Genomic integration of Wnt/β-catenin and BMP/Smad1 signaling coordinates foregut and hindgut transcriptional programs

Mariana L. Stevens1, Praneet Chaturvedi1, Scott A. Rankin1, Melissa Macdonald1, Sajjeev Jagannathan2, Masashi Yukawa2, Artem Barski2 and Aaron M. Zorn1,*

KEY WORDS: BMP, Smad1, Wnt, β-Catenin, Foregut, Hindgut, ChIP-seq, RNA-seq, Xenopus, Endoderm, Mesoderm

INTRODUCTION

Embryonic development of the digestive and respiratory systems is controlled by a reiterated series of growth factor interactions between the epithelium and mesenchyme (reviewed by Zorn and Wells, 2009). Our understanding of how combinatorial signals orchestrate organogenesis in animal models has been the foundation for designing strategies to direct the differentiation of human pluripotent stem cells (hPSCs) into organoids for disease modeling and regenerative medicine (reviewed by Lancaster and Knoblich, 2014). Despite our growing knowledge of which growth factors act where and when during organogenesis, how combinatorial signals are integrated at the genomic level to coordinate gene expression through cis-regulatory DNA elements, and by a crosstalk whereby Wnt negatively regulates BMP ligand expression in the foregut. These results extend our understanding of gastrointestinal organogenesis and of how Wnt and BMP might coordinate genomic responses in other contexts.

ABSTRACT

Digestive system development is orchestrated by combinatorial signaling interactions between endoderm and mesoderm, but how these signals are interpreted in the genome is poorly understood. Here we identified the transcriptionomes of Xenopus foregut and hindgut progenitors, which are conserved with mammals. Using RNA-seq and ChIP-seq we show that BMP/Smad1 regulates dorsal-ventral gene expression in both the endoderm and mesoderm, whereas Wnt/β-catenin acts as a genome-wide toggle between foregut and hindgut programs. Unexpectedly, β-catenin and Smad1 binding were associated both with transcriptional activation and repression, with Wnt-repressed genes often lacking canonical Tcf DNA binding motifs, suggesting a novel mode of direct repression. Combinatorial Wnt and BMP signaling was mediated by Smad1 and β-catenin co-occupying hundreds of cis-regulatory DNA elements, and by a crosstalk whereby Wnt negatively regulates BMP ligand expression in the foregut. These results extend our understanding of gastrointestinal organogenesis and of how Wnt and BMP might coordinate genomic responses in other contexts.

Here we identify the transcriptional program of FG and HG progenitors in *Xenopus laevis* embryos, which is largely conserved with mammals. RNA-seq and ChIP-seq revealed how BMP and Wnt signals coordinate spatially restricted FG and HG gene expression, with β-catenin and Smad1 co-binding CRMs of hundreds of key cell identity regulators. We identify a Wnt-BMP crosstalk in the FG and, unexpectedly, find that many genes inhibited by BMP or Wnt are associated with Smad1 or β-catenin binding, suggesting direct repression. These findings advance our understanding of how combinatorial signaling is interpreted in the genome during gastrointestinal development and serve as a paradigm for other developmental and disease contexts.

**RESULTS**

**Transcriptional programs of FG and HG progenitors**

In order to examine how BMP and Wnt regulate early gut development at the genomic level, we first defined the FG and HG transcriptomes. Taking advantage of the size and abundancy of *X. laevis* embryos, we microdissected the ventral FG endoderm, FG mesoderm, HG endoderm and HG mesoderm tissues from 50 sibling embryos at stage NF20 (22 h post fertilization, ∼6-7 somites) and performed RNA-seq (Fig. 1A,B). This time point is similar to E8.5 in mice, when the FG is being patterned and is still plastic. Differential expression analysis of FG (endoderm plus mesoderm) versus HG (endoderm plus mesoderm) identified 906 FG-enriched genes and 987 HG-enriched genes, whereas endoderm (FG plus HG) compared with mesoderm (FG plus HG) identified 3439 endoderm-enriched and 4829 mesoderm-enriched transcripts (log2 FC ≥1, FDR ≤5%) (Fig. 1A-C; Fig. S1A, Tables S1 and S2).

