

# Supporting Information

Bertrand et al. 10.1073/pnas.1014235108

## SI Materials and Methods

**Syntenic Conservation Analysis.** Conserved syntenic regions in the genomes of *Monodelphis domestica*, *Canis familiaris*, *Gallus gallus*, *Xenopus tropicalis*, and *Tetraodon nigroviridis*, which are available in the Ensembl database (<http://www.ensembl.org/index.html>), and of *Branchiostoma floridae*, which is available at the US Department of Energy Joint Genome Institute (JGI) Web site (<http://genome.jgi-psf.org/Brafl1/Brafl1.home.html>), were obtained with CASSIOPE (9). When no statistically significant conserved regions containing an FGF gene were found in amphioxus, we searched its genome for orthologs of the genes that were found as part of the conserved regions in at least two vertebrates. Phylogenetic analyses to verify the orthology of each gene within the conserved regions were performed using RaxML version 7.0.4 with the WAG + G + I model, 100 bootstrap replicates, and the rapid bootstrapping algorithm (10). All the accession numbers for genes found in the conserved syntenic regions are given in Table S2.

**Immunostaining.** Embryos were fixed in 4% (wt/vol) paraformaldehyde as for in situ hybridization but were thereafter kept in PBS + 0.1% Tween 20 at 4° C until use. Three 10-minute washes in PBS + 0.1% Triton X100 were performed, followed by a 1-h incubation in PBS + 1% Triton X100. Embryos were blocked in PBS + 0.1% Triton X100 + 5% (vol/vol) sheep serum + 0.2% BSA for several hours. They were then incubated in blocking solution with antibodies diluted at a ratio of 1/250 [monoclonal antiacetylated-tubulin produced in mouse (T7451; Sigma) and antiphospho-histone H3 (Ser10) (06-570; Millipore)] overnight at 4° C. They were subsequently washed six times for 1 h in PBS +

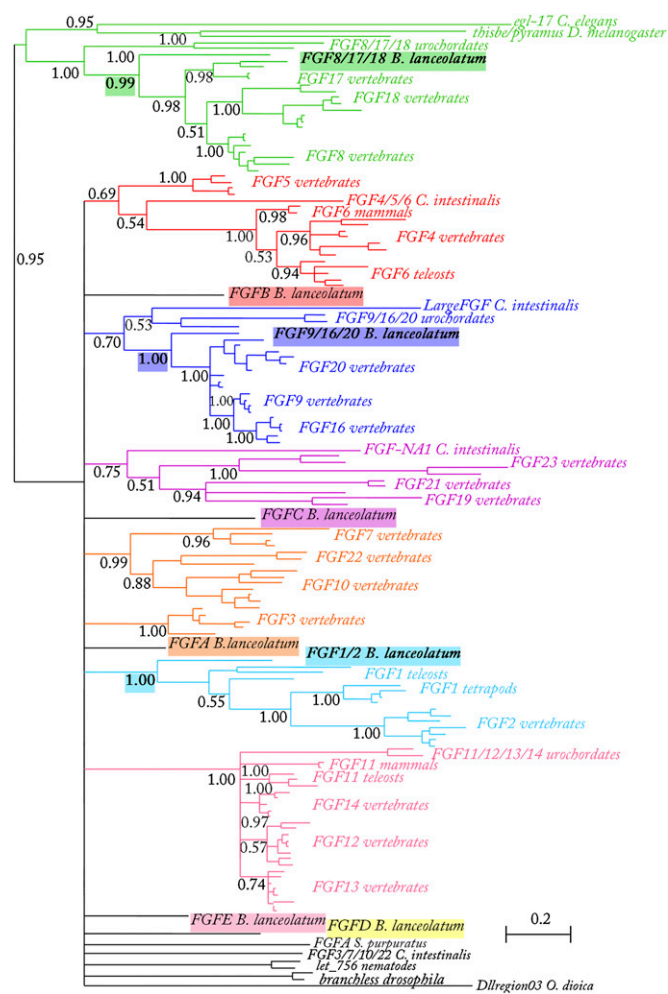
0.1% Triton X100 and blocked again for several hours. They were then incubated overnight at 4° C in the blocking solution containing secondary antibodies coupled to FITC or Texas Red at a ratio of 1/250. Embryos were then washed three times for 10 min in PBS + 0.1% Tween 20 and mounted in glycerol with 2.5% DABCO (Sigma) for photographs.

