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,[4] Paul Piccinelli,[1] Tiago Faial,[1,3,5] Matthew W.B. Trotter,[5] Michael J. Gilchrist,[1] and James C. Smith[1,2,3,*]
[1]Division of Systems Biology, National Institute for Medical Research, London NW7 1AA, UK
[2]Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge CB2 1QN, UK
[3]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
[4]Division of Physical Biochemistry, National Institute for Medical Research, London NW7 1AA, UK
[5]Anne 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.)
http://dx.doi.org/10.1016/j.celrep.2013.08.012

SUMMARY

The design of effective cell replacement therapies requires detailed knowledge of how embryonic stem cells form primary tissues, such as mesoderm or neurectoderm 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 genome-wide 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 ensures 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 continuation of gastrulation, during which period primary tissues, such as neurectoderm 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 genome-wide 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 genome-wide 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

A scaffold 73 (JGI4.1)

980kb | 990kb | 1000kb |

X. tropicalis gastrula (st. 11-12.5)

800 — Xbra
400 — Xbra biological replicates
100 — RNAPII
100 — poly(A) RNA

LOC100495414 LOC733709

B

number of binding sites

120
80
40

Xbra

-20 -10 TSS TES +10 +20 kb

1-5kb 9.3%
<1kb 5.8%
5' UTR 0.5%
CDS 0.6%
3' UTR 0.3%
>5kb 37.9%
intragenic 26.8%
>5kb 16.5%
intron 25.3%
<1kb 0.6%
1-5kb 3.2%

C

Frequency

v1 40% T T C A C A C T 200 / 0
v2 44% 200 / 0
82%
v3 32% 200 / 0
v4 41% 200 / 0

-200 0 +200 bp

D

20 kb

gastrula st. 11-12.5
6 — Xbra
150 — poly(A) RNA

early tailbud st. 19-20
6 — Xbra
300 — poly(A) RNA

mespa mesp2 mespb

E

kDa
150
75
50
37
25
20
15
10

Xbra-His$_{10}$

F

Normalised Response

1
0.8
0.6
0.4
0.2
0

0 50 100 150
[Xbra] (nM)

1
0.8
0.6
0.4
0.2
0

0 1500 3000
[Xbra] (nM)

v1 K$_d$ = 14 ± 2 nM
v1^{T2A} K$_d$ > 10 μM
v1^{C5G} K$_d$ > 10 μM
v2 K$_d$ = 1.1 ± 0.2 μM
v4 K$_d$ = 2.9 ± 0.6 μM

G

Xbra-His$_{10}$ | X. tropicalis (stage)

1ng 500pg 200pg 50pg | 9 12

+ — — Xbra RNA

Xbra

12 6 2.4 0.6 0 ~3.8 x10^9 molecules (per embryo)
~9.5 x10^5 molecules/cell (average)
~2.9 μM (nuclear concentration)

α-tubulin

H

Xbra protein

st. 12

I

Gastrula-staged DNA occupancy levels of Xbra target genes

Frequency

250
200
150
100
50
0

20 10 9 4 3 2
8 7 6 5

0 500 1000
sum (-log p) / target gene

1 *LOC100495414*
2 *foxl1**, *tbx3*, *zic3*
3 *snai1*
4 *hes7.2*, **pax3**
5 *agrn**
6 **foxc1**, *zcchc6**, **LOC733709**
7 *sall1*, *lmod3**
8 **xbra**, *sox2*
9 *clybl**, *znf703*, *pkcdl**
10 *vps13b.1*, *efnb1*, *cdh2**
11 *meis3*
12 *sema3a*, **foxb1**
13 *znf608*, *LOC100145154**
14 *epha4**, *cdx2*, *dact1*, *gtdc1*, **msgn1**, *sh3d19**
15 **not**, *tanc2**, *hipk1*, *meox2*, *fn1*
16 *kiaa1217**, *e2f8*, *lef1*
17 **wnt11**, *fgf16*, **mespa**, *dhrs3*, *fzd10*, *eif4h**
18 *plekhn1*, *zfp36l2*, **mespb**
19 *irx1*, *spata5l**, **aldh1a2**, **myf5**, *eef1a1*
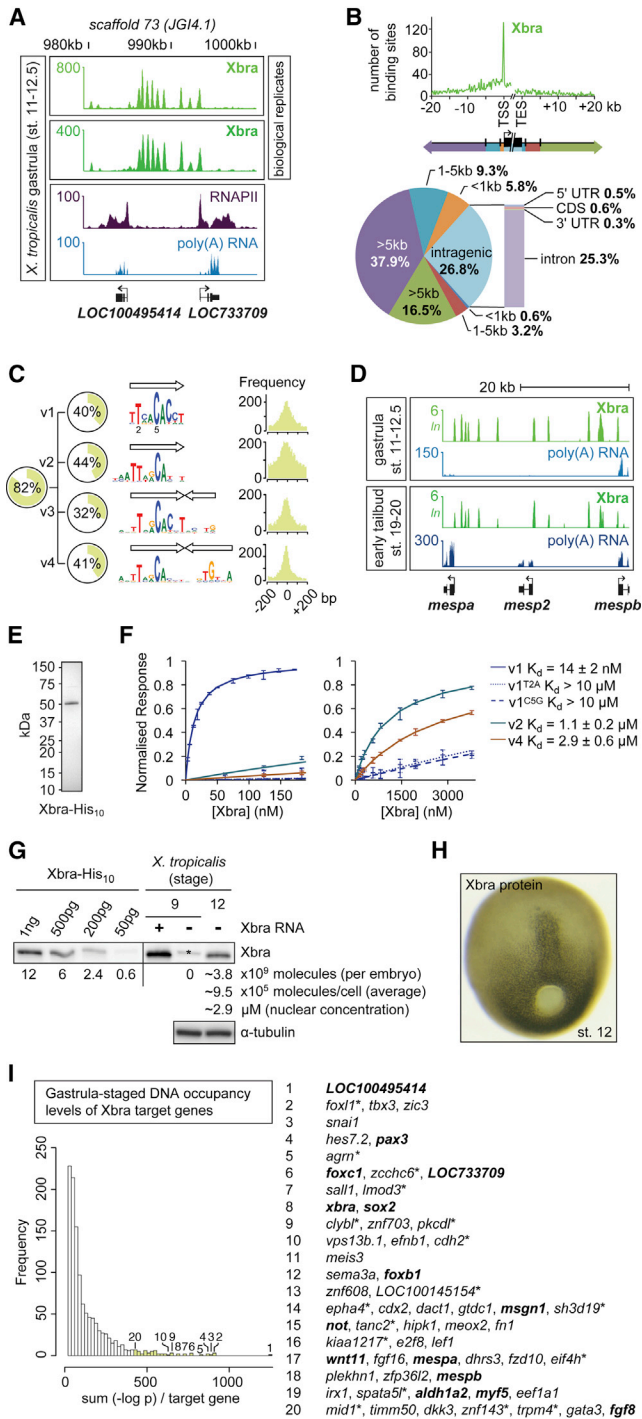20 *mid1**, *timm50*, *dkk3*, *znf143**, *trpm4**, *gata3*, **fgf8**