These gene lists contained over 98% of well-known FG and HG transcripts (n=74) manually curated from mouse and differentiated hPSCs, confirming the extensive conservation (Fig. S1B, Table S3). The FG endoderm-enriched set included transcription factors *hhex* and *pdx1*, while the FG mesoderm-enriched set included key regulators of cardiac (*hand2*), myeloid (*spib* and *cebpa*) and endothelial (*fli1*) lineages (Fig. 1D). Gene Ontology (GO) analysis showed epithelial, vasculature and circulatory system development among the top FG-enriched terms (Fig. S1C). The HG transcriptome was enriched for GO terms related to A-P patterning and included the key intestinal regulators *cdx1*, 2 and 4 in the HG endoderm and the homeobox genes *hox5-11* in the HG mesoderm sets (Fig. 1D; Fig. S1B,C, Table S1).

**Fig. 1. Transcriptional program of FG and HG progenitors correlates with differential BMP and Wnt signaling.** (A) Fate map showing that FG progenitors (yellow) give rise to lungs, liver, pancreas and stomach, whereas HG progenitors (green) give rise to intestine (Chalmers and Slack, 2000). (B) Experimental design. RNA-seq was performed on FG endoderm, FG mesoderm, HG endoderm and HG mesoderm explants dissected from stage NF20 *Xenopus* embryos. Differentially expressed transcripts were identified by pairwise comparisons of FG (endoderm and mesoderm) versus HG (endoderm and mesoderm), as well as endoderm (FG and HG) versus mesoderm (FG and HG) tissue (log2 FC ≤−1 or ≥1, FDR ≤5%). (C) Venn diagram showing the intersection of two separate differential expression analyses: FG versus HG and endoderm versus mesoderm, showing mutually exclusive lists of transcripts with enriched expression. (D) Heatmap clustering of the 906 FG-enriched and 987 HG-enriched genes showing expression in the indicated tissues with representative FG (orange) and HG (green) genes listed on the right. (E) *In situ* hybridization of sagittal bisected stage NF20 embryos with *hhex* and *ventx2.1* marking the FG and HG domains, respectively. (F) BMP and Wnt activity shown by pSmad1 (red) and nuclear (n) β-catenin (red) immunostaining in NF20 embryos. Nuclei staining with DAPI (green). (G) pSmad1 is high in the ventral and low in the dorsal FG and HG, whereas nβ-catenin is low in the FG and high in the HG. FG/fg, foregut; HG/hg, hindgut; endo, endoderm; meso, mesoderm. Scale bars: 100 µm.
BMP and Wnt regulate gut tube patterning

Consistent with a role in patterning, GO terms related to Wnt and BMP were enriched in the FG and HG datasets (Fig. S1C). Examination of BMP and Wnt pathway components in the total RNA-seq data indicated that BMP ligands (bmp2, 4 and 7), receptors (bmpr1a, bmpr2) and BMP target genes (id3 and scl) were variably expressed in both FG and HG tissues, with no obvious A-P difference (Fig. S1D). By contrast, Wnt ligands were enriched in the HG, while Wnt antagonists (dkk1, sfrp2 and sfrp3) were restricted to the FG endoderm (Fig. S1D).

We next compared the levels of nuclear (n) β-catenin and phosphorylated (p) Smad1 immunostaining in the FG and HG domains, which are marked by hhex and ventx2, respectively (Fig. 1E) (McLin et al., 2007). Consistent with previous reports, the gastrula NF11 anterior mesendoderm (presumptive FG) had low levels of both pSmad1 and nβ-catenin, whereas the posterior mesendoderm (presumptive HG) had high levels of both (Schohl and Fagotto, 2002). After gastrulation, at stage NF20, nβ-catenin was still low in the FG and high in the HG; however, we observed an upregulation of pSmad1 levels in the ventral (but not dorsal) FG (Fig. 1F), consistent with the known de novo expression of bmp2/4/7 in the pre-cardiac FG mesoderm (Kenny et al., 2012). Thus, during FG-HG patterning, BMP and Wnt are differentially active along orthologous axes; pSmad1 is high in the ventral and low in the dorsal FG and HG, whereas nβ-catenin is low in the FG and high in the HG (Fig. 1G).

We next examined the impact of BMP or Wnt inhibition on progenitor patterning by treating embryos with the BMP receptor inhibitor DMH1 (40 μM) or by inhibiting Wnt in a heat shock-inducible dkk1 transgenic line Tg(hsp70:dkk1) (Lin and Slack, 2008). We added DMH1 or performed heat shock between stages NF12 and 20, during FG-HG patterning, and after gastrulation to avoid disruption of axial patterning. BMP inhibition reduced both hhex and ventx2.1, whereas Wnt inhibition expanded hhex and reduced ventx2.1 expression (Fig. S1E). To test whether these changes in patterning impacted subsequent organogenesis, we analyzed organ lineages at NF35. BMP inhibition between NF12 and 20 resulted in an expansion of the dorsal esophageal marker sox2 and loss of liver (nr1h5), lung (nkx2-1) and heart (nkx2-5) marker expression. By contrast, Wnt inhibition expanded the liver (nr1h5) and reduced the intestinal (dlarmatin) markers (Fig. S1E). Thus, spatially restricted Wnt and BMP pattern the FG and HG progenitors.