**Pharmacological Treatments.** SU5402 and U0126 were dissolved at 10<sup>-2</sup> M in DMSO. A range of concentrations, ranging from 10 to 250 μM, was tested. At 10 μM, there is no effect of SU5402 treatment, whereas at >100 μM, development is completely arrested. The highest concentration at which we observe a specific effect is 50 μM, and at this dose, we observe a complete loss of the expression of two orthologs of vertebrate FGF signaling target genes: *Dusp6/7/9* and *ER81/Erm/Pea3* (Fig. S8). We therefore subsequently performed our experiments using 50 μM SU5402. For U0126, the effect on embryogenesis is specific between 10 and 25 μM; therefore, we used 25 μM except when specified otherwise in the text.

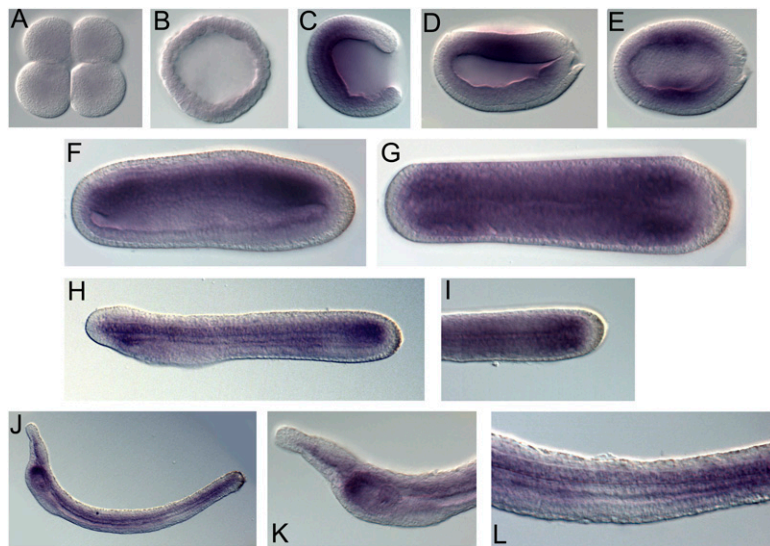
**Sections.** Sections ranging from 1 to 1.5 μm in thickness were performed after embedding in Epon resin and subsequently stained using Ponceau Red or Richardson Blue.

**Accession Numbers.** Accession numbers of the sequences used for probe synthesis are as follows: *FGF1/2* (EU606032.1), *FGF8/17/18* (EU606035.1), *FGF9/16/20* (EU606036.1), *FGFA* (EU606033.1), *FGFB* (EU606034.1), *FGFC* (EU606038.1), *FGFD* (HM854710), *FGFE* (EU606037.1), *FGFR* (HM854709), *Snail* (HM359129), and *Delta* (HM359124).

