R, D' Auria G, Pedone C, Paolillo L, Ruvo M, Marasco D, Falcigno L. 2009. Structural insights into the interaction between the Cripto CFC domain and the ALK4 receptor. *J Pept Sci.* **15**(3):175–183.

Christiaen L, Bourrat F, Joly JS. 2005. A modular cis-regulatory system controls isoform-specific pitx expression in ascidian stomodaeum. *Dev Biol.* **277**(2):557–566.

Chu J, Ding J, Jeays-Ward K, Price SM, Placzek M, Shen MM. 2005. Non-cell-autonomous role for Cripto in axial midline formation during vertebrate embryogenesis. *Development* **132**(24):5539–5551.

Ding J, Yang L, Yan YT, Chen A, Desai N, Wynshaw-Boris A, Shen MM. 1998. Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature* **395**(6703):702–707.

Duboc V, Röttinger E, Lapraz F, Besnardeau L, Lepage T. 2005. Left-right asymmetry in the sea urchin embryo is regulated by nodal signaling on the right side. *Dev Cell.* **9**(1):147–158.

Eimon PM, Harland RM. 2002. Effects of heterodimerization and proteolytic processing on Derrière and Nodal activity: implications for mesoderm induction in Xenopus. Development; **129**:3089-3103

Ermakov AS. 2013. Establishment of visceral left-right asymmetry in mammals: the role of ciliary action and leftward fluid flow in the region of Hensen's Node. *Ontogenez* **44**(5):341–356.

Foley SF, Van Vlijmen HW, Boynton RE, Adkins HB, Cheung AE, Singh J, Sanicola M, Young CN, Wen D. 2003. The CRIPTO/FRL-1/CRYPTIC (CFC) domain of human Cripto. Functional and structural insights through disulfide structure analysis. *Eur J Biochem.* **270**(17):3610–3608.

Grande C. 2010. Left-right asymmetries in Spiralia. *Integr Comp Biol.* **50**(5):744–755.

Grande C, Martín-Durán JM, Kenny NJ, Truchado-García M, Hejnol A. 2014. Evolution, divergence and loss of the Nodal signalling pathway: new data and a synthesis across the Bilateria. *Int J Dev Biol.* **58**(6–8):521–532.

Grande C, Patel NH. 2009. Nodal signalling is involved in left-right asymmetry in snails. *Nature* **457**(7232):1007–1011.

Gray PC, Vale W. 2012. Cripto/GRP78 modulation of the TGF-β pathway in development and oncogenesis. *FEBS Lett.* **586**(14):1836–1845.

Hamada H, Mano C, Watanabe D, Saijoh Y. 2002. Establishment of vertebrate left-right asymmetry. *Nat Rev Genet*. **3**(2):103–113.

Hamada H, Tam P. 2020. Diversity of left-right symmetry breaking strategy in animals. Review. *F1000 Res*. **9**:123.

Henry JJ, Collin R, Perry KJ. 2010. The slipper snail, *Crepidula*: an emerging lophotrochozoan model system. *Biol Bull*. **218**(3): 211–229.

Henry JJ, Perry KJ. 2008. MAPK activation and the specification of the D quadrant in the gastropod mollusc, *Crepidula fornicata*. *Dev Biol*. **313**(1):181–195.

Joseph EM, Melton DA. 1998. Mutant Vg1 ligands disrupt endoderm and mesoderm formation in Xenopus embryos. *Development* **125**:2677–2685.

Kannan S, De Santis M, Lohmeyer M, Riese DJ II, Smith GH, Hymes N, Seno M, Brandt R, Bianco C, Persico G, *et al.* 1997. Cripto enhances the tyrosine phosphorylation of Shc and activates mitogen-activated protein kinase (MAPK) in mammary epithelial cells. *J Biol Chem*. **272**(6):3330–3335.