BMP-regulated FG and HG transcriptomes

To identify BMP-regulated transcripts, we isolated FG and HG explants (containing both endoderm and mesoderm) from DMH1- or vehicle-treated NF20 embryos and analyzed these by RNA-seq (Fig. 2A). Transcripts with reduced expression upon DMH1 treatment compared with controls were classified as normally activated by BMP (n=697), whereas increased expression upon DMH1 treatment indicated that they are repressed by BMP (n=1063) (log2 FC ≥1, FDR <5%) (Fig. 2B; Fig. S2A,B; Tables S2 and S4). Eight genes showed variable regulation by DMH1 and were not included in further analysis (Fig. S2B), resulting in 1760 unique BMP-regulated genes. Approximately 17% (155/906) of FG-enriched genes were activated by BMP, whereas 20% (185/906) were repressed. Of the HG-enriched transcripts, 10% (97/987) were activated by BMP and 9% (89/987) were repressed (Fig. 2B). BMP-activated genes were enriched for GO terms related to cardiovascular, blood vessel and digestive system development, whereas BMP-repressed genes were enriched for skeletal and renal system, indicative of dorsal gene expression, normally low in FG and HG explants (Fig. S2C).

BMP signaling was required to maintain the expression of key posterior homeobox genes cdx2, cdx4, ventx2 and msx1 in both the HG endoderm and mesoderm (Fig. 2C-E). The role of BMP was more complex in the FG. Approximately 85% of the BMP-regulated FG endoderm genes appear to be repressed by BMP (Fig. 2C,D), in contrast to the FG mesoderm where BMP was required for expression of ~60% of the BMP-regulated genes, including known regulators of the heart (tbx20 and hand2) and myeloid (sphb and cebpα) lineages (Fig. 2D).

BMP regulates dorsal-ventral patterning of the early FG

Although most BMP-regulated FG endoderm genes were repressed, a few required BMP for their expression, including hhex and sfrp5 (Fig. 2D,E), which are implicated in liver and ventral pancreas development (Li et al., 2008; McLin et al., 2007). Together with the observation that the ventral FG has higher nuclear pSmad1 levels than the dorsal FG (Fig. 1F), this prompted us to examine dorsal-ventral (D-V) patterning in more detail.

Since some digestive organs, including the esophagus, stomach, intestine and pancreas originate from both ventral and dorsal endoderm cells (Chalmers and Slack, 2000), we hypothesized that ventral and dorsal transcripts might be differentially regulated by BMP. Unsupervised clustering of BMP-regulated transcripts in control and DMMH1-treated FG explants along with isolated dorsal explants (which contain a thin layer of dorsal endoderm) revealed that the DMMH1-treated FG had an expression profile similar to dorsal tissue, suggesting that BMP represses dorsal lineages (Fig. S2D). In situ hybridization confirmed that DMMH1 treatment reduces the expression of ventral FG genes sfrp5 (endoderm), cebpα (myeloid) and nks2-5 (cardiac), while causing the expansion of dorsal endoderm transcripts mmx1 and hrg and the paraxial mesoderm gene foxc2. Injection of BMP2 protein into the FG had the opposite effect, causing an expansion of ventral genes and repression of dorsal markers (Fig. 2E; Fig. S2E). Thus, in the FG between stages NF12 and 20, BMP promotes ventral (presumptive liver, lung, cardiac) and represses dorsal (presumptive esophagus, kidney and paraxial mesoderm) fates.

Smad1 chromatin binding to BMP-regulated HG and FG genes

In efforts to identify direct BMP target genes we performed Smad1 ChIP-seq on NF20 embryos and identified 7976 Smad1 peaks, located within 20 kb upstream or downstream of 5252 genes (Fig. 3A; Table S5). This represents ~18% of the 45,099 predicted genes in the allotetraploid X. laevis genome (Session et al., 2016). Smad consensus DNA binding sites, as well as Gata and Tbx motifs, were enriched in Smad1 peaks (Fig. 3B). Specificity of the Smad1 antibody was confirmed by reduction of Smad1 binding to known BMP target genes msx1, id3 and ventx2.1 (Karaulanov et al., 2004) upon DMMH1 treatment (Fig. S3A,B). Of the Smad1 peaks, 27% overlapped with p300 ChIP-seq peaks indicative of active transcription, and ~56% of the Smad1-associated genes exhibited detectable expression in NF20 embryos (Fig. S3C,D).

