Spemann organizer transcriptome induction by early beta-catenin, Wnt, Nodal, and Siamois signals in *Xenopus laevis*

Yi Ding, Diego Lopez, Eric A. Sosa, Gabriele Colozza, Yuki Moriyama, Maria D. J. Benitez, Kelvin Zhang, Daria Merkurjev, and Edward M. De Robertis

Howard Hughes Medical Institute, University of California, Los Angeles, CA 90095-1662; Department of Biological Chemistry, University of California, Los Angeles, CA 90095-1662; Department of Medicine, University of California, Los Angeles, CA 90095-1662; Department of Microbiology, University of California, Los Angeles, CA 90095-1662; and Department of Human Genetics, University of California, Los Angeles, CA 90095-1662

Contributed by Edward M. De Robertis, February 24, 2017 (sent for review January 17, 2017; reviewed by Juan Larraín and Stefano Piccolo)

**Significance**

We present a genome-wide study of the signals responsible for the early induction of the body axis in the following experimental conditions: beta-catenin morpholino; Wnt, Siamois, and Cerberus mRNAs; LiCl treatment; and dorsal-ventral regenerating half-embryos bisected at gastrula. Comparing 46 RNA-seq libraries, we uncovered the genetic networks that initiate dorsal-ventral patterning and Spemann’s organizer formation. We defined an early beta-catenin signature that has only minor overlap with recently published late zygotic Wnt signatures. The relation of these early steps of development to endomesodermal germ layer induction was studied by overexpressing the growth factor antagonist Cerberus. This study offers a rich resource for understanding the earliest inductive events in the body plan of a model vertebrate embryo.

With the recent completion of the *Xenopus laevis* genome (1), it has now become practical to study global transcriptional changes at the earliest stages of development in this classical model organism: for example, by analyzing the transcriptomes of the dorsal and ventral regions of the gastrula embryo (2). Recent studies analyzed the zygotic Wnt pathway in *X. laevis* and *Xenopus tropicalis* embryos, defining late gene signatures transcriptionally activated by endogenous Wnt8 expressed in the ventro-lateral marginal zone (3, 4). This zygotic Wnt signal occurs during mid- to late-gastrulation, promotes the formation of ventral-posterior tissues, and inhibits head development. For example, when Wnt8 DNA constructs that are expressed only at gastrula are injected into embryos (5), head structures are reduced. Reciprocally, overexpression of Wnt antagonists such as Dickkopf1 (Dkk1) (6) and Frzb1 (7) results in enlarged head structures.

The opposite phenotype is elicited by microinjection of Wnt mRNAs during early cleavage, generating embryos with spectacular secondary axes containing complete heads when injected into ventral blastomeres (8). When injected radially into embryos, Wnt8 mRNA leads to a dorsalized phenotype consisting entirely of head structures without trunks and a radial Spemann organizer (9–11). Similar dorsalizing effects are obtained by incubating embryos in lithium chloride (LiCl) solution at the 32-cell stage (12). LiCl mimics the early Wnt signal by inhibiting the enzymatic activity of glycogen synthase kinase 3 (GSK3) (13), an enzyme down-regulated by the canonical Wnt signal (14, 15). The molecular nature of the endogenous dorsal signal in *Xenopus* eggs remains elusive, but it is clear that its main effect is to cause the stabilization of beta-catenin during cleavage on the dorsal side of the embryo (16, 17). A beta-catenin antisense morpholino (β-CatMO) is one of the most potent and reliable reagents in *X. laevis* embryology, and its microinjection results in embryos lacking Spemann organizer tissue and dorsal development (18).

Dorsal beta-catenin stabilization requires a cortical rotation of the egg cortex toward the animal pole and the displacement of maternal cytoplasmic determinants containing Dishevelled (Dvl) toward the animal pole (19, 20). Interestingly, the endogenous cytoplasmic determinants are refractory to inhibition by Dkk1 and Frzb mRNAs, which, however, readily inhibit microinjected...
$Wnt$ mRNA (21, 22). The combination of maternal $\beta$-catenin, which induces the homeobox transcriptional activator $Siamois$/Twin (10), with high zygotic Nodal/TGF-$\beta$ (called $Xnr1/2/5/6$ in $Xenopus$) signaling on the dorsal side induces formation of Spemann organizer tissue (23). The organizer in turn secretes BMP antagonists such as Chordin and Noggin, Wnt antagonists such as Dkk1, Frzb1 and Crescent, and multivalent inhibitors such as Cerberus (24). These proteins shape dorsal-ventral (D-V) and anterior-posterior (A-P) patterning (25, 26). Therefore, understanding the signaling pathways that induce the transcriptional activation of Spemann organizer tissue is of fundamental importance for understanding vertebrate morphogenesis.

In the present study, we applied genome-wide transcriptome analyses to investigate the induction of the Spemann organizer by the maternal $Wnt/\beta$-catenin signal. We used $\beta$-catenin MO-, $Wnt8$ mRNA-, $Siamois$ mRNA-, and $Cerberus$ mRNA-injected embryos and LiCl-treated embryos. We also performed RNA-seq of un.injected embryos bisected into dorsal and ventral fragments at midblastula and allowed to develop for 5 h before isolating RNA at early gastrula. This procedure maximizes dorsal-ventral (D-V) differences and, in addition, tests the regenerative potential of the embryo (26). A total of 46 RNA-seq libraries were examined, comparing levels of the expression of 43,673 transcripts at late blastula (stage 9) and early gastrula (stage 10.5) of $X. laevis$. The results define an early $\beta$-catenin signature that has very little overlap with the zygotic Wnt signature and generate a comprehensive genome-wide catalog of the earliest developmental transcriptional changes in a vertebrate embryo.