Kelber JA, Shani G, Booker EC, Vale WW, Gray PC. 2008. Cripto is a noncompetitive activin antagonist that forms analogous signaling complexes with activin and nodal. *J Biol Chem*. **283**(8): 4490–4500.

Kenny NJ. 2014. The evolution and development of left/right asymmetry in the Lophotrochozoa [PhD thesis]. Oxford University, UK.

Kenny NJ, Namigai EK, Dearden PK, Hui JH, Grande C, Shimeld SM. 2014. The lophotrochozoan TGF-β signalling cassette – diversification and conservation in a key signalling pathway. *Int J Dev Biol*. **58**(6–8):533–549.

Kenny NJ, Truchado-García M, Grande C. 2016. Deep, multi-stage transcriptome of the schistosomiasis vector *Biomphalaria glabrata* provides platform for understanding molluscan disease-related pathways. *BMC Infect Dis*. **16**(1):618.

Klauzinska M, Castro NP, Rangel MC, Spike BT, Gray PC, Bertolette D, Cuttitta F, Salomon D. 2014. The multifaceted role of the embryonic gene Cripto-1 in cancer, stem cells and epithelial-mesenchymal transition. *Semin Cancer Biol*. **29**:51–58.

Kurita Y, Wada H. 2011. Evidence that gastropod torsion is driven by asymmetric cell proliferation activated by TGF-B signaling. *Bio Lett*. **7**(5):759–762.

Kuroda R, Endo B, Abe M, Shimizu M. 2009. Chiral blastomere arrangement dictates zygotic left-right asymmetry pathway in snails. *Nature* **462**(7274):790–794.

Levin M, Pagan S, Roberts DJ, Cooke J, Kuehn ME, Tabin CJ. 1997. Left/right patterning signals and the independent regulation of different aspects of situs in the chick embryo. *Dev Biol*. **189**: 57–67.

Liguori GL, Borges AC, D'Andrea D, Liguoro A, Gonçalves L, Salgueiro AM, Persico MG, Belo JA. 2008. Cripto-independent Nodal signaling promotes positioning of the A-P axis in the early mouse embryo. *Dev Biol*. **315**(2):280–289.

Lobikin M, Wangc G, Xuc J, Hsiehd YW, Chuangd CF, Lemirea JM, Levin M. 2012. Early, nonciliary role for microtubule proteins in left–right patterning is conserved across kingdoms. *PNAS* **109**(31):12586–12591.

Luo K. 2017. Signaling cross-talk between TGF-β/Smad and other signaling pathways. *Cold Spring Harb Perspect Biol*. **9**(1): a022137.

Marasco D, Saporito A, Ponticelli S, Chambery A, De Falco S, Pedone C, Minchiotti G, Ruvo M. 2006. Chemical synthesis of mouse cripto CFC variants. *Proteins* **64**(3):779–788.

Marlétaz F, Peijnenburg KTCA, Goto T, Satoh N, Rokhsar DS. 2019. A new spiralian phylogeny places the enigmatic arrow worms among gnathiferans. *Curr Biol*. **29**:312–318.e3.

Martín-Durán JM, Vellutini BC, Hejnol A. 2016. Embryonic chirality and the evolution of spiralian left–right asymmetries. *Phil Trans R Soc B*. **371**:20150411.

Massagué J. 2012. TGF-β signalling in context. *Nat Rev Mol Cell Biol*. **13**(10):616–630.

Minchiotti G, Manco G, Parisi S, Lago CT, Rosa F, Persico MG. 2001. Structure-function analysis of the EGF-CFC family member Cripto identifies residues essential for nodal signalling. *Development* **128**(22):4501–4510.

Namwanje M, Brown CW. 2016. Activins and inhibins: roles in development, physiology, and disease. *Cold Spring Harb Perspect Biol*. **8**(7):pii:a021881.

Palmquist K, Davidson B. 2017. Establishment of lateral organ asymmetries in the invertebrate chordate, *Ciona intestinalis*. *EvoDevo*. **8**:12.

