**In Vivo T-Box Transcription Factor Profiling Reveals Joint Regulation of Embryonic Neuromesodermal Bipotency**

George E. Gentsch,1,2,3,* Nick D.L. Owens,1 Stephen R. Martin,5 Paul Piccinelli,1 Tiago Faial,1,3,5 Matthew W.B. Trotter,5 Michael J. Gilchrist,1 and James C. Smith1,2,3,*

1Division of Systems Biology, National Institute for Medical Research, London NW7 1AA, UK
2Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge CB2 1QN, UK
3Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
4Division of Physical Biochemistry, National Institute for Medical Research, London NW7 1AA, UK
5Anne McLaren Laboratory for Regenerative Medicine, Cambridge CB2 0SZ, UK
*Correspondence: ggentsc@nimr.mrc.ac.uk (G.E.G.), director@nimr.mrc.ac.uk (J.C.S.)

**SUMMARY**

The design of effective cell replacement therapies requires detailed knowledge of how embryonic stem cells form primary tissues, such as mesoderm or neuromesoderm that later become skeletal muscle or nervous system. Members of the T-box transcription factor family are key in the formation of these primary tissues, but their underlying molecular activities are poorly understood. Here, we define in vivo genomewide regulatory inputs of the T-box proteins Brachyury, Eomesodermin, and VegT, which together maintain neuromesodermal stem cells and determine their bipotential fates in frog embryos. These T-box proteins are all recruited to the same genomic recognition sites, from where they activate genes involved in stem cell maintenance and mesoderm formation while repressing neurogenic genes. Consequently, their loss causes embryos to form an oversized neural tube with no mesodermal derivatives. This collaboration between T-box family members thus enables the continuous formation of correctly proportioned neural and mesodermal tissues in vertebrate embryos during axial elongation.

**INTRODUCTION**

As the vertebrate embryo elongates along its anteroposterior axis, primary tissues are produced in a continuous fashion to form trunk and tail. This process is thought to occur as a continuum of gastrulation, during which period primary tissues, such as neuromesoderm and mesoderm, emerge for the first time. Recent research concluded that axial elongation is driven mainly by neuromesodermal stem cells at the caudal end of the embryo that go on to form the posterior nervous system and mesodermal derivatives, such as skeletal muscle and notochord (Davis and Kirschner, 2000; Gont et al., 1993; Tzouanacou et al., 2009).

Prominent among the genes that influence the fate of early embryonic cells are members of the T-box transcription factor (TF) family, including Brachyury (also known as T), Eomesodermin (Eomes), and VegT. For example, mouse embryos that lack Brachyury fail to form mesoderm posterior to somites 8–12 (Chesley, 1935). Previous analyses of Brachyury (Xbra), Eomes, and VegT in the Xenopus embryo have focused on their expression patterns, their powerful transactivation activities, and their ability to cause isolated ectodermal tissue to activate mesoderm-specific genes (Showell et al., 2004). However, the way in which T-box TFs exert such profound effects in vertebrate embryos during normal development remains poorly understood.

By combining genomewide chromatin profiling, gain- and loss-of-function experiments, and quantification of DNA binding dynamics, we now provide mechanistic insights into the T-box-mediated cell fate switches that cause neural and mesodermal tissues to form in the correct proportions along the rostrocaudal axis of the vertebrate embryo. Mesodermal cell fate is defined by multiple T-box TFs, and their combined loss restricts the fates of neuromesodermal stem cells such that the embryo forms excess neural tissue at the expense of mesoderm.

**RESULTS**

**Xbra Is Stably Recruited to Motif Variants in Early Development**

To discover how T-box TFs regulate primary tissue formation in vivo, a genomewide binding map was first created for Xbra in *X. tropicalis* gastrula embryos (stages 11–12.5) by chromatin immunoprecipitation coupled to deep sequencing (ChIP-seq) (Figure 1A). A protocol (Extended Experimental Procedures) was developed to efficiently extract and shear chromatin from crosslinked *Xenopus* and zebrafish embryos (Figures S1A–S1D). During gastrulation, Xbra expression occurs predominantly in nascent mesoderm and in the forming notochord (Figures 4A and S5A). Peak calling (false discovery rate [FDR] < 1%) identified ~5,500 Xbra binding sites (Table S1) across the nearly fully sequenced (~88%) genome of *X. tropicalis* (JGI4.1). More than half of Xbra binding was detected upstream of...
2,700 Ensembl genes, determined according to their shortest distances from Xbra binding sites (Figure 1B). However, a significant number of genes showed binding at lower rather than higher levels (Figure 1I), suggesting that many of them are not regulated by this TF in a way that achieves biological relevance (Biggin, 2011). Most Xbra binding occurred within 400 bp of the transcription start site (TSS), with more than a quarter within gene bodies, mostly in introns (Figure 1B). A de novo search for enriched motifs at Xbra binding sites identified four related motif variants (v1–v4), which, with some overlapping coverage, together account for 82% of binding sites detected at the gastrula stage, suggesting that they are involved in Xbra binding (Figure 1C). However, we note that some peaks do not include any of these variants and that many recognition sites elsewhere in the genome are not occupied by Xbra (Figure S1E). This suggests that Xbra binding in vivo is influenced by other factors, such as tethered binding to other proteins and chromatin accessibility (Neph et al., 2012). The 9 bp motif v1 resembles the consensus sequence TVWCACCH selected by Xbra in vitro (Conlon et al., 2001), but, like motifs v2–v4, it includes an additional 5’ thymine that is likely to make hydrophobic contact with a loop of the Xbra T-domain (Müller and Herrmann, 1997). All discovered motif variants include an adenine preceded by a cytosine, with the corresponding guanine being the main contact point for the T-domain in the major groove of dsDNA (Müller and Herrmann, 1997). Motif v2 retains a strong preference for the initial pair of thymines of v1 and the cytosine followed by an adenine, whereas v3 and v4 comprise partial and almost complete palindromes. Motifs v3 and v4 are more degenerate than v1 and v2, except for the main contact bases of the T-box motif. The affinities between native full-length Xbra protein (Figure 1E) and DNA motifs v1, v2, and v4 were confirmed in vitro by surface plasmon resonance (Figure 1F). Motif variant v1 showed the strongest affinity for Xbra, with half of the available...
sites occupied at an Xbra concentration of ~14 nM (the dissociation constant, \(K_d\)). Base changes at the most strongly conserved positions 2 (alanine for thymine) or 5 (guanine for cytosine) of v1 caused this affinity to drop ~1,000-fold. The average nuclear concentration of Xbra at the midgastrula stage was quantified as ~2.9 \(\mu M\) (Figures 1G and 1H). This is 200 times greater than the \(K_d\) of v1, suggesting that the great majority of accessible v1 motifs are likely to be bound at any time. Motifs v2 and v4 show lower affinities for Xbra of ~1.1 \(\mu M\) and ~2.9 \(\mu M\), respectively, suggesting that their occupancies are more sensitive to changes in Xbra concentration. However, it cannot be excluded that Xbra-associating proteins further influence the stability of these interactions in vivo.

