eLIFE elifesciences.org

RESEARCH ARTICLE

The transcription factor Pitx2 positions the embryonic axis and regulates twinning

Angela Torlopp[1†‡], Mohsin A F Khan[1†§], Nidia M M Oliveira[1], Ingrid Lekk[1], Luz Mayela Soto-Jiménez[1,2¶], Alona Sosinsky[3#], Claudio D Stern[1*]

[1]Department of Cell and Developmental Biology, University College London, London, United Kingdom; [2]Programa de Ciencias Genómicas, Universidad Nacional Autónoma de México, Morelos, Mexico; [3]Institute of Structural and Molecular Biology, Birkbeck College, University of London, London, United Kingdom

***For correspondence:** c.stern@ucl.ac.uk

[†]These authors contributed equally to this work

Present address: [‡]Cardiovascular Research Institute, University of California, San Francisco, San Francisco, United States; [§]Cardiology and Thorax Surgery, Faculty of Medical Sciences, University of Groningen, Groningen, Netherlands; [¶]Stark Lab, Research Institute of Molecular Pathology, Vienna, Austria; [#]Clinical Genome Informatics Facility, Imperial Centre for Translational and Experimental Medicine, Imperial College, London, United Kingdom

Competing interests: The authors declare that no competing interests exist.

Funding: See page 21

Received: 22 June 2014
Accepted: 14 November 2014
Published: 12 December 2014

Reviewing editor: Marianne E Bronner, California Institute of Technology, United States

Abstract Embryonic polarity of invertebrates, amphibians and fish is specified largely by maternal determinants, which fixes cell fates early in development. In contrast, amniote embryos remain plastic and can form multiple individuals until gastrulation. How is their polarity determined? In the chick embryo, the earliest known factor is cVg1 (homologous to mammalian growth differentiation factor 1, GDF1), a transforming growth factor beta (TGFβ) signal expressed posteriorly before gastrulation. A molecular screen to find upstream regulators of cVg1 in normal embryos and in embryos manipulated to form twins now uncovers the transcription factor Pitx2 as a candidate. We show that Pitx2 is essential for axis formation, and that it acts as a direct regulator of cVg1 expression by binding to enhancers within neighbouring genes. Pitx2, Vg1/GDF1 and Nodal are also key actors in left–right asymmetry, suggesting that the same ancient polarity determination mechanism has been co-opted to different functions during evolution.
DOI: 10.7554/eLife.03743.001

Introduction

In most invertebrates and anamniote vertebrates (fishes and amphibians), embryonic polarity is first established by localisation of maternal determinants in the cytoplasm and/or cortex of the fertilised egg. This generates differences between the blastomeres that will form by cell division from the egg, and which will culminate in specifying the orientation of the embryonic axes (***Wilson, 1898***). Separation of the first two blastomeres can lead to twinning: the formation of genetically identical, complete individuals (***Driesch, 1892***). Separation of blastomeres after the four-cell stage, however, does not generate twins; in most cases it interferes with development of even a single embryo owing to the removal of important determinants that have by then segregated to different cells. This is known as the *mosaic* mode of development. Among the vertebrates, amniotes (birds and many mammals, and possibly also reptiles) have a remarkably extended capacity to give rise to twins. Some species of the armadillo genus Dasypus generate quadruplets or octuplets from a single fertilisation event, as a result of two or more sequential 'splitting' events of the embryo at a stage when it is already highly multicellular (***Newman and Patterson, 1910***; ***Loughry et al., 1998***; ***Enders, 2002***; ***Eakin and Behringer, 2004***). Conjoined ('Siamese') twins occur in mammals including humans (***Chai and Crary, 1971***; ***Vanderzon et al., 1998***; ***Kaufman, 2004***) and are also seen in reptiles (***Cunningham, 1937***) and birds (***Ulshafer and Clavert, 1979***); most of these are thought to arise from splitting of the embryo relatively late in development (***Kaufman, 2004***). Perhaps the most dramatic example is seen in the chick, where cutting an embryo into fragments at the blastoderm stage (when the embryo contains as many as 20,000–50,000 cells) can lead to each fragment generating a complete embryo; up to eight embryos

eLife digest In warm-blooded animals, including chickens and humans, a single embryo can give rise to several separate individuals (identical twins). Some species of armadillos routinely give birth to quadruplets in this way—and in experiments, up to eight identical chick embryos can be produced by cutting one embryo into smaller pieces (a type of 'experimental twinning'). This ability of a developing embryo to subdivide into separate individuals ends when the embryo starts to form its first midline structure, called the 'primitive streak'. This is the first line of symmetry and defines where the head–tail axis will later develop.

The steps that establish the axes of the embryo in birds and mammals, and the factors that prevent further splitting of the embryo to form twins after this point, are only just beginning to be understood. In chick embryos, the production of a protein called cVg1 is the first known step and precedes the development of a line of symmetry. A similar protein is produced in mammalian embryos and both proteins are members of an important family of signalling proteins.

Now, Torlopp, Khan et al. have used a combination of techniques to search for other proteins that that control the production of the cVg1 protein. Genes that are active in the region of the embryo that will express cVg1 later in development were identified, both in normal embryos and during the process of experimental twinning. This search revealed Pitx2 as a protein that acts to switch on the expression of the gene that encodes cVg1. When the Pitx2 protein is removed, the embryonic axis forms from the opposite side.

Next, Torlopp, Khan et al. searched the chicken genome to identify stretches of DNA around the cVg1 gene where proteins that regulate gene expression might bind. Six potential sites were found, including four to which Pitx2 can bind. Further experiments confirmed that two of these regulatory sequences encourage the expression of the cVg1 gene at its correct position in the embryo.

Pitx2 and related proteins were known to be involved with the development of left–right symmetry later in development; the findings of Torlopp, Khan et al. reveal, unexpectedly, that these proteins are also involved in first establishing the position at which the midline of the embryo will arise. It remains unclear what prevents most embryos from forming twins. But Torlopp, Khan et al.'s findings could help to explain some strange observations, made long ago, about left–right asymmetry in identical twins. For example, they could help explain why one of the twins in an identical twin pair is more likely to be left-handed than an individual in the general population, and why the direction of whorls of hair on the back of the head is often mirrored between identical twins.

DOI: 10.7554/eLife.03743.002

have been generated from a single blastoderm by experimental splitting, right up to the time of appearance of the primitive streak (**Lutz, 1949**; **Spratt and Haas, 1960**). The ability of higher vertebrate embryos to retain a *regulative* model of development until such a late stage strongly suggests that localisation of maternally inherited determinants is not an essential component of the mechanisms specifying embryo polarity (**Stern and Downs, 2012**). Moreover, since a single blastoderm can generate multiple embryos, mechanisms must exist that suppress this ability in regions of the embryo that do not normally initiate axis formation (**Bertocchini and Stern, 2002**; **Bertocchini et al., 2004**).

In chick embryos, the earliest symmetry breaking event known is the localised expression of *cVg1*, the chick orthologue of mammalian growth differentiation factor 1 (*GDF1*)—a member of the transforming growth factor beta (TGFβ) superfamily of secreted proteins—encoding a Nodal/Activin-type molecule that signals through Smad2/3 (**Weeks and Melton, 1987**; **Thomsen and Melton, 1993**; **Kessler and Melton, 1995**; **Seleiro et al., 1996**; **Shah et al., 1997**; **Kessler, 2004**; **Birsoy et al., 2006**; **Chen et al., 2006**; **Andersson et al., 2007**). Before primitive streak stages, *cVg1* is expressed in the posterior marginal zone (PMZ), an extraembryonic region adjacent to where the primitive streak will form; misexpression of *cVg1* in other (anterior or lateral) parts of the marginal zone is sufficient to induce a complete axis from adjacent embryonic cells (**Seleiro et al., 1996**; **Shah et al., 1997**; **Skromne and Stern, 2001**, **2002**). The mechanisms that position *cVg1* in the PMZ are unknown. Moreover, when a blastoderm is cut in half at right angles to the future primitive streak axis, *cVg1* expression spontaneously initiates in the marginal zone adjacent to the cut edge, in either the right or left side at equal frequency, foreshadowing the appearance of the primitive streak a few hours later (**Bertocchini et al., 2004**). This observation shows that the mechanisms that position *cVg1* are active in the

blastoderm stage embryo. Here we take advantage of these observations to design a molecular screen for new genes involved in the earliest stages of specifying embryo polarity; together with bioinformatic analysis and embryological experiments we identify the transcription factor *Pitx2* as a direct and essential regulator of *cVg1* expression both during normal development and in embryonic regulation (induced twinning).

Results

A molecular screen to identify upstream regulators of cVg1 uncovers Pitx2

To search for putative upstream regulators of *cVg1*, we took advantage of two of its properties: that it is expressed in the PMZ at early stages of development and that when a blastoderm is cut in half at right angles to the axis of the future primitive streak, *cVg1* expression is initiated stochastically on either the left or the right corner (adjacent to the cut edge) of the isolated anterior half (***Bertocchini et al., 2004***). We therefore performed two screens. First, we dissected the PMZ and an equivalent anterior explant (anterior marginal zone, AMZ) from 40 embryos (in triplicate) and analysed their transcriptomes using Affymetrix microarrays (***Figure 1A–B***, ***Figure 1—figure supplement 1***). At this stage of development, it is not possible to predict the polarity of the embryo with complete certainty. To prevent contamination of the samples, we designed a verification strategy by which the predicted posterior and anterior explants were collected, the rest of the embryo immediately fixed and then processed for in situ hybridisation for *cVg1*, developing the colour reaction for long enough to detect residual *cVg1* expression around the posterior explant site. From each set of 40 embryos, approximately 36 had been dissected correctly; the explants from the remainder (3 × 4) were discarded (***Figure 1—figure supplement 1***). Each set of verified PMZs and AMZs (3 × 36 of each) was then pooled and run on Affymetrix 30K chicken microarrays.

Next, we used a similar strategy for isolated anterior halves of embryos. It was previously reported that *cVg1* starts to be expressed in either the left or right corner of the cut anterior half around 6 hr after bisection (***Bertocchini et al., 2004***). We chose to collect the left and right corners of the marginal zone adjacent to the cut edge 7 hr following bisection to ensure that surrounding *cVg1* expression could be detected after excision of the fragment. To confirm that all embryos had been cut at exactly right angles to the future axis, the posterior half of each embryo was fixed immediately after cutting and subjected to in situ hybridisation for *cVg1*. The anterior half was cultured for 7 hr, the left and right corners of the margin were dissected, and the remainder was fixed and processed for *cVg1* expression (with extended reaction time to detect weak expression). We estimated that 70 explants would be needed for microarray analysis of each sample; this was done in triplicate (***Figure 1C–D***; ***Figure 1—figure supplement 2***). Samples were designated '*cVg1*-like' or '*cVg1*-unlike' based on this and pooled accordingly. This strategy randomised any left–right asymmetric genes unrelated to axial polarity and regulation, and enriched those for the cells in which *cVg1* was just starting to be expressed de novo in one sample, and their contralateral equivalents (not expressing *cVg1*) in the other. RNA from the pooled explants (approximately 3 × 63 of each type) was analysed using Affymetrix chicken microarrays.

The intersection between the two datasets from the above screens was used to identify genes co-regulated with *cVg1*, as well as those that are enriched in equivalent regions *not* expressing *cVg1*, both in normal embryos and during regulation (***Figure 1B,D,E*** and ***Figure 1—figure supplement 3***). Using a threshold of just 1.2-fold change and $p < 0.05$, this strategy identified 122 sequences (corresponding to 85 genes) with putative 'cVg1-like' expression (a c*Vg1*-synexpression group) and 78 sequences (52 genes) expressed more highly in the 'Vg1-unlike' explants (cVg1 negative) . A list of the top common genes ranked by fold change is shown in ***Table 1***. Comparison of the top 'Vg1-like' candidates from whole embryos (PMZ vs AMZ) with their counterparts from half-embryos shows highly significant correlation (Spearman's rank Rho = 0.73; p = 0.00036). Confirming that the screen was performed appropriately, *cVg1* (incorrectly annotated as *GDF3* instead of *GDF1* in the current version of the chicken genome, Galgal4) itself appears among the top genes: it is upregulated 4.3 fold, with a p value of 0.00006 (rank 11) in whole embryos, and 1.71, p = 0.0062 in the isolated anterior half (rank 21). Among all genes, *Pitx2* immediately stands out as the best candidate, being very strongly co-regulated with *cVg1* and the top transcription factor on the list. In the PMZ of whole embryos *Pitx2* is upregulated almost 10-fold compared to the AMZ explants (three different probes, ranking two, three and eight on the list; p = 0.0001, 0.00019 and 0.003 respectively), whereas in the cut