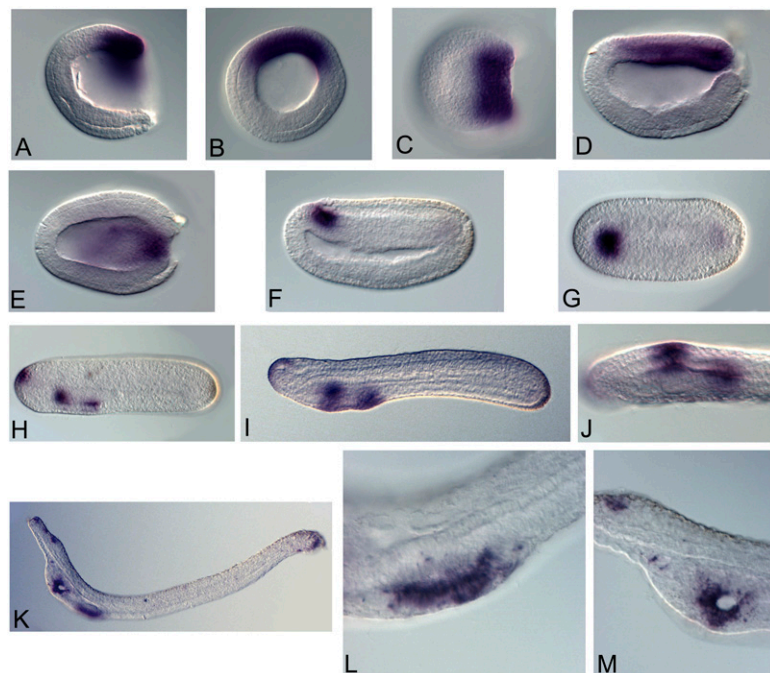
1. Letunic I, Doerks T, Bork P (2009) SMART 6: Recent updates and new developments. *Nucleic Acids Res* 37(Database issue):D229–D232.
2. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc Natl Acad Sci USA* 95: 5857–5864.
3. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*, Chapter 2:Unit 2.3.
4. Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO\_WIN: Two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* 12:543–548.
5. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
6. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
7. Yu JK, Meulemans D, McKeown SJ, Bronner-Fraser M (2008) Insights from the amphioxus genome on the origin of vertebrate neural crest. *Genome Res* 18: 1127–1132.
8. Meulemans D, Bronner-Fraser M (2007) Insights from amphioxus into the evolution of vertebrate cartilage. *PLoS ONE* 2:e787.
9. Rascol VL, et al. (2009) CASSIOPE: An expert system for conserved regions searches. *BMC Bioinformatics* 10:284.
10. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57:758–771.



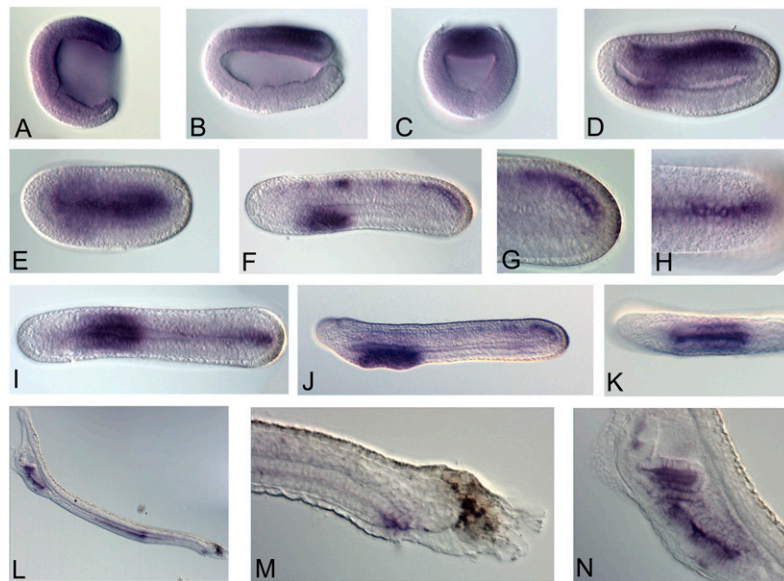
**Fig. S1.** Phylogenetic analysis of the FGF family. Only FGF domains, as predicted by the online SMART software (<http://smart.embl-heidelberg.de/>) (1, 2), were used for the alignment. Sequences were aligned automatically using ClustalX (3), with manual correction in Seaview (4). Bayesian inference trees were inferred using MrBayes 3.1.2 (5, 6), with the WAG + G + I model. Two independent runs of 1 million generations each, sampled every 100 generations with two chains, were performed. A 50 majority rule consensus tree was calculated using a 250,000-generation burn-in.