The gastrula-stage genome-wide Xbra binding profile was then compared with that of early tail bud embryos at stages 19 and 20, when Xbra expression is confined to the notochord and the caudal end of the embryo (Figure 4A). This comparison indicated that at least 94% of the Xbra binding sites are maintained (overlap ≤ 100 bp) beyond gastrulation (Figures 1D, S1F, and S1G), with at least 97% of target genes being bound at both stages (Figure S1H). However, DNA occupancy levels of target genes did alter slightly between gastrula and early tail bud stages (Figure S1I). These results are consistent with the notion that gastrula and early tail bud embryos contain the same kind of Xbra-expressing cells, including neuromesodermal stem cells (Davis and Kirschner, 2000; Gont et al., 1993; Tzouanacou et al., 2009).

Xbra Balances Mesodermal over Neural Cell Fates

Of the genes that are bound by Xbra, a few have particularly high DNA occupancies within 10 kb of the TSS (Figures 1I and S1J), and we asked whether these are regulated by this T-box TF. The X. tropicalis genome contains two nearly syn-expressed Brachyury paralogues, Xbra and Xbra3, and their activities were inhibited by use of splice- and translation-blocking antisense morpholino oligonucleotides (Figures S2A–S2C). In the course of these experiments, we discovered that the ChIP-grade Xbra antibody does not detect Xbra3 (Figure S2Cii), suggesting that the Xbra binding profiles do not include Xbra3 binding events. Knockdown of Xbra caused truncation of the embryonic body axis, whereas depletion of Xbra3 had little discernible effect. Depletion of both gene products caused a more severe truncation of the tail than did depletion of Xbra alone (Figures S2D and S2E).

Depleted embryos were transcriptionally profiled against controls by RT-qPCR at early neurula, midtail bud, and early tadpole stages. Analysis of 78 of our putative Xbra target genes revealed that slightly less than half (37) were affected at one or more of these stages by at least 1.5-fold (FDR < 10%; Figures 2A and S2F) in either the single or double knockdowns.

![Figure 2. Brachyury Paralogues Xbra and Xbra3 Balance Mesodermal over Neural Fate and Prime Mesoderm for Differentiation](image-url)

**Figure 2. Brachyury Paralogues Xbra and Xbra3 Balance Mesodermal over Neural Fate and Prime Mesoderm for Differentiation**

(A) Differential expression profile (n = 3) of Xbra target genes. Transcriptional fold changes upon Xbra, Xbra3, or Xbra/Xbra3 KD determined at stage 13 (neurula), 26 (midtail bud), and 32 (early tadpole) by RT-qPCR (i), logarithmized, clustered, and visualized as heat map. Gene-associated total Xbra binding levels (ii) detected at stages 11–12.5 (gastrula) by ChIP-seq (n = 2).

(B) WMISH (i) of control and Xbra/Xbra3 KD embryos for selected Xbra target genes at stages 12–12.5 (late gastrula) and 24–25 (midtail bud). Arrowheads and line indicate loss of posterior mesoderm (myf5, msgn1, mespa), formation of irregular, anterior somites (actc1), and ectopic or elevated expression within tail bud and dorsal nervous system (pax3). The scale bar represents 0.5 mm. (ii)

Number of actc1+ somites formed by stage 24 and 25 in control and Xbra/ Xbra3 KD embryos (n = 9).