Of the total 1760 BMP-regulated genes in the FG or HG (from Fig. 2B), ~35% (615/1760) were associated with Smad1 binding, a statistically significant enrichment based on a hypergeometric test (HGT) [2.7-fold enrichment (FE), P<0.05] (Fig. 3A). These included 48% of the BMP-activated HG genes (47/97; 3.7 FE, HGT P<0.05), such as cdx2, ventx2.1 and msx1, consistent with direct activation by Smad1 (Fig. 3C,D). By contrast, Smad1 binding was only associated with 17% (26/155) of BMP-activated FG genes, such as hand2 (mesoderm) and hhex (endoderm) (Fig. 3C,D), suggesting that most BMP-activated FG genes are indirect targets. Unexpectedly, Smad1 binding was more closely associated with
BMP-repressed genes (n=410) than with BMP-activated genes (n=205). Indeed, ∼56% of BMP-repressed FG genes (103/185; 4.3 FE, HGT P<0.05) and ∼49% of BMP-repressed HG genes (44/89; 3.8 FE, HGT P<0.05) were associated with Smad1 peaks, suggesting direct BMP repression, including dorsal genes such as foxc2 and mnx1 (Fig. 3D). Analysis of p300 co-occupancy on different classes of Smad1 peaks did not, however, show a statistically significant enrichment on activated versus repressed genes, presumably because the ChIP was performed on whole embryos containing both expressing and non-expressing cells. Altogether, about one-third of the BMP-regulated FG and HG transcriptomes was associated with Smad1 binding, suggesting direct transcriptional activation and repression.

**Canonical Wnt promotes HG and represses FG transcriptomes**

To determine the genomic basis of Wnt-mediated A-P patterning, we performed RNA-seq on FG and HG explants (endoderm plus...
mesoderm) from NF20 embryos where Wnt/β-catenin activity was either activated by BIO treatment or inhibited with the secreted Wnt antagonist Dkk1 expressed from Tg(hsp70:dkk1) from stages NF12-20 (Fig. 4A). Transcripts repressed by Dkk1 or induced by BIO, relative to controls, were considered to be Wnt activated, whereas transcripts upregulated by Dkk1 or repressed by BIO were considered to be Wnt repressed (Fig. S4A). Differential expression analysis identified 959 Wnt-activated transcripts and 2032 Wnt-repressed transcripts (log2 FC ≥ 1, FDR ≤ 5%); 41 transcripts showed variable regulation by BIO and Dkk1 and were not included in further analysis (Fig. 4B; Fig. S4A, Tables S2 and S6). Thus, we classified a total of 2991 Wnt-regulated genes. Interestingly, over half of the FG-enriched genes (496/906) were repressed by Wnt, whereas only 3% (28/906) were activated. HG-enriched genes exhibited the opposite behavior, with 25% (247/987) being Wnt activated and only 6% (57/987) repressed (Fig. 4B). GO term enrichment showed A-P patterning among the top terms for Wnt-activated genes, whereas Wnt-repressed genes were enriched for terms related to the circulatory system, consistent with the known role of Wnt in repressing early cardiac fate (Fig. S4B). Unsupervised clustering of FG and HG transcript levels in the various experimental conditions revealed that BIO-treated FG samples cluster with HG samples, indicating a genome-scale switch from FG to HG upon Wnt stimulation (Fig. 4C), which is also clearly shown in the scatter plots (Fig. 4D; Fig. S4C,D).

Wnt had the same impact on endoderm and mesoderm transcripts — repressing in the FG and activating in the HG — regardless of germ layer. Wnt-repressed FG transcripts included hhex and sfrp2 in the FG endoderm, as well as spib and hand2 in the FG mesoderm. HG transcripts activated by Wnt included ventx2.1 and cdx2 in the HG endoderm, as well as hoxa9 and hoxd10 in the HG mesoderm (Fig. 4E).

In situ hybridization confirmed that Dkk1 expanded expression of the FG genes gata4, cebpα and tbx1, while reducing the HG transcripts cdx2 and msx1, with BIO treatment having the opposite effect (Fig. 4F). Thus, Wnt/β-catenin acts as a genome-wide switch, promoting the HG transcriptional program and repressing the FG program.