Perry KJ, Lyons DC, Truchado-Garcia M, Fischer AH, Helfrich LW, Johansson KB, Diamond JC, Grande C, Henry JQ. 2015. Deployment of regulatory genes during gastrulation and germ layer specification in a model spiralian mollusc crepidula. *Dev Dyn*. **244**(10):1215–1248.

Ravisankar V, Singh TP, Manoj N. 2011. Molecular evolution of the EGF-CFC protein family. *Gene* **482**(1–2):43–50.

Raya A, Belmonte JCI. 2004. Unveiling the establishment of left–right asymmetry in the chick embryo. *Mech Dev*. **121**:1043–1054.

Schier AF. 2009. Nodal morphogens. *Cold Spring Harb Perspect Biol*. **1**(5):a003459.

Schier AF, Neuhauss SC, Helde KA, Talbot WS, Driever W. 1997. The one-eyed pinhead gene functions in mesoderm and endoderm formation in zebrafish and interacts with no tail. *Development* **124**(2):327–342.

Schiffer SG, Foley S, Kaffashan A, Hronowski X, Zichittella AE, Yeo CY, Miatkowski K, Adkins HB, Damon B, Whitman M, *et al.* 2001. Fucosylation of Cripto is required for its ability to facilitate nodal signaling. *J Biol Chem*. **276**:37769–37778.

Shen MM. 2007. Nodal signaling: developmental roles and regulation. *Development* **134**(6):1023–1034.

Shen MM, Schier AF. 2000. The EGF-CFC gene family in vertebrate development. *Trends Genet*. **16**(7):303–309.

Shi S, Ge C, Luo Y, Hou X, Haltiwanger RS, Stanley P. 2007. The threonine that carries fucose, but not fucose, is required for Cripto to facilitate Nodal signaling. *J Biol Chem*. **282**(28):20133–20141.

Shibazaki Y, Shimizu M, Kuroda R. 2004. Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. *Curr Biol*. **14**(16):1462–1467.

Soukup V, Yong LW, Lu TM, Huang SW, Kozmik Z, Yu JK. 2015. The Nodal signaling pathway controls left-right asymmetric development in amphioxus. *EvoDevo* **6**(1):1–23.

Su WY, Xiong H, Fang JY. 2010. Natural antisense transcripts regulate gene expression in an epigenetic manner. *Biochem Biophys Res Commun*. **396**(2):177–181.

Sun Y, Li D, Zhang R, Peng S, Zhang G, Yang T, Qian A. 2017. Strategies to identify natural antisense transcripts. *Biochimie* **132**:131–151.

van den Biggelaar J, Edsinger-Gonzales E, Schram FR. 2002. The improbability of dorso-ventral axis inversion during animal evolution, as presumed by Geoffroy Saint Hilaire. *Contrib Zool*. **71**(1–3):29–36.

Watanabe K, Salomon DS. 2010. Intercellular transfer regulation of the paracrine activity of GPI-anchored Cripto-1 as a Nodal co-receptor. *Biochem Biophys Res Commun*. **403**(1):108–113.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. **18**(5):691–699.

Yan YT, Gritsman K, Ding J, Burdine RD, Corrales JD, Price SM, Talbot WS, Schier AF, Shen MM. 1999. Conserved requirement for EGF-CFC genes in vertebrate left-right axis formation. *Genes Dev*. **13**(19):2527–2537.

Yan YT, Liu JJ, Luo YEC, Haltiwanger RS, Abate-Shen C, Shen MM. 2002. Dual roles of Cripto as a ligand and coreceptor in the nodal signaling pathway. *Mol Cell Biol*. **22**:4439–4449.

Yoshida K, Saiga H. 2008. Left–right asymmetric expression of Pitx is regulated by the asymmetric Nodal signaling through an intronic enhancer in *Ciona intestinalis*. *Dev Genes Evol*. **218**:353–360.