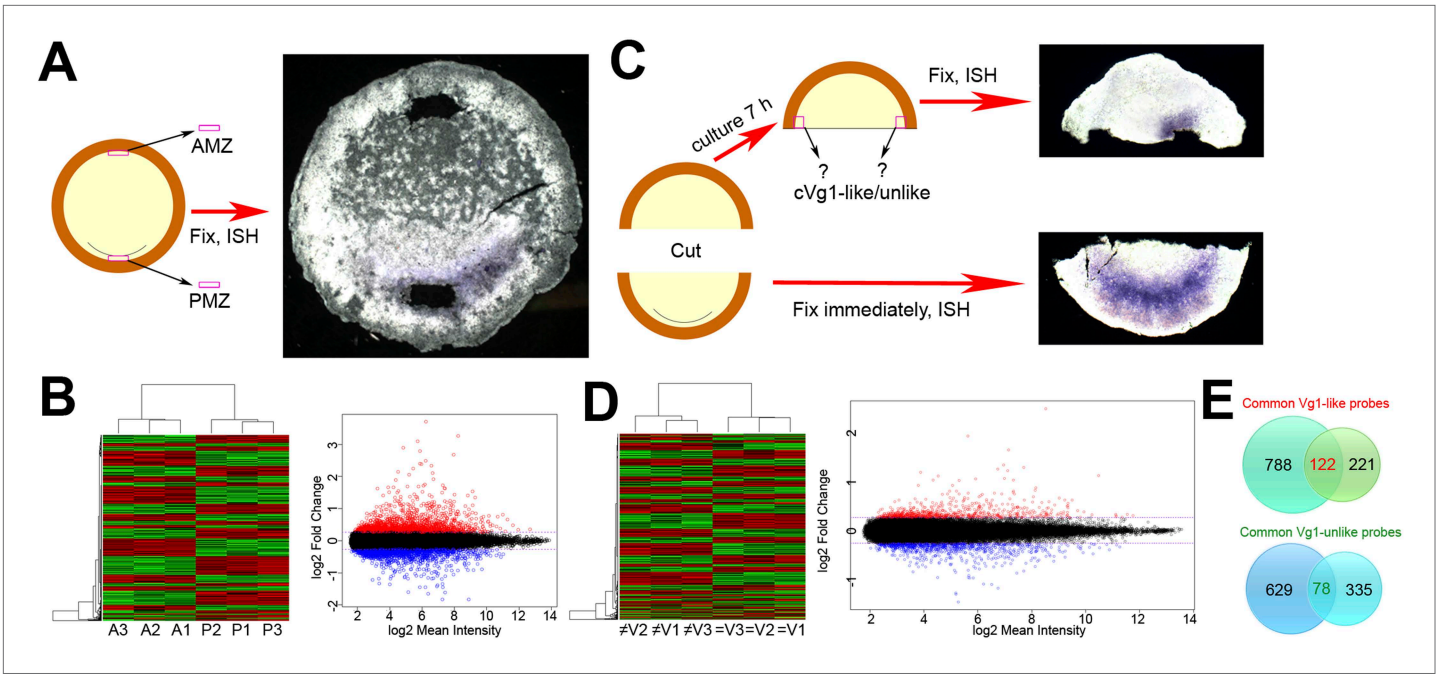

Figure 1. Microarray screens for upstream regulators of cVg1. (**A**) Diagram of the first screen: the posterior marginal zone (PMZ) and anterior marginal zone (AMZ) were dissected from embryos at stage XI–XII; the remaining embryo was then fixed and stained for *cVg1* by in situ hybridisation (ISH) to confirm that the explants had been obtained from the correct regions. This was done from 40 embryos for each of three biological replicates, which were then run on microarrays. The diagram is accompanied by an example of an embryo after ISH. All 120 embryos are shown in ***Figure 1—figure supplement 1***. (**B**) Hierarchical clustering of differentially expressed genes for this experiment, and a plot of where *cVg1*-like probes (enriched in PMZ) are displayed in red and *cVg1*-unlike ('downregulated') probes shown in green across triplicate samples (A1–A3 for AMZ, P1–P3 for PMZ). The scatter plot relates normalised $\log_2$ mean signal intensities and $\log_2$ fold changes of probes from both samples (AMZ and PMZ). Probes identified as upregulated in the PMZ with a $\log_2$ fold change cut-off of 0.263 (linear fold change 1.2) are displayed in red and those identified as downregulated in the PMZ with the same cut-off are displayed in blue. (**C**) Diagram of the second screen. An embryo at stage XI–XII was cut in half at a right angle to the future midline; the posterior half was fixed for ISH with *cVg1* to confirm the orientation (an example is shown), and the isolated anterior half cultured for 7 hr. At this point, a small explant was dissected from the marginal zone adjacent to the left and right side of the cut, and the remaining anterior half-embryo fixed for ISH with *cVg1* (an example is shown). This allowed identification of the '*cVg1*-like' and '*cVg1*-unlike' explants, which were then pooled appropriately. This was done for 70 embryos for each of three biological replicates; all 210 posterior and anterior fragments are shown in ***Figure 1—figure supplement 2*** after ISH for *cVg1*. (**D**) Hierarchical clustering of the probes expressed differentially in this assay, and corresponding scatter plot; details similar to (**B**) ≠V1, ≠V2, and ≠V3 correspond to each of the triplicate samples that do not express *cVg1* and = V1, =V2, and = V3 correspond to explants that express *cVg1*. (**E**) Venn diagrams showing the intersection of upregulated and downregulated probes common to both the PMZ and isolated anterior cut halves. A total of 122 upregulated probes and 78 downregulated probes were found to be common in both experiments using both p value and fold change as the criteria. The complete dataset has been submitted to ArrayExpress where it has been assigned the Accession number E-MTAB-3116.

DOI: 10.7554/eLife.03743.003

The following figure supplements are available for figure 1:

Figure supplement 1. The 120 embryos used for the first screen (AMZ vs PMZ), after excision of the explants and ISH for *cVg1*.

DOI: 10.7554/eLife.03743.004

Figure supplement 2. The 210 embryos used for the second screen (lateral marginal zone from isolated anterior half embryos, cultured for 7 hr, then sorted into *cVg1*-like or *cVg1*-unlike expression).

DOI: 10.7554/eLife.03743.005

Figure supplement 3. Analysis of microarray datasets for posterior marginal zone (PMZ) and isolated anterior cut halves.

DOI: 10.7554/eLife.03743.006

halves it is upregulated by about 2.4 fold (three probes ranking four, six and eight on the list; p between 0.002–0.008).

To confirm the microarray results, we examined 53 of the differentially expressed genes by whole-mount in situ hybridisation at pre-primitive streak stages X-XIII (***Eyal-Giladi and Kochav, 1976***) (23 of these are shown in ***Figure 2***). Apart from *Pitx2* three other genes show a similar expression to *cVg1* at stage XII: *Elk3* (an Ets-domain protein also known as SRF accessory protein-2), *PKDCC* (protein kinase domain containing cytoplasmic protein) and *LITAF* (lipopolysaccharide-induced tumor necrosis

factor-alpha). Others are expressed in cells adjacent to the PMZ (and are therefore likely to represent early axial cells), such as *ADMP*, *Brachyury* (*T*), *Mixl1*, *Tbx6*, *FGF8*, and *CHRD*, or are expressed much later (stage XIV), such as *DENND5B*. A final group is virtually undetectable, such as *Thrombopoietin*, *Ovoinhibitor* and *PMEPA*. All of these rank lower than *Pitx2* (see above and ***Table 1***, ***Table2***, ***Table3***, ***Table 4***): *Elk3* ranks 17th in whole embryos and 43rd in cut halves, *PKDCC* ranks 7th in whole embryos and 33rd in the anterior half, and *LITAF* ranks 18–19th in whole embryos and 17th and 25th in anterior halves. *Pitx2* is therefore the strongest candidate as a putative regulator of *cVg1*. The '*cVg1*-unlike' genes (***Figure 2P–W***) give less obvious information. Comparison of the top genes identified from differential expression in whole embryos (stronger in AMZ than in PMZ) with their counterparts in the corners of anterior halves reveals weak correlation (Spearman's rank Rho between −0.02 and 0.17; p = 0.44–0.93). These genes include those encoding extracellular matrix proteins as well as glucose-, glutamate-, glycine-, GABA- and LDL transporters and receptors, the transcriptional repressor ID3, and BASP1, which has been reported to act as a transcriptional co-suppressor for WT1 (***Carpenter et al., 2004***), among others. In situ hybridisation for these genes does not show enrichment in the AMZ or any other obvious pattern consistent with a putative role as an inhibitor of c*Vg1* expression in the PMZ at the appropriate stages of development (***Figure 2P–W***). *Pitx2* therefore remains as the most likely candidate.

To determine the temporal relationship between *Pitx2* and *cVg1* in whole embryos and during embryonic regulation, we compared their expression in time-course. In normal embryos *cVg1* is first detected at around stage XI (***Bertocchini and Stern, 2012***). We detected *Pitx2* in the PMZ by the

Table 1. Genes identified from the screens, sorted according to their ranking in PMZ>AMZ in whole embryos

		Posterior vs anterior			Anterior half, Vg1-like		
Gene symbol	**Probe ID**	**Fold change**	**p Value**	**Rank**	**Fold change**	**p Value**	**Rank**
ADMP	Gga.354.1.S1_at	**13.095**	**0.0080**	1	**3.771**	**0.0686**	2
PITX2	Gga.3398.2.S1_a_at	**9.586**	**0.0001**	2	**2.162**	**0.0069**	8
PITX2	Gga.3398.1.S1_a_at	**8.528**	**0.0001**	3	**2.429**	**0.0087**	4
THPO	GgaAffx.21801.1.S1_at	**7.398**	**0.0009**	4	**2.272**	**0.0267**	7
ST6-GAL2	Gga.14379.1.S1_at	**6.896**	**0.0744**	5	**2.506**	**0.0003**	3
T	Gga.3772.1.S1_a_at	**6.388**	**0.0867**	6	**5.544**	**0.0251**	1
PKDCC	Gga.12157.1.S1_at	**5.618**	**0.0006**	7	**1.572**	**0.2048**	33
PITX2	Gga.3398.1.S1_at	**5.263**	**0.0030**	8	**2.320**	**0.0026**	6
MIXL1	Gga.426.1.S1_s_at	**5.197**	**0.0160**	9	**2.378**	**0.0185**	5
n/a	Gga.2705.1.S1_at	**5.019**	**0.00002**	10	**1.415**	**0.0237**	63
GDF3	Gga.4324.2.S1_a_at	**4.309**	**0.00006**	11	**1.710**	**0.0062**	21
PMEPA1	Gga.6268.1.S1_at	**4.141**	**0.0805**	12	**2.032**	**0.0103**	12
PMEPA1	GgaAffx.12721.1.S1_at	**3.777**	**0.0007**	13	**1.745**	**0.0003**	19
TBX6	Gga.466.1.S1_at	**3.694**	**0.0004**	14	**2.030**	**0.0210**	13
FGF8	Gga.661.1.S1_at	**3.277**	**0.0052**	15	**1.690**	**0.0859**	22
Ovoinhibitor	Gga.6976.1.S1_at	**3.213**	**0.0015**	16	**1.467**	**0.0633**	50
ELK3	Gga.4498.1.S1_s_at	**3.094**	**0.0096**	17	**1.503**	**0.0010**	43
LITAF	Gga.3383.1.S2_at	**3.066**	**0.0031**	18	**1.670**	**0.0110**	25
LITAF	Gga.3383.1.S1_at	**2.895**	**0.0019**	19	**1.822**	**0.0027**	17
n/a	Gga.13092.1.S1_at	**2.823**	**0.0025**	20	**1.406**	**0.0284**	66

List of the top 20 common upregulated probes expressed in both the PMZ of whole embryo and isolated anterior cut halves. Entries in red are probes that pass a fold change cut-off of 1.2 as well as a p value cut-off of 0.05; those in blue pass the fold change cut off of 1.2 but not the p value cut-off of 0.05; and those in black pass the p value cut-off but not the fold change. Common genes are ranked according to the fold change of genes expressed in the PMZ (Spearman's rank Rho = 0.72, p = 0.00048).

DOI: 10.7554/eLife.03743.007

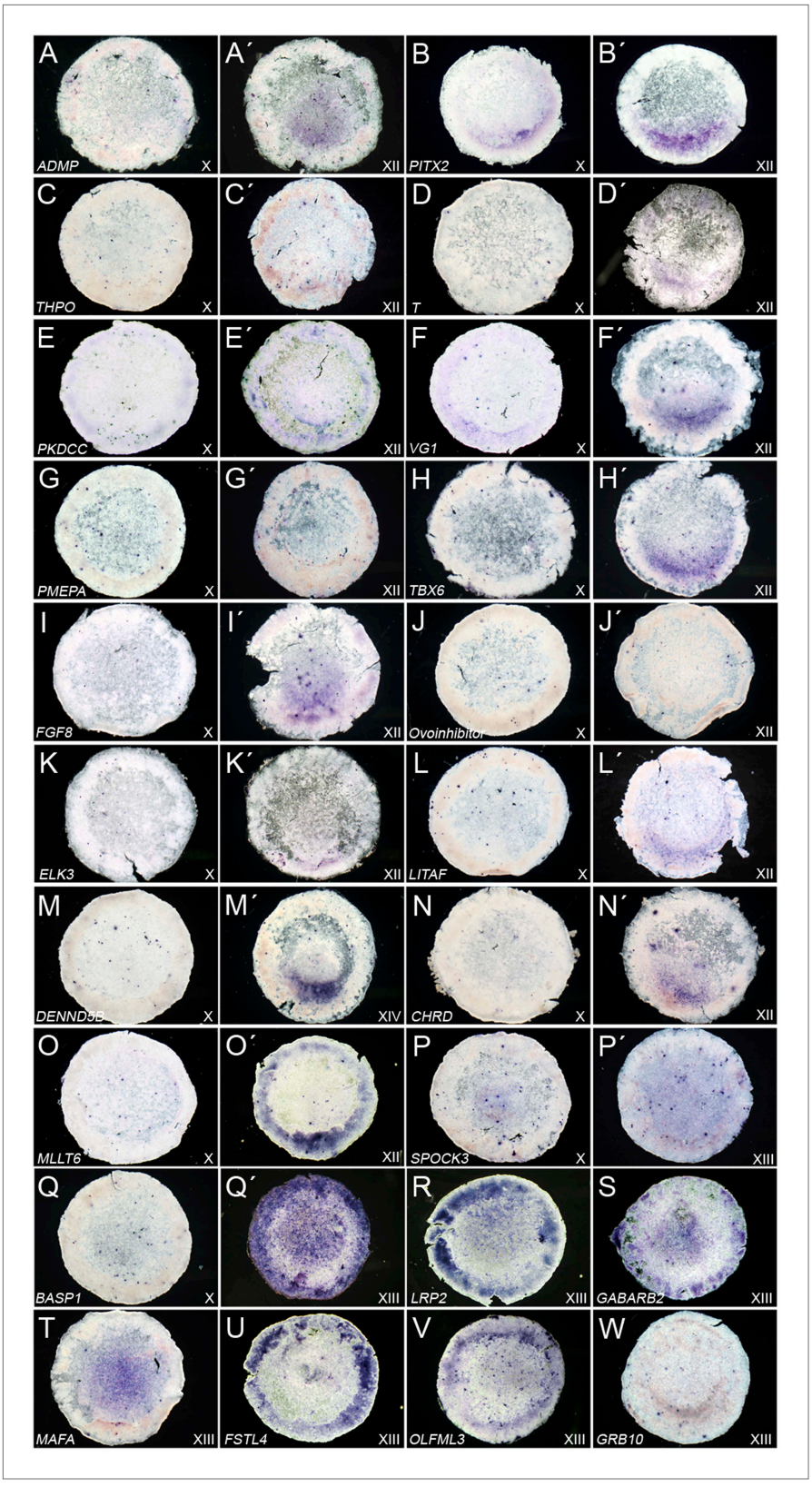

Figure 2. Expression of '*cVg1*-like' and '*cVg1*-unlike' genes, verified by in situ hybridisation. Embryos at stage X–XIII (the earliest stage at which differential expression was detected is shown) were processed using in situ hybridisation for genes co-regulated with *cVg1* ('*cVg1*-like', ***Table 1A–B*** (**A**–**O**) and genes expressed at lower level

Figure 2. Continued on next page

Figure 2. Continued

in the *cVg1*-positive region than in its counterpart ('*cVg1*-unlike', ***Table 1C–D***) (**P**–**W**), from the two microarray screens (see ***Table 1C–D***). The expression of 23 genes (15 '*cVg1*-like' and 8 '*cVg1*-unlike') is shown here.
DOI: 10.7554/eLife.03743.008

time of laying, stage X (***Figure 3A***), where it remains until early streak stages (***Figure 3B–F***). In isolated anterior halves, we increased the sensitivity of the assay by developing the NBT/BCIP colour reaction for several days to ensure that even weak expression could be detected. With this strategy we detected *cVg1* expression 4–5 hr after cutting (***Figure 3L–P***), 1–2 hr earlier than in previous reports (***Bertocchini et al., 2004***), while *Pitx2* appeared even earlier, just 3 hr after embryo bisection (***Figure 3G–K***). Taken together, these results implicate *Pitx2* as a good candidate for an upstream regulator of *cVg1* expression: it is a transcription factor, it is expressed in the same domain as *cVg1* in whole embryos and in bisected embryo marginal zone, and it is expressed before *cVg1*.