**Results**

Defining Spemann Organizer Induction by $\beta$-Catenin. As shown in Fig. 1A, the dorsal maternal determinant signal can be blocked by microinjection of $\beta$-CatMO at the two- to four-cell stage. This signal can be mimicked by microinjection of $Wnt8$ mRNA, LiCl treatment, or microinjection of mRNA for the downstream effector $Siamois$. The early $\beta$-catenin signal results in the induction of the TGF-$\beta$-related molecules $Xnr5/6$, and a mixture of organizer growth factor inhibitors on the dorsal side that oppose signals provided by BMP and the zygotic (late) Wnt signaling pathways on the ventral side of the embryo (24, 25) (Fig. 1A). As indicated in Fig. 1B, the effect of $\beta$-catenin loss of function or $Wnt8$ mRNA gain of function was analyzed at early gastrula by RNA-seq in three independent experiments. In addition, inhibition of GSK3 by LiCl at the 32-cell stage and overexpression of $Siamois$ mRNA were analyzed in duplicate, and the effect of blocking endomesoderm induction with $Cerberus$ mRNA was analyzed in triplicate (Fig. 1B). As controls, we used uninjected embryos because previous work indicates that they behave like mock-injected embryos (27). We also analyzed transcripts from

![Fig. 1. RNA-seq libraries for transcriptome analysis of early $\beta$-catenin targets.](image-url)

Fig. 1. RNA-seq libraries for transcriptome analysis of early $\beta$-catenin targets. (A) Diagram showing the early $\beta$-catenin-dependent gene network on the dorsal side and the zygotic $Wnt8$ signal that arises later on the ventral side. The embryonic manipulations that were subjected to RNA-seq in the present study are indicated in red. (B) For RNA-seq, embryos were injected with $\beta$-catenin MO, $Wnt8$, $Siamois$, or $Cerberus$ mRNA, treated with LiCl, or cut into dorsal and ventral halves at stage 8. All embryos were harvested at stage 10.5 for RNA extraction and RNA-seq. Some embryos were also analyzed at blastula stage 9. (C–G) Phenotypes of embryos subjected to manipulations in this study. Note that $\beta$-CatMO–injected embryos lacked all traces of a dorsal axis whereas $Wnt8$, $Siamois$ mRNA microinjected, and LiCl-treated embryos were all radially dorsalized. $Cerberus$ mRNA–microinjected embryos lacked endomesoderm and consisted mostly of anterior neural tissue and cement gland. Ventral half-embryos lacked dorsal axes whereas dorsal half-embryos developed a complete axis. (Magnification: 8x.) (H) Scatter plot comparing the log$_2$ of fold change (FC) in RPKM upon $\beta$-catenin knockdown with that of dorsal halves/ventral halves. Shown here are 40,157 transcripts; red dots indicate transcripts decreased by $\beta$-catenin MO (by at least 1.4-fold in average of three experiments) and enriched in dorsal halves compared with ventral halves (by at least 1.5-fold in average of five experiments). Full datasets are provided in Dataset S1. Cer, Cerberus; D, dorsal half; St., stage; V, ventral half.
dorsal and ventral half-embryos (bisected at stage 8 and allowed to regenerate for 5 h until stage 10.5) in quintuplicate (Fig. 1B). In all experiments, sibling embryos were allowed to develop until tailbud stage, and gastrula stage 10.5 RNA samples were used for RNA-seq only when completely dorsalized or ventralized phenotypes, such as those in Fig. 1C–G, were obtained. The RPKM (reads per kilobase per million mapped reads) for all annotated X. laevis genes are provided in Dataset S1, and the raw data for all 46 libraries are available at the GEO repository (accession no. GSE93195).

β-Catenin binds to thousands of gene promoters during early Xenopus development (3, 4). However, the initial accumulation of mRNA although some transcripts are reproduced in three independent replicates, but also were reproducibly enriched dorsally in four of five D-V experiments. The D-V quintuplicate samples identified most of the known Spemann organizer genes, as well as the classical ventrally enriched genes (Fig. S1D).

The early dorsal β-catenin signature identified by these experiments is listed in Fig. 2A. Of these 123 dorsal genes, 86 were reported to be bound by β-catenin in X. tropicalis (4) (Dataset S2). There were also 124 genes that were decreased by β-CATMO but not dorsally enriched, of which 34 bound β-catenin in ChIP experiments (Dataset S3) (4). The early dorsal signature includes many secreted factors and transcription factors known to be induced genes was performed using Gene Set Enrichment Analysis (GSEA) (29). All genes with RPKM values above 1 (in corresponding control embryos, ~16,000 genes) were ranked according to their average fold change. The early β-catenin gene set was greatly enriched within the genes ranked according to their levels of induction by Wnt8 mRNA (indicated as vertical black lines in Fig. 4D). Similarly, LiCl and Siamois-induced gene ranking showed enrichment among many secreted factors and transcription factors known to be dorsally required for endomesodermal differentiation (Fig. 4B). The expression of the early β-catenin signature across the different conditions sequenced was reproducible, as indicated by Pearson correlation matrix analysis (Fig. 4C). In addition, principal component analysis (PCA) showed that replicates and Spemann organizer-inducing conditions clustered together, confirming reproducible expression of the early β-catenin signature (Fig. S4). Dataset S2 provides the RPKM levels for the early β-catenin signature genes in each library.

Analysis of the early β-catenin signature enrichment among Wnt8 mRNA-, LiCl-, Siamois mRNA-, and Cerberus mRNA-induced genes was performed using Gene Set Enrichment Analysis (GSEA) (29). All genes with RPKM values above 1 (in corresponding control embryos, ~16,000 genes) were ranked according to their average fold change. The early β-catenin gene set was greatly enriched within the genes ranked according to their levels of induction by Wnt8 mRNA (indicated as vertical black lines in Fig. 4D). Similarly, LiCl- and Siamois-induced gene ranking showed enrichment among many secreted factors and transcription factors known to be dorsally required for endomesodermal differentiation (Fig. 4B). The expression of the early β-catenin signature across the different conditions sequenced was reproducible, as indicated by Pearson correlation matrix analysis (Fig. 4C). In addition, principal component analysis (PCA) showed that replicates and Spemann organizer-inducing conditions clustered together, confirming reproducible expression of the early β-catenin signature (Fig. S4). Dataset S2 provides the RPKM levels for the early β-catenin signature genes in each library.