Figure 1. Xbra Is Stably Recruited to Mono- and Dimeric Motif Variants in the *X. tropicalis* Genome during Early Embryogenesis

(A) Excerpt of normalized Xbra binding at gastrula stage. RNAPII and poly(A) RNA profile from Akkers et al. (2009).

(B) Genomic distribution of Xbra binding sites (FDR ≤ 1%) relative to the start (TSS) and end (TES) of transcription of nearest target genes.

(C) De novo motif discovery analysis of Xbra-bound regions with coverage, sequence logo, and positional distribution for each T-box motif variant (v1–v4). Arrow indicates monomeric binding site.

~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

(D) Comparison of Xbra binding at gastrula and early tail bud stage near *mesp* gene cluster with poly(A) RNA profile from Akkers et al. (2009) and this study.

(E) Coomassie staining of 0.5 μg purified Xbra-His$_{10}$ run on a SDS-polyacrylamide gel.

(F) Surface plasmon resonance diagrams (normalized response versus Xbra concentration) including K$_d$ values for the interaction between native Xbra protein and different DNA motifs (v1, v2, and v4). Superscript T2A and C5G refer to base changes introduced in v1.

(G) Quantification of Xbra protein levels in midgastrula embryos (stage 12) by western blotting with standard curve of purified Xbra-His$_{10}$ as indicated. Positive control, pregastrula embryo (stage 9) injected with RNA encoding untagged Xbra. Protein extracts equivalent to two embryos at stages 9 (negative control) and 12 were loaded. Asterisk marks nonspecific band seen at stage 9. The same band is present at the same intensity in the absence of Xbra at stage 12 (data not shown), and its intensity was therefore subtracted from the Xbra band for quantification. Further calculations (molecules/cell) and nuclear concentrations (μM) are based on an estimated 4,000 Xbra-positive cells at stage 12 ([H]; Cooke, 1979), a nuclear envelope (sphere) surface of ~300 μm^2 (Levy and Heald, 2010), and 90% of Xbra being nuclear. Loading control, α-tubulin.

(H) Whole-mount immunohistochemistry of Xbra protein in a midgastrula embryo (stage 12).

(I) Histogram of nearest gene-associated Xbra binding levels as detected by ChIP-seq. The asterisk indicates genes with nearest Xbra binding >10 kb from TSS. Genes in bold are mentioned elsewhere in this study.