Chromatin binding of β-catenin is associated with activation of HG and repression of FG genes

To determine how Wnt regulates A-P patterning at the genomic level, we performed β-catenin ChIP-seq on stage NF20 control and BIO-treated FG and HG explants (Fig. S5A). We merged sequence files from the four different ChIPs to call β-catenin peaks. This identified 16,303 β-catenin peaks associated with 11,007 genes (±20 kb) (Fig. 4E). In situ hybridization confirmed that Dkk1 expanded expression of the FG genes gata4, cebpα and tbx1, while reducing the HG transcripts cdx2 and msx1, with BIO treatment having the opposite effect (Fig. 4F). Thus, Wnt/β-catenin acts as a genome-wide switch, promoting the HG transcriptional program and repressing the FG program.

Chromatin binding of β-catenin is associated with activation of HG and repression of FG genes
Of the 11,007 genes associated with β-catenin peaks only 1243 (∼11%) were Wnt regulated based on our RNA-seq data (Fig. 5A), suggesting that ∼89% of all genomic β-catenin binding events in the FG or HG tissue were not associated with Wnt-regulated transcription, similar to recent findings in the Xenopus gastrula (Nakamura et al., 2016). Of all Wnt-regulated genes, ∼42% (1243/2991; 1.5 FE, HGT \(P<0.05\)) were associated with β-catenin peaks, including 73% (180/247; 2.7 FE, HGT \(P<0.05\)) of the HG-enriched Wnt-activated genes (Fig. 5Aa), suggesting direct regulation. Unexpectedly, 42% (208/496; 1.6 FE, HGT \(P<0.05\)) of all FG-enriched Wnt-repressed genes were also associated with β-catenin peaks (Fig. 5Ab), suggesting direct repression. This was surprising.
since β-catenin is usually thought to stimulate transcription and thus we expected that Wnt would indirectly repress FG genes. We next assessed whether activation or repression correlated with changes in the recruitment of β-catenin to chromatin through differential peak enrichment analysis of control versus BIO-treated explants (Fig. 5B). As expected, 88% (159/180) of the Wnt-activated HG genes had increased β-catenin binding in BIO-treated FG tissue, as exemplified by cdx2 (Fig. 5Ba′,C). Surprisingly, among the 208 Wnt-repressed FG genes (Fig. 5Ab) we observed different behaviors: 59 genes were associated with increased β-catenin
binding upon BIO treatment (Fig. 5Bb’), 43 genes experienced reduced β-catenin binding (Fig. 5Bb”), while 91 Wnt-repressed genes had no change in β-catenin (Fig. 5B). For example, β-catenin binding near the sfrp2 promoter was increased upon BIO treatment, whereas the nkd2-3 peak was reduced (Fig. 5C). Plotting the average tag density of these different classes of peaks confirmed the significant changes in β-catenin recruitment upon BIO treatment (Fig. 5D; Fig. S5E,F). These data suggest that, in the context of Wnt-repressed genes, elevated nuclear β-catenin levels (as a consequence of BIO treatment) do not necessarily correlate with increased chromatin recruitment.

We next performed p300 ChIP-seq in control and BIO-treated FG and HG tissue to determine how recruitment of the HAT co-activator complex correlates with changes in β-catenin binding (Fig. 5E; Table S7). Examination of individual genes such as cdx2 (Fig. 5C), as well as average tag density analysis, confirmed that p300 was recruited to β-catenin peaks of HG Wnt-activated genes upon BIO treatment (Fig. 5E; Fig. S5E). For both classes of Wnt-repressed FG genes, with either increased β-catenin (e.g. sfrp2) or reduced β-catenin (e.g. nkd2-3) recruitment upon BIO, we observed a trend of reduced p300 recruitment to β-catenin peaks, consistent with repression (Fig. 5C), although this was not statistically significant (Fig. 5E; Fig. S5F). To further investigate how β-catenin recruitment might activate some genes and repress others, we performed motif enrichment analysis on different classes of peaks. Tcf was the most enriched motif in HG Wnt-activated peaks. Surprisingly, Tcf DNA binding sites were not enriched in β-catenin peaks from FG Wnt-repressed genes; rather, these were enriched for Gata, Sox and TEAD motifs (Fig. 5F), suggesting that other DNA-binding proteins might form co-repressor complexes with β-catenin.