Cerberus Inhibits Most Endomesodermal Organizer Genes. Cerberus is a protein secreted by the head organizer that functions as a multivariate growth factor antagonist by inhibiting the Nodal/TGFβ, BMP, and Wnt pathways (30, 31). To investigate the convergence of these pathways with the early β-catenin signature, we overexpressed Cerberus mRNA in X. laevis embryos, and RNA-seq libraries were prepared at early gastrula in three independent experiments. Most genes within the early dorsal β-catenin signature were inhibited, presumably indicating a requirement for endomesodermal differentiation (Fig. 4B, in blue). However, a few genes in the signature were increased by Cerberus mRNA, including Xnr5, Xnr6, and many neural genes such as Hes1, Zic1, Zic4, Otx2, Irx1, Nkx6-2, and Ttki/Tbrda2A (Fig. 4B). GSEA analysis confirmed that many components of the β-catenin gene set were inhibited by Cerberus mRNA although some transcripts still ranked highly in the list (Fig. 4G).
Early dorsal β-Catenin Signature

**Secreted Factors**
- ADMP  bone morphogenetic protein 2 preproprotein
- AGR2  PREDICTED: anterior gradient protein 2 homolog isoform X1
- BMP2  bone morphogenetic protein 2 preproprotein
- CER1  cerberus precursor
- CHRD  chordin precursor
- CPE  carboxypeptidase E preproprotein
- CRESCENT  Crescent / secreted frizzled-related protein 5 precursor
- DKK1  dickkopf-related protein 1 precursor
- FGF20  fibroblast growth factor 20
- FRZB  secreted frizzled-related protein 3 precursor
- KAZALD1  MgIgα / kazal-type serine protease inhibitor domain-containing
- NOG  noggin precursor
- PKDC  protein kinase domain-containing protein, cytoplasmic precursor
- PRSS27  serine protease 27 precursor
- SHISA2  protein shisa-2 homolog precursor
- XNR3  Xnr3 / Nodal homolog precursor
- XNR5  Xnr5 / Nodal homolog precursor
- XNR6  Xnr6 / Nodal homolog precursor

**Transcription Factors and Regulators**
- BHLHE22  class E basic helix-loop-helix protein 22
- CRX  homeobox protein OTX2 isoform b
- DMX1  Otx3 / diencephalon/mesencephalon homeobox protein 1
- EGR1  early growth response protein 1
- FOXB1  forkhead box protein B1
- FOXD3  forkhead box protein D3
- GATA4  transcription factor GATA-4
- GATA5  transcription factor GATA-5 isoform X1
- GSC  homeobox protein goosecoid
- HES1  transcription factor HES-1
- HHEX  homeotaxially-expressed homeobox protein HHEX
- HLX  H2O-like homeobox protein
- IRX1  irxoi-class homeobox protein IRX-1
- LEF1  lymphoid enhancer-binding factor 1 isoform 2
- LHX1  LIM/homeobox protein Lhx1
- LMO1  rhomboid-1 isoform a
- LMO4  PREDICTED: LIM domain transcription factor LMO4 isoform X1
- MYF5  myogenic factor 5
- NKX6-2  homeobox protein Nkx-6.2
- OLG4  oligodendrocyte transcription factor 3
- OSR1  PREDICTED: protein odd-skipped-related 1 isoform X1
- OTX2  homeobox protein OTX2 isoform b
- SEBOX  homeobox protein SEBOX
- SIAMOIS  Siamois (Sia) homeobox protein
- SKOR1  SKI family transcriptional corepressor 1
- T  brachyury protein isoform 1
- ZIC1  zinc finger protein ZIC 1
- ZIC2  zinc finger protein ZIC 2
- ZIC3  zinc finger protein ZIC 3
- ZIC4  zinc finger protein ZIC 1

**Transmembrane Proteins and Receptors**
- ADAM19  disintegrin and metalloproteinase domain-containing protein 19
- CHRNA4  neuronal acetylcholine receptor subunit alpha-4 isoform 1 precursor
- CNRIP1  CB1 cannabinoid receptor-interacting protein 1 isoform X2
- EFNB2  ephrin-B2 precursor
- FZD8  frizzled-8 precursor
- LPAR6  lysophosphatidic acid receptor 6
- LRIG3  leucine-rich repeats and immunoglobulin-like domains protein 3 isoform 2
- PCDH9  PREDICTED: protocadherin-9 isoform X2
- PDGFRα  PREDICTED: platelet-derived growth factor receptor alpha isoform X1
- PRRT1  prion-like transmembrane protein 1
- SLC2A1  solute carrier family 2, facilitated glucose transporter member 1
- SLC34A2  sodium-dependent phosphate transport protein 2B isoform a
- SKOR1  PREDICTED: leucine zipper putative tumor suppressor 1 isoform X1
- TMEM150B  transmembrane protein 150B isoform X3

**Enzyme/Kinases**
- ALDH1A2  retinal dehydrogenase 2 isoform 1
- APOBEC2  probably C'-U-editing enzyme APOBEC-2
- ARHGAP39  rho GTP-activating protein 39
- DENND2C  DENN domain-containing protein 2C isoform 1
- DNASE1L3  deoxyribonuclease gamma isoform 1 precursor
- HS6ST1  heparan-sulfate 6-O-sulfotransferase 1
- NAT8B  probably N-acetyltransferase 8B
- PLEKHG5  pleckstrin homology domain-containing family G member 5
- ST3GL5  lactosylceramide-alpha-2,3-sialyltransferase isoform 2