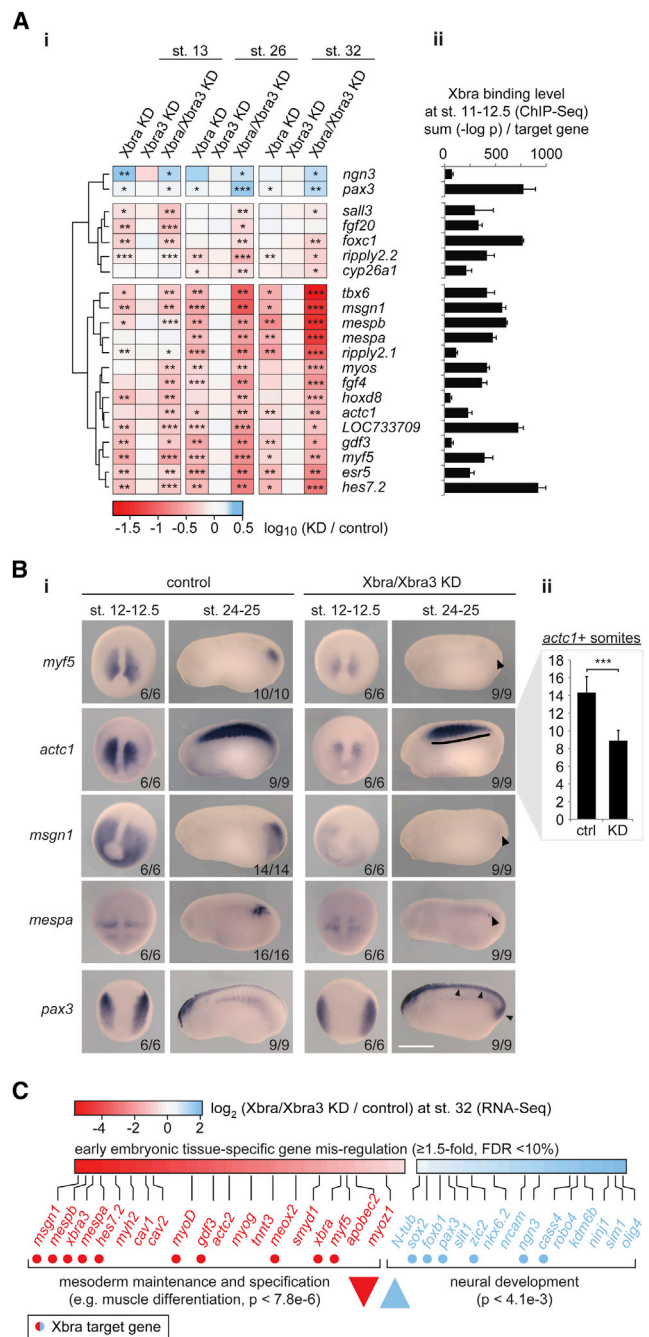See also Figure S1 and Table S1.

A

i

st. 13   st. 26   st. 32

(columns repeated: Xbra KD / Xbra3 KD / Xbra/Xbra3 KD for each stage)

ii

Xbra binding level
at st. 11–12.5 (ChIP-Seq)
sum (-log p) / target gene

0   500   1000

Genes (top to bottom):
ngn3, pax3, sall3, fgf20, foxc1, ripply2.2, cyp26a1, tbx6, msgn1, mespb, mespa, ripply2.1, myos, fgf4, hoxd8, actc1, LOC733709, gdf3, myf5, esr5, hes7.2

-1.5 -1 -0.5 0 0.5   log₁₀ (KD / control)

B

i

control   Xbra/Xbra3 KD   ii

st. 12–12.5 | st. 24–25 | st. 12–12.5 | st. 24–25

myf5   6/6   10/10   6/6   9/9
actc1   6/6   9/9   6/6   9/9
msgn1   6/6   14/14   6/6   9/9
mespa   6/6   16/16   6/6   9/9
pax3   6/6   9/9   6/6   9/9

actc1+ somites
18
16
14
12
10
8
6
4
2
0
     ctrl   KD
***

C

log₂ (Xbra/Xbra3 KD / control) at st. 32 (RNA-Seq)

-4 -2   0   2

early embryonic tissue-specific gene mis-regulation (≥1.5-fold, FDR <10%)

msgn1, mespb, xbra3, hes7.2, myf2, cav1, cav2, myoD, gdf3, actc2, myog, tnnt3, meox2, smyd1, xbra, myf5, apobec2, myoz1 | N-tub, sox2, foxb1, pax3, slit2, kcl2, nkx6.2, ncam, ngn3, cass4, robo4, kdm6b, ninj1, olig4

mesoderm maintenance and specification
(e.g. muscle differentiation, p < 7.8e-6)

neural development
(p < 4.1e-3)

Xbra target gene

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.

---

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 μM (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 μM and ~2.9 μM, 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-quantitative PCR (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

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); *msgn1* (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 (*msgn1* and *foxc1*), and nascent somites (*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 transcriptome-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; $p < 7.8 \times 10^{-6}$) of genes expressed in muscle, including *myh2* (5.9-fold) and *tnnc1* (3.6-fold). Upregulated genes enriched for neural differentiation ($p < 4.1 \times 10^{-3}$) 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 ($\geq$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., *actc+* 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
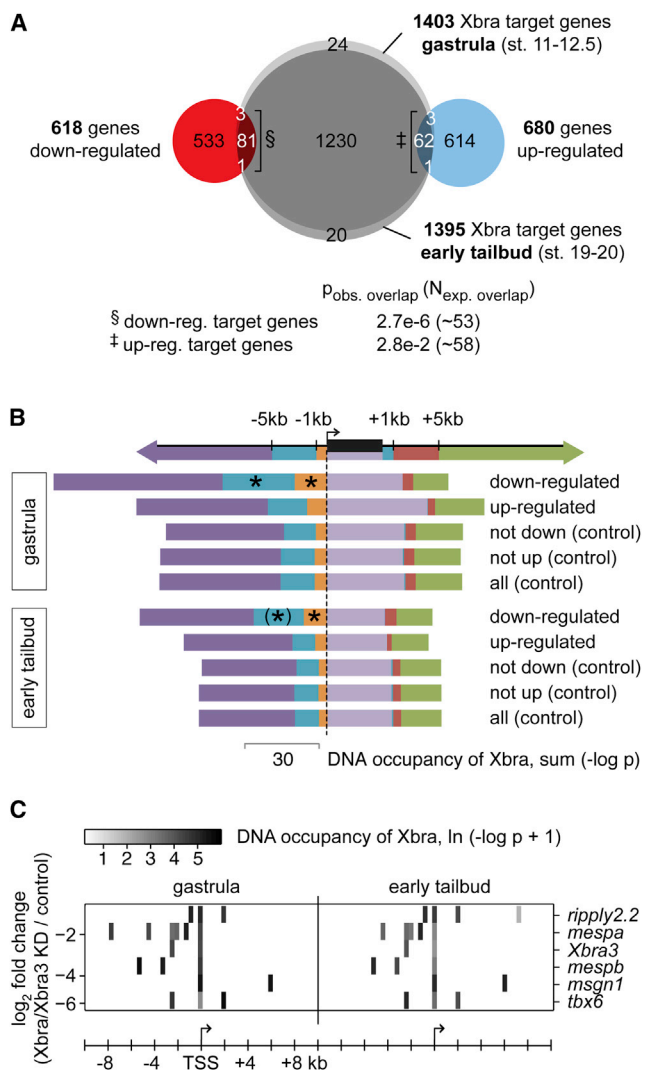
A