Pitx2 is required for axis development and embryonic regulation

To determine whether Pitx2 is important for embryonic regulation and for controlling *cVg1* expression, we used targeted electroporation of morpholino oligonucleotides (MOs). When a translation-blocking *Pitx2*-MO was targeted to the right edge of an isolated anterior half embryo, the frequency of axis formation shifted to the opposite side (***Figure 4A–D***; ***Figure 4—source data 1A,B***): embryos

Table 2. Genes identified from the screens, sorted according to their ranking in *cVg1*-like > -unlike in isolated anterior halves

		Anterior half, Vg1-like			Posterior vs anterior		
Gene symbol	Probe ID	Fold change	p Value	Rank	Fold change	p Value	Rank
T	Gga.3772.1.S1_a_at	5.544	0.0251	1	6.388	0.0867	6
ADMP	Gga.354.1.S1_at	3.771	0.0686	2	13.095	0.0080	1
ST6GAL2	Gga.14379.1.S1_at	2.506	0.0003	3	6.896	0.0744	5
PITX2	Gga.3398.1.S1_a_at	2.429	0.0087	4	8.528	0.0001	3
MIXL1	Gga.426.1.S1_s_at	2.378	0.0185	5	5.197	0.0160	9
PITX2	Gga.3398.1.S1_at	2.320	0.0026	6	5.263	0.0030	8
THPO	GgaAffx.21801.1.S1_at	2.272	0.0267	7	7.398	0.0009	4
PITX2	Gga.3398.2.S1_a_at	2.162	0.0069	8	9.586	0.0001	2
DENND5B	Gga.16679.1.S1_at	2.151	0.0026	9	2.341	0.1952	27
WNT8A	Gga.886.1.S1_at	2.078	0.0042	10	1.931	0.0259	53
CHRD	Gga.490.1.S1_at	2.064	0.0739	11	2.267	0.0959	30
PMEPA1	Gga.6268.1.S1_at	2.032	0.0103	12	4.141	0.0805	12
TBX6	Gga.466.1.S1_at	2.030	0.0218	13	3.694	0.0004	14
HOXB1	Gga.18352.1.S1_at	2.015	0.0298	14	1.749	0.0029	73
AREGB	GgaAffx.6867.1.S1_at	1.891	0.0526	15	1.640	0.1195	85
SOHO-1	Gga.770.1.S1_at	1.884	0.0256	16	1.417	0.1282	106
LITAF	Gga.3383.1.S1_at	1.822	0.0027	17	2.895	0.0019	19
MLLT6	Gga.11449.1.S1_at	1.762	0.0194	18	2.218	0.0091	32
PMEPA1	GgaAffx.12721.1.S1_at	1.745	0.0003	19	3.777	0.0007	13
FAM19A1	Gga.11944.2.S1_a_at	1.737	0.0948	20	2.102	0.0956	41

List of the top 20 common upregulated probes expressed in both the PMZ of whole embryo and isolated anterior cut halves. Entries in red are probes that pass a fold change cut-off of 1.2 as well as a p value cut-off of 0.05; those in blue pass the fold change cut off of 1.2 but not the p value cut-off of 0.05; and those in black pass the p value cut-off but not the fold change. Common genes are ranked according to the fold change of genes in the anterior cut halves (Spearman's rank Rho = 0.73, p = 0.00036).
DOI: 10.7554/eLife.03743.009

Table 3. Genes identified from the screens, sorted according to their ranking in AMZ > PMZ in whole embryos

		Posterior vs anterior			Anterior half, Vg1-like		
Gene symbol	Probe ID	Fold change	p Value	Rank	Fold change	p Value	Rank
LPL	Gga.4248.1.S1_at	−3.578	0.0350	1	−1.217	0.0191	69
UPK1B	Gga.17532.1.S1_at	−2.573	0.0662	2	−1.519	0.0046	15
UPK1B	Gga.17532.1.S1_s_at	−2.550	0.0383	3	−2.133	0.0003	2
SPOCK3	GgaAffx.6009.1.S1_at	−2.531	0.0284	4	−1.461	0.0174	24
SLC1A3	GgaAffx.25896.1.S1_at	−2.333	0.2168	5	−1.338	0.0272	47
SLC5A1	Gga.8594.1.S1_at	−2.284	0.2814	6	−2.253	0.0857	1
BASP1	Gga.3179.1.S1_at	−2.276	0.0916	7	−1.170	0.0079	75
ID3	Gga.4048.1.S1_at	−2.243	0.0004	8	−1.422	0.0463	34
ATP13A4	GgaAffx.12270.1.S1_at	−2.172	0.0486	9	−1.202	0.0148	71
DIO3	Gga.552.1.S1_at	−2.149	0.0051	10	−1.656	0.0376	12
UPK1B	Gga.12930.1.S1_at	−2.134	0.0640	11	−2.123	0.0014	3
COL4A2	Gga.3104.1.S1_at	−2.115	0.1024	12	−1.286	0.0059	58
SLC1A3	GgaAffx.25896.1.S1_s_at	−2.089	0.3192	13	−1.241	0.0363	60
LRP2	GgaAffx.23355.1.S1_at	−2.052	0.0637	14	−1.306	0.0144	55
EFNB2	Gga.13001.1.S1_s_at	−2.049	2.6527	15	−1.360	0.0220	42
---	Gga.15960.1.S1_at	−2.005	0.0007	16	−1.437	0.0203	29
---	Gga.18649.1.A1_at	−1.991	0.0092	17	−1.358	0.0068	43
GABRA1	Gga.17167.1.S1_at	−1.955	0.1060	18	−1.487	0.0978	20
GLRA3	GgaAffx.6806.1.S1_at	−1.951	0.0001	19	−1.672	0.0209	10
DIO2	Gga.1819.1.S1_at	−1.908	0.1217	20	−1.457	0.0025	25

List of the top 20 downregulated probes common to both the PMZ of whole embryo and isolated anterior cut halves. Entries in red are probes that pass a fold change cut-off of −1.2 as well as a p value cut-off of 0.05; those in blue pass the fold change cut off of −1.2 but not the p value cut-off of 0.05; and those in black pass the p value cut-off but not the fold change cut-off. Common genes are ranked according to the fold change of genes in the PMZ (Spearman's rank Rho = −0.02, p = 0.93).

DOI: 10.7554/eLife.03743.010

electroporated with control-MO at the right edge had random *cVg1* expression on the right, left or neither edges in equal proportion. In contrast, *Pitx2*-MO on the right generated a majority (9/10, p = 0.05) of embryos with no expression and the remaining embryo with expression on the opposite side 7 hr post-electroporation (***Figure 4B***, ***Figure 4—source data1B***). After 16 hr (***Figure 4—source data1B***) embryos electroporated with control-MO at the right edge formed a primitive streak expressing *Brachyury* randomly; with *Pitx2*-MO on the right (***Figures 4B***), 8/10 embryos formed a streak on the left side and 2 had two streaks (one from each corner), but no embryo formed a single streak arising from the MO-transfected side (hypergeometric exact test with 2 × 3 contingency table, p = 0.026). Equivalent results were observed with MO electroporations on the left (***Figure 4C–D***; ***Figure 4—source data1A,B***): control-MO embryos transfected on the left and examined for *cVg1* expression had random expression on the right, whereas for *Pitx2*-MO *cVg1* expression was biased towards the left. After 16 hr, control embryos transfected on the left formed a *Bra*-expressing streak on the right, again randomly, while *Pitx2*-MO embryos formed a streak on the right in 2/8 cases, none on the left, 1/8 on both sides and 5/8 with no streak (p = 0.003) (***Figure 4D***, ***Figure 4—source data1B***).

In intact embryos electroporated at stage X, *Pitx2*-MO affected *cVg1* expression and streak formation (***Figure 4E–F***; ***Figure 4—source data 1C–D***). Control-MO targeted to the PMZ did not alter *cVg1* expression after 5 hr or streak formation after 12–16 hr (***Figure 4E***; ***Figure 4—source data1C,D***). With *Pitx2*-MO, *cVg1* expression was affected after 5 hr (0/6 embryos expressing; p < 0.001). Embryos started to recover, however, at later time points: at 12 hr, 3/12 had a normal streak, 3/12 had a

Table 4. Genes identified from the screens, sorted according to their ranking in Vg1-unlike > Vg1-like in isolated anterior halves

		Anterior half, Vg1-like			Posterior vs anterior		
Gene symbol	Probe ID	Fold change	p Value	Rank	Fold change	p Value	Rank
SLC5A1	Gga.8594.1.S1_at	−2.253	0.0857	1	−2.284	0.2814	6
UPK1B	Gga.17532.1.S1_s_at	−2.133	0.0003	2	−2.554	0.0383	3
UPK1B	Gga.12930.1.S1_at	−2.123	0.0014	3	−2.134	0.0642	11
EDNRB	Gga.3306.1.S1_s_at	−1.935	0.1491	4	−1.773	0.2348	32
FSTL4	Gga.13574.1.S1_at	−1.878	0.0152	5	−1.496	0.0014	51
GABARB2	Gga.17131.1.S1_at	−1.814	0.0491	6	−1.586	0.0647	45
KCNAB1	Gga.4971.1.S1_at	−1.763	0.0185	7	−1.302	0.0238	72
MYLK	Gga.6776.1.S1_at	−1.758	0.0126	8	−1.822	0.1777	30
CBLN4	GgaAffx.4848.1.S1_s_at	−1.702	0.0424	9	−1.641	0.0564	40
GLRA3	GgaAffx.6806.1.S1_at	−1.672	0.0209	10	−1.951	0.0001	19
SLC14A2	Gga.7955.1.S1_at	−1.663	0.0094	11	−1.897	0.0561	22
DIO3	Gga.552.1.S1_at	−1.656	0.0376	12	−2.149	0.0051	10
KCNMA1	Gga.19342.1.S1_at	−1.599	0.2068	13	−1.384	0.0097	63
GRB10	GgaAffx.8324.2.S1_at	−1.534	0.0569	14	−1.213	0.0236	76
UPK1B	Gga.17532.1.S1_at	−1.519	0.0046	15	−2.573	0.0662	2
CALD1	GgaAffx.21386.1.S1_s_at	−1.518	0.0054	16	−1.727	0.1106	33
MAFA	Gga.974.1.S1_at	−1.518	0.2597	17	−1.855	0.0144	29
OLFML3	Gga.1150.2.S1_a_at	−1.509	0.0035	18	−1.382	0.0149	62
---	Gga.18986.1.S1_at	−1.487	0.0109	19	−1.522	0.0312	48
GABRA1	Gga.17167.1.S1_at	−1.487	0.0978	20	−1.955	0.1067	18

List of the top 20 downregulated probes common to both the PMZ of whole embryo and isolated anterior cut halves. Entries in red are probes that pass a fold change cut-off of −1.2 as well as a p value cut-off of 0.05; those in blue pass the fold change cut off of −1.2 but not the p value cut-off of 0.05; and those in black pass the p value cut-off but not the fold change cut-off. Common genes are ranked according to the fold change of genes in the anterior cut halves (Spearman's rank Rho = 0.17, p = 0.44).

DOI: 10.7554/eLife.03743.011

displaced streak, 3/12 had two streaks, and the remaining 3/12 had no streaks (p = 0.1—not significantly different). By 16 hr, the majority of the embryos were normal (7/8; the remaining embryo had a displaced streak) (***Figure 4F***; ***Figure 4—source data 1D***). These results suggest that while knockdown of *Pitx2* affects *cVg1* expression, embryos tend to recover by 12–16 hr. We reasoned that functional redundancy with another *Pitx* gene, or compensatory upregulation of such a gene in response to *Pitx2* knockdown, could account for this recovery. To test this, we examined *Pitx1* expression in normal embryos at stages X–XII and in embryos electroporated with *Pitx2*-MO. *Pitx1* is barely detectable in the PMZ at stage X–XII (***Figure 4—source data 1***). After *Pitx2*-MO electroporation, expression increased considerably (6/6 embryos; ***Figure 4—source data 1***). We therefore repeated the targeting experiments in whole embryos using a mixture of two MOs targeting the translation start site of *Pitx2* and an internal splice junction of *Pitx1*, respectively. *Pitx1+2*-MOs caused loss of *cVg1* at 5 hr in 6/6 cases (p = 0.005). By 12 hr 1/9 embryos had a normal streak, 3/9 had a displaced streak, 3/9 had two streaks and 2/9 had none (p = 0.043). By 16 hr no recovery was observed: 4/4 embryos had double streaks, neither arising from the targeted site (p = 0.029; ***Figure 4G***; ***Figure 4—source data 1E,F***). This effect could be rescued by supplying *Pitx2* alone: co-electroporation of *Pitx1+2*-MO together with a *Pitx2* expression construct lacking the MO target sequence led to normal axis formation: in 5/6 cases *cVg1* expression was restored after 5 hr incubation and by 16 hr 12/16 embryos displayed a normal streak (***Figure 4H***; ***Figure 4—source data 1G–H***). Likewise in isolated anterior halves, the effects of *Pitx1+2*-MO (***Figure 4I***; ***Figure 4—source data 1I,J***) could be rescued by *Pitx2*. After

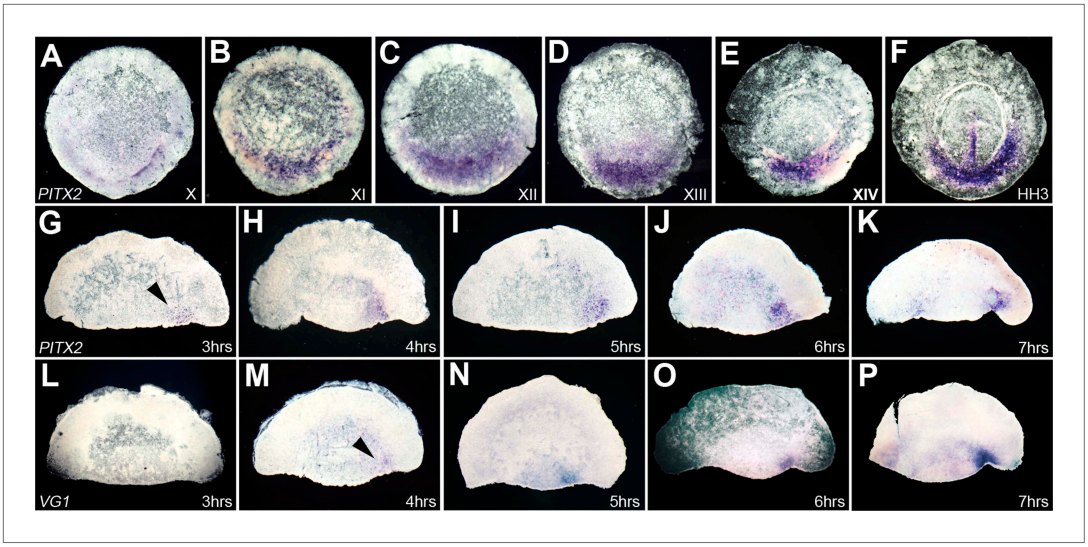

Figure 3. Time-course of expression of *Pitx2* in whole embryos and in isolated anterior halves, and comparison with *cVg1*. (**A**–**F**) Time-course of *Pitx2* expression in whole embryos. Expression is restricted to the posterior marginal zone (PMZ) already at the time of laying (stage X); this intensifies over the subsequent stages. At primitive streak stages (HH2–3), expression is seen in the primitive streak itself as well as in the PMZ (F). (**G**–**P**) Comparison of the time-course of *Pitx2* (**G**–**K**) and *cVg1* (**L**–**P**) upregulation in isolated anterior halves. *cVg1* is first detectable 4–5 hr after cutting (**M**), whereas *Pitx2* can be detected an hour earlier (**G**). Both *Pitx2* and *cVg1* expression appears randomly on the left or the right corner of the marginal zone adjacent to the cut edge. The series chosen for this figure shows upregulation on the right side of all embryo fragments.