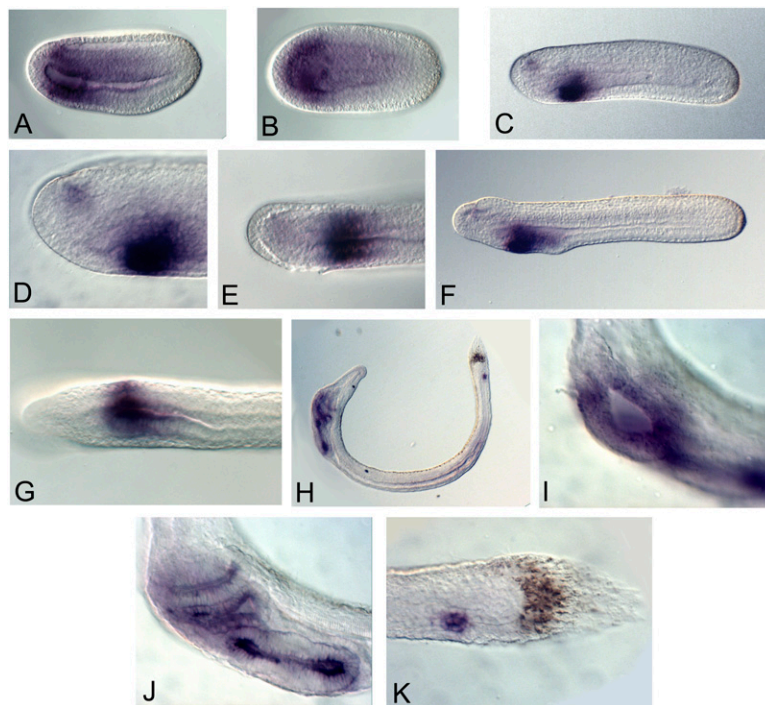
**Fig. 52.** FGFR expression pattern. The embryos were underincubated in the staining solution to detect the embryonic territories that express FGFR at a higher level. A probe for the TK domain was synthesized (accession no. HM854709). (A and B) Four-cell and blastula stage embryos showing no or very low expression of FGFR. (C) Gastrula stage embryo showing a higher level of expression in the anterior mesendoderm. (D and E) Lateral and dorsal views of a late gastrula/early neurula stage embryo. At this stage, expression is higher in the paraxial mesoderm. (F and G) Lateral and dorsal views of a midlate neurula stage embryo showing a higher FGFR expression level in the mesoderm, particularly in the most anterior and posterior somites. (H) Late neurula before the mouth opens showing a higher expression level in the notochord, the posterior somites, and the anterior pharyngeal endoderm. (I) Dorsal view of the posterior part of the specimen shown in H. (J) Larva showing a higher expression level in the notochord and the anterior pharyngeal endoderm. In the gut, the iliocolonic region is less strongly labeled than the other regions. (K) Enlargement of the anterior part of the specimen shown in J. (L) Enlargement at the level of the iliocolonic region of the specimen shown in J. Side views are shown except when specified. Anterior is to the left, and dorsal is to the top.



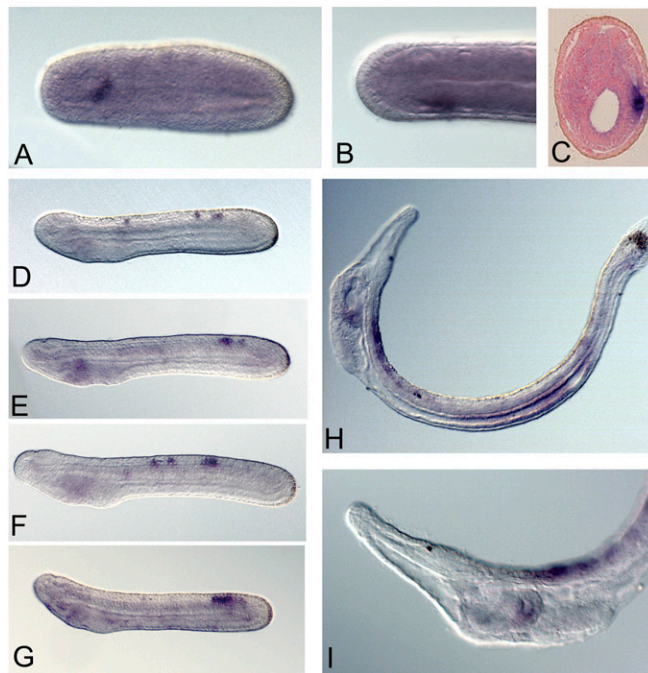
**Fig. 53.** FGF8/17/18 expression pattern. (A–C) Lateral, blastopore, and dorsal views of a gastrula stage embryo showing expression in the dorsal posterior mesendoderm. (D and E) Lateral and dorsal views of a late gastrula/early neurula stage embryo showing expression in the dorsal mesendoderm, with a higher level in the posterior part as described by Yu et al. (7). (F and G) Lateral and dorsal views of an early midneurula stage embryo showing a very high level of FGF8/17/18 expression in the anterior neural plate and a low level of FGF8/17/18 expression in the posterior mesoderm. (H) Lateral view of a midlate neurula stage embryo with expression in the anterior epidermis and in a ventral and a lateral spot in the pharyngeal endoderm as described by Meulemans and Bronner-Fraser (8). (I) Late neurula before the mouth opens showing expression in the anterior epidermis and in two regions of the pharyngeal endoderm corresponding to the mouth and first gill slit anlagen. (J) Ventral view of the anterior part of the specimen shown in I. (K) Larva showing expression of FGF8/17/18 in the cerebral vesicle, around the mouth, in the first gill slit, and in the posterior wall of the tailbud. (L) Enlargement of the specimen shown in K at the level of the first gill slit. (M) Enlargement of the specimen shown in K at the level of the mouth. Side views are shown except when specified. Anterior is to the left, and dorsal is to the top.