1403 Xbra target genes
gastrula (st. 11-12.5)

24

618 genes
down-regulated

533 | 3 | 81 | § | 1230 | ‡ | 3 | 62 | 614

680 genes
up-regulated

20

1395 Xbra target genes
early tailbud (st. 19-20)

$p_{obs.\ overlap}$ ($N_{exp.\ overlap}$)
§ down-reg. target genes   2.7e-6 (~53)
‡ up-reg. target genes   2.8e-2 (~58)

B

-5kb -1kb +1kb +5kb

gastrula

* * down-regulated
up-regulated
not down (control)
not up (control)
all (control)

early tailbud

(*) * down-regulated
up-regulated
not down (control)
not up (control)
all (control)

30 DNA occupancy of Xbra, sum (-log p)

C

DNA occupancy of Xbra, ln (-log p + 1)
1 2 3 4 5

gastrula   early tailbud

log_2 fold change
(Xbra/Xbra3 KD / control)

ripply2.2
mespa
Xbra3
mespb
msgn1
tbx6

-8 -4 TSS +4 +8 kb

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 $p_{gastrula}$] or sum [−log $p_{tailbud}$] ≥ 25) and genes misregulated at stage 32 (≥1.5-fold; FDR < 10%) following Xbra/Xbra3 KD. Fisher's exact test indicates probability of observed overlap ($p_{obs.\ overlap}$) and expected number of overlap ($N_{exp.\ overlap}$) 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 (≥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.

TF binding profiles, the three T-box TFs differed in their DNA occupancies of particular sites (see peaks in Figures 4B and S5F).
Further analysis revealed that over 80% of the binding positions of nuclear Smad2/Smad3 (Yoon et al., 2011), which medi-

ates transforming growth factor β (TGF-β) signaling and targets responsive *cis*-regulatory elements on the genome, overlap with those of the T-box TFs during gastrulation (Figures 4B–4F). 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*, *gdf3*, *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 *szl*, *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 *sox17b* 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).

T-box TFs Recruit RNA Polymerase II to Define Neuromesodermal Bipotency
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 (*delta2*, *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
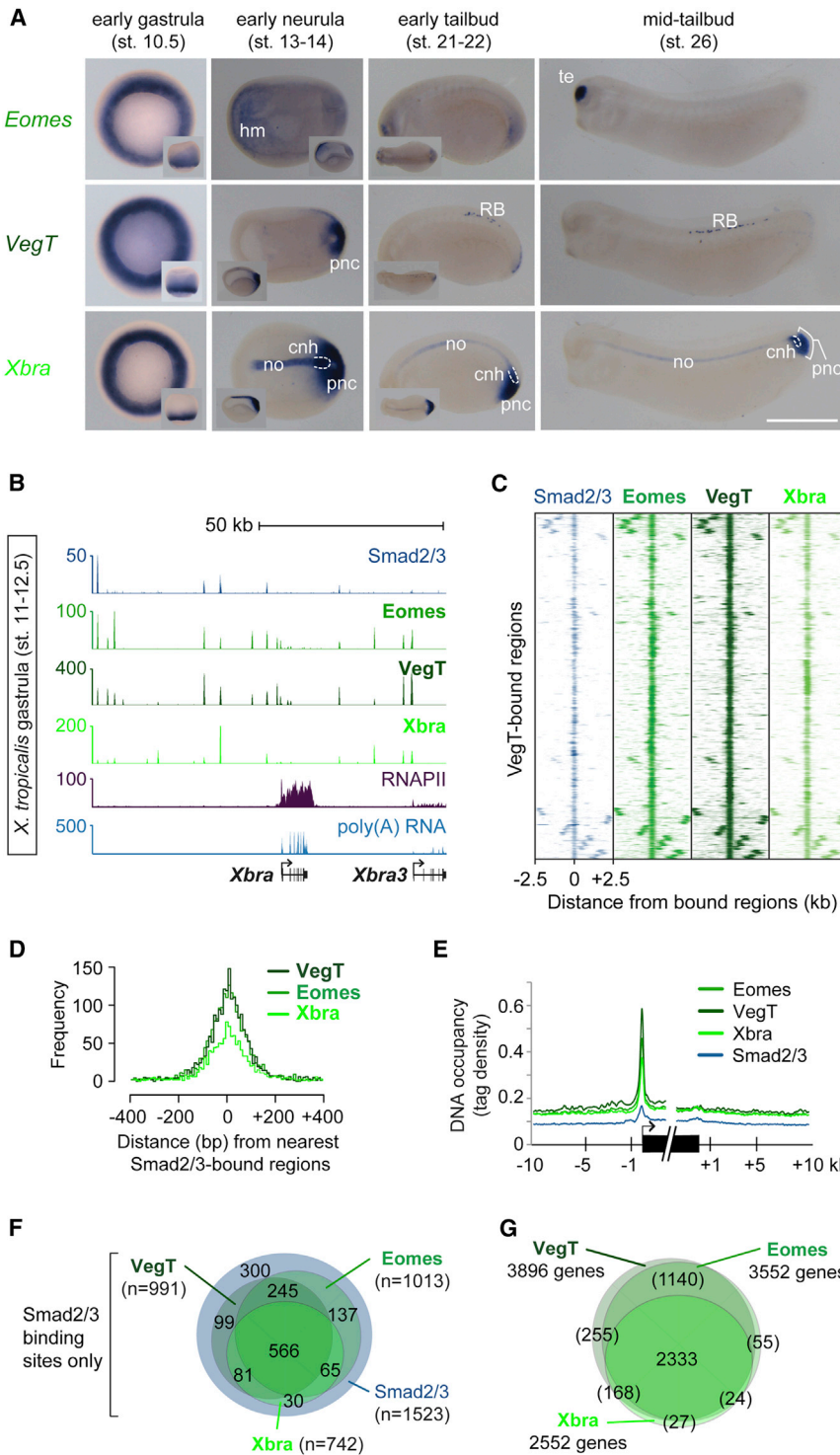
Figure 4. Eomes, VegT, and Xbra Occupy the Same Genomic Recognition Sites and TGF-β Responsive Regulatory Elements