DOI: 10.7554/eLife.03743.012

electroporation on the right, *cVg1* expression was seen on the right in 4/10, on the left in 3/10, and in neither in 3/10 cases (***Figure 4J***; ***Figure 4—source data 1K***). After 16 hr a primitive streak developed on the right in 3/8 cases, on the left in 2/8 and no streak in 3/8 cases (***Figure 4—source data l***).

In conclusion, Pitx2 is required for expression of *cVg1* in the PMZ as well as for formation of the normal primitive streak. In isolated anterior halves, Pitx2 is required both for *cVg1* expression and for the later formation of a primitive streak. Knockdown of *Pitx2* is followed by upregulation of the related transcription factor *Pitx1*, which can partly compensate for the loss of *Pitx2*.

A bioinformatics approach to uncover candidate regulatory regions for cVg1

The *cVg1* gene (erroneously annotated as *GDF3* in the chick genome; its orthologue is human *GDF1*, as confirmed by synteny; see ***Figure 5***) is located on chicken chromosome 28. As a parallel approach to the above to identify putative upstream regulators, we applied a recently described pipeline (***Khan et al., 2013***), starting with prediction of conserved, constitutive CTCF-binding sites (CTCF is an 11 zinc-finger transcriptional repressor protein that co-localizes with cohesin and acts to delimit chromatin loops) (***Holwerda and de Laat, 2013***; ***Merkenschlager and Odom, 2013***) around this locus which could act as insulators. This was followed by algorithms to identify conserved motifs in non-coding regions that are order independent, modified from the Enhancer Discovery using only Genomic Information (EDGI) tool described for Drosophila (***Sosinsky et al., 2007***).

The insulator-predicting software identifies strongly conserved CTCF-binding sites about 200 kB upstream and about 100 kB downstream of *cVg1*/*GDF3* (***Figure 5***). The *cVg1*/*GDF3* gene itself is bicistronic, the upstream exons encoding *Lass1*/*CERS1* and the last two exons containing the cVg1 sequence (***Wang et al., 2007***). Several other genes lie in this region, including *COPE*, *DDX49* and *HOMER3* upstream and *UPF1* downstream. If the predicted CTCF-binding sites are indeed insulators, we would expect these genes to be co-regulated with *cVg1*. To test this, we examined their expression. Strikingly, all genes examined (*CERS1*/*LASS1*, *COPE*, *HOMER3*, and *UPF1*) are expressed in a similar domain of the PMZ as *cVg1* (***Figure 5—figure supplement 1***).

We then applied the DREiVe tool (Discovery of Regulatory Elements in Vertebrates, the vertebrate version of EDGI) (***Sosinsky et al., 2007***; ***Khan et al., 2013***), to identify de novo conserved

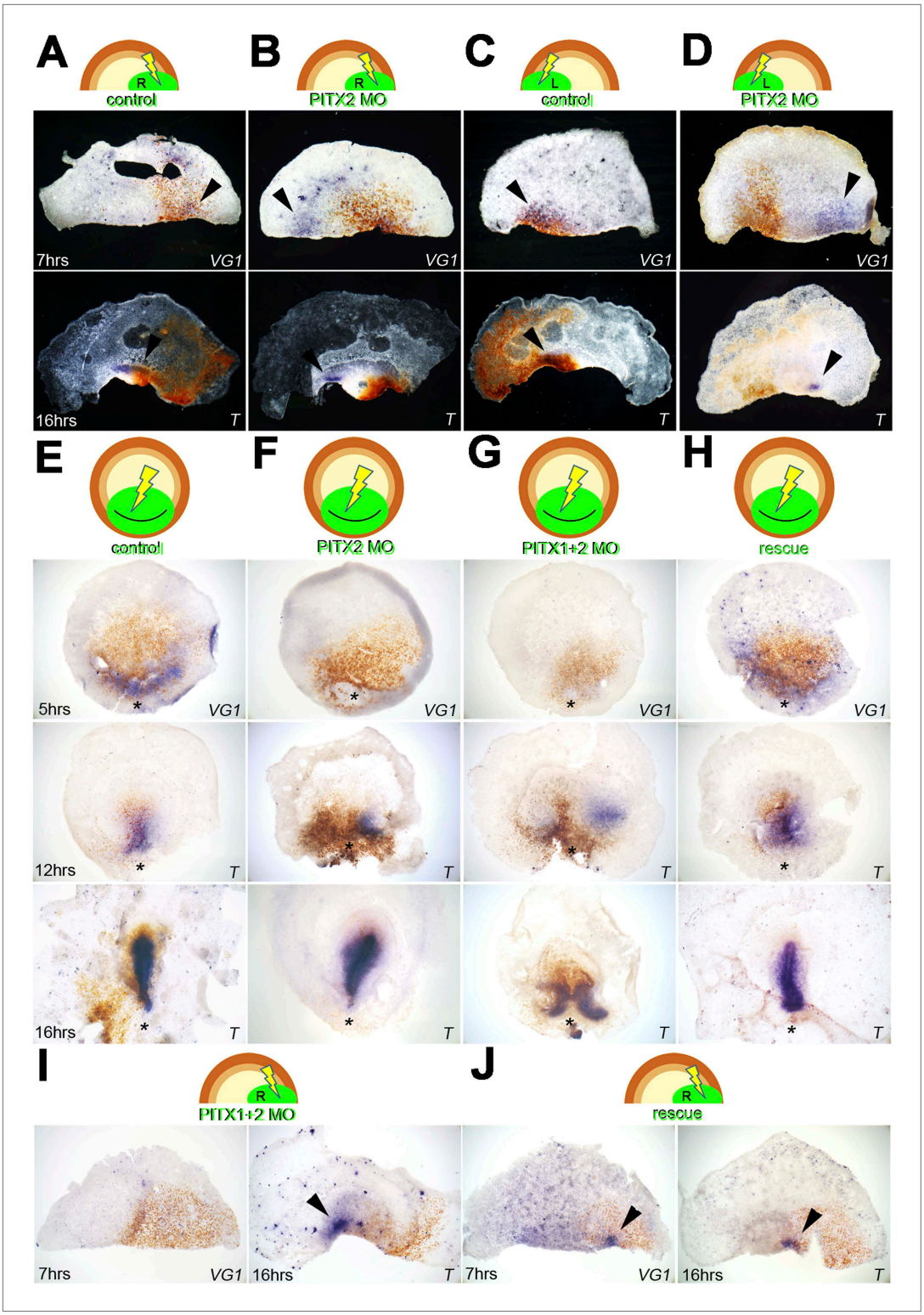

Figure 4. Pitx2 is required for *cVg1* expression and for axis formation. (**A**–**D**) A morpholino (MO) targeting Pitx2 applied to one side of the marginal zone of an isolated anterior half shifts axis formation to the opposite marginal zone. (**A**–**B**) The experiment done on the right side (A = control-MO, B=Pitx2-MO), (**C**–**D**) the equivalent done on the left marginal zone. The top row shows diagrams of the experiment, the middle row shows embryo fragments 7 hr after electroporation/cutting, hybridised for *cVg1* (purple) and stained with anti-fluorescein (in the MO, brown),

Figure 4. Continued on next page

Figure 4. Continued

and the lower row shows fragments processed for *Brachyury* (*T*) in purple and fluorescein in brown. (**E**–**F**) In whole embryos, *Pitx2*-MO has a transient effect. At 5 hr, *cVg1* expression is lost, at 12 hr a primitive streak (*T*-expressing) is sometimes seen especially at the edge of the electroporated domain, but by 16 hr embryos appear essentially normal. (**E**) Shows embryos electroporated with control-MO, F are embryos transfected with Pitx2-MO. (**G–H**). Embryos electroporated with MOs targeting both Pitx2 and Pitx1 do not recover: at 5 hr, no *cVg1* is seen; at 12–16 hr there is a high proportion of embryos with either no streak or two *T*-expressing streaks arising from outside the MO-electroporated domain (as shown in **G**). This effect can be rescued by co-electroporation of Pitx2 with the mixture of *Pitx1*-MO/*Pitx2*-MO (**H**). (**I**–**J**) Likewise, in isolated anterior half-embryos, the effects of electroporation of the *Pitx1*-MO/*Pitx2*-MO combination (**I**) can be rescued by co-electroporation with *Pitx2* alone (**J**): *cVg1* expression is now seen on the electroporated side. Arrowheads point to sites of expression.

DOI: 10.7554/eLife.03743.013

The following source data and figure supplement are available for figure 4:

Source data 1. Numbers of embryos displaying different types of results in the Pitx loss-of-function experiments illustrated in the main Figure.

DOI: 10.7554/eLife.03743.014

Figure supplement 1. *Pitx1* is expressed posteriorly in normal embryos, and upregulated in Pitx2 morphants.

DOI: 10.7554/eLife.03743.015

sequence motifs around this region. This identifies six domains (designated E1–E6), ranging in size from 600–3000 bases, located within the introns of *CERS1*/*LASS1* and of the neighbouring *Homer3* in chick (***Figure 5***). Analysis of these regions using Position Frequency Matrices from JASPAR and TRANSFAC databases together with the algorithms Matrix-Scan (from RSAT) and Clover predicts four of these regions (E1, E3, E5, and E6) to contain one or more putative binding sites for *Pitx2* and/or the related factor *Pitx1* (***Figure 5***). The power of DREiVe as a tool for discovering regulatory elements is highlighted by the observation that it is able to identify homologous non-coding regions in the human genome, where the syntenic region (on chromosome 19) is not only inverted but also the orthologous elements are found in a different order, within introns of different neighbouring genes (***Figure 5***). In mouse (chromosome 8) the arrangement is similar to chick (***Figure 5***).

Testing candidate regulatory regions

To determine whether any of the predicted regions bind Pitx2 in the PMZ of normal embryos, we conducted chromatin-immunoprecipitation experiments (ChIP), assessing precipitation by real-time quantitative polymerase chain reaction (qPCR) analysis using primers targeting the predicted regions. We compared chromatin from the AMZ and PMZ (***Figure 6***). A monoclonal antibody against Pitx2 precipitated chromatin from the PMZ more effectively than from the AMZ for all predicted enhancers that contained consensus Pitx1/2-binding sites (especially E3, E5, and E6) but not those that do not (E2 and E4). These findings suggest that Pitx2 is differentially bound to putative enhancer sites E3, E5, and E6 in the PMZ of normal embryos. We also tested each of the six putative enhancer regions for acetylation of Lys-27 of Histone-3 (H3K27ac), which is associated with active enhancers (***Creyghton et al., 2010***). Enhancers E5 and E6 showed the greatest differential activity in the PMZ relative to the AMZ. Together, these results suggest that E3, E5, and E6 are the most likely enhancers driving expression in the PMZ.

To test whether these putative enhancers do indeed direct transcription in the correct endogenous domain, we generated reporter constructs based on a vector designed by the group of Kondoh (***Uchikawa et al., 2003***). Each construct contained one candidate enhancer (E1–E6), a minimal promoter (TK), and a reporter fluorescent protein (EGFP or RFP) and was electroporated either in a very broad domain including the PMZ and lateral marginal zones of normal embryos, or encompassing the entire cut edge (including both corners) of bisected embryos and the anterior half subsequently cultured. A ubiquitous reporter (pCAβ with either EGFP or RFP) was co-electroporated with each construct to reveal the extent of the electroporated domain (***Figures 7 and 8***). Embryos were then photographed live to reveal the electroporated and expressing domains, then fixed and processed for in situ hybridisation for *cVg1* to determine whether the side with reporter activity corresponds to the *cVg1* expressing region. In whole embryos, E3 and E5 are most efficient in driving expression of the reporter in the PMZ (***Figure 7***). The same two enhancers also drive expression in the *cVg1*-expressing side in isolated anterior halves (***Figure 8***).

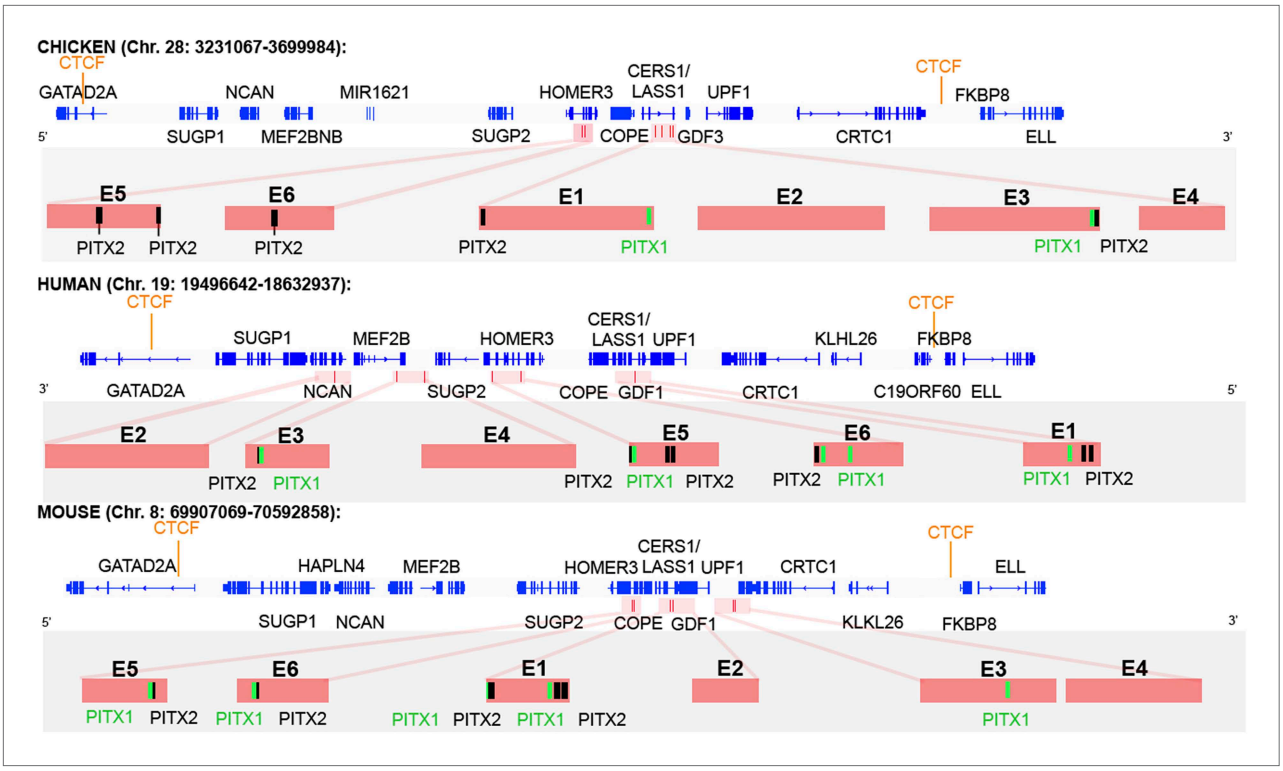

Figure 5. CTCF insulator analysis, enhancer identification and synteny of the *cVg1* locus. The chicken *cVg1* locus with computationally predicted conserved CTCF-binding sites in chick, human, and mouse is shown (genes represented in blue). This putative insulator region lies ~200 kB upstream and ~100 kB downstream of *cVg1/GDF3* and harbours other genes such as *CERS1/Lass1*, *COPE*, *DDX49* and *HOMER3* upstream and *UPF1* downstream of *cVg1*. Six putative enhancer regions (E1–E6), predicted using DREiVe, are displayed in pink. In chick, E1 (galGal4 genomic coordinates: *chr28:3502783-3504834*) and E2 (*chr28:3504993-3508041*) lie in the first intron of the bicistronic *CERS1/Lass1* gene. E3 (*chr28:3510154-3511032*) and E4 (chr28:3,511,413-3,511,725) lie in intron 4 of *CERS1/Lass1* and E5 (*chr28:3471200-3471520*) and E6 (*chr28:3,471,946-3,472,230*) respectively lie in introns 1 and 2 of *HOMER3*. E1 and E3 each contain conserved Pitx1 (black) and Pitx2 (green) binding sites. E2 and E4 do not contain any Pitx sites and E5 and E6 contain Pitx2-binding sites but no Pitx1-binding sites. The orthologous regions of human (chromosome 19) and mouse (chromosome 8) genomes are also shown. Note that the corresponding human region has been inverted, and that although all six elements are found within it, these appear in different order and are associated with different introns and intergenic regions than in chick and mouse.