**Fig. 54.** FGF9/16/20 expression pattern. (A) Gastrula stage embryo with a higher level of FGF9/16/20 expression in the posterior dorsal ectoderm. (B and C) Lateral and blastopore views of a late gastrula stage embryo showing expression in the neural plate. (D and E) Lateral and dorsal views of an early midneurula stage embryo with labeling in the neural plate and the pharyngeal endoderm. (F–I) Views of a midlate neurula embryo. (F) Labeling is visible in the neural tube and in the pharyngeal endoderm. (G) Enlargement of the posterior part of the embryo shown in F. Dorsal view (I) and enlargement of the posterior part (H). (J) Late neurula embryo before the mouth opens showing expression in the pharynx and neural tube. (K) Ventral view of the anterior part of the specimen shown in J. (L) Larva showing FGF9/16/20 expression in the club-shaped gland, the first gill slit, the midgut, and the anus. Enlargement at the level of the tail (M) and at the level of the pharynx (N) of the larva shown in L. Side views are shown except when specified. Anterior is to the left, and dorsal is to the top.



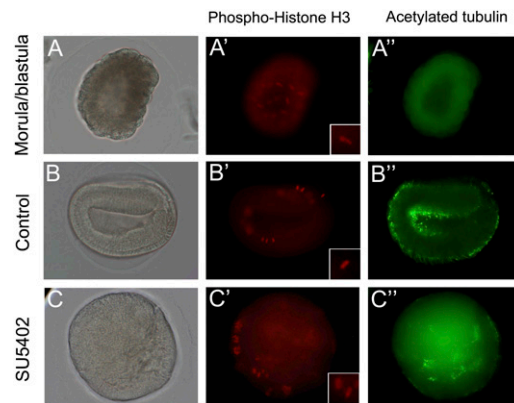
**Fig. 55.** FGFA expression pattern. (A and B) Lateral and dorsal views of an early midneurula stage embryo with expression in the anterior neural plate and in the pharyngeal endoderm. (C) Midlate neurula stage embryo showing FGFA expression in the cerebral vesicle and the ventral pharyngeal endoderm. (D and E) Enlarged lateral and ventral views of the specimen shown in C. (F) Late neurula before the mouth opens with labeling in the cerebral vesicle and in the pharyngeal endoderm. (G) Ventral view of the pharyngeal region of the specimen shown in F. (H) Larva showing FGFA expression around the mouth, in the endostyle, in the club-shaped gland, in the first gill slit, and in the anus. (I) Enlargement of the pharyngeal region focused on the mouth. (J) Enlargement of the pharyngeal region focusing on the right part. (K) Enlargement of the posterior part of the larva shown in H. Side views are shown except when specified. Anterior is to the left, and dorsal is to the top.