**Others and Unknowns**
- #N/A #1  Xelaev18042095m
- #N/A #2  Xelaev1803555m
- #N/A #3  Xelaev1800592m
- #N/A #4  Xelaev18004073m
- #N/A #5  MGC151585 protein Xenopus laevis
- #N/A #6  Xelaev18002038m
- #N/A #7  LOC101733854.1
- FAM110C  PREDICTED: protein FAM110C isoform X1
- FAM98A  protein FAM98A
- KRT12  keratin, type 1 cytoskeletal 12
- L2TS1  PREDICTED: leucine zipper putative tumor suppressor 1 isoform X1
- NCK2  PREDICTED: cytoplasmic protein NCK2 isoform X1
- RAB20  ras-related protein Rab-20
- RN1  regulator of G-protein signaling 1
- SNCG  gamma-synuclein
- SPNS2  protein spinster homolog 2
- SRRD  SRRR1-like protein

![Fig. 2](https://www.pnas.org/cgi/doi/10.1073/pnas.1700766114)

**Fig. 2.** The early dorsal β-catenin gene signature. (A) Genes were categorized as follows: secreted factors, transcriptional factors and regulators, transmembrane protein and receptors, enzymes/kinases, and others and unknowns. Human gene symbols and full names are shown (2), except when *Xenopus*-specific names exist. Because the subtetraploid *X. laevis* has long and short homologs for many genes, duplicate names have been removed from this list. For unknown genes, the gene IDs from the JGI9 genome are indicated. (B) Venn diagram showing that the early β-catenin gene signature is distinct from the late zygotic Wnt signature, with only three genes in common.

When the entire transcriptome of *Cerberus* mRNA-injected embryos was examined, many more genes were significantly repressed than were induced (Fig. 5A). This observation was consistent with the inhibition of mesodermal and endodermal differentiation marked by *Xbra* and *Sox17β*, respectively (Fig. 5B). It is noteworthy that expression of *xWnt8*, the major Wnt
that signals from ventrolateral mesoderm during Xenopus gastrulation, is essentially eliminated in Cerberus mRNA-injected embryos (Fig. SB and Dataset S1). Cerberus overexpression represses, and induces, a wide range of genes of both dorsal and ventral origin by stage 10.5 (Fig. 5C and Datasets S5 and S6).

Previous RNA-seq studies used the Nodal/TGF-β receptor inhibitor SB431542 to define Nodal target genes (27). We observed that Cerberus, a Nodal antagonist, strongly repressed 107 genes, of which 49 overlapped with genes repressed by SB431542 (Dataset S5). In GSEA analyses, a set of 107 genes repressed by Cerberus were found predominantly at the bottom of the correlation of those inhibited by β-Catenin (Fig. 5D). Cerberus-repressed genes also correlated with genes arranged according to their D-V enrichment (Fig. 5E). When the Cerberus-repressed gene set was probed against rankings of transcripts induced by Wnt8 or Siamois, they appeared at either the top or bottom of the ranked list, presumably reflecting the dual inhibition of Nodal and BMP pathways by this multivalent antagonist (Fig. 5F and G). We conclude that many, but not all, of the early β-catenin–induced genes require endomesoderm germ layer specification by stage 10.5.

The Early β-Catenin Signature Is Established at Gastrula. In this study, we concentrated on the analysis of genes expressed at early gastrula because it is at this stage that Spemann organizer tissue is fully functional. We also analyzed gene expression at the earlier stage 9 (blastula, about 3.5 h before harvesting RNA from siblings used for gastrula studies) in some experiments and provide the complete RPKM data for stage 9 libraries in Dataset S7. As shown in Fig. 6A, most of the 125 genes in the early β-catenin signature were not activated at gastrula (shown as blue at stage 9 in the RPKM heat map). However, a subset of β-catenin targets had higher expression at blastula (in red) relative to gastrula stage 10.5 (Fig. 6A, boxed on top). Accordingly, most genes in the early β-catenin signature were unaffected by β-CatMO at stage 9 (Fig. 6B). Results from all experimental conditions tested at blastula and gastrula (control, β-CatMO, Wnt8 mRNA, Cerberus mRNA, LiCl, and Siamois mRNA) were allowed to cluster hierarchically by row according to RPKM expression as shown in Fig. 6C. At blastula stage 9, most genes in the signature were unaffected by β-catenin depletion or dorsalizing treatments (Fig. 6C, in blue). However, a group of genes did respond to Wnt-mimicking agents and were inhibited by β-CatMO at stage 9 (boxed at top of Fig. 6C and listed in Dataset S8), which included the classical Wnt targets Siamois (10) and Xnr3 (32). In addition, many organizer-specific genes were induced at stage 9, such as Xnr5, Xnr6, Admp, Cerberus, Noggin, and Zic2. The relative delay in the expression of the majority of the early β-catenin signature genes expressed at gastrula is likely explained by the need of the early β-catenin signal to activate Siamois as well as Xnr5/6 expression to achieve full organizer expression (33, 34), as indicated in the pathway outlined earlier in Fig. 1A.

Discussion

In Xenopus, fertilization triggers a cortical rotation that transports Dvl-containing cytoplasmic organelles, called maternal dorsal determinants, along microtubules toward prospective dorsal regions (19, 35). The rotation brings cytoplasmic determinants from vegetal cytoplasm in contact with the marginal zone of the egg before first embryonic cleavage (36, 37). This process determines the site at which the dorsal side of the embryo is formed by stabilizing maternal β-catenin protein, which accumulates in dorsal nuclei as early as the two- to four-cell stage (16, 17). The full molecular composition of the cytoplasmic determinant and the mechanism by which β-catenin is stabilized remain unknown although the participation of maternal Wnt11 and Dvl proteins, perhaps in combination with the formation of multivesicular bodies, has been proposed (22, 38). It is likely that the elusive cytoplasmic determinant is protein in nature because recent RNA-seq analyses of early cleavage stages (eight-cell) of X. laevis or X. tropicalis embryos have not detected any asymmetrically expressed mRNAs along the D-V axis (39, 40). The early β-catenin stabilization, in combination with zygotic Siamois and Nodal signaling, induces Spemann organizer formation at the gastrula stage (11). There has been considerable progress
in identifying the components of the Spemann organizer (2, 41–43). However, a global examination of the transcriptomic landscapes elicited by the early nuclear β-catenin signal and associated pathways that induce formation of Spemann organizer tissue was lacking.