(A) Developmental series of WMISHs for *Eomes*, *VegT*, and *Xbra*. cnh, chordoneural hinge; hm, head (prechordal) mesoderm; no, notochord; pnc, posterior wall of neurenteric canal; RB: Rohon-Beard cells; te, telencephalon. The scale bar represents 0.5 mm.

(B) Gastrula-staged snapshot of Eomes, VegT, Xbra, and Smad2/Smad3 (Yoon et al., 2011) binding near *Xbra* and *Xbra3*.

(C) Heat maps represent DNA occupancies of Smad2/Smad3, Eomes, VegT, and Xbra relative to all VegT-bound regions during gastrulation.

(D) Histograms based on pairwise calculations of shortest distances between T-box TF and Smad2/Smad3 binding positions (FDR ≤ 1%) during gastrulation.

(E) Metagene model shows DNA occupancy level of Eomes, VegT, Xbra, and Smad2/Smad3 relative to the start and end of nearest target genes.

(F) Venn diagrams for Smad2/Smad3 binding positions overlapping (distance ≤ 100 bp) with the binding position of Eomes, VegT, and/or Xbra (FDR ≤ 1%).

(G) Venn diagrams for Eomes, VegT, and Xbra target genes. The stringency was relaxed up to $p \leq 0.1$ unless DNA occupancy was found equivalent to $p \leq 10^{-25}$ when comparing binding between two T-box TFs. DNA occupancies ($p \leq 10^{-25} + 10^{-25} < p \leq 0.1 =$ total) detected as follows: Eomes (2407 + 1145 = 3552); VegT (3628 + 268 = 3896); Xbra (1379 + 1173 = 2552). The extent of overlap between genes targeted by different T-box TFs might be greater than indicated in brackets, because peaks ($p \leq 0.1$) sometimes failed to be detected.

See also Figure S5 and Table S1.