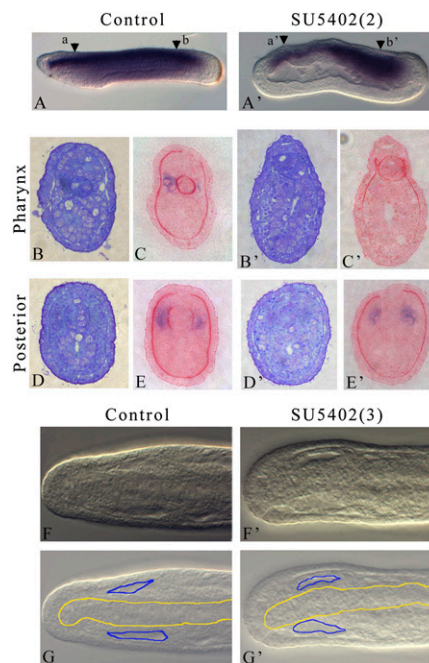
**Fig. S6.** FGFE expression pattern. (A) Midneurula stage embryo with restricted labeling in the first left somite. (B) Dorsal view of the anterior region of the specimen in A. (C) Section of the specimen shown in A at the level of the restricted labeling. (D–G) Late neurula stage embryos before the mouth opens. Dynamic expression is visible in some specific neurons. (H) Larva showing expression in the neural tube, the gut, and the club-shaped gland. (I) Enlargement of the anterior region of the larva shown in H. Side views are shown except when specified. Anterior is to the left, and dorsal is to the top.



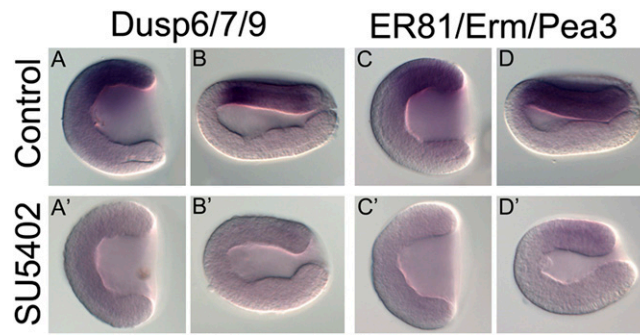
**Fig. S7.** FGFC expression pattern. (A) Late neurula stage embryo showing a higher expression level of FGFC in the anterior pharynx, in the midgut, and in the tailbud. (B) Enlargement of the anterior region of the specimen shown in A. (C) Larva with high FGFC expression level in the club-shaped gland, in the anterior most part of the pharynx, in the preoral pit, and in part of the endostyle. (D and E) Enlargement of the anterior part of the specimen shown in C focused on the mouth and on the right side, respectively. Side views are shown except when specified. Anterior is to the left, and dorsal is to the top.



**Fig. 58.** Embryos treated with SU5402 at the four- to eight-cell stage are not arrested in their development. Immunostaining experiments were performed for WT morula/blastula stage embryos and for embryos treated with SU5402 at the four- to eight-cell stage and fixed at the early midneurula stage, as well as for the corresponding control embryos. In the morula/blastula embryos (A), all the cells are dividing, as shown by antiphospho-histone H3 immunostaining (A'), but there are still no cilia, as demonstrated by the absence of labeling after antiacetylated-tubulin immunostaining (A''). In the treated embryos (C), only some patches of cells are labeled using the antiphospho-histone H3 antibody (C'), as in the control embryos (B, B'). Enlarged images of labeled nuclei are shown at the lower right in A'–C' (boxes). The antiacetylated-tubulin staining shows that in treated embryos, cilia are well developed (C''), as in control embryos (B'') (SI Materials and Methods).



**Fig. 59.** Somite morphology after treatments with SU5402. (A, A') Sections of embryos after in situ hybridization using a probe for *MLC*, stained with Richardson Blue and Ponceau Red. In late neurula stage control embryos, the somites are clearly visible and express *MLC* in the pharyngeal (B and C, sections at the level of a in A) and posterior (D and E, sections at the level of b in A) regions. In the SU5402-treated (treatment 2; Fig. 3A) late neurula stage embryos, no somites are formed in the pharyngeal region (B' and C', sections at the level of a' in A'), whereas in the posterior region, somite morphology and *MLC* expression are normal (D' and E', sections at the level of b' in A'). (F, F' and G, G') differential optical contrast images of the anterior part of late neurula stage embryos, dorsal views. The anterior somite cavities are clearly visible in the control embryos (F and G) as well as in SU5402-treated embryos (treatment 3; Fig. 3A) (F' and G'). (G, G') Somite cavities are encircled in yellow and notochord in blue.