The Early β-Catenin Dorsal Gene Set. In this study, we set out to determine the transcriptional output of the early maternal β-catenin signal. To this end, we sequenced the transcriptome of β-CatMO–injected embryos at gastrula (stage 10.5) in three independent experiments and compared them with their uninjected controls. We used β-CatMO as a loss-of-function reagent because it blocks dorsal development with 100% penetrance (18). Secreted Wnt antagonists (such as Dkk1 and Frzb) are ineffective at blocking the endogenous early β-catenin cytoplasmic determinant, yet inhibit the zygotic Wnt8 ventral-lateral signal that mediates A-P patterning (24). We found 247 transcripts that reproducibly required β-catenin, but, of these, to our surprise, only about half were dorsal genes, which suggested that many β-catenin transcriptional targets were not dependent on the dorsal cytoplasmic determinant signal. In fact, recent ChIP-seq studies indicate that endogenous β-catenin binds to 5,000 genes in the *X. tropicalis* genome, but only a small fraction of these are regulated by Wnt at any one stage (4).

To specifically focus on the genetic program activated by β-catenin on the dorsal side, we sequenced libraries from regenerating dorsal and ventral half-embryos bisected at mid-gastrula and cultured until early gastrula in five independent experiments. This approach resulted in much higher D-V differences than those observed in previous studies using dissected dorsal and ventral blastopore lips (2) because the BMP and late Wnt8 signals in the ventral half were unopposed by diffusing organizer-secreted antagonists (26). Transcripts were filtered so that only those decreased by β-catenin knockdown and dorsally enriched resulted in a set of 123 genes, which were named the early dorsal β-catenin signature (Fig. 2D). The correlation of the early dorsal β-catenin signature genes with *Wnt8* , LiCl, *Siamois*-induced genes, and *Cerberus*-repressed genes. (A) Heat map showing the fold change of early β-catenin signature genes in three β-catenin MO and five dorsal halves with respect to control and ventral half-embryo libraries, respectively. Fold changes (FC) over controls were used as inputs, and unsupervised hierarchical clustering of columns and rows was performed. Note that hierarchical clustering of rows identified most classic known Spemann organizer genes on the top of the heat map. (B) Heat map displaying regulation of genes composing the early β-catenin signature by *Wnt8* mRNA, LiCl, and *Siamois* or *Cerberus* mRNAs. (C) Correlation matrix of β-CatMO–, *Wnt8*- , *Cerberus-* , and *Siamois*-injected embryos and LiCl-treated embryos with the early dorsal β-catenin signature. Correlation scores were calculated as Pearson correlation coefficients (PCCs) and were color-coded as shown in the scale bar on the bottom of the panel. Note that the β-CatMO (in blue) anticorrelates with all conditions except *Cerberus* mRNA-injected embryos. These results show that the early β-catenin signature obtained via RNA-seq is reproducible and readily identifies dorsalizing and ventralizing conditions. (D–G) Gene Set Enrichment Analysis (GSEA) showed that the early β-catenin gene set was significantly enriched at the top of the *Wnt8*, LiCl, and *Siamois*-induced gene rankings, and at the bottom of the *Cerberus*-induced ranking. Cer, *Cerberus* con, control; D1/2, dorsal half; Sia, *Siamois* ; V1/2, ventral half.

![Fig. 4. Correlation of the early dorsal β-catenin signature with Wnt8, LiCl, Siamois-induced genes, and Cerberus-repressed genes.](image-url)
gene signature identifies early dorsal \( \beta \)-catenin genes that require endomesodermal induction. This study used 16,729 mRNAs between control and Cerberus mRNA-injected embryos. The average log fold change in expression of transcripts in Cerberus mRNA-injected embryos over control embryos is plotted on the y axis; the average gene expression in RPKM of uninjected control embryos is on the x axis. This MA-plot confirms the proper normalization of the RNA-seq data because the horizontal nature of the line across the zero baseline shows that there is no systematic bias. (B) Cerberus overexpression greatly inhibits Xbra, xSox17\(1\), and xWnt8 mRNA expression when examined by in situ hybridization. (Magnification: 10×.) (C) Heat map showing how gene signatures repressed or induced by Cerberus mRNA were differentially regulated by various dorsalizing or ventralizing conditions. Note that a large block of Spemann organizer genes (Dataset S1) Cerberus overexpression greatly inhibits Xbra mRNA induction and were examined for the ranking of Cerberus-repressed genes (vertical lines). Cer, Cerberus; con, control; Ctrl, control; DV, dorsal/ventral ratio; Sia, Siamois.

A Rich Genome-Wide Embryonic Resource. We have compiled the RNA-seq results in complete RPKM files of expression levels for the 43,673 genes annotated in the Xenopus genome (1) comprising a total of 46 RNA-seq libraries. This rich resource is now accessible to the Xenopus community or anyone interested in the complete transcriptional repertoire of Spemann organizer-inducing signals in vertebrate embryos. The embryonic transcriptional response should be considered globally because, for each action in the organizer, there is a reaction in the ventral side (45).

We also examined the timing of the activation of the early \( \beta \)-catenin signature. By comparing the transcriptional activation of the early \( \beta \)-catenin signature. The Cerberus data also defined global gene signatures that require endoderm and mesoderm formation and increased neural genes within the early dorsal \( \beta \)-catenin signature. The Cerberus data also defined global gene signatures repressed and induced by this mRNA that may help decipher the physiological role of this growth factor antagonist.