asymmetry, such as *gdf3* (Hanafusa et al., 2000), and some in retinoic acid signaling, such as *aldh1a2* 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.,
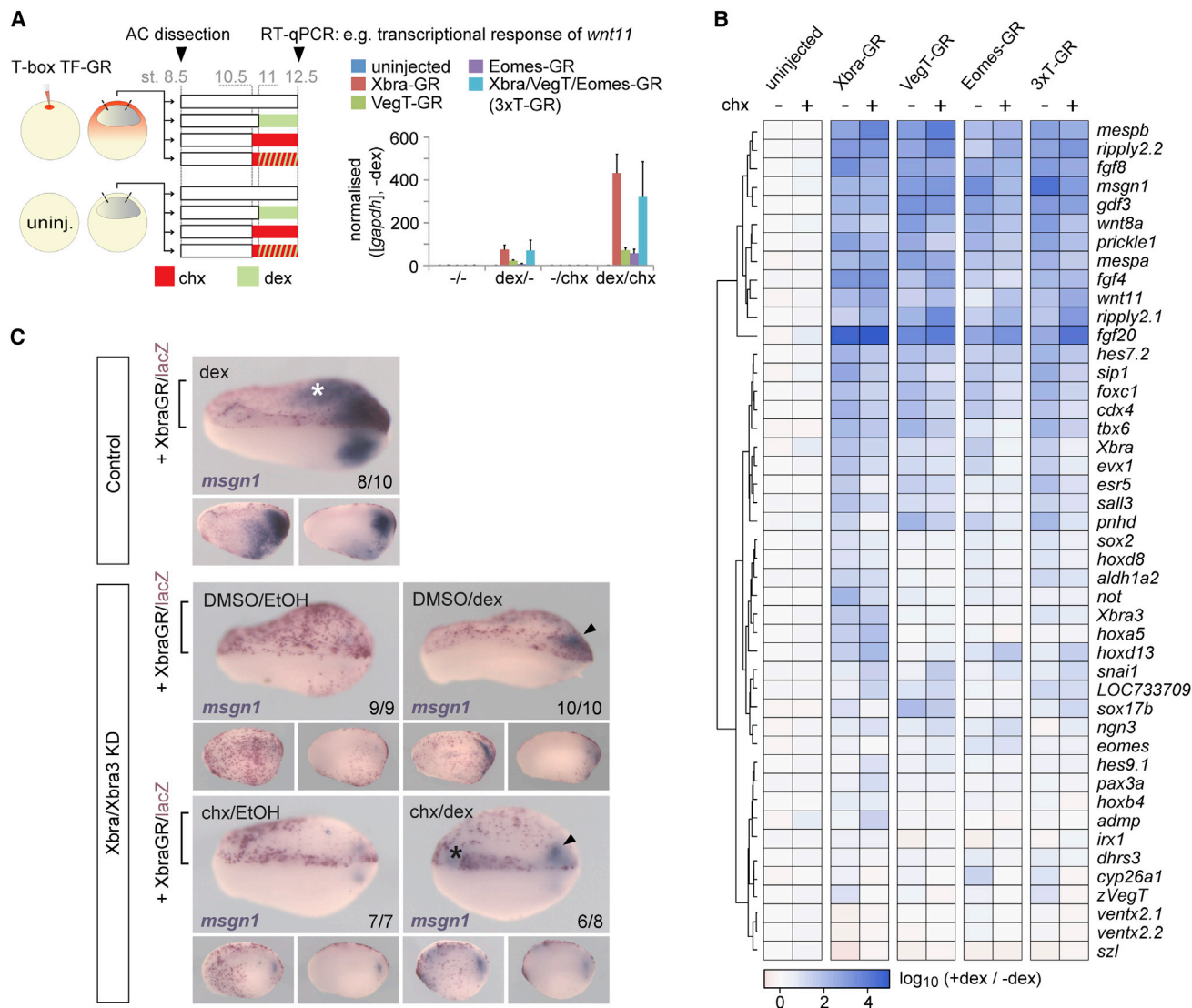
T-box TFs impaired gastrulation and abolished the formation of mesoderm and of its derivatives, such as muscle, heart, blood, and pronephros (Figure S7B).

Among target genes whose expression was significantly downregulated in such embryos were some involved in left-right

Figure 5. Eomes, VegT, and Xbra Can Activate Directly the Same T-box TF-dependent Target Genes

(A) Animal cap (AC) assay to detect direct regulation (i.e., in the presence of chx) of gene transcription by RT-qPCR using dex-inducible fusion constructs (Xbra-GR, VegT-GR, Eomes-GR) individually or in combination (3xT-GR). Data normalized to *gapdh* and the uninduced sample (−dex). The error bars represent SD of biological duplicates.

(B) Heat map representation of clustered transcriptional response ratios (+dex/−dex) of T-box TF target genes to the activity of T-box TFs with (−chx) or without (+chx) de novo protein synthesis.

(C) Protein synthesis-independent rescue of *msgn1* transcription in the tail bud (arrowheads) of Xbra/Xbra3-depleted embryos (stages 22–23) by activated Xbra-GR, whose RNA was unilaterally injected together with *lacZ* lineage tracer RNA.