**Fig. S10.** SU5402 treatment at 50  $\mu\text{M}$  abolishes expression of Dusp6/7/9 and ER81/Erm/Pea3. Embryos were treated from blastula to gastrula or to early neurula stage with 50  $\mu\text{M}$  of SU5402. In situ hybridization experiments were performed using probes for Dusp6/7/9 (HM359125) and ER81/Erm/Pea3 (HM359126) on treated and control embryos. (A–D) At the gastrula and early neurula stages, Dusp6/7/9 and ER81/Erm/Pea3 are expressed in the dorsal mesendoderm and ectoderm in control embryos. (A'–D') In SU5402-treated embryos, expression is totally absent at both stages even after several days of incubation in the staining solution. Dusp6/7/9 and ER81/Erm/Pea3 are the orthologs of well-known FGF pathway target genes in vertebrates. This experiment leads us to suggest, first, that both genes are FGF signaling target genes and, second, that our SU5402 treatment at 50  $\mu\text{M}$  is able to inhibit the pathway completely.

**Table S1. Characteristics of eight predicted amphioxus FGF proteins**

Name	Protein domain	Exon–intron
Amphioxus FGF1/2	FGF 35–163, 1.86e-44	1–205 // 206–310 // 311–501
Amphioxus FGF8/17/18	Signal peptide 1–31/FGF 55–181, 1.39e-25/low complexity 200–217	1–41 // 42–80 // 81–262 // 263–369 // 370–654
Amphioxus FGF9/16/20	Low complexity 4–19/FGF 48 179, 5.35e-68	1–244 // 245–346 // 347–591
Amphioxus FGFA	Signal peptide 1–16/FGF 56–189, 2.90e-55/low complexity 206–215	1–265 // 266–369 // 370–651
Amphioxus FGFB	Signal peptide 1–24/FGF 55–185, 4.75e-51	1–98 // 99–262 // 263–564
Amphioxus FGFC	Signal peptide 1–20/FGF 60–193, 1.70e-38/low complexity 210–232	1–279 // 280–385 // 386–699
Amphioxus FGFD	Signal peptide 1–26/FGF 51–199, 6.54e-31/low complexity 202–213	1–258 // 259–393 // 394–687
Amphioxus FGFE	FGF 61–192, 9.71e-57	1–283 // 284–386 // 387–582

Protein domains of the eight amphioxus FGFs, predicted by SMART (<http://smart.embl-heidelberg.de/>), are indicated (including the e value for the FGF domain). The exon-intron structure of the coding sequence of the eight amphioxus FGFs is also indicated.

**Table S2. Accession number of the genes in the conserved syntenic regions**

Vertebrates paralogues, abbreviations	Vertebrates paralogues, names	Homo sapiens	Monodelphis domestica	Canis familiaris	Gallus gallus	Xenopus tropicalis	Tetraodon nigroviridis	Proorhtholog name	Branchiostoma floridae
FGF1/2	BBS7	ENSG000000138686	ENSMODG00000018927	ENSCAFG00000004128	ENSGALG00000011880	ENSGALG00000011880	ENSTNIG0000000013049	BBS7	113523
	GRXCR1	ENSG00000005203	ENSMODG00000024820	ENSCAFG00000003537	ENSGALG00000023032	ENSGALG00000023032	ENSTNIG0000000015736	GRXCR1/2	63706
	GRXCR2	ENSG00000004928	ENSMODG00000023740	ENSCAFG00000006236	ENSGALG00000023774	ENSGALG00000023774	ENSTNIG0000000018212	KIAA0141	117215
	KIAA0141	ENSG00000001791	ENSMODG00000010456	ENSCAFG00000003988	TBLASTN	TBLASTN	ENSTNIG0000000014206	SPATA5	63613, 117082
	SPATA5	ENSG00000005375	ENSMODG00000018895	ENSCAFG00000003988	ENSGALG00000011833	ENSGALG00000011833	ENSTNIG0000000013064	SPRY	117102
	SPRY1	ENSG00000004056	ENSMODG00000018890	ENSCAFG00000003979					
	SPRY2								
	SPRY3								
	SPRY4	ENSG00000007678	ENSMODG00000023756	ENSCAFG00000006250	ENSGALG00000007336	ENSGALG00000007336	ENSTNIG0000000018206	C19orf29	121714
FGF3/7/10/22	C19orf29	ENSG00000005298	ENSMODG00000006837	ENSCAFG00000009208					
	LMNA							Nuclear Lamin	121709
	LMNB1	ENSG00000003368			ENSGALG00000006083	ENSGALG00000006083			
	LMNB2	ENSG00000006619	ENSMODG00000004824	ENSCAFG00000009389	ENSGALG00000000470	ENSGALG00000000470			
	PIAS1	ENSG00000003800		ENSCAFG00000007463	ENSGALG00000007970	ENSGALG00000007970	ENSTNIG0000000009549	PIAS1/2/3/4	194147, 279826
	PIAS2				ENSGALG00000001843	ENSGALG00000001843			
	PIAS3								
	PIAS4	ENSG00000005229	ENSMODG00000000765	ENSCAFG00000009154	ENSGALG00000001836	ENSGALG00000001836	ENSTNIG0000000003243		