We also examined the timing of the activation of the early \( \beta \)-catenin signature. By comparing the transcriptional activation of the early \( \beta \)-catenin signature. The Cerberus data also defined global gene signatures repressed and induced by this mRNA that may help decipher the physiological role of this growth factor antagonist. The widespread dorsal gene activation by overexpression of the homeobox protein Siamois was particularly interesting because it implied that most of the early maternal \( \beta \)-catenin signal is mediated through this Xenopus-specific transcription factor (10, 46, 47). The X. laevis genome contains six active Siamois/Twin genes, which ensure a rapid response to the maternal \( \beta \)-catenin signal (48). We note that Siamois overexpression tended to generate more robust RPKM inductions than Wnt8 or LiCl, presumably because this transcription factor reached higher concentration levels after direct injection (Dataset S1), even though the dorsalized embryonic phenotypes were comparable (Fig. 1E). Microinjection of Cerberus mRNA repressed genes that require endoderm and mesoderm formation and increased neural genes within the early dorsal \( \beta \)-catenin signature. The Cerberus data also defined global gene signatures repressed and induced by this mRNA that may help decipher the physiological role of this growth factor antagonist.

We also examined the timing of the activation of the early \( \beta \)-catenin signature. By comparing the transcriptional activation of the early \( \beta \)-catenin signature. The Cerberus data also defined global gene signatures repressed and induced by this mRNA that may help decipher the physiological role of this growth factor antagonist.
whereas all genes within the dorsal β-catenin signature are activated by stage 10.5, a small cluster of these genes were expressed at higher levels at stage 9 (Fig. 6). It is known that in Xenopus β-catenin bound to TCF-3 plays an important role in recruiting epigenetic regulators even before zygotic transcription (49). Nuclear β-catenin recruits Prmt2 (a protein-histone arginine methyltransferase) that causes an epigenetic mark called H3R8Me2 (dimethylation of histone H3 at Arginine 8), which deposits an epigenetic memory that poises promoters for transcription up to six cell cycles before Wnt target genes are first transcribed (49). The maternal β-catenin/Wnt signal, followed by zygotic Siamois and Nodal signaling, provides the initial step in the process that gives rise to the formation of the Spemann organizer in the dorsal marginal zone at the gastrula stage (11, 23). Spemann organizer tissue then orchestrates dorsal ventral patterning by secreting a mixture of growth factor antagonists (24).

Comparison of the Early and Late Wnt Signatures. The cellular responses to Wnt ligands in early development are stage-dependent, which represents a difficulty when trying to assign a general transcriptional signature for the pathway. There are two distinct waves of Wnt signaling in Xenopus early development. Maternal Wnt/β-catenin signaling drives the formation of the Spemann organizer in dorsal regions whereas zygotic Wnt8,

Fig. 6. Comparison of the expression of the early dorsal β-catenin gene signature at blastula and gastrula stages. (A) Heat map showing that the majority of the early dorsal β-catenin gene signature had higher expression levels (in red) in WT control (Cont.) embryos at stage 10.5 rather than at blastula stage 9. RPKMs of the genes constituting the early β-catenin signature genes at blastula stage 9 or gastrula stage 10.5 (results from five and six sequenced libraries, respectively) were normalized by row. (B) The majority of genes in the early β-catenin signature were unaffected by β-CatMO at stage 9; the effect of β-catenin MO becomes evident at stage 10.5. (C) Heat map comparing RPKM levels of the early dorsal β-catenin signature in 46 libraries at blastula and gastrula stages. Only a minority of the β-catenin signature were inhibited by β-CatMO or activated by Wnt-like signals at blastula stage 9. RPKM values of all libraries were allowed to cluster automatically according to rows. The cluster delimited by a box at the top indicates transcripts that are capable of being induced or repressed by dorsalizing or ventralizing treatments at blastula stage 9. A list of these genes is provided in Dataset S8 and includes many known early Wnt targets, such as Siamois, Xnr3, Xnr5, Xnr6, Noggin, and ADMP. The scale bar at the bottom indicates normalized RPKM values. In conclusion, most genes in the early β-catenin signature become sensitive to β-catenin depletion only at gastrula stage, presumably indicating the requirement of the endodermal and mesodermal germ layer induction for Spemann organizer development. Cer, Cerberus; Cont, control; St., stage.
expressed at mid- to late-gastrula in ventro-lateral mesoderm, promotes the development of the posterior region of the embryo.

Recently, two studies have defined the late zygotic Wnt8 gene set using a combination of RNA-seq and ChiP-Seq (3, 4). Kjøbler and Harland (3) used the Wnt inhibitor Dkk1 to define the transcriptome activated by Wnt signal at mid- to late-gastrula in X. laevis. A total of 82 genes constituted their late Wnt transcriptome. Nakamura et al. (4) used X. tropicalis to screen for target genes repressed by a morpholino specifically targeting Wnt8. As controls, they rescued these effects with Wnt8 DNA and also selected only those genes bound by endogenous β-catenin at TCF/LEF sites. These more stringent criteria identified fewer genes (4), but all were included in the late signature from the other laboratory (3). Strikingly, it was found that β-catenin bound to 5,195 genes at gastrula, most of which are not regulated by Wnt at this stage; β-catenin recruitment does not imply transcriptional regulation, which must depend on additional developmental contexts (4).

That the early and late Wnt signatures would differ was expected (11). However, the extreme degree of divergence we found was surprising. We identified only 3 genes (Irx, Myf5, and Sebox) that overlapped between our early β-catenin signature of 123 genes and the late Wnt signature of 82 genes (3) (Fig. 2B). It is currently unknown what provides this “switch” from maternal to late β-catenin/Wnt signaling although epigenetic mechanisms or stage-specific TCF proteins may be involved (49, 50). It has been proposed that β-catenin recruitment may occur during early embryogenesis. A DNA regulatory signature for Wnt-regulated expression in other tissues at much later stages (4). In this regard, cross-talk between signaling pathways such as BMP, Nodal, and FGF may be important. In addition, as yet unknown lineage markers, as occurs with the recruitment of β-catenin in hematopoietic progenitor cells (51), might be involved in triggering differential transcriptional programs. It is generally accepted that the main role of β-catenin in gastrula is to signal dorsally and establish the Spemann organizer. However, in the present study it was found that many other genes require β-catenin for transcription but are not dorsally enriched, perhaps indicating as yet unknown transcriptional roles for β-catenin in early development. Studies are needed to investigate the effects of β-catenin/Wnt signaling at later embryonic stages: for example, to examine the deployment of the Hox complexes that are not yet active at the stages studied here.