DOI: 10.7554/eLife.03743.016

The following figure supplement is available for figure 5:

Figure supplement 1. Expression of genes adjacent to *cVg1/GDF3* on chick chromosome 28.

DOI: 10.7554/eLife.03743.017

To test whether the Pitx-binding sites are required within these enhancers to direct expression to the appropriate domain, we generated reporter constructs for enhancers E1, E3, E5, and E6 containing mutations in each Pitx1- or Pitx2-binding site as well as constructs where all Pitx-binding sites were mutated. In whole embryos, mutations in either of the Pitx-binding sites (***Figure 7—figure supplement 1A–B***) or in both binding sites (***Figure 7—figure supplement 1C***) of E1 still showed GFP expression without selectivity, in the middle of the embryo. For E3 and E5, however, mutations in either a single or both binding sites of each reporter completely abolished expression in the PMZ (***Figure 7—figure supplement 1D–G***). Un-mutated reporter E6, which did not show any activity (see above), was not altered by a mutation of its Pitx2-binding site (***Figure 7—figure supplement 1H***). Similar results were found in isolated anterior half-embryos: mutations of any of the Pitx-binding sites in E3 or E5 completely eliminated expression in the *cVg1*-expressing corner of the isolated half (***Figure 8—figure supplement 1***).

Taken together, these results suggest that *cVg1* expression is regulated directly by Pitx2/Pitx1 binding to an enhancer (E5) within an intron of the *HOMER3* gene, adjacent to *cVg1/GDF3*, and to an intronic enhancer (E3) within the *Lass1/CERS1* locus. The ChIP experiments suggest that Pitx2 binding to an additional intronic enhancer in *HOMER3* (E6) may also be functional in the PMZ, although

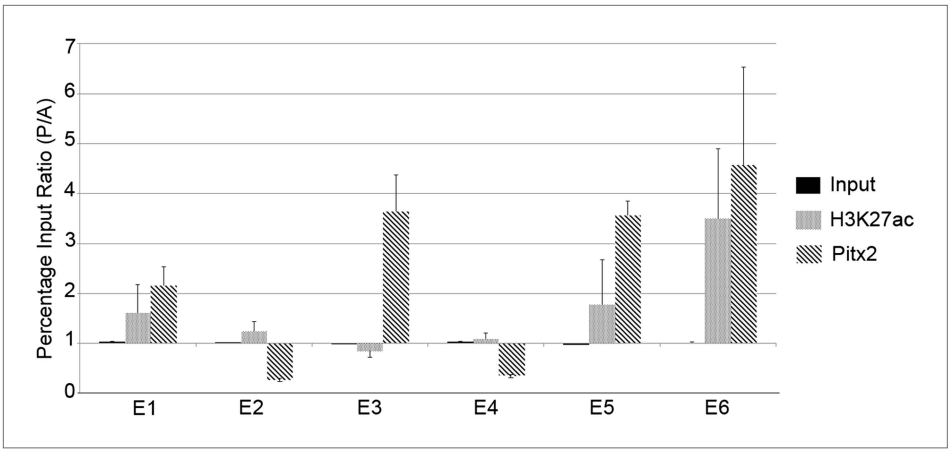

Figure 6. Chromatin immunoprecipitation to test for active histone marks and Pitx2 binding to predicted enhancers. Relative immunoprecipitation around each of the putative six enhancers by an antibody to Pitx2 (diagonal hatching) or an antibody to acetyl-lysine-27 of Histone-H3 (grey shading), expressed as a ratio of the amount precipitated from posterior and anterior marginal zone (PMZ and AMZ) chromatin. Primers were used to target each of the putative enhancers and precipitated chromatin measured by the quantitative polymerase chain reaction (qPCR). Each data bar represents the average of at least three true biological replicates and the error bars indicate standard error of the mean. Amplification from input DNA from each of the same samples is also shown (solid black shading). Note that enhancers that contain Pitx2-binding sites (E1, E3, E5, and E6) are precipitated much more strongly from the PMZ than the AMZ.

DOI: 10.7554/eLife.03743.018

this was not seen when a reporter construct containing this element was electroporated into the PMZ. The same enhancers (E3 and E5) are involved in controlling *cVg1* expression in the normal PMZ as in the portion of the lateral marginal zone where *cVg1* is upregulated a few hours after isolation of a portion of embryo. These results implicate the transcription factor Pitx2 as the earliest gene described to date that regulates the position of the embryonic axis as well as embryonic regulation/regeneration and twinning.

Discussion

The PMZ of the chick embryo is the equivalent of the Nieuwkoop centre of amphibians: it can induce a complete axis including the organiser from neighbouring cells, without making a cellular contribution to the axis (***Azar and Eyal-Giladi, 1979***; ***Khaner and Eyal-Giladi, 1989***; ***Bachvarova et al., 1998***). The Nodal/Activin-related TGFβ superfamily member *cVg1* (homologous to mammalian *GDF1*) is expressed in the PMZ. When ectopically applied to another region of the marginal zone, *cVg1* is sufficient to initiate formation of a complete embryonic axis from adjacent embryonic (area pellucida) cells (***Seleiro et al., 1996***; ***Shah et al., 1997***). To act, cVg1 requires canonical Wnt, which seems to be provided mainly by *cWnt8C*, expressed all around the marginal zone (***Skromne and Stern, 2001***). A target of cVg1 and Wnt is *Nodal*, transcribed in area pellucida cells next to the *cVg1+Wnt* expression domain (***Skromne and Stern, 2002***). The anterior end of the embryo also has an early identity, defined by the expression of GATA binding protein 2 (GATA2) (***Sheng and Stern, 1999***; ***Bertocchini and Stern, 2012***). However, unlike cVg1, GATA2 is not a sufficient determinant of polarity and at best only acts as a bias (***Bertocchini and Stern, 2012***).

While amphibian and teleost embryos lose their ability to generate complete, independent embryos if fragmented after the first few cell divisions, amniotes have huge regulative capacity. Dividing a chick blastoderm even just before primitive streak formation (when the embryo may have as many as 50,000 cells) into up to eight fragments can lead to formation of as many complete embryos (***Lutz, 1949***; ***Spratt and Haas, 1960***). In an isolated anterior half, a visible primitive streak (and expression of *Brachyury* and *Snail2*) can be detected about 12 hr after cutting. This appears randomly from either the left or right area pellucida adjacent to the cut edge of the marginal zone (***Spratt and Haas, 1960***), preceded at least 6 hr earlier by expression of *cVg1* in the right or left marginal zone (***Bertocchini et al., 2004***). Blocking cVg1 on one side of the marginal zone of an isolated anterior half will cause the axis to arise

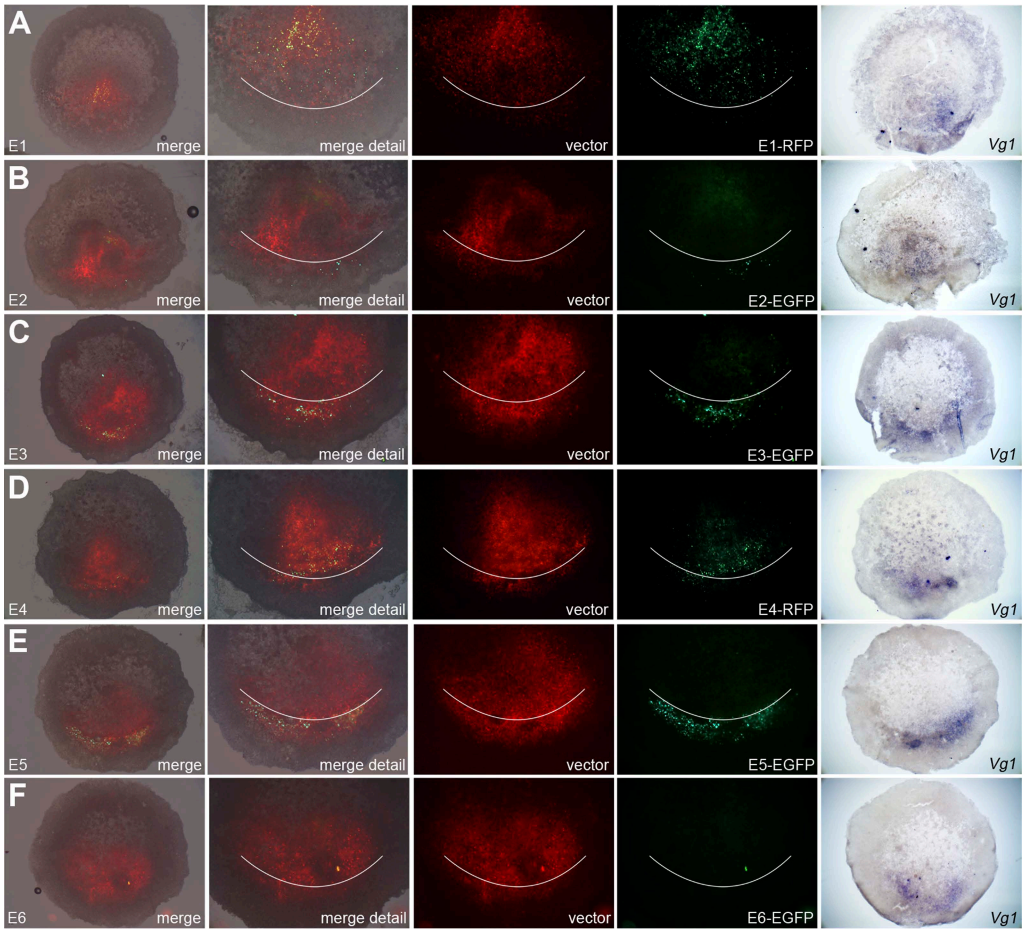

Figure 7. Enhancers E3 and E5 drive expression in the posterior marginal zone of whole embryos. Embryos were electroporated with a construct containing a candidate enhancer (E1–E6), a minimal promoter (TK) and a fluorescent reporter (GFP or RFP), together with a ubiquitous marker (pCAβ-EGFP or DS-RedExpress) to reveal the electroporated area. After 5–9 hr culture the embryos were observed by fluorescence (first 4 columns) and then fixed and processed to reveal *cVg1* expression by in situ hybridisation (last column). The position of Koller's sickle is marked with a curved white line. Enhancers E3 and E5 faithfully recapitulate *cVg1* expression in the posterior marginal zone (PMZ), whereas E1 drives expression inside the embryo (but not in the PMZ) and the remaining enhancers show little or no detectable activity. In all cases, the electroporated area appears red and the activity of the specific enhancer construct in green.

DOI: 10.7554/eLife.03743.019

The following figure supplement is available for figure 7:

Figure supplement 1. Effects of mutations in Pitx-binding sites on activity of cVg1 enhancers in whole embryos.

DOI: 10.7554/eLife.03743.020

from the opposite side. A similar manipulation in the PMZ will cause a streak to arise from outside the MO-electroporated domain (***Bertocchini et al., 2004***; ***Bertocchini and Stern, 2012***). Thus, *cVg1* expression in the marginal zone is both necessary and sufficient to initiate formation of a primitive streak.

There was no information, however, about what positions *cVg1* expression in the PMZ of normal embryos or in isolated fragments. This led us to undertake a screen for upstream regulators. We took two complementary approaches: a molecular screen, designed to identify genes co-expressed with *cVg1* both in the normal PMZ and in the *cVg1*-expressing edge of the marginal zone in isolated anterior halves at the time when *cVg1* first becomes detectable, and a bioinformatics-based approach, predicting and analysing putative enhancers of *cVg1* situated on chromosome 28. A particular difficulty of the molecular screen was that it is impossible to know, in a single embryo, which of the two