**Table S2. Cont.**

Vertebrates paralogues, abbreviations	Vertebrates paralogues, names	<i>Homo sapiens</i>	<i>Monodelphis domestica</i>	<i>Canis familiaris</i>	<i>Gallus gallus</i>	<i>Xenopus tropicalis</i>	<i>Tetraodon nigroviridis</i>	Proortholog name	<i>Branchiostoma floridae</i>
FGF4/5/6	C12orf4	ENSG00000004	ENSMODG000000	ENSACFG00000001	ENSGALG000000	ENSXETG000000	ENSTNIG000000008	C12orf4	126642, 283171
	Uncharacterized protein C12orf4	7621	18299	5354	17289	05966	813		
CD9	CD9 molecule	ENSG00000001	ENSMODG000000	ENSACFG00000001	ENSGALG000000		ENSTNIG000000019	CD9/81/TSPAN2	96598
		0278	18053	5172	17274		223, ENSTNIG000000012		
CD81	CD81 molecule	ENSG00000001		ENSACFG00000001	ENSGALG000000		ENSTNIG000000013		
		0651		0121	06546		685		
TSPAN2	Tetraspanin 2	ENSG00000011	ENSMODG000000	ENSACFG00000001	ENSGALG000000		ENSTNIG000000011	FoxM	232011
FoxM1	Forkhead box M1	1206	18349	5642	013420		394		
TSPAN4	Tetraspanin 4	ENSG00000021		ENSACFG00000002	ENSGALG000000			TSPAN4/9/CD53	231999
		4063		5574	06837				
TSPAN9	Tetraspanin 9	ENSG00000001	ENSMODG000000	ENSACFG00000001	ENSGALG000000		ENSTNIG000000019		
		1105	18335	5466	14346		258, ENSTNIG000000012		
CD53	CD53 molecule	ENSG00000010		ENSACFG00000000	ENSGALG000000		ENSTNIG00000004023	CCNJ/JL	93075
CCNJ	Cyclin J	7443		8454	06955				
CCNJL	Cyclin J-like	ENSG00000013	ENSMODG000000		ENSGALG000000		ENSTNIG00000004651		
		5083	07987		01455				
GPAM	Glycerol-3-phosphate acyltransferase, mitochondrial	ENSG00000011	ENSMODG000000	ENSACFG00000001	ENSGALG000000			GPAM	126875, 126887
		9927	10326	0868	08795				
GPAT2	Glycerol-3-phosphate acyltransferase 2, mitochondrial		ENSMODG000000						
			08088						
HPSE	Heparanase	ENSG00000017	ENSMODG000000	ENSACFG00000000	ENSGALG000000			HPSE1/2	931000
HPSE2	Heparanase 2	2987	03602	9386	07508				
PEBP4	Phosphatidylinositol 4-OH kinase class 4	ENSG00000013	ENSMODG000000	ENSACFG00000000	ENSGALG000000			PEBP4	126894
		4020	09235	9270					
RHOBTB1	Rho-related BTB domain containing 1	ENSG00000007	ENSMODG000000	ENSACFG00000001	ENSGALG000000			RHOBTB1/2	93088
		2422	17097	2932	03073				
RHOBTB2	Rho-related BTB domain containing 2	ENSG00000000	ENSMODG000000	ENSACFG00000000					
		8853	09229	9245					