(C) Tissue-specific Gene Ontology (GO) term analysis of differentially expressed genes (≥1.5-fold; FDR < 10%) in transcriptome-wide study of control and Xbra/Xbra3 KD embryos at stage 32. Statistical significance (p) according to Mann-Whitney U test using PANTHER classification system (Mi et al., 2010). All error bars, SD of indicated biological replicates (n). ≥1.5-fold transcriptional misregulation: *, FDR < 10%; **, FDR < 1%; ***, FDR < 0.1%. See also Figures S2, S3, and S4 and Table S2.
knockdown of Xbra and Xbra3. Loss of just Xbra yielded results that resembled those of the double knockdown but were less severe and in line with the weaker phenotype, suggesting that Xbra and Xbra3 act in a functionally redundant manner. Among the most significantly downregulated genes in Xbra/Xbra3 knockdown embryos were seven involved in the maintenance and specification of paraxial mesoderm and the initiation of somitogenesis at the posterior end of the embryo: tbx6 (Chapman and Papaioannou, 1998); msn1 (Yoon and Wold, 2000); mespa (Sparrow et al., 1998); mespb; ripply2.1 (Chan et al., 2006); hes7.2/esr4; and esr5 (Jen et al., 1999). The disruption of posterior mesoderm formation (see arrowheads in Figures 2B and S3) was confirmed by whole-mount in situ hybridization (WMISH) of several of our Xbra target genes, such as Xbra itself and markers of muscle (myf5, myoD, and actc1), notochord (not), paraxial mesoderm (msn1 and foxc1), and somitomes (delta2, mespa, and esr5). The Xbra/ Xbra3 loss-of-function phenotypes, including reduced numbers of actc1+ or myh1+ somites, may derive in large part from the loss of these gene products (Figures 2B and S3). The Xbra/ Xbra3-dependent target gene LOC733709 (Figure S3), whose sequence and expression pattern is similar to that of esr5 (Jen et al., 1999), may also be a component of the segmentation clock. We also note that the loss of Brachyury function causes significant misregulation of Xbra target genes in gastrula and early neurula embryos (Figures 2A, 2B, and S3), but this is not sufficient to completely disrupt the formation of anterior somites or to cause obvious morphological defects before the tail bud stage (Figure S2E).

Some Xbra target genes were identified as being upregulated in embryos lacking Xbra/Xbra3. These include pax3 and ngn3 (Figure 2A), both of which pattern the dorsal spinal cord (Bang et al., 1997; Nieber et al., 2009), and indeed pax3 showed increased expression in the posterior neural tube and ectopic expression in the tail bud (see arrowheads in Figure 2B), the source, at the posterior wall of the neurenteric canal, of paraxial mesoderm.

To substantiate Brachyury-dependent down- and upregulation, respectively, of mesoderm-specific and neurogenic genes, control and Xbra/Xbra3 knockdown embryos were subjected to transcription-ome-wide profiling (RNA-seq) at the early tadpole stage (Figure S4A; Table S2), when the knockdown phenotype was most pronounced and RT-qPCR suggested that transcriptional misregulation might be most dramatic. Loss of Xbra/ Xbra3 caused misregulation of 1,568 (FDR < 10%) out of 16,760 genes (9.4%), with about half downregulated and half upregulated (Figure S4B). Among the downregulated genes were Xbra3 (35.8-fold), the notochord markers cav1 (5.1-fold) and cav2 (4.5-fold), and an overrepresented group (Mann- Whitney U test; p = 7.8 × 10−6) of genes expressed in muscle, including myh2 (5.9-fold) and tnn1 (3.6-fold). Upregulated genes enriched for neural differentiation (p = 4.2 × 10−5) included Xbra targets, such as sox2 (1.5-fold), foxb1 (1.7-fold), pax3 (1.8-fold), zic2 (1.9-fold), and ngn3 (2.2-fold; Figures 2C and S4C). These results confirm that loss of mesodermal identity, including muscle (Figure 2C) and notochord (Figure S4D), is accompanied by elevated expression of several neural genes, some of which are Xbra targets.

DNA Occupancy Pattern of Xbra Correlates with Gene Activation

Loss of Brachyury caused downregulation of some target genes and upregulation of others. We asked whether the level and position of binding might discriminate between these genes (≥1.5-fold; FDR < 10%) and unaffected target genes.

To this end, gastrula and early tail bud Xbra binding profiles were compared with Xbra/Xbra3 loss-of-function analyses at the early tadpole stage (Figure 3A). Downregulated and upregulated genes both overlapped to a small but significant extent with Xbra binding profiles (Fisher’s exact test; p < 0.05), with more downregulated target genes than upregulated. Similarly, compared to other target gene sets, downregulated genes showed a higher, statistically significant (p < 0.05) binding level than upregulated genes at both proximal (<1 kb) and intermediate distances (1–5 kb) from their TSSs (Figures 3B; Table S3). This binding pattern was particularly prominent at target genes, such as ripply2.2, mespa/b, Xbra3, and tbx6, whose transcription was strongly activated by Xbra/Xbra3 (Figure 3C). This group of genes also showed a slight enrichment for motif v1 when compared to all other gene sets, suggesting that affinity may play a role in regulating transcription (data not shown).

T-box TF Family Members Bind and Regulate Overlapping Genes

Embryos lacking Xbra/Xbra3 gastrulate normally (Figure S2E) and form mesodermal structures anterior to somites 8–12 (e.g., actc1+ somites in Figure 2B). Other T-box TFs may complement Xbra/Xbra3 to allow the formation of these anterior tissues. To test this, we extended our study to include Eomes and zygotic VegT, whose expression patterns around the blastopore and the posterior wall of the neurenteric canal resemble that of Xbra during gastrulation and neurulation, with the exception of the chordoneural hinge and notochord (Figures 4A and S5A).