Conclusion

This work presents an early dorsal β-catenin signature of 123 genes and provides the complete transcriptional landscape of treatments that massively stimulate or inhibit formation of the Spemann organizer in X. laevis. The genome-wide transcriptomic signatures associated with β-catenin loss of function, with dorsal and ventral half-embryos, and with Wnt8, LiCl, Siamois, and Cerberus phenotypes and the specific differences and similarities among them provide a rich substratum for investigations into the self-organizing nature of the earliest gene regulation events during development of a vertebrate embryo. Materials and Methods

Embryo Manipulations, mRNA Injections, and Whole-Mount in Situ Hybridization. X. laevis were purchased from Nasco. Embryos were generated through in vitro fertilization and cultured in 0.1× Marc’s modified Ringer’s (MMR) and staged according to Nieuwkoop and Faber (52). A total of 24 ng of mor pholino oligonucleotide against Xenopus β-catenin (18) was injected four times for the four-cell stage. For in vitro mRNA analysis, pC2S-xWnt8, pC2S-xCerberus, and pC2S-xSiamois were linearized with NotI and transcribed with SP6 RNA polymerase using the Ambion mMessage mMachine kit. Synthetic mRNAs were injected into the marginal zone of each whole embryo at the four-cell stage as follows: 3 pg of xWnt8; 100 pg of xCerberus, and 20 pg of xSiamois. All of the injected embryos were cultured until stage 9 or 10.5. For LiCl treatment, 32-cell stage embryos were treated with 0.3 M LiCl in 0.1× MMR for 7 min, washed three times with 0.1× MMR, and cultured until stage 9 or 10.5. For dorsal and ventral half-embryos, regularly cleaving stage 8 midblastula embryos were bisected into dorsal and ventral halves (44) and cultured until gastrula stage 10.5. For hybridization probe synthesis, pC2S-xWnt8 was linearized with BamH1 and transcribed with T3 RNA polymerase; pC2S-xBra with SalI and transcribed with SP6 RNA polymerase; pC2S-xSox17 with ApaI and transcribed with SP6 RNA polymerase. Whole-mount in situ hybridizations were performed as described previously (2, 53, 54). All RNA-seq data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database (accession No. GSE93195). Gene set enrichment analysis (GSEA) was performed using the GSEA software from the Broad Institute (software.broadinstitute.org/gsea/index.jsp) (29, 55), with a permutation-based calculation based Kolmogorov-Smirnov nonparametric rank test (10,000 permutations). For Weighted Gene Coexpression Network Analysis (WGCNA) (28), a gene coexpression network was constructed using the 2,000 transcripts that showed greatest variations upon microinjection of β-CatMO and their expression values throughout the 29 libraries sequenced at stage 10.5 (Dataset S1). To find modules of highly correlated genes, average linkage hierarchical clustering was performed, and modules were colored by immune colors.

Heat Maps. Heat maps were generated in Multiexperimenter Viewer (MeV) (56). For Fig. 6, RPMKs were used as inputs, and rows/genes were normalized using the mean and SD of the row to which the value belongs, using the following formula: Value = [(Value) – (Mean)/(SD)]+ 1|; hierarchical clustering using Manhattan distance was performed. The heat map in Fig. 5 used fold change as input and was similarly normalized.

Mean Log Ratios-Average Plot and PCA Analysis. Mean log ratios-average plots (MA-plots) were generated in R. For Fig. 5A, the log fold change (FC) in expression of transcripts between control embryos corresponding to β-CatMO embryos was plotted versus the average RPMK expression of transcripts in control embryos. In Fig. 4A, the logFC in expression of transcripts between Cerberus embryos to control embryos was plotted against the average RPMK levels of expression in control embryos. Transcripts with an average RPMK below 1 across all conditions were eliminated. Principal component analysis (PCA) in Fig. 5A was generated in R by comparing logFC in all libraries for transcripts identified in our early β-catenin signature.

Scatter and Box Plots. The scatter plot matrix in Fig. 1H contains 15,300 transcripts from dorsal and ventral half RNA-seq libraries. The logFC in average expression of transcripts between dorsal and ventral half-embryos (five experiments) was plotted versus the logFC of control embryos with β-CatMO embryos (three experiments). Transcripts highlighted in red are those with fold changes higher than 1.5 and 1.4 in the dorsal/ventral (D/V) and con/β-CatMO conditions, respectively. Box plots in Fig. 5A and Fig. 3 were generated in R. The middle line in the box indicates the median, the box edges indicate the 25th/75th percentiles, and the whiskers indicate min and max values. The statistical significance of differences in gene expression levels between pairwise sets of genes was tested using the Mann–Whitney test and indicated as follows: *p≤0.05, **p≤0.01, and ***p≤0.005.

Cloning and qRT-PCR. To clone full-length Loc100170590, forward and reverse PCR primers were designed according to the genomic sequence deposited in the Xenbase database (www.xenbase.org/). The oligos also contained upstream sequences for Gateway-mediated cloning. PCR was performed on cDNA obtained from gastrula stage X. laevis embryos, resulting in an amplification product migrating at the expected size (about 1.5 kb). The PCR product was purified, cloned in a pDONR221 vector and subsequently in a home-made Gateway-compatible pCS2 vector suitable for antisense probe and in vitro mRNA synthesis. Primers for cloning were as follows: Fwd, GGGGACGACGTTGGTACACAAAAAGCTG; Rev, GGGGACGACGTTGGTACACAAAAAGCTG. Primers for qRT-PCR were as follows: Fwd, GCTTTCGAGTCTTTTTGAC; Rev, CAGACATGAGAGATGG.
ACKNOWLEDGMENTS. We thank Prof. Guoping Fan (UCLA) for suggestions in library preparation, and members of the E.M.D.R. laboratory for comments on the manuscript. We thank Dr. Pengpeng Liu for bioinformatics instruction. E.A.S. was supported by the MARC program (NIH T34 GM008563-21) and M.D.J.B was supported by NIH Training Grant R25 GM005552. This work was supported by the Norman Sprague Endowment and the Howard Hughes Medical Institute, of which E.M.D.R. is an investigator.