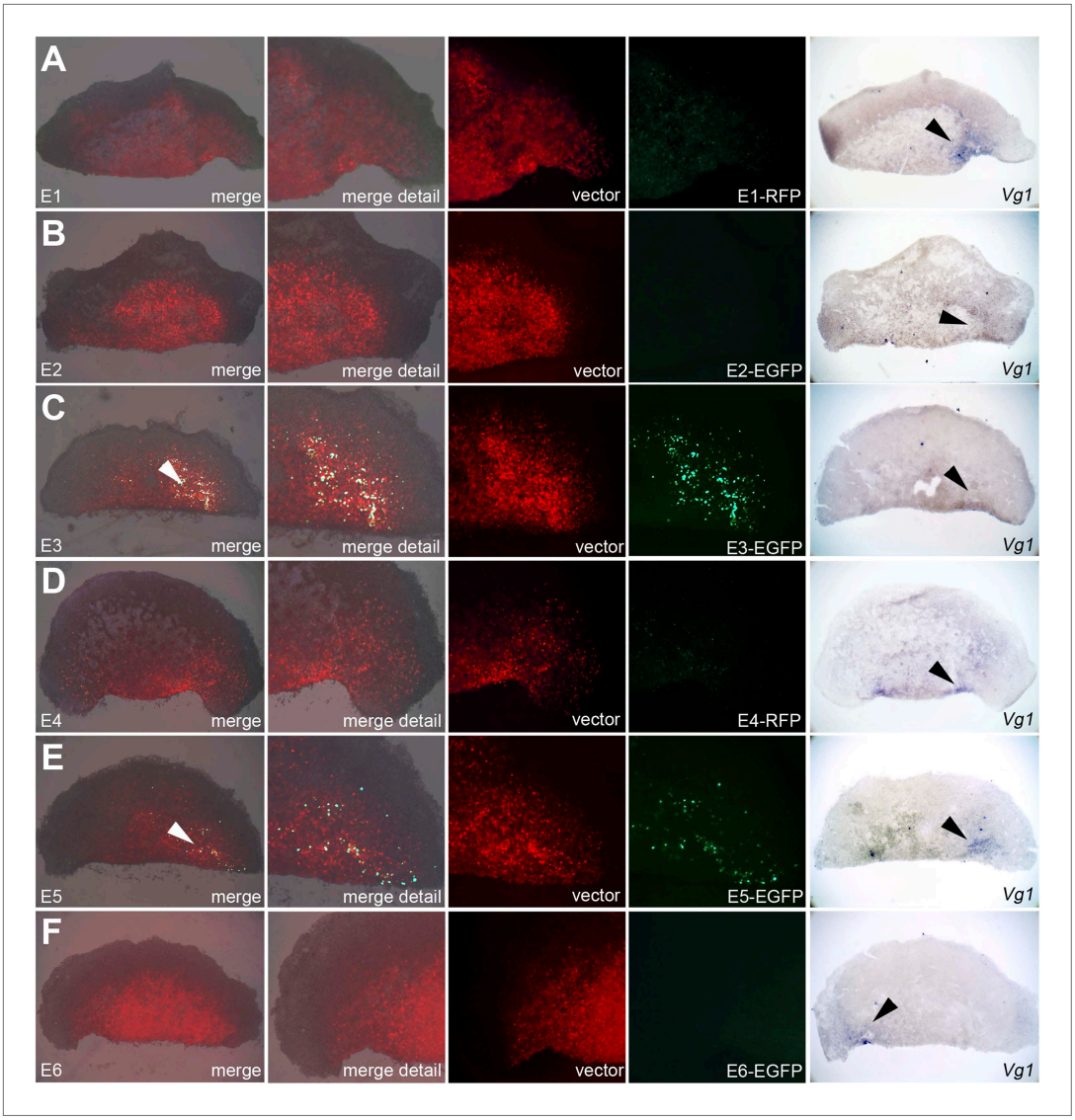

Figure 8. Enhancers E3 and E5 drive expression in the *cVg1*-expressing corner of the marginal zone at the cut edge of isolated anterior half-embryos. Embryos were electroporated with the same vectors as described in ***Figure 7***, then bisected. The anterior half was then cultured for 5–7 hr and viewed under fluorescence (first 4 columns), then fixed and processed for cVg1 expression (last column). Enhancers E3 and E5 drive expression of the reporter at the *cVg1*-expressing edge of isolated anterior half-embryos. Note that unlike what is found in whole embryos, Enhancer E1 does not appear to drive expression in the area pellucida of the isolated anterior half.

DOI: 10.7554/eLife.03743.021

The following figure supplement is available for figure 8:

Figure supplement 1. Effects of mutations in Pitx-binding sites on activity of *cVg1* enhancers in isolated anterior halves.

DOI: 10.7554/eLife.03743.022

edges of the isolated anterior half will express *cVg1*. This required analysis of each individual fragment just after excising left and right edge explants, then pooling them appropriately according to whether they derived from the *cVg1*-expressing side or the opposite edge. This has the additional advantage of removing any intrinsic left–right differences that may exist in the embryo at this stage. Selecting genes that are co-expressed with *cVg1* when this is first upregulated at one edge of an isolated fragment, and also co-expressed with *cVg1* in the normal PMZ, turned out to be a powerful strategy to reduce the number of relevant genes, thereby avoiding the 'cherry-picking' approaches often associated with molecular screens. This combination of molecular screens and bioinformatic analysis converged towards a single strong candidate: the transcription factor *Pitx2*. The regulatory role of Pitx2

on *cVg1* expression was then confirmed by loss-of-function experiments showing that Pitx2 is required for *cVg1* activation and axis formation both in normal development and during embryonic regulation, and that it binds directly to four non-coding regions around the *cVg1* locus (E1, E3, E5, and E6), two of which (E3 and E5) are sufficient to direct expression specifically to the PMZ and *cVg1*-expressing edge of an isolated anterior fragment. Moreover, mutation of any of the Pitx-binding sites in enhancers E3 and E5 abolishes the activity of these enhancers in the PMZ. These findings strongly implicate Pitx2 as a direct and essential regulator of *cVg1* expression.

Pitx transcription factors are characterised by possessing a Lysine residue at position 50 of the homeodomain, an unusual property shared with the vertebrate genes *Goosecoid* and *OTX-1/2* and with the founder gene of the family, Drosophila *Bicoid*. In Drosophila, Bicoid is a critical specifier of 'anterior' identity and essential for setting up head–tail polarity of the early embryo. Although it is tempting to speculate that a *Bicoid*/*Pitx* system may have an ancient function in the specification of head–tail polarity, the *Bicoid* gene does not appear to have direct orthologues in species other than schizophoran flies (closely related to Drosophila), so this may be either a coincidence or convergent evolution.

The four key components that form part of this gene regulatory network initiating axis formation, *Pitx2*, *cVg1/GDF1*, *Tbx6,* and *Nodal*, are also involved a little later in development (from the late primitive streak stage) in specifying left–right asymmetry in different vertebrate classes (**Levin et al., 1995**; **Hyatt et al., 1996**; **Hyatt and Yost, 1998**; **Logan et al., 1998**; **Meno et al., 1998**; **Piedra et al., 1998**; **Ryan et al., 1998**; **St Amand et al., 1998**; **Yoshioka et al., 1998**; **Zhu et al., 1999**; **Rankin et al., 2000**; **Wall et al., 2000**; **Levin, 2005**; **Raya and Izpisua Belmonte, 2006**; **Tanaka et al., 2007**; **Hadjantonakis et al., 2008**). Indeed, these are the main conserved components of the left–right pathway among different vertebrates. These observations raise the possibility that the left–right pathway may have evolved by co-opting a more ancient mechanism for initiating formation of the gastrular axis. This is supported by the finding that a *Nodal*/*Pitx2* loop is involved in specifying both left–right asymmetry and mesendoderm formation (oral-aboral polarity) in the sea urchin, a non-vertebrate deuterostome (**Duboc et al., 2004**; **Hibino et al., 2006**; **Warner et al., 2012**). *Pitx2*, *Vg1*, *Nodal,* and *Tbx6* are also involved in early mesendoderm development in anamniotes, although *Vg1* is maternal (**Thomsen and Melton, 1993**; **Kessler and Melton, 1995**; **Faucourt et al., 2001**).

In the mouse, double mutants for Pitx1 and Pitx2 lead to early lethality (at pre- or peri-implantation stages) of the embryo (**Marcil et al., 2003**). Only a single embryo was ever recovered that had survived to E10–E12 (Drouin, personal communication). Expression of these transcription factors has to date only been studied in detail at later stages of mouse development (**L'Honore et al., 2007**; **Lanctot et al., 1997**) and it will therefore be interesting to see if they are indeed expressed as in the chick at pre-primitive streak stages. Certainly, the strong conservation of the active enhancers (E3 and E5, including the Pitx-binding sites contained therein) near mouse and human GDF1 suggests that this is likely to be a conserved feature of birds and mammals, and despite the fact that the peculiar geometry of rodent embryos at these early stages has led to some differences in the processes leading to axis development (**Stern and Downs, 2012**). Monozygotic twins do not seem to occur commonly in the mouse and it is possible that these very small, cup-shaped embryos do not survive to term if more than one primitive streak appears within a single blastocyst. Together, these findings suggest that the retention of a regulative mode of development at late stages by amniote embryos (that allows the formation of the types of monozygotic twins that arise relatively late in development, including Siamese twins) evolved through novel uses of an ancient pathway, involved in both mesendoderm development and left–right asymmetry, but in slightly different ways.

Our study also brings forward the time at which the earliest responses to cutting an embryo can be detected. First, we can now detect *cVg1* expression at one edge of a cut anterior fragment 4–5 hr after cutting, about 2 hr earlier than the 6 hr reported previously (**Bertocchini et al., 2004**); *Pitx2* appears even earlier, 3 hr after cutting. But this also begs the question of what lies upstream of *Pitx2*. At some early point in the cascade, the regulators will no longer be controlled at the transcriptional level and it will be considerably more difficult to identify them. The present finding that *Pitx2* is upregulated locally just 3 hr after cutting an embryo suggests that its regulators may not be differentially expressed mRNAs, but other asymmetries. Our studies allow us to predict the *Pitx2*/*cVg1*(*GDF1*)/*Nodal* pathway as a possible candidate to explain the obligate quadruplets of armadillos (**Newman and Patterson, 1910**; **Loughry et al., 1998**; **Enders, 2002**; **Eakin and Behringer, 2004**) and/or the high incidence of monozygotic and conjoined twins in certain human populations (**Cox, 1963**; **Hamamy et al., 2004**; **Forsberg et al., 2010**). Answering these questions represent interesting future challenges.

Materials and methods

Embryos, manipulation and RNA in situ hybridisation

Fertile Brown Bovan Gold hens' eggs (Henry Stewart, UK) were incubated for 1–16 hr to obtain stages X–XIII (**Eyal-Giladi and Kochav, 1976**) and stage 4 (**Hamburger and Hamilton, 1951**) (HH). Embryo manipulation was performed in Tyrode's solution. Anterior halves were obtained by cutting embryos with a hair loop. Unlike previous studies (**Bertocchini et al., 2004**), here we did not use a strip of anterior area opaca to seal the cut edge. Embryos and fragments were set up in modified New culture (**New, 1955**; **Stern and Ireland, 1981**) and incubated at 38°C as required. Whole mount in situ hybridisation was performed as described previously (**Stern, 1998**; **Streit and Stern, 2001**). The probes used were: chick *cVg1* (**Shah et al., 1997**), *Brachyury* (**Kispert et al., 1995a**; **Kispert et al., 1995b**; **Knezevic et al., 1997**), and *Pitx2* (**Logan et al., 1998**; **Zhu et al., 1999**).

Microarray screens, analysis and verification of candidate genes

Two microarray screens were performed with tissues collected from stages X–XII (**Eyal-Giladi and Kochav, 1976**). A first screen was performed with triplicates of 40 pieces of PMZ and AMZ, dissected and individually frozen from whole embryos (**Figure 1A**, **Figure 1—figure supplement 1**). A second screen was done with triplicates of 70 pieces of the left and right corners of the marginal zone, dissected and frozen individually from anterior embryo halves that had been cultured for 7 hr (**Figure 1C**, **Figure 1—figure supplement 2**). After dissection, whole embryos and anterior halves were fixed in 4% PFA for in situ hybridisation with *cVg1*. In both screens, in situ hybridisation was carried out for an extended period to detect low *cVg1* expression adjacent to the excised pieces and confirm the orientation of the embryo (see Results); the orientation was ambiguous or incorrect in about 10% of the embryos, and explants obtained from them were therefore not included. The remaining validated PMZ and AMZ samples, *cVg1*-like and non-*cVg1*-like samples were pooled in TRIzol reagent (Ambion, Invitrogen, UK). RNA was prepared and run for a complete Affymetrix analysis by ARK-Genomics. 500 ng of total RNA was required for the standard 3' IVT-Express protocol. Each label was quality checked through all stages of amplification and preparation for hybridisation on Affymetrix 30K chicken microarrays. Microarray raw data were analysed using Bioconductor in R (**Gentleman et al., 2004**). Raw datasets were normalised using the Robust Multi-array Average (RMA) method (**Irizarry et al., 2003**). Differentially expressed genes were then identified using the Limma package in R (**Smyth, 2005**) with a fold change threshold of 1.2 and $p < 0.05$. This strategy identified 122 sequences (corresponding to 85 genes) with putative *cVg1*-like expression ('*cVg1*-like') and 78 sequences (52 genes) expressed in the *cVg1*-negative explants ('*cVg1*-unlike'). The complete dataset was deposited with ArrayExpress where it can be accessed under Accession number E-MTAB-3116.

A selection of genes from the top up- and downregulated genes was verified by in situ hybridisation: *ADMP*, *PITX2*, *THPO*, *PKDCC*, *PMEPA*, *TBX6*, *FGF8*, *Ovoinhibitor*, *ELK3*, *LITAF*, *PITX1*, *CAMK1G*, *DENND5B*, *MLLT6*, *Homer3*, *SEMA5B*, *PIK3R1*, *GALNTL1*, *SH2D4A*, *GLI3*, *PDLIM5*, *VANGL1*, *LEF1*, *FGF18*, *LOC768709*, *CXCL14*, *PLCB1*, *ZBTB1*, *MED15*, *SPOCK3*, *BASP1*, *LRP2*, *GABRB2*, *PDGFA*, *Autotaxin*, *SALL1*, *MAFA*, *ESRRG*, *FOXD2*, *FSTL4*, *MALT1*, *OLFM3*, *GNAZ* and *GRB10*; when not already available, probes were generated from the chick EST collection (**Boardman et al., 2002**).

Insulator analysis

Computational prediction of CTCF insulator elements was performed as previously described (**Khan et al., 2013**). A PERL script (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3664090/#SD1) was used to scan chromosome 28 (location of *GDF3*/*cVg1*) from the galGal4 build of the chicken genome for occurrences of CTCF-binding sites with stringent parameters (False Discovery Rate, FDR 0%). Equivalent regions of human chromosome 19 (hg19 genome build, location of homologous GDF1) and mouse chromosome 8 (mm10 genome build, location of homologous GDF1) were also scanned for CTCF-binding sites with the same FDR parameter. The nearest conserved CTCF-binding sites harbouring the same set of genes both upstream and downstream of GDF3/GDF1 in all three species were then identified and these domains were defined as regions bounded by putative insulators.

Enhancer prediction

Enhancer prediction was carried out using the software package Discovery of Regulatory Elements in Vertebrates (DREiVe) (**Khan et al., 2013**), the vertebrate version of EDGI (**Sosinsky et al., 2007**).

Genomic coordinates for the predicted insulator region in human were used as the reference to predict order-independent conserved patterns of DNA sequences shared between human and any seven of the following species: horse (Equcab2), cow (Bostau4), rabbit (Orycun2), guinea pig (Cavpor3), mouse (mm9), opossum (Mondom5), platypus (Taegut1), chicken (Galgal3), and lizard (Anocar1). Parameters used included motif density of 6 matching nucleotides within a window length of 8 bp, where the minimum number of matching nucleotides in the motifs was set at 12 bp. The maximal cluster length (maximum length of predicted enhancers) was set at 3000 bp with a sequence conservation score cut-off of 2. This set of parameters successfully predicted a series of conserved blocks, designated E1–E4 (see ***Figure 5***), in human, chicken, mouse, and other species. Transcription factor binding site analysis of these predicted enhancers was carried out using the matrix-scan algorithm from the RSAT workbench (http://rsat.ulb.ac.be/rsat/) (***Thomas-Chollier et al., 2011***). Position frequency matrices from both Jaspar (http://jaspar.cgb.ki.se/) and Transfac (http://www.gene-regulation.com/pub/databases.html) libraries were used in matrix-scan where the background model estimation method was based on a Markov order of 0. Organism-specific 'upstream no-orf' background sequences were used (galGal4) with a pseudo-frequency of 0.01 and an upper p-value threshold of $1e^{-4}$. As a complementary approach, a modified methodology of Clover (http://cagt.bu.edu/page/Clover_about) was used to detect enhancers that shared order-independent transcription factor binding sites rather than DNA patterns. This approach uncovered two additional putative enhancers, E5 and E6.