DNA occupancies of Eomes and VegT were determined by ChiP-seq. Despite their different loss-of-function phenotypes and in vitro affinities for DNA sequences (Conlon et al., 2001; Fukuda et al., 2010), Eomes and VegT were recruited to the same genomic sites as Xbra during gastrulation (Figures 4B, 4C, and S5D) such that most, if not all, target genes were occupied by at least two of these T-box TFs (Figure 4G), suggesting that all three recognize the same binding motifs in vivo (Figures S5B and S5C). A comparison of Brachyury and Eomes (Teo et al., 2011) binding in mesoderm and definitive endoderm derived from human embryonic stem cells reached a similar conclusion (Figure S5E). However, we have no evidence for competition between T-box TFs for individual T-box recognition sites, because Eomes and VegT binding did not increase at Xbra/Xbra3-depleted sites during gastrulation (Figure S5F). This might explain why embryos could not fully compensate for gene misregulation caused by the loss of Brachyury (Figures 2A, 2B, and S2F), and it suggests that there may be only limited overlap of T-box protein expression in single cells or poor accessibility for other T-box TFs at Xbra/Xbra3-depleted sites. Interestingly, the loss of Xbra/Xbra3 caused a significant reduction of DNA occupancy of VegT at some sites, suggesting that some VegT binding is Xbra/Xbra3-dependent. Despite the great similarity of T-box
TF binding profiles, the three T-box TFs differed in their DNA occupancies of particular sites (see peaks in Figures 4B and S5F). This supports the idea that Smads and T-box TFs may act together (Teo et al., 2011) to regulate target gene expression underlying primary cell fate decisions at these stages. There is little enrichment for Smad2/Smad3 motifs at bound sites (Figures S5B and S5C), suggesting that T-box proteins and perhaps members of other TF families contribute to the recruitment of Smad2/Smad3 to their binding sites (Mullen et al., 2011).

The genome-wide binding characteristics of Eomes, VegT, and Xbra suggest that they regulate the same genes. To test this possibility directly, we used an animal cap assay. Gene activation by hormone-inducible versions (glucocorticoid receptor [GR]) of Eomes, VegT, and Xbra was analyzed in the presence and absence of the protein synthesis inhibitor cycloheximide (chx) to ask whether induction was direct (Figure 5A). Forty-five target genes were analyzed (Figures 5B). Most target genes that were downregulated in embryos lacking Brachyury (including mespb, ripply2.2, fgf8, msgn1, gdf5, mespa, fgf4, ripply2.1, fgf20, hes7.2, foxc1, and esr5) were activated directly by all three T-box TFs. Target genes that were upregulated in such embryos (such as zic, pax3, and ngn3) were not activated or were only weakly so. There were some differences in the inducing activities of the T-box TFs, however. For example, not and Xbra3 were significantly induced only by Xbra, and indeed, Eomes and VegT repressed their expression (Figure 5B). Similarly, tbx6 and LOC733709 were activated by Xbra and VegT but not Eomes, and the endodermal marker sqx17b was preferentially induced by VegT. These differences may arise through the differential recruitment of transcriptional cofactors by the different T-box TFs. Analogous experiments within the whole embryo revealed that Xbra-GR can partially restore msgn1 expression within the tail bud of Xbra/Xbra3-depleted embryos both in the presence and in the absence of de novo protein synthesis (see arrowhead in Figure 5C). Exogenous Xbra activity was also able to drive ectopic msgn1 transcription in mesodermal and, less frequently, nonmesodermal tissues (see asterisks in Figure 5C).

**Figure 3. Strongly Activated Target Genes Show Preferential Xbra Binding to Promoter-Proximal and Intermediate Upstream Regions**

(A) Venn diagram of genes targeted by Xbra at gastrula and/or early tail bud stages (sum $[-\log p_{\text{gastrula}}]$ or sum $[-\log p_{\text{tailbud}}] \geq 25$) and genes misregulated at stage 32 ($\geq 1.5$-fold; FDR < 10%) following Xbra/Xbra3 KD. Fisher’s exact test indicates probability of observed overlap ($p_{\text{obs. overlap}}$) and expected number of overlap ($N_{\text{exp. overlap}}$) based on random draws of gene sets from 16,760 genes (for which differential expression was calculated in Table S2). (B) Xbra binding at gastrula and early tail bud stage across down-or upregulated target genes ($\geq 1.5$-fold; FDR < 10%) compared with control sets of target genes. The asterisk indicates significantly ($p < 0.05$) enriched binding compared to controls according to a one-tailed Mann-Whitney U test (Extended Experimental Procedures). Brackets indicate loss of statistical significance ($p \sim 0.2$) when zero DNA occupancies were excluded. See also Table S3. (C) Heat map representation of Xbra binding near strongly Xbra/Xbra3-dependent target genes at gastrula and early tail bud stages.

The similar binding profiles and regulatory capacities of Xbra, Eomes, and VegT encouraged us to explore potential collaboration between these T-box TFs in paraxial mesoderm formation. This was achieved by simultaneous knockdown of the gene products by previously verified antisense morpholino oligonucleotides (Fukuda et al., 2010). Loss-of-function of Eomes, zygotic VegT, or both, in addition to Xbra/Xbra3, caused a downregulation of mesodermal target genes in the trunk that exceeded that observed following knockdown of Xbra/Xbra3 alone (Figures 6A–6F and S6A–S6E). VegT, whose zygotic expression persists at the caudal end of the embryo until the early tail bud stage (Figure 4A), contributes more than Eomes to the ongoing process of presomitic mesoderm specification (msgn1 and foxc1), somitogenesis (defa2, mespa, LOC733709, and esr5), determination (myf5 and myoD), and differentiation (actc1) of skeletal muscle (Figures 6A, 6B, 6D–6F, S6A, and S6C–S6E). The loss of all
T-box TFs impaired gastrulation and abolished the formation of mesoderm and of its derivatives, such as muscle, heart, blood, and pronephros (Figure S7B).