Figures 6H, S6F, and S6I). 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
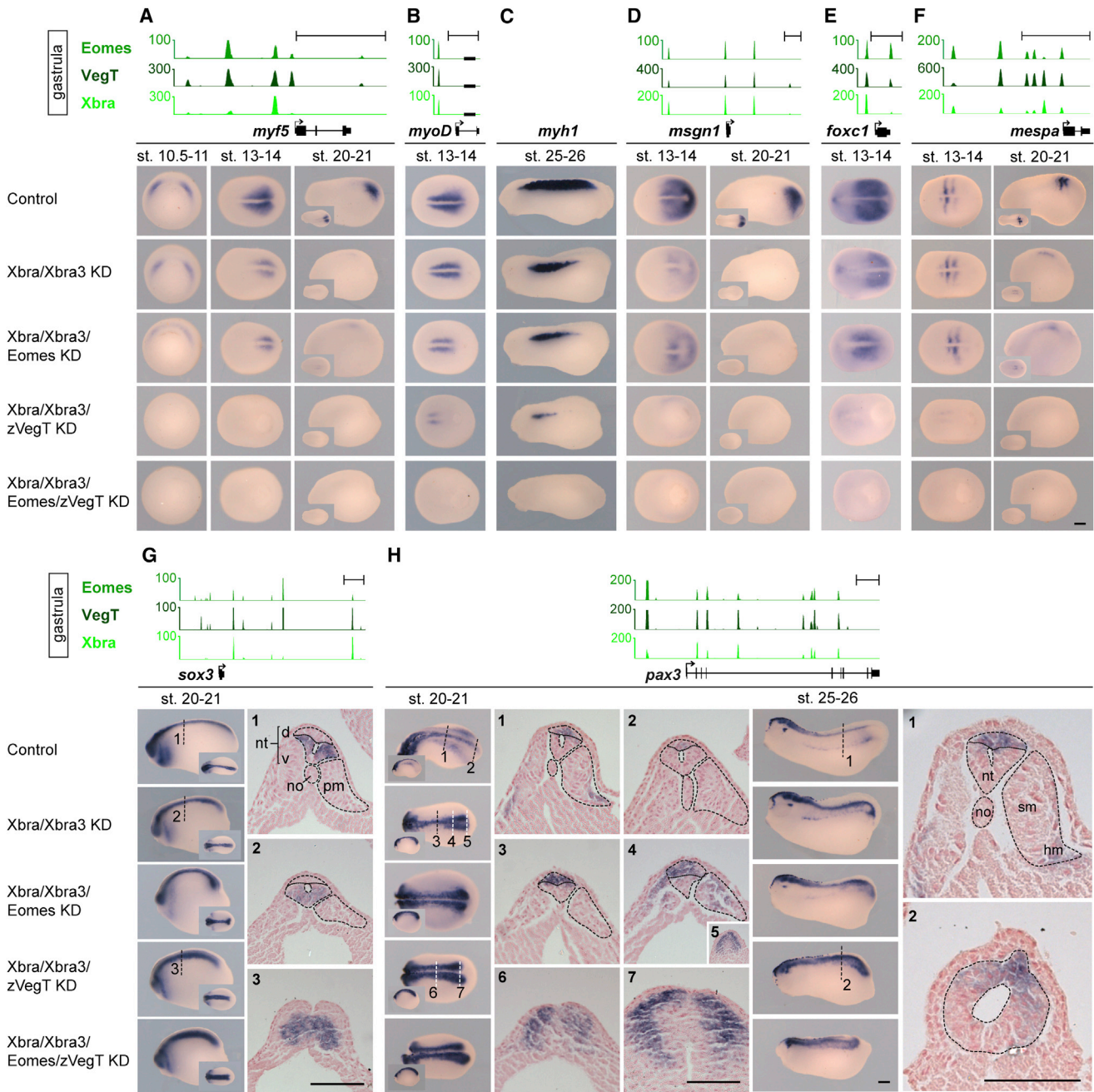
Figure 6. Combined Loss of Eomes, VegT, and Xbra Reveals Their Collaboration to Determine Neuromesodermal Bipotency and Promote Mesodermal Differentiation during Axial Elongation

Gastrula-staged snapshots of Eomes, VegT, and Xbra binding near mesoderm-specific genes *myf5*, *myoD*, *msgn1*, *foxc1*, *mespa* (A, B, and D–F), and neurogenic genes *sox3* and *pax3* (G and H). WMISH on control and indicated KD embryos for target genes (A, B, and D–H) and muscle-specific differentiation marker *myosin heavy chain 1*, *myh1* (C). Cross-sections at positions of *sox3* and *pax3* WMISH as indicated. no, notochord; nt, neural tube (d, dorsal; v, ventral); hm, hypaxial muscle; pm, paraxial mesoderm; sm, skeletal muscle. The scale bar represents 0.2 mm. See also Figures S6 and S7.

the absence of any axial or paraxial mesoderm (Figure 6H). This neuromesodermal 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
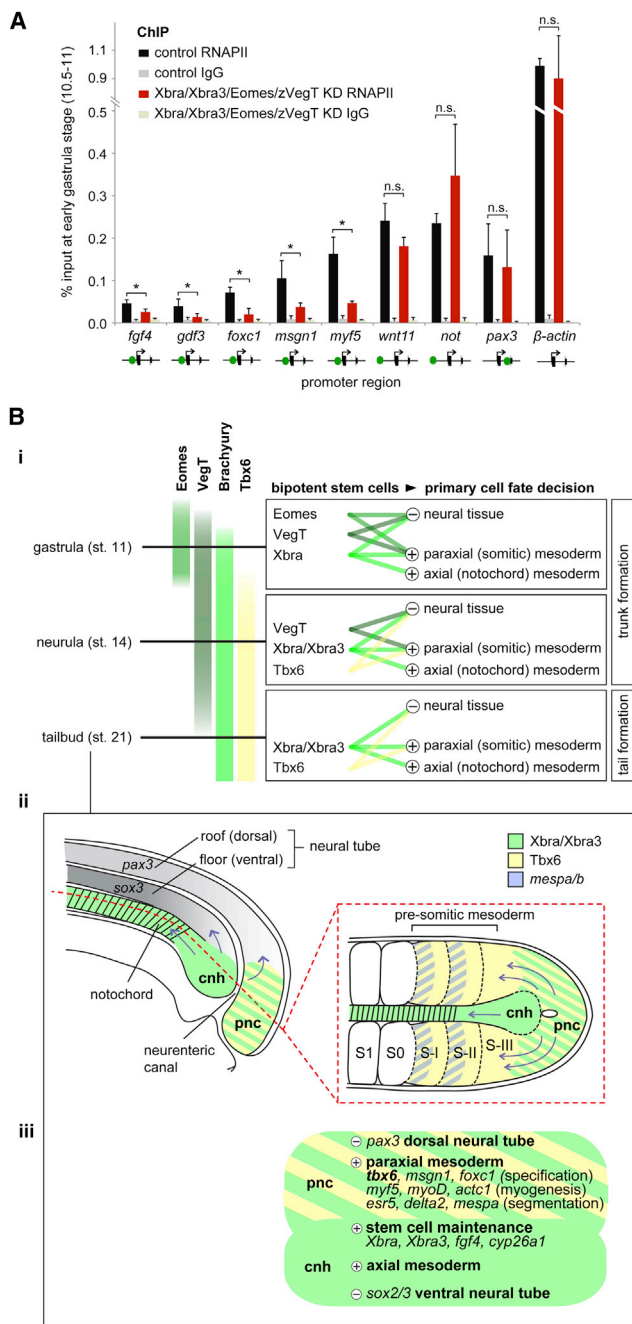
Figure 7. T-box TF-Dependent Recruitment of RNAPII and Model for the Way in which Stage-Dependent Combinations of T-box TFs Define and Instruct Bipotential Stem Cells to Be Recruited to Neural and Mesodermal Tissues and Prime Mesoderm for Differentiation