Electroporation, morpholinos and DNA constructs

Fluorescein-labelled MOs against *Pitx2*, *Pitx1* and a standard control-MO (Gene Tools, Philomath, Oregon, USA) were delivered to young embryos by electroporation as described (***Voiculescu et al., 2007***; ***Voiculescu et al., 2008***). *Pitx2*-MO was designed to target the translational start site: (5′-CAAGTTTACGGCAGTTGGACTCCAT-3′). A splice blocking *Pitx1*-MO was designed to target the start of exon-2: (5′-CTCTCTTTTTCTACGGTGGGATGTT-3′). The coding sequence of chicken *Pitx2* (***Logan et al., 1998***) was cloned into pCAβ-IRES-GFP to generate a *Pitx2* expression construct. The expression of *Pitx2* protein was confirmed by western blot with a Pitx2-antibody (Abcam, UK, ab55599). Candidate enhancers 1–6 (E1–E6) were amplified from chick genomic DNA using the following primer pairs: (E1)-forward (5′-CAGCCCCAGGCAGACAGGGCTGCAGGGAAGAAGGGG-3′), reverse (5′-ACGGGACCCCCAGCCCTGCAGGATGCTGCCCGGGGT-3′); (E2)-forward (5′-GGAGGTACCATAA TTCATGCTTTCTGGGCTCGGGAC-3′), reverse (5′-GTACTCGAGCAAAGGTATCCCAGACCCTGCT GTC-3′); (E3)-forward (5′- GGAGGTACCACTCATTTGGTTTTAGCATTAATAAAC-3′), reverse (5′- GTACT CGAGCTGCCAGGGCAGAGGGAGCAGGGTG-3′); (E4)-forward (5′-CAGGGGATGAAGGGGGTGTTG GGGATCAAGCTCTTC-3′), reverse (5′-CAGTCTGCTACAATCCCTTCCCATGGATTCCTGGGG -3′); (E5)-forward (5′-GTGAGGTACCCTGTTCAGTC-3′), reverse (5′- AACTCGAGGTACAAGCTCTGC -3′); (E6)-forward (5′-GGTAGAGACCTGGTACCAGTAG-3′), reverse (5′-GAAGGGAGCTCGAGTGTCAC -3′). The amplified fragments were subcloned into pGEMT-easy (Promega, Madison, Wisconsin, USA), excised with inserted *KpnI* and *XhoI* sites and cloned into ptkEGFP (***Uchikawa et al., 2003***) or ptkRFP (kind gifts of H Kondoh). All constructs were checked by sequencing. 3 μg of each enhancer construct were co-electroporated with pCAβ-RFP, DS-RedExpress (Invitrogen, UK) or pCAβ-IRES-GFP as appropriate.

Mutations were introduced into each of the Pitx1- or Pitx2-binding sites of the four enhancers (E1, E3, E5, and E6) that contained such sites, with base changes highlighted in red (***Table 5***). The PCR primers used for site-directed mutagenesis are shown in ***Table 6***. For Enh5, site 1 refers to the 5′ Pitx2, while site 2 is the 3′ Pitx2-binding site (see ***Figure 5***). Site-directed mutagenesis was performed as described (***Liu and Naismith, 2008***). Briefly, 100 ng of template Enhancer DNA was PCR amplified with 2 μM of each mutant primer pair, 200 μM dNTPs, 2 μl Phusion high-fidelity polymerase, and 5 μl 10X buffer in a total

Table 5. Construction of mutations introduced to destroy Pitx-binding sites

Enh PITX site	Mutated bases
E1 PITX2	GCTGG**TC**AT**G**ACTTCTT
E1 PITX1	TGGTT**TC**GA**G**TTGAAAG
E3 PITX2	GGGTG**TC**AT**G**AATGGCC
E3 PITX1	AGGGT**TC**GA**G**TAATGGC
E5 PITX2 1	TAAAA**TC**CT**G**AAACTGC
E5 PITX2 2	AGAAA**TC**AT**G**AATAAAT
E6 PITX2	TGTCT**TC**AT**G**ACTCCTC

The mutations (highlighted in red) introduced into the sequences of Pitx-binding sites in enhancers E1, E3, E5 and E6.

DOI: 10.7554/eLife.03743.023

Table 6. Primers used for construction of mutations in Pitx-binding sites

Primer sequence	Primer name
aggggtttggttTCgaGttgaaagcgtgtacttctcacccattgaaactgccaggtctgt	Enh1PitX1MutF
ctttcaaCtcGAaaccaaaccccctcccatcacacacacaaactctactgcttcttcaaacca	Enh1PitX1MutR
agggggctggTCatGacttcttgcaggtgccccaaaggcagggcca	Enh1PitX2MutF
aagaagtCatGAccagccccctacccgaccctgctgcacagggaat	Enh1PitX2MutR
ggccattaCtcGAaccctgctccctctgccctggcag	Enh3PitX2MutF
ctgccagggcagagggagcagggtTCgaGtaatggcc	Enh3PitX2MutR
agtgcagtttCagGAttttactttatagttttttattcctacttgccatgaaagtggagaacat	Enh5PitX2MutF
taaaaTCctGaaactgcacttaacatagcgattacaaaacactcttggttagagtagcaacacaa	Enh5PitX2MutR
tgggctttcctatttattCatGAtttcttacagtagacaagaagattcacatttgtttgtgtttcatagctagag	Enh5PitX22MutF
tgtaagaaaTCatGaataaataggaaagcccaaatgttacaagcttcattgggcctccattgctggaagaaaaga	Enh5PitX22MutR
ataacaaaggggaggagtCatGAagacattgattctaatttgttctacatttcagtttattagcaaagtgaca	Enh6PitX2MutF
aatcaatgtctTCatGactcctccccttttgttatttcattttatgtgctttctaattcatcttcaattaaga	Enh6PitX2MutR

This table shows the primer combinations used to generate the mutations indicated in ***Table 5***.
DOI: 10.7554/eLife.03743.024

volume of 50 µl. The PCR programme was as follows: 94°C 3 min, 94°C 1 min, 52°C 1 min, 68°C 8, 12, or 24 min (500 bp/min), final extension 68°C 1 hr, then 4°C. 1/25th of each reaction was run on an Agarose gel to verify amplification, after which the remainder of the reaction was digested with DpnI to remove the parent template for 1 hr at 37°C. Another 1/25th of the DpnI digest was transformed into DH5α cells and clones selected and amplified in culture for DNA extraction and sequence verification of the introduced mutations.

ChIP and qPCR

Micro ChIP experiments were performed with pools of 25 pieces of PMZ and their anterior counterparts (AMZ) from whole pre-streak stage X–XI embryos. Chromatin was cross-linked, extracted and sonicated to obtain 200–1000 bp fragments. Immunoprecipitation was done with rabbit anti-H3K27ac (Abcam, UK), mouse anti-PITX2 (Abcam, ab55599), antibodies and rabbit IgG (Millipore, Merck Millipore UK) and mouse IgG (Millipore, Merck Millipore UK) bound to magnetic Dynabeads (Life Technologies, Carlsbad, California, USA). The purified DNA was used as template for qPCR analyses with a Bio-Rad iCycler and Sybr Fast Q-PCR mix (Kapa Biosystems, Wilmington, Massachusetts, USA). The following primer pairs were used: enhancer1—forward primer1 E1F1 (5′-gctctatcccgattccctgtgca-3′) reverse primer1 E1R1 (5′- ggtagggttcacttcattagggatg -3′, forward primer2 E1F5 (5′-cactctgtgtccggagaatgctac-3′) reverse primer2 E1R5 (5′- caatgggtgagaagtacacgctttc-3′ ; enhancer2—forward primer1 E2F1 (5′- atgaagagcgcaagagtggcaaag-3′) reverse primer1 E2R1 (5′-tgggaaattccgacctggaagcag-3′, forward primer2 E2F3 (5′- agctgagctgtttaacggtggatc-3′) reverse primer2 E2R3 (5′- agctgaccgctgctagttctct-3′); enhancer 3—forward primer E3F2 (5′- cctagctattacactctgctcttcc-3′) reverse primer E3R2 (5′- CAGTAGAGACGGAACCAGAACct-3′); enhancer 4—forward primer1 E4F1 (5′- gcttcttggtgctggaactgagaat-3′) reverse primer1 E4R1(5′- atggagagacacagtcagccacga-3′, forward primer2 E4F2 (5′- agctgcccaatgctctgaaagaag -3′) reverse primer2 E4R2 (5′- ggaatggaaaggtgaggattcatgg -3′); enhancer 5—forward primer E5F3 (5′- CGTGAGGCAGTCTGTTCAGT-3′) reverse primer E5R3 (5′-GCTATGAAACACAAACAAATGTGAA-3′); enhancer 6—forward primer E6F3 (5′- CACTGGGGTCCTGATGTAGTG-3′), reverse primer E6R3 (5′- AGAAGGGAGCACAAATGTCA-3′); Vg1 -0.5 ATG—forward primer Vg1F1 (5′-agtgggtgctgattgctgtctgtg-3′) reverse primer Vg1R1 (5′-ctcccatcccttctgcatctccata-3′). All qPCR reactions were run in triplicate and contained an input control for normalisation of chromatin concentration, and the appropriate IgG as control. The amplification products were quantified using the ΔΔCt method (***Livak and Schmittgen, 2001***).

Acknowledgements

This study was funded by an Advanced Investigator Grant from the European Research Council (ERC) to CDS ("GEMELLI"). LMS-J was a thesis student from the Genomic Science Program (UNAM, Campus

Cuernavaca, Mexico). IL is a student on the Wellcome Trust-funded 4-year PhD Programme in Developmental and Stem Cell Biology at UCL. We are grateful to Jacques Drouin for a generous gift of PITX-antibodies, to Hisatoh Kondoh for the TK-GFP and TK-RFP reporter vectors, to Jacques Drouin and Aurore L'Honoré for sharing unpublished information and to Irene De Almeida, Rosemary Bachvarova, Patrick Lemaire, Bruno Reversade and Andrea Streit for helpful comments on the manuscript and advice.

Additional information

Funding

Funder	Grant reference number	Author
European Research Council	GEMELLI 249892	Angela Torlopp, Mohsin A F Khan, Nidia M M Oliveira, Claudio D Stern
Wellcome Trust	WT083345MA	Ingrid Lekk, Claudio D Stern

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

AT, NMMO, LMS-J, Acquisition of data, Analysis and interpretation of data; MAFK, Conception and design, Analysis and interpretation of data; IL, Perfected the method for ChIP, performed the initial experiments and analysed the results; AS, Conceived and designed the DREiVe analysis used to predict conserved enhancers across species, Analysis and interpretation of data, Contributed unpublished essential data or reagents; CDS, Conceived and designed the project including experimental design of the screen, directed the team and wrote the paper, Analysis and interpretation of data

Author ORCIDs

Claudio D Stern, http://orcid.org/0000-0002-9907-889X

Additional files

Major dataset

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Khan M, Torlopp A, Stern CD	2014	Pitx2 positions the embryonic axis and regulates twinning	E-MTAB-3116	Publicly available at Array Express (https://www.ebi.ac.uk/arrayexpress/).

References

Andersson O, Bertolino P, Ibáñez CF. 2007. Distinct and cooperative roles of mammalian Vg1 homologs GDF1 and GDF3 during early embryonic development. *Developmental Biology* **311**:500–511. doi: 10.1016/j.ydbio.2007.08.060.

Azar Y, Eyal-Giladi H. 1979. Marginal zone cells–the primitive streak-inducing component of the primary hypoblast in the chick. *Journal of Embryology and Experimental Morphology* **52**:79–88.

Bachvarova RF, Skromne I, Stern CD. 1998. Induction of primitive streak and Hensen's node by the posterior marginal zone in the early chick embryo. *Development* **125**:3521–3534.

Bertocchini F, Skromne I, Wolpert L, Stern CD. 2004. Determination of embryonic polarity in a regulative system: evidence for endogenous inhibitors acting sequentially during primitive streak formation in the chick embryo. *Development* **131**:3381–3390. doi: 10.1242/dev.01178.

Bertocchini F, Stern CD. 2002. The hypoblast of the chick embryo positions the primitive streak by antagonizing nodal signaling. *Developmental Cell* **3**:735–744. doi: 10.1016/S1534-5807(02)00318-0.

Bertocchini F, Stern CD. 2012. Gata2 provides an early anterior bias and uncovers a global positioning system for polarity in the amniote embryo. *Development* **139**:4232–4238. doi: 10.1242/dev.081901.