Among target genes whose expression was significantly downregulated in such embryos were some involved in left-right asymmetry, such as gdf3 (Hanafusa et al., 2000), and some in retinoic acid signaling, such as aldha2 and cyp26a1 (Deimling and Drysdale, 2009; Martin and Kimelman, 2010; Figures S6H, S6K, and S6L). Some target genes involved in morphogenesis and in the maintenance of mesodermal tissue, such as wnt11 (Tada and Smith, 2000) and fgf8 (Schulte-Merker and Smith, 1995), showed slight downregulation toward the end of gastrulation, whereas not and ventx2.1 remained robustly expressed during gastrulation, even in the absence of all T-box TFs, suggesting that other factors are required for their regulation (Figures S6F, S6G, S6I, and S6J). In line with statistical tests outlined in Figure 3B, visual inspection of all three T-box TF binding profiles indicated that promoter-proximal binding might determine whether target genes are strongly induced by T-box TFs in vivo (e.g., Figures 6A, 6D, 6E, and S6H) or not (e.g.,...
Indeed, ChIP analysis of T-box TF-depleted early gastrula embryos (stages 10.5–11) confirmed that the recruitment of RNA polymerase II (RNAPII) depends on T-box TFs only at mesodermal target genes, which feature promoter-proximal binding of T-box TFs, such as *fgf4*, *gdf3*, *foxc1*, *msgn1*, and *myf5* (Figure 7A). In contrast, mesodermal or neural target genes without promoter-proximal occupancy of T-box TFs, such as *wnt11*, *not*, and *pax3*, did not show any significant reductions in RNAPII deposition upon T-box TF knockdown.

The upregulation of neurogenic target genes in embryos lacking Xbra/Xbra3 was enhanced by the loss of Eomes and zygotic VegT (Figures 6G and 6H). Cross-sections through the tail buds of such embryos demonstrated the transition of mesodermal to neural identity in the chordoneural hinge (*sox3*) and posterior wall of the neurenteric canal (*pax3*; Figures 6G, 6H, and 7B), whereas the emergence of supernumerary N-tubulin-positive primary neurons (Figure S7A) provided further evidence of increased neural differentiation. Thus, embryos lacking all T-box TFs in the tail bud formed an oversized neural tube in
the absence of any axial or paraxial mesoderm (Figure 6H). This neumesodermal conversion occurred without significant apoptosis in the tail bud, as shown by terminal deoxynucleotidyl transferase deoxyuridine triphosphate nick end labeling (TUNEL) assays of whole-mount embryos lacking T-box TFs (Figure S7C).

**DISCUSSION**

Our results provide several lines of evidence that the T-box TFs Eomes, VegT, and Xbra/Xbra3 (and probably Tbx6) together constitute genetic regulatory inputs that define bipotential
stem cells at the caudal end of the frog embryo and instruct their continuous and correct recruitment to neural and mesodermal tissues (Figure 7B). First, the combinatorial loss of T-box TFs causes embryos to generate more neural cells at the expense of mesoderm without significant induction of programmed cell death. The residual generation of somitic mesoderm (the first 8–12 somites) observed in vertebrates lacking Brachyury (Chesley, 1935; Martin and Kimelman, 2008; this study) occurs through the early action of remaining T-box TFs such as VegT and Eomes (Figure 7Bi). The loss of a T-box TF collective, including Eomes, zygotic VegT, and Brachyury (which strongly activates the expression of Tbx6) abolishes neuromesodermal bipotency and causes caudal cells to form neural tissue with complete loss of mesoderm. Second, gain-of-function experiments confirm that Eomes, VegT, and Xbra can directly activate most T-box TF-dependent target genes regulating neuromesodermal stem cell maintenance and posterior mesoderm development. And third, genome-wide binding profiles reveal that Eomes, VegT, and Xbra are recruited to the same mono- and dimeric recognition sites during gastrulation and that Xbra maintains its binding profile at least throughout the early phases of axial elongation. A nuclear concentration of Xbra quantified as ~2.9 \( \mu \text{M} \) at the midgastrula stage suggests that there is a high occupancy rate of accessible recognition sites, especially at the most stringently conserved motif, whose dissociation constant is ~14 nM.

The observation that T-box TFs are recruited to the same sites bears on the interpretation of experiments involving dominant-interfering TF constructs. For example, an Xbra-EnR construct may well inhibit the function of Eomes, VegT, and Tbx6, as well as that of Xbra. Collaborations between TFs may also be encountered in other TF families, including the Sox proteins, several of which define the nervous system (Bergsland et al., 2011).