**Smad1 and β-catenin converge on common CRMs to coordinate transcription**

We next investigated the extent to which BMP and Wnt cooperate to regulate FG and HG transcription. A comparison of BMP-regulated and Smad1-bound genes with Wnt-regulated and β-catenin-bound genes revealed 229 coordinately regulated genes (Fig. 6A; Fig. S6A). Of these, 33 genes had a total of 111 non-overlapping Smad1 and β-catenin peaks, suggesting regulation through separate CRMs, while the other 196 genes had a total of 369 overlapping β-catenin and Smad1 peaks, suggesting co-occupied CRMs (Fig. 6B). This included 21 genes activated by both pathways, which were primarily regulatory factors known to confer HG fate such as cdx2, ventx2 and mxl1 (Fig. 6C-D). Sixty-two genes with overlapping β-catenin and Smad1 peaks were Wnt activated but BMP repressed, many of which had a paraxial and intermediate mesoderm signature, whereas 55 genes were repressed by both Wnt and BMP including the FG genes irx1-3, sfrp2 and thx1. Finally, a group of 56 genes were BMP activated and Wnt repressed, including the ventral FG endoderm gene hhex, cardiac genes such as nkd2-3, as well as bmp4, bmp7 and the BMP target gene sgl (Fig. 6C-D; Table S8). Although comparable FG and HG datasets are not available in mammals, examination of public ChIP-seq data from human ESC and transformed cell lines (Estaras et al., 2015; Tsankov et al., 2015; Watanabe et al., 2014) revealed that ≈47% (92/194) of the human orthologs have β-catenin and/or SMAD1 peaks in regions of the genome syntenic to the co-occupied peaks in Xenopus (Fig. S7, Table S9), suggesting considerable conservation.

Motif analysis of these distinct classes of co-regulated genes revealed differential enrichment for Tcf, Smad, Gata, Lhx and/or Foxa sites (Fig. 6E), suggesting that different transcription factors probably influence whether β-catenin and Smad1 recruitment to chromatin results in transcriptional activation or repression. As a proof-of-principle that co-bound CRMs are coordinately regulated by Wnt and BMP, we tested the hhex distal region (lacking a Tcf motif) and a cdx2 intronic region (that contains a Tcf motif) in embryo injection luciferase reporter assays (Fig. 6F). The cdx2:luc reporter was more active in the HG than the FG, as expected, whereas the hhex:luc reporter was active in both the FG and the HG, indicating that this CRM alone was not sufficient to confer FG-restricted expression. As predicted, both hhex:luc and cdx2:luc were repressed upon DMH1 treatment (Fig. 6F), consistent with both genes being BMP activated. On the other hand, BIO suppressed the hhex:luc construct but stimulated cdx2:luc activity, consistent with Wnt repressing the FG gene hhex and activating the HG gene cdx2 (Fig. 6F). These data indicate that the binding of Smad1 and β-catenin to the same CRMs coordinates Wnt- and BMP-responsive gene expression to pattern the FG and HG progenitors.

The observation that the bmp4 and bmp7 loci are repressed by Wnt and activated by BMP, and that they both have CRMs with overlapping β-catenin and Smad1 peaks suggested an additional layer of Wnt-BMP signaling crosstalk (Fig. S6B). In situ hybridization confirmed that Wnt negatively regulates expression of bmp4/7 as well as bmp2 in the FG and, to a lesser extent, in the HG. By contrast, BMP signaling was required to maintain robust bmp2/4/7 expression in both the FG and HG (Fig. S6C-D). Interestingly, in the gastrula (the stage when the BIO and DMH1 treatment is started), bmp2 is expressed in the anterior mesendoderm (future FG), a low Wnt environment, whereas bmp4/7 are expressed in the ventral/posterior mesendoderm (future HG), a high Wnt environment (Hoppler and Moon, 1998). Taken together, this suggests that low Wnt is required to maintain bmp2 in the presumptive FG, which then initiates a known BMP positive-feedback loop to promote bmp4/7 expression in the FG (Karavanov et al., 2004; Kenny et al., 2012) (Fig. 7A).

Overall, these data reveal a complex BMP-Wnt gene regulatory network that coordinates A-P and D-V patterning of the FG and HG progenitors (Fig. 7A), with β-catenin and Smad1 converging on CRMs to regulate both transcriptional activation and repression (Fig. 7B; Table S10).