Supporting Information

Ding et al. 10.1073/pnas.1700766114

Fig. S1. Transcriptome analysis of β-CatMO–injected embryos and dorsal and ventral halves. (A) MA-plot comparing gene expression between uninjected control and β-catenin MO–injected embryos including 16,729 transcripts. The average (three independent experiments) log
2 fold change (FC) in expression of transcripts in control (con) embryos over β-CatMO–injected embryos (con/β-CatMO) is plotted on the ordinate; the average mean gene expression (in RPKM) in control embryos is represented on the abscissa. Red indicates β-CatMO–repressed transcripts (1.4-fold minimum decrease); blue indicates β-CatMO–induced transcripts (1.4-fold increased); gray dots indicate all other transcripts. Note that β-CatMO represses more genes than it activates, which is consistent with the widespread role of β-catenin as a transcriptional coactivator. (B) Heat map showing all transcripts that were repressed 1.4-fold by β-catenin MO in three independent pairs of libraries, and the dorsal to ventral (D/V) FC in five independent dorsal and ventral half libraries bisected at midblastula stage 8 and allowed to regenerate for 5 h until gastrula. Rows indicate FC of transcripts as indicated in the scale bar at the right of the panel. The various conditions are indicated at the top of the columns. Hierarchical clustering of rows clustered classical Spemann organizer genes on the top region. Genes within the dashed line box were repressed by β-CatMO but not enriched in the dorsal side; these were removed from our early β-catenin signature because they represent genes that require β-catenin independently of the maternal dorsal signal. (C) Venn diagram illustrating overlap of genes between β-CatMO–inhibited genes and dorsally enriched genes. Of 247 genes reduced by β-catenin MO and 237 genes enriched in dorsal side, 123 genes overlapped and defined our early dorsal β-catenin signature. In these stringent conditions, which require results above a certain threshold in all three independent experiments, some β-catenin–regulated genes may miss the cutoff. The complete stage 10.5 RPKM data are presented in Dataset S1, and, from this dataset, any genes that narrowly missed the cutoff can be retrieved. (D) Box plots of RPKMs of whole embryo (WE) and dorsal (D) and ventral (V) halves in five independent RNA-seq experiments. Note that, as expected, Chrd, Nog, and Gsc were significantly enriched in regenerating dorsal halves and that Bambi, Szl, and Ventx were enriched in ventral halves, which supports the quantitative nature of the RNA-seq analyses.
Fig. S2. qRT-PCR validation and sequence homology of LOC100170590. (A) β-catenin–dependent regulation and D-V enrichment of LOC100170590 was validated by qRT-PCR analysis. β-CatMO decreased LOC100170590 mRNA levels compared with uninjected WT controls (WT), and expression in dorsal half-embryos (D half) was significantly higher than in ventral half-embryos (V half). (B) Amino acid sequence comparison between *Xenopus laevis* Loc100170590 and *Latimeria chalumnae* XP_005988953.1 (a coelacanth fish) revealed a 27% degree of identity. The coelacanth gene has been annotated as possible Dapper homolog-like 2, but XLoc100170590 lacks critical Dapper-like features, such as a PDZ domain and a leucine zipper. The sequence alignment was generated by using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/) followed by BoxShade (embnet.vital-it.ch/software/BOX_form.html). Black and gray shaded boxes indicate identical and similar amino acid residues, respectively. There were no significant homologies with mammalian genomes.
Fig. S3. Weighted Gene Coexpression Network Analysis (WGCNA) of stage 10.5 libraries. WGCNA unsupervised hierarchical clustering identified a distinct cluster (brown cluster, red arrow) that contained 106 of the 123 genes that compose the early β-catenin signature. Hierarchical clustering of marbling score-related genes and visualization of gene modules was performed. The colored bars are directly consistent with the module (color) for the clusters of genes.

Fig. S4. The expression of the 123-gene early dorsal β-catenin signature examined via principal component analysis (PCA) to analyze dimensionality in six experimental conditions: Cerberus/control (Cer/con), β-catenin morpholino/control (β-CatMO/con), Sia/~/control (Sia/con), lithium chloride/control (LiCl/con), dorsal/ventral regenerating halves (Dor/Ven), and Wnt8/control (Wnt/con). Each axis represents a principal component (PC), with the first one having the most variation. PCA analysis confirmed the reproducibility of the early β-catenin signature by clustering replicate conditions together and distinguishing ventralizing (β-catenin depletion) and dorsalizing treatments. The Cerberus mRNA-injected condition was located between the β-CatMO and the rest of the sequenced conditions due to its ability to inhibit both Nodal and BMP target genes.
Dataset S1. List of RPKM values in our stage 10.5 RNA-seq libraries showing the 43,668 transcripts annotated in the *X. laevis* JGI9.1 genome, including human gene symbols

Dataset S2. Early dorsal β-catenin signature (123 genes) expression values (in RPKM) throughout all stage 10.5 libraries sequenced

Dataset S3. Genes decreased by β-catenin but not dorsally enriched

Dataset S4. Dorsal/β-catenin cluster defined by WGCNA

Dataset S5. Genes repressed by Cerberus overexpression at stage 10.5

Dataset S6. Genes induced by Cerberus overexpression at stage 10.5

Dataset S7. List of RPKM values in our stage 9 RNA-seq libraries showing the 37,006 transcripts annotated in the *X. laevis* JGI9.1 genome, including human gene symbols

Dataset S8. Subset of the early β-catenin signature that is capable of being induced at blastula