Despite the thousands of T-box TF binding sites detected in our genome-wide study, only a minority of these binding events seem to cause biologically significant changes in transcription.
We show that genes that are strongly activated by T-box TFs show significantly enriched binding of T-box TFs to promoter-proximal and intermediate upstream elements. Deletion analysis of the Xbra target fgrf4 suggests that both binding locations are required for appropriate gene expression (Casey et al., 1998). Our experiments indicate that promoter-proximal binding may be important for T-box TFs to recruit RNAPII and induce mesoderm-specific transcription. Recent work in the Drosophila embryo emphasizes the importance of promoters in recruiting RNAPII paused for robust and tissue-specific gene expression (Lagha et al., 2013). The patterns of DNA occupancy and differential expression analyses reveal T-box TF regulatory inputs that define neuromesodermal bipotency and prime mesoderm for further differentiation as follows: (1) stem cell maintenance by autoregulation of Xbra/Xbra3 via fibroblast growth factor (FGF) signaling (Schulte-Merker and Smith, 1995) and regulation of retinoic acid levels by cyp26a1 (Martin and Kimelman, 2010) beyond the bud stage; (2) specification of paraxial mesoderm by Tbx6 (Chapman and Papaioannou, 1998), msgn1 (Wittler et al., 2007), and foxc1 (Wilm et al., 2004); (3) myogenic differentiation by myf5, myoD, myos, and actct1; (4) patterning of presomatic mesoderm by delta2 (Jen et al., 1997), esr4, esr5 (Jen et al., 1999), LOC773709 (this study), mespa, mespb (Sparrow et al., 1998), mesp2, ripply2.1, and ripply2.2 (Kawamura et al., 2005; Kondow et al., 2006); and (5) continuous protection from neuralization by repression of neurogenic genes, such as sox3 and pax3. Interestingly, these neural markers show as much high-density T-box TF binding as some activated target genes mentioned above. These binding clusters may define neuromesodermal bipotency, as they have recently been defined as superenhancers conferring cell identity (Whyte et al., 2015). Overexpression of sox3 and pax3 in anatomical positions of the chordoneural hinge and the posterior wall of the neurenteric canal reflects the local shift of neuromesodermal identity in T-box TF-depleted embryos, such that the loss of axial and paraxial mesoderm is accompanied by the gain of ventrolateral and dorsal neural tissue, respectively (Figure 7Bii).

The way in which T-box TFs suppress transcription of neurogenic genes is not known, although protein phosphorylation (Hwang et al., 2005) and coregulation of a repressor complex (Kawamura et al., 2008) might both be involved in turning these activators into repressors. Indirect repression via T-box TF-dependent signaling pathways, such as FGF, retinoic acid, and Wnt, may also be involved in determining the fate of neuromesodermal stem cells (Li and Storey, 2011; Olivera-Martinez et al., 2012). Notably, some mesoderm-specific targets, such as mespa and ripply2.1, are only activated in presomatic mesoderm (S-I/S-II), where expression of Eomes, VegT, and Xbra3 is low or virtually absent. It is possible that these T-box TFs act as “place-holders” for Tbx6, which continues to be expressed in newly emerging paraxial (presomitic) mesoderm and can activate ripply2.1 expression (Hitachi et al., 2009; Figure 7Bii). Interestingly, presomitic mesoderm retains neuromesodermal plasticity, because Tbx6 mutant mice form two supernumerary neural tubes at the expense of paraxial mesoderm (Chapman and Papaioannou, 1998). In contrast, Eomes, VegT, and Brachyury define neuromesodermal bipotency at an earlier stage, and thus their loss leads to one oversized neural tube.

Together, our experiments demonstrate that a T-box TF collective controls the emergence and fates of the bipotential neuromesodermal stem cells at the caudal end of the vertebrate embryo. The presence of these TFs causes cells to differentiate as mesoderm, and their absence permits them to fulfill their neural potential. Our work provides mechanistic insights into the way in which T-box TFs act together to regulate neuromesodermal fate, and this will inform attempts to define the differentiation pathways of embryonic and induced pluripotent stem cells.

**EXPERIMENTAL PROCEDURES**

**Embryo Culture**

In vitro-fertilized X. tropicalis and X. laevis embryos were cultured in 5% Marc’s modified Ringer (MMR) at 20°C–26°C or 10% normal amphibian medium (NAM) at 14°C–25°C, respectively. Embryos were staged according to Nieuwkoop and Faber (1994). For details on knockdown (and hormone-inducible) overexpression experiments, see below and the Extended Experimental Procedures. All Xenopus studies complied fully with the UK Animals (Scientific Procedures) Act 1986 as implemented by the University of Cambridge and the MRC National Institute for Medical Research.

**Dexamethasone-Inducible GR Assays**

For animal cap assays, X. laevis embryos were injected with 400 pg RNA encoding Xbra-GR (Tada et al., 1997), VegT-GR (White et al., 2002), and/or Eomes-GR (Extended Experimental Procedure). Animal caps were dissected at the blastula stage and cultured in 75% NAM at 20°C until sibling embryos reached stage 10.5. Half of the control and injected caps were then preincubated for 30 min in 10 μM chx and then left untreated or treated with 2 μM dexamethasone (dex), 10 μM chx, or both for about 3 hr until sibling embryos reached stage 12.5. We used a similar experimental set-up for the rescue of msgn1 transcription in Xbra/Xbra3-depleted embryos (Extended Experimental Procedures). Animal caps and embryos without dex and/or chx treatments were incubated with dex and chx solvents ethanol and DMSO.

**Transcriptome-wide Single and Differential Expression Analysis**

Total RNA to the amount of two X. tropicalis embryos (~3 μg) was processed according to the TruSeq protocol (Illumina). Libraries were read paired-end along 55 bases on the HiSeq 2000 machine (Illumina). Bowtie 0.12.7 (Langmead, 2010) was used with the parameters -a-best -v 3 -y -iO -X 10000 to align the reads to the Ensembl JGI4.1 transcriptome. Any read pair that aligned to multiple transcripts of different genes was discarded, and any read pair that mapped to one or many transcripts of the same gene was counted once. Quantitative readouts from RNA-seq experiments were analyzed with DESeq (Anders and Huber, 2010). To display the RNA-seq profile as a track on the University of California Santa Cruz (UCSC) genome browser, reads were mapped to the genome of X. tropicalis as outlined for Chip-seq profiles. The maximal distance between each read pair was set to 100 kb to allow paired reads to map to the genome across large introns. Resultant compressed binary version of sequence alignment/map files were converted to the bedGraph format using the bedTool function genomeCoverageBed.