Birsoy B, Kofron M, Schaible K, Wylie C, Heasman J. 2006. Vg1 is an essential signaling molecule in Xenopus development. *Development* **133**:15–20. doi: 10.1242/dev.02144.
Boardman PE, Sanz-Ezquerro J, Overton IM, Burt DW, Bosch E, Fong WT, Tickle C, Brown WR, Wilson SA, Hubbard SJ. 2002. A comprehensive collection of chicken cDNAs. *Current Biology* **12**:1965–1969. doi: 10.1016/S0960-9822(02)01296-4.
Carpenter B, Hill KJ, Charalambous M, Wagner KJ, Lahiri D, James DI, Andersen JS, Schumacher V, Royer-Pokora B, Mann M, Ward A, Roberts SG. 2004. BASP1 is a transcriptional cosuppressor for the Wilms' tumor suppressor protein WT1. *Molecular and Cellular Biology* **24**:537–549. doi: 10.1128/MCB.24.2.537-549.2004.
Chai CK, Crary DD. 1971. Conjoined twinning in rabbits. *Teratology* **4**:433–444. doi: 10.1002/tera.1420040407.
Chen C, Ware SM, Sato A, Houston-Hawkins DE, Habas R, Matzuk MM, Shen MM, Brown CW. 2006. The Vg1-related protein Gdf3 acts in a Nodal signaling pathway in the pre-gastrulation mouse embryo. *Development* **133**:319–329. doi: 10.1242/dev.02210.
Cox ML. 1963. INCIDENCE AND AETIOLOGY OF MULTIPLE BIRTHS IN NIGERIA. *The Journal of Obstetrics and Gynaecology of the British Commonwealth* **70**:878–884. doi: 10.1111/j.1471-0528.1963.tb04995.x.
Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of USA* **107**:21931–21936. doi: 10.1073/pnas.1016071107.
Cunningham B. 1937. *Axial Bifurcation in Serpents, An Historical Survey of Serpent Monsters Having Part of the Axial Skeleton Duplicated*. Durham, N.C: Duke University Press.
Driesch H. 1892. Entwicklungsmechanische Studien. I. Der Werth der beiden ersten Furchungszellen in der Echinodermenentwicklung. Experimentelle Erzeugen von Theil- und Doppelbildung. Z. wissensch. *Zoology* **53**:160–184.
Duboc V, Röttinger E, Besnardeau L, Lepage T. 2004. Nodal and BMP2/4 signaling organizes the oral-aboral axis of the sea urchin embryo. *Developmental Cell* **6**:397–410. doi: 10.1016/S1534-5807(04)00056-5.
Eakin GS, Behringer RR. 2004. Gastrulation in other mammals and humans. In: Stern CD, editor. *Gastrulation: from Cells to Embryo*. New York: Cold Spring Harbor Press. pp. 275–287.
Enders AC. 2002. Implantation in the nine-banded armadillo: how does a single blastocyst form four embryos? *Placenta* **23**:71–85. doi: 10.1053/plac.2001.0753.
Eyal-Giladi H, Kochav S. 1976. From cleavage to primitive streak formation: a complementary normal table and a new look at the first stages of the development of the chick. I. General morphology. *Developmental Biology* **49**:321–337. doi: 10.1016/0012-1606(76)90178-0.
Faucourt M, Houliston E, Besnardeau L, Kimelman D, Lepage T. 2001. The pitx2 homeobox protein is required early for endoderm formation and nodal signaling. *Developmental Biology* **229**:287–306. doi: 10.1006/dbio.2000.9950.
Forsberg CW, Goldberg J, Sporleder J, Smith NL. 2010. Determining zygosity in the Vietnam era twin registry: an update. *Twin Research and Human Genetics* **13**:461–464. doi: 10.1375/twin.13.5.461.
Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**:R80. doi: 10.1186/gb-2004-5-10-r80.
Hadjantonakis AK, Pisano E, Papaioannou VE. 2008. Tbx6 regulates left/right patterning in mouse embryos through effects on nodal cilia and perinodal signaling. *PLOS ONE* **3**:e2511. doi: 10.1371/journal.pone.0002511.
Hamamy HA, Ajlouni HK, Ajlouni KM. 2004. Familial monozygotic twinning: report of an extended multi-generation family. *Twin Research* **7**:219–222. doi: 10.1375/136905204774200479.
Hamburger V, Hamilton HL. 1951. A series of normal stages in the development of the chick embryo. *Journal of Morphology* **88**:49–92. doi: 10.1002/jmor.1050880104.
Hibino T, Nishino A, Amemiya S. 2006. Phylogenetic correspondence of the body axes in bilaterians is revealed by the right-sided expression of Pitx genes in echinoderm larvae. *Development, Growth & Differentiation* **48**:587–595. doi: 10.1111/j.1440-169X.2006.00892.x.
Holwerda SJ, de Laat W. 2013. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **368**:20120369. doi: 10.1098/rstb.2012.0369.
Hyatt BA, Lohr JL, Yost HJ. 1996. Initiation of vertebrate left-right axis formation by maternal Vg1. *Nature* **384**:62–65. doi: 10.1038/384062a0.
Hyatt BA, Yost HJ. 1998. The left-right coordinator: the role of Vg1 in organizing left-right axis formation. *Cell* **93**:37–46. doi: 10.1016/S0092-8674(00)81144-7.
Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**:249–264. doi: 10.1093/biostatistics/4.2.249.
Kaufman MH. 2004. The embryology of conjoined twins. *Child's Nervous System* **20**:508–525. doi: 10.1007/s00381-004-0985-4.
Kessler DS. 2004. Activin and Vg1 and the search for embryonic inducers. In: Stern CD, editor. *Gastrulation: from cells to embryo*. New York: Cold Spring Harbor Press. pp. 505–520.
Kessler DS, Melton DA. 1995. Induction of dorsal mesoderm by soluble, mature, Vg1 protein. *Development* **121**:2155–2164.

Khan MA, Soto-Jimenez LM, Howe T, Streit A, Sosinsky A, Stern CD. 2013. Computational tools and resources for prediction and analysis of gene regulatory regions in the chick genome. *Genesis* **51**:311–324. doi: 10.1002/dvg.22375.
Khaner O, Eyal-Giladi H. 1989. The chick's marginal zone and primitive streak formation. I. Coordinative effect of induction and inhibition. *Developmental biology* **134**:206–214. doi: 10.1016/0012-1606(89)90090-0.
Kispert A, Koschorz B, Herrmann BG. 1995a. The T protein encoded by Brachyury is a tissue-specific transcription factor. *The EMBO Journal* **14**:4763–4772.
Kispert A, Ortner H, Cooke J, Herrmann BG. 1995b. The chick Brachyury gene: developmental expression pattern and response to axial induction by localized activin. *Developmental Biology* **168**:406–415. doi: 10.1006/dbio.1995.1090.
Knezevic V, De Santo R, Mackem S. 1997. Two novel chick T-box genes related to mouse Brachyury are expressed in different, non-overlapping mesodermal domains during gastrulation. *Development* **124**:411–419.
L'Honoré A, Coulon V, Marcil A, Lebel M, Lafrance-Vanasse J, Gage P, Camper S, Drouin J. 2007. Sequential expression and redundancy of Pitx2 and Pitx3 genes during muscle development. *Developmental biology* **307**:421–433. doi: 10.1016/j.ydbio.2007.04.034.
Lanctôt C, Lamolet B, Drouin J. 1997. The bicoid-related homeoprotein Ptx1 defines the most anterior domain of the embryo and differentiates posterior from anterior lateral mesoderm. *Development* **124**:2807–2817.
Levin M. 2005. Left-right asymmetry in embryonic development: a comprehensive review. *Mechanisms of Development* **122**:3–25. doi: 10.1016/j.mod.2004.08.006.
Levin M, Johnson RL, Stern CD, Kuehn M, Tabin C. 1995. A molecular pathway determining left-right asymmetry in chick embryogenesis. *Cell* **82**:803–814. doi: 10.1016/0092-8674(95)90477-8.
Liu H, Naismith JH. 2008. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology* **8**:91. doi: 10.1186/1472-6750-8-91.
Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**:402–408. doi: 10.1006/meth.2001.1262.
Logan M, Pagan-Westphal SM, Smith DM, Paganessi L, Tabin CJ. 1998. The transcription factor Pitx2 mediates situs-specific morphogenesis in response to left-right asymmetric signals. *Cell* **94**:307–317. doi: 10.1016/S0092-8674(00)81474-9.
Loughry WJ, Prodohl PA, McDonough CM, Avise JC. 1998. Polyembryony in armadillos. *American Scientist* **86**:274–279. doi: 10.1511/1998.3.274.
Lutz H. 1949. Sur la production expérimentale de la polyembryonie et de la monstruosité double chez les oiseaux. *Archives d'anatomie Microscopique Et De Morphologie Expérimentale* **39**:79–144.
Marcil A, Dumontier E, Chamberland M, Camper SA, Drouin J. 2003. Pitx1 and Pitx2 are required for development of hindlimb buds. *Development* **130**:45–55. doi: 10.1242/dev.00192.
Meno C, Shimono A, Saijoh Y, Yashiro K, Mochida K, Ohishi S, Hamada H. 1998. lefty-1 is required for left-right determination as a regulator of lefty-2 and nodal. *Cell* **94**:287–297. doi: 10.1016/S0092-8674(00)81472-5.
Merkenschlager M, Odom DT. 2013. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* **152**:1285–1297. doi: 10.1016/j.cell.2013.02.029.
New DAT. 1955. A new technique for the cultivation of the chick embryo in vitro. *Journal of Embryology and Experimental Morphology* **3**:326–331.
Newman HH, Patterson JT. 1910. The development of the nine-banded armadillo from primmitive streak till birth, with special reference to the question of specific polyembryony. *Journal of Morphology* **21**:359–423. doi: 10.1002/jmor.1050210303.
Piedra ME, Icardo JM, Albajar M, Rodriguez-Rey JC, Ros MA. 1998. Pitx2 participates in the late phase of the pathway controlling left-right asymmetry. *Cell* **94**:319–324. doi: 10.1016/S0092-8674(00)81475-0.
Rankin CT, Bunton T, Lawler AM, Lee SJ. 2000. Regulation of left-right patterning in mice by growth/differentiation factor-1. *Nature Genetics* **24**:262–265. doi: 10.1038/73472.
Raya A, Izpisua Belmonte JC. 2006. Left-right asymmetry in the vertebrate embryo: from early information to higher-level integration. *Nature Reviews Genetics* **7**:283–293. doi: 10.1038/nrg1830.
Ryan AK, Blumberg B, Rodriguez-Esteban C, Yonei-Tamura S, Tamura K, Tsukui T, de la Peña J, Sabbagh W, Greenwald J, Choe S, Norris DP, Robertson EJ, Evans RM, Rosenfeld MG, Izpisua Belmonte JC. 1998. Pitx2 determines left-right asymmetry of internal organs in vertebrates. *Nature* **394**:545–551. doi: 10.1038/29004.
Seleiro EA, Connolly DJ, Cooke J. 1996. Early developmental expression and experimental axis determination by the chicken Vg1 gene. *Current Biology* **6**:1476–1486. doi: 10.1016/S0960-9822(96)00752-X.
Shah SB, Skromne I, Hume CR, Kessler DS, Lee KJ, Stern CD, Dodd J. 1997. Misexpression of chick Vg1 in the marginal zone induces primitive streak formation. *Development* **124**:5127–5138.
Sheng G, Stern CD. 1999. Gata2 and Gata3: novel markers for early embryonic polarity and for non-neural ectoderm in the chick embryo. *Mechanisms of Development* **87**:213–216. doi: 10.1016/S0925-4773(99)00150-1.
Skromne I, Stern CD. 2001. Interactions between Wnt and Vg1 signalling pathways initiate primitive streak formation in the chick embryo. *Development* **128**:2915–2927.
Skromne I, Stern CD. 2002. A hierarchy of gene expression accompanying induction of the primitive streak by Vg1 in the chick embryo. *Mechanisms of Development* **114**:115–118. doi: 10.1016/S0925-4773(02)00034-5.
Smyth GK. 2005. Limma: linear models for microarray data. In Gentleman RC, Carey V, Dudoit S, Irizarry R, Huber W. Editors, *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer. 397–420.

Sosinsky A, Honig B, Mann RS, Califano A. 2007. Discovering transcriptional regulatory regions in Drosophila by a nonalignment method for phylogenetic footprinting. *Proceedings of the National Academy of Sciences of USA* **104**:6305–6310. doi: 10.1073/pnas.0701614104.
Spratt NT, Haas H. 1960. Integrative mechanisms in development of the early chick blastoderm. I. Regulative potentiality of separated parts. *Journal of Experimental Zoology* **145**:97–137. doi: 10.1002/jez.1401450202.
St Amand TR, Ra J, Zhang Y, Hu Y, Baber SI, Qiu M, Chen Y. 1998. Cloning and expression pattern of chicken Pitx2: a new component in the SHH signaling pathway controlling embryonic heart looping. *Biochemical and Biophysical Research Communications* **247**:100–105. doi: 10.1006/bbrc.1998.8740.
Stern CD. 1998. Detection of multiple gene products simultaneously by in situ hybridization and immunohistochemistry in whole mounts of avian embryos. *Current Topics in Developmental Biology* **36**:223–243. doi: 10.1016/S0070-2153(08)60505-0.
Stern CD, Downs KM. 2012. The hypoblast (visceral endoderm): an evo-devo perspective. *Development* **139**:1059–1069. doi: 10.1242/dev.070730.
Stern CD, Ireland GW. 1981. An integrated experimental study of endoderm formation in avian embryos. *Anatomy and Embryology* **163**:245–263. doi: 10.1007/BF00315703.
Streit A, Stern CD. 2001. Combined whole-mount in situ hybridization and immunohistochemistry in avian embryos. *Methods* **23**:339–344. doi: 10.1006/meth.2000.1146.
Tanaka C, Sakuma R, Nakamura T, Hamada H, Saijoh Y. 2007. Long-range action of Nodal requires interaction with GDF1. *Genes & Development* **21**:3272–3282. doi: 10.1101/gad.1623907.
Thomas-Chollier M, Defrance M, Medina-Rivera A, Sand O, Herrmann C, Thieffry D, van Helden J. 2011. RSAT 2011: regulatory sequence analysis tools. *Nucleic Acids Research* **39**:W86–W91. doi: 10.1093/nar/gkr377.
Thomsen GH, Melton DA. 1993. Processed Vg1 protein is an axial mesoderm inducer in Xenopus. *Cell* **74**:433–441. doi: 10.1016/0092-8674(93)80045-G.
Uchikawa M, Ishida Y, Takemoto T, Kamachi Y, Kondoh H. 2003. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Developmental Cell* **4**:509–519. doi: 10.1016/S1534-5807(03)00088-1.
Ulshafer RJ, Clavert A. 1979. The use of avian double monsters in studies on induction of the nervous system. *Journal of Embryology and Experimental Morphology* **53**:237–243.
Vanderzon DM, Partlow GD, Fisher KR, Halina WG. 1998. Parapagus conjoined twin Holstein calf. *The Anatomical record* **251**:60–65. doi: 10.1002/(SICI)1097-0185(199805)251:13.0.CO;2-9.
Voiculescu O, Bertocchini F, Wolpert L, Keller RE, Stern CD. 2007. The amniote primitive streak is defined by epithelial cell intercalation before gastrulation. *Nature* **449**:1049–1052. doi: 10.1038/nature06211.
Voiculescu O, Papanayotou C, Stern CD. 2008. Spatially and temporally controlled electroporation of early chick embryos. *Nature Protocols* **3**:419–426. doi: 10.1038/nprot.2008.10.
Wall NA, Craig EJ, Labosky PA, Kessler DS. 2000. Mesendoderm induction and reversal of left-right pattern by mouse Gdf1, a Vg1-related gene. *Developmental Biology* **227**:495–509. doi: 10.1006/dbio.2000.9926.
Wang B, Shi G, Fu Y, Xu X. 2007. Cloning and characterization of a LASS1-GDF1 transcript in rat cerebral cortex: conservation of a bicistronic structure. *DNA sequence* **18**:92–103. doi: 10.1080/10425170601060947.
Warner JF, Lyons DC, McClay DR. 2012. Left-right asymmetry in the sea urchin embryo: BMP and the asymmetrical origins of the adult. *PLOS Biology* **10**:e1001404. doi: 10.1371/journal.pbio.1001404.
Weeks DL, Melton DA. 1987. A maternal mRNA localized to the vegetal hemisphere in Xenopus eggs condes for a growth factor related to TGF-beta. *Cell* **51**:861–867. doi: 10.1016/0092-8674(87)90109-7.
Wilson EB. 1898. Considerations on cell-lineage and ancestral reminiscence, based on a reexamination of some points in the early development of Annelids and Polyclades. Lectures Marine Biol. Lab. Woods Hole, MA. **1898**:21–42.
Yoshioka H, Meno C, Koshiba K, Sugihara M, Itoh H, Ishimaru Y, Inoue T, Ohuchi H, Semina EV, Murray JC, Hamada H, Noji S. 1998. Pitx2, a bicoid-type homeobox gene, is involved in a lefty-signaling pathway in determination of left-right asymmetry. *Cell* **94**:299–305. doi: 10.1016/S0092-8674(00)81473-7.
Zhu L, Marvin MJ, Gardiner A, Lassar AB, Mercola M, Stern CD, Levin M. 1999. Cerberus regulates left-right asymmetry of the embryonic head and heart. *Current Biology* **9**:931–938. doi: 10.1016/S0960-9822(99)80419-9.