**DISCUSSION**

In embryonic development and hPSC differentiation, combinatorial growth factor signaling controls cell fate decisions, but how these signals coordinate gene expression at the genomic level is poorly understood. We investigated how BMP and Wnt signaling, which are differentially active along orthologous embryonic axes, are interpreted in the genome to coordinate the FG and HG transcriptomes (Fig. 7A). BMP/Smad1 activity, which is high in the ventral and low in the dorsal gut tube, regulates D-V identity in the FG and promotes the expression of key HG genes. On the other hand, low Wnt/β-catenin activity is required for FG fate, while high Wnt induces HG and represses FG transcription. We defined Smad1- and β-catenin-bound CRMs of many Wnt target and BMP target genes, including lineage-specifying transcription factors. Our data suggest a much more complicated regulatory landscape than previously appreciated, with Smad1 or β-catenin binding correlated with either activation or repression, and combinatorial Wnt and BMP signaling converging on hundreds of common CRMs to regulate FG- and HG-specific responses (Fig. 7; Table S10). Although there are many aspects of the model to be resolved with detailed cis-regulatory analysis, this provides a framework for understanding how binding of β-catenin and Smad1 in the genome control early gut patterning.
We found that combinatorial Wnt and BMP signaling coordinates A-P and D-V patterning in the *Xenopus* embryo in a manner that is almost identical to recent hPSC differentiation protocols in which BMP_{low}/Wnt_{low} specify general FG endoderm but BMP_{high}/Wnt_{low} promote ventral FG lineages (Green et al., 2011; Loh et al., 2014). BMP_{high}/Wnt_{low} also promotes cardiac, myeloid and endothelial fates, while repressing intermediate and paraxial mesoderm fates, which require BMP_{low}/Wnt_{high} (Loh et al., 2016). Finally, BMP_{high}/Wnt_{high} promotes intestinal fate (Spence et al., 2011).

Our ChIP-seq results suggest a model for how this differential Wnt and BMP signaling regulates FG and HG transcription (Fig. 7B), and identified hundreds of putative β-catenin- and Smad1-bound CRMs associated with known and novel target genes. In a comparison with recent β-catenin ChIP-seq of *Xenopus* gastrula, we found that 82% (695/849) of *X. laevis* (Kjolby and Harland, 2016) and 43% (1970/4529) of *X. tropicalis* (Nakamura et al., 2016) genes associated with β-catenin occupancy at the gastrula were also in our dataset. These included HG homeobox genes such as *cdx2, ventx1*, *hoxd1* and *msx1*, consistent with Wnt promoting the expression of these posterior genes starting at gastrulation. We also identified several hundred genes co-regulated by Wnt and BMP that exhibit overlapping β-catenin and Smad1 peaks, many of which appear to be conserved in humans based on public ChIP-seq data from hPSCs and transformed cell lines.

![Fig. 6. Smad1 and β-catenin converge on common CRMs.](image-url)

(A) Overlap of BMP-regulated/Smad1-bound genes with Wnt-regulated/β-catenin-bound genes. (B) Schematic of 33 genes with distinct β-catenin and Smad1 peaks and of 196 genes with overlapping β-catenin and Smad1 peaks. Right panel shows Smad1 and β-catenin read density in the corresponding peaks. WE, whole embryo. (C) Venn diagram showing the 196 genes with overlapping Smad1 and β-catenin peaks categorized based on activation (act) or repression (rep) by BMP and Wnt signaling. FG genes are in yellow and HG genes are in green. (D) Genome browser view of β-catenin and Smad1 peaks on *hhex* and *cdx2* with illustration of CRMs tested in luciferase constructs. Red boxes indicate overlapping Smad1 and β-catenin peaks. (E) Motif enrichment analysis of gene sets based on BMP/Wnt regulation. (F) Luciferase assays of reporter constructs with CRMs depicted in D. FG cells were injected in the C1 blastomere and HG cells in the C4 blastomere of 32-cell stage embryos. Error bars represent s.d. of three biological replicates. *P<0.05, Student’s t-test.
(Benahmed et al., 2008; Gaunt et al., 2003; Tsankov et al., 2015; Watanabe et al., 2014), suggesting a conserved paradigm.

Most BMP-activated genes in the FG were not associated with Smad1 binding, suggesting indirect regulation. Despite this, a small cohort of genes encoding lineage-specifying transcription factors, such as nkh2-3, nkh2-5, hand2, spib, gata5, and hhex, were associated with Smad1-bound CRMs. In the case of Nkh2-5 this Smad1-bound CRM is conserved in mammals (Lien et al., 2002), and public ChIP-seq data from hPSCs indicates that SMAD1 binding can also occur at the human HHEX loci (Tsankov et al., 2015) (Fig. S7). This suggests that in the FG, BMP/Smad1 initiate a cascade of transcription factors that promote cardiac, myeloid and hepatic fates. Gata motifs were enriched in our Smad1 peaks, suggesting that they might be key components of this regulatory cascade. Gata factors are known to be required for cardiac and FG development and can physically interact with Smad1 to regulate target gene transcription (Benchabane and Wrana, 2003; Brown et al., 2004; Haworth et al., 2008; Rossi et al., 2001; Trompouki et al., 2011; Xuan et al., 2012).