**WMSH**

WMSH was carried out as described in Monsoro-Burq (2007) with digoxigenin-labeled probes (Extended Experimental Procedures). For sectioning, embryos were dehydrated and embedded in paraffin. Ten micrometer sections were counterstained with Nuclear Fast Red.

**ChiP**

This protocol (Extended Experimental Procedures), designed to process whole Xenopus embryos (which also proved to be applicable to zebrafish embryos), evolved from Lee et al. (2006). Key changes were made to the removal of residual fixative and the extraction of crosslinked nuclei from embryos prior to sonication to facilitate solubilization and efficient shearing of chromatin.
ChIP-seq Analysis

Sequencing reads were mapped to the X. tropicalis genome assembly JGI4.1 using CLC Bio Genomics Workbench default settings. Nonspecific and ambiguous matches were ignored. ChIP-seq peaks were identified using MACS 2.0.4 (Zhang et al., 2008). Genomic coordinates of peaks are summarized in Table S1. Binding (pile-up of reads in Figure 1A; peak p values in all other figures) and transcript profiles were visualized on the UCSC genome browser. The nearest genes to peaks were found by ranking distances between peak summits and TSS of Ensembl genes (JGI4.1) using MySQL 5.6.2. Homer (Heinz et al., 2010) was used to perform metagene analysis and create tag and motif density maps. R, Excel, Cluster3, and JavaTreeview were subsequently used to combine different ChIP/RNA-seq data sets and visualize data as histograms, Venn diagrams, or heat maps. De novo motif analysis was performed with cisFinder (Sharov and Ko, 2009). See Extended Experimental Procedures for further details.

Surface Plasmon Resonance

The affinity of Xbra binding to different DNA motifs was measured on an Octet RED biolayer interferometer. Biotinylated DNA oligonucleotides (Extended Experimental Procedures) were immobilized on streptavidin biosensors at concentrations in the range 0.5–0.7 μg/ml. Binding of native Xbra protein (Extended Experimental Procedures) at concentrations of 3 nM to 3.8 μM was measured at 25 °C in a 5–10 min association step. The buffer contained 10 mM sodium phosphate pH 7.4, 150 mM NaCl, 0.005% Tween-20, and 0.1 mg/ml BSA. The (relative) amount of Xbra bound to the sensors was calculated from the amplitude of the response at the end of each association step. Equilibrium dissociation constants (K_d) were determined by fitting the response as a function of the Xbra concentration.

ACCESSION NUMBERS

The Gene Expression Omnibus accession numbers of the ChIP-seq and RNA-seq datasets reported in this paper are GSE48560 and GSE48663.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.08.012.

AUTHOR CONTRIBUTIONS

G.E.G. conceived the study and carried out most experiments and postsequencing analysis. N.D.L.O., P.P., M.W.B.T., and M.J.G. helped with the processing of sequencing data. T.F. contributed Brachyury binding data from human embryonic stem cell derivatives. J.C.S. and S.R.M. carried out whole-mount immunohistochemistry for Xbra protein and SPR, respectively. G.E.G. is grateful to M. Thompson, A. Bernardo, K. Dingwell, and M. Wu for critical reading of the manuscript. J.C.S. and G.E.G. were supported by the Wellcome Trust and by the UK Medical Research Council (programme number U117597140).

ACKNOWLEDGMENTS

We thank J. Gurdan for antibodies, the Zimmerman lab for X. tropicalis embryos, A. Sesay for processing Illumina libraries, C. Brenner for implementing the X. tropicalis genome into Homer, T. Mohun for plasmids, R. Mahmood for histology, and A. Ramos for technical advice on surface plasmon resonance (SPR) design. We thank the Smith lab for discussions and advice, and G.E.G. is grateful to M. Thompson, A. Bernardo, K. Dingwell, and M. Wu for critical reading of the manuscript. J.C.S. and G.E.G. were supported by the Wellcome Trust and by the UK Medical Research Council (programme number U117597140).

Received: May 3, 2013
Revised: July 11, 2013
Accepted: August 6, 2013
Published: September 19, 2013

REFERENCES

Akkers, R.C., van Heeringen, S.J., Jacobi, U.G., Janssen-Megens, E.M., Françoïs, K.−J., Stunnenberg, H.G., and Veenstra, G.J.C. (2009). A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos. Dev. Cell 17, 425–434.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Bang, A.G., Papapololu, N., Kintner, C., and Goulding, M.D. (1997). Expression of Pax-3 is initiated in the early neural plate by posteriorizing signals produced by the organizer and by posterior non-axial mesoderm. Development 124, 2075–2085.

Bergsland, M., Ramsköld, D., Zouuter, C., Klum, S., Sandberg, R., and Muhr, J. (2011). Sequentially acting Sox transcription factors in neural lineage development. Genes Dev. 25, 2453–2464.

Biggin, M.D. (2011). Animal transcription networks as highly connected, quantitative continua. Dev. Cell 21, 611–626.

Casey, E.S., O’Reilly, M.A., Conlon, F.L., and Smith, J.C. (1998). The T-box transcription factor Brachyury regulates expression of eFGF through binding to a non-palindromic response element. Development 125, 3887–3894.

Chan, T., Satow, R., Kitagawa, H., Kato, S., and Asasahima, M. (2006). Letterline, a novel Xenopus laevis gene, regulates differentiation of preemiotic mesoderm during somitogenesis. Zool. Sci. 23, 689–697.

Chapman, D.L., and Papaioannou, V.E. (1998). Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. Nature 391, 695–697.

Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. J. Exp. Zool. 70, 429–459.

Conlon, F.L., Fairclough, L., Price, B.M., Casey, E.S., and Smith, J.C. (2001). Determinants of T-box protein specificity. Development 128, 3749–3758.

Cooke, J. (1979). Cell number in relation to primary pattern formation in the embryo of Xenopus laevis. I. The cell cycle during new pattern formation in response to implanted organizers. J. Embryol. Exp. Morphol. 51, 165–182.

Davis, R.L., and Kirschner, M.W. (2000). The fate of cells in the tailbud of Xenopus laevis. Development 127, 255–267.

Deimling, S.J., and Drysdale, T.A. (2009). Retinoic acid regulates anterior-posterior patterning within the lateral plate mesoderm of Xenopus. Mech. Dev. 126, 913–923.

Fukuda, M., Takahashi, S., Haramoto, Y., Onuma, Y., Kim, Y.J., Yeo, C.Y., Ishiura, S., and Asashima, M. (2010). Zygotic VegT is required for Xenopus paraxial mesoderm formation and is regulated by nodal signaling and Eomesoderm. Int. J. Dev. Biol. 54, 81–92.

Gont, L.K., Steinbeisser, H., Blumberg, B., and de Robertis, E.M. (1993). Tail formation as a continuation of gastrulation: the multiple cell populations of the Xenopus tailbud derive from the late blastopore lip. Development 119, 991–1004.

Hanafusa, H., Masuyama, N., Kusakabe, M., Shibuya, H., and Nishida, E. (2000). The TGF-beta family member demière is involved in regulation of the establishment of left-right asymmetry. EMBO Rep. 1, 32–39.

Heinz, S., Benner, C., Spahn, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589.

Hitachi, K., Danno, H., Tazumi, S., Aihara, Y., Uchiyama, H., Okabayashi, K., Kondow, A., and Asasahima, M. (2009). The Xenopus Bowline/Ripply family proteins negatively regulate the transcriptional activity of T-box transcription factors. Int. J. Dev. Biol. 53, 631–639.

Hwang, E.S., Szabo, S.J., Schwartzberg, P.L., and Glimcher, L.H. (2005). T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. Science 307, 430–433.

Jen, W.C., Wettstein, D., Turner, D., Chitnis, A., and Kintner, C. (1997). The Notch ligand, X-Delta-2, mediates segmentation of the paraxial mesoderm in Xenopus embryos. Development 124, 1169–1178.
Embryos. Genes Dev. 1999. Periodic repression of Notch pathway genes governs the segmentation of Xenopus embryos. Genes Dev. 13, 1486–1499.

Kawamura, A., Koshida, S., Hijioka, H., Ohbayashi, A., Kondoh, H., and Takada, S. (2005). Groucho-associated transcriptional repressor ripply1 is required for proper transition from the presomitic mesoderm to somites. Dev Cell 9, 735–744.

Kawamura, A., Koshida, S., and Takada, S. (2008). Activator-to-repressor conversion of T-box transcription factors by the Groucho/TLE-associated mediators. Mol. Cell. Biol. 28, 3236–3244.

Kondow, A., Hitachi, K., Ikegame, T., and Asashima, M. (2006). Bowline, a novel protein localized to the presomitic mesoderm, interacts with Groucho/TLE-associated mediators. Mol. Cell. Biol. 26, 176–189.

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. Curr. Protoc. Bioinformatics Chapter 11, Unit 11.7.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. 1, 729–748.

Levy, D.L., and Heald, R. (2010). Nuclear size is regulated by importin a and Ntf2 in Xenopus Cell 143, 288–298.

Li, R.A., and Storey, K.G. (2011). An emerging molecular mechanism for the neural vs mesodermal cell fate decision. Cell Res. 21, 708–710.

Martin, B.L., and Kimelman, D. (2008). Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. Dev. Cell 15, 121–133.

Martin, B.L., and Kimelman, D. (2010). Brachyury establishes the embryonic mesodermal progenitor niche. Genes Dev. 24, 2778–2783.

Mi, H., Dong, Q., Muruganujan, A., Gaudet, P., Lewis, S., and Thomas, P.D. (2010). PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Res. 38(Database issue), D204–D210.

Monsoro-Burq, A.H. (2007). A rapid protocol for whole-mount in situ hybridization on Xenopus embryos. CSH Protoc. 2007, pdb prot4809.

Mullen, A.C., Orlando, D.A., Newman, J.J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-β signalling. Cell 147, 565–576.

Müller, C.W., and Herrmann, B.G. (1997). Crystallographic structure of the T-domain-DNA complex of the Brachyury transcription factor. Nature 389, 884–888.

Neph, S., Vierstra, J., Stengachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature 489, 83–90.

Niebler, F., Pieler, T., and Henningfeld, K.A. (2009). Comparative expression analysis of the neurogenins in Xenopus tropicalis and Xenopus laevis. Dev. Dyn. 238, 451–458.