For most BMP- and Wnt-activated genes, our data support the canonical model of Smad1 and β-catenin associating with DNA-binding proteins (Tcf in the case of β-catenin) to recruit HAT co-activator complex and stimulate transcription. Unexpectedly, ∼55% of BMP-repressed and 73% of Wnt-repressed FG genes associated with Smad1 and β-catenin peaks, respectively, suggesting direct repression. Motif analysis indicated that these peaks were not enriched in consensus Tcf or Smad DNA binding sites; Wnt-repressed genes were enriched for Gata, Sox and TEAD motifs, whereas BMP-repressed genes were enriched for Gata, Lhx and Nkh. Transcriptional repression by Smad or β-catenin has only been documented in a handful of cases. For example, Smad1 can bind to Nkh2-3 on DNA and recruit Sin3/HDAC1 co-repressors to inhibit reporter construct expression in response to BMP signaling (reviewed by Blitz and Cho, 2009; Kim and Lassar, 2003; Marty et al., 2000), consistent with the presence of Nkh motifs in many of our Smad1 peaks (Fig. S3E). In mammalian and Drosophila cells, β-catenin can recruit co-repressor complexes to repress Cdh1 and dpp (bmp) transcription, respectively (Jamora et al., 2003;
and disease contexts. Insight into how Wnt and BMP interact in other developmental lineages (Blythe et al., 2010; Loh et al., 2014; Tsankov et al., 2011). Differentiation and gene expression analyses was carried out with RU486 (Riuso et al., 2014) with log fold change (FC) ≥ 1 or ≤ −1, P < 0.05 and false discovery rate (FDR) ≤ 5%. ChIP was carried out as previously described (Akkers et al., 2012; Blythe et al., 2009) with 25-50 whole embryos or 100 FG or HG explants using the following antibodies: anti-Smad1 (Innovitrogen, 38-5400), anti-β-catenin (Life Technologies, 127200) and anti-p300 (Santa Cruz, sc-585 X). Libraries were sequenced with ~30 million reads/library. Reads were mapped to the X. laevis genome assembly v9.1 using Bowtie 2 at default thresholds (Langmead and Salzberg, 2012). ChIP-seq peaks were called using MACS2 at default thresholds (Zhang et al., 2008). Irreproducibility discovery rate (IDR) was calculated with standard thresholds (Li et al., 2011). Publicly available data for SMAD1 (GSM1507534), and β-catenin (GSM1579346 and GSM1303695) (Estaras et al., 2015; Tsankov et al., 2015; Watanabe et al., 2014) were processed with Bowtie and MACS2 using the BioWardrobe Toolkit at default thresholds (Kartashov and Barski, 2015). For further details of RNA-seq analysis, ChIP and ChIP-seq analysis see the supplementary Materials and Methods.

Acknowledgements
We specially thank Dr Ken Cho and Jin Cho for the ChIP-seq training, Dr Gert Veenstra’s lab for ChIP-seq advice and Dr Matt Weirauch for valuable discussions. We are grateful to the A.M.Z. and J. M. Wells labs and Cincinnati Children’s Endoembryonic Club for helpful discussions. We thank Brian Gebelein and Malcolm Fisher for comments on the manuscript.

Competing interests
A.B. is a co-founder of Datrium, LLC, which provides software development and bioinformatics support services, including installation of BioWardrobe.

Author contributions
M.L.S. and A.M.Z. designed the study, interpreted the data, and wrote the manuscript. M.L.S., S.A.R., M.M. and A.M.Z. performed Xenopus experiments, S.J., M.Y. and A.B. prepared the ChIP-seq libraries. P.C. and M.L.S. performed the bioinformatics analysis.

Funding
This work was funded by National Institutes of Health (NIH) grant R01 DK070858 to A.M.Z. M.L.S. was supported by NIH grant T32 HL007752-2 and a University of Cincinnati Research Council Summer Fellowship. NGS library construction and sequencing were supported by National Institute of General Medical Sciences grant DP2 GM119134 and National Center for Advancing Translational Sciences grant 5UL1TR001425-02 to A.B. This project was supported in part by NIH grant P30 DK078392. Deposited in PMC for release after 12 months.

Data availability
All RNA-seq and ChIP-seq datasets from this study are available in NCBI Gene Expression Omnibus (GEO) under accession number GSE87654.

Supplementary information
Supplementary information available online at http://dev.biologists.org/lookup/doi/10.1242/dev.145789.supplemental

References
microarray analysis identifies a CTNNB1 target signature enriched in intestinal stem cells and colon cancer. PLoS ONE 9, e92317.