(A) RNAPII deposition at promoters of mesodermal (*fgf4*, *gdf3*, *foxc1*, *msgn1*, *myf5*, *wnt11*, *not*), neural (*pax3*), and house-keeping (*β-actin*) genes in control and T-box TF-depleted embryos at early gastrula stage (10.5–11) determined by ChIP-qPCR. Proximal or distal (upstream or intronic) binding of T-box TFs to indicated gene promoter is symbolized with green dot. The error bars represent SD of biological triplicates. Two-tailed Student's t test: *p < 0.1; n.s., not significant (p ≥ 0.1). IgG, immunoglobulin G.

(B) Model: (i) The different spatial and temporal patterns of T-box TFs cause neuromesodermal stem cells to be defined by Eomes, VegT, and Xbra during

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 μM 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-En[R] 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.

gastrulation; VegT, Xbra, Xbra3, and Tbx6 during neurulation; and Xbra, Xbra3, and Tbx6 during tail bud stages. These combinations of T-box TFs also ensure that the correct ratio of mesodermal over neural tissue is formed during trunk and tail formation by activating mesoderm-specifying genes and repressing neurogenic genes. The development of axial (notochord) mesoderm depends mainly on Xbra/Xbra3, due to their exclusive expression among these T-box TFs in the chordoneural hinge and developing notochord. Other mesodermal derivatives, such as heart, may similarly depend on combinations of T-box TFs (omitted from model). (ii) Schematic diagram of a sagittal section and a horizontal section (red dashed line) through the posterior region of an early tail bud embryo illustrating the expression of *Xbra/Xbra3*, *Tbx6*, and *mespa/b* and the recruitment of mesodermal and neural cells (blue arrows) from the stem niche (chordoneural hinge and posterior wall of the neurentic canal). Most cells of the chordoneural hinge give rise to the notochord and the ventrolateral horns of the neural tube, whereas cells in the posterior wall of the neurentic canal contribute to paraxial (presomitic) mesoderm and the dorsal roof of the neural tube. cnh, chordoneural hinge; pnc, posterior wall of neurentic canal; S1, first somite; S0, newly forming somite; S-I/II/III, presomitic mesoderm. (iii) Genetic regulatory inputs of T-box TFs in early tail bud embryos with several functional nodes being active in different domains (cnh, pnc) of the tail bud: stem cell maintenance; specification of somitic mesoderm; myogenic differentiation; patterning of presomitic mesoderm; notochord formation; and protection from neuralization.

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 *fgf4* 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 meso-derm-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 tail 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 *actc1*; (4) patterning of presomitic 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., 2013). 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 7Biii).

The way in which T-box TFs suppress transcription of neurogenic genes is not known, although protein phosphorylation (Hwang et al., 2005) and corecruitment 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 presomitic 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 "placeholders" 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–28°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 -I 0 -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.

WMISH

WMISH 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

ACKNOWLEDGMENTS

REFERENCES

Akkers, R.C., van Heeringen, S.J., Jacobi, U.G., Janssen-Megens, E.M., François, 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., Papalopulu, 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., Zaouter, 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 Asashima, M. (2006). Ledgerline, a novel Xenopus laevis gene, regulates differentiation of presomitic mesoderm during somitogenesis. Zoolog. 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 Eomesodermin. 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 derrière is involved in regulation of the establishment of left-right asymmetry. EMBO Rep. *1*, 32–39.

Heinz, S., Benner, C., Spann, 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 Asashima, 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.

Jen, W.C., Gawantka, V., Pollet, N., Niehrs, C., and Kintner, C. (1999). Periodic repression of Notch pathway genes governs the segmentation of Xenopus embryos. Genes Dev. *13*, 1486–1499.

Kawamura, A., Koshida, S., Hijikata, 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 Ripply family of 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 in Xenopus. Int. J. Dev. Biol. *50*, 473–479.

Lagha, M., Bothma, J.P., Esposito, E., Ng, S., Stefanik, L., Tsui, C., Johnston, J., Chen, K., Gilmour, D.S., Zeitlinger, J., and Levine, M.S. (2013). Paused Pol II coordinates tissue morphogenesis in the Drosophila embryo. Cell *153*, 976–987.

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 α 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., Lovén, 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 signaling. 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., Stergachis, 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.

Nieber, 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.

Nieuwkoop, P.D., and Faber, J. (1994). Normal table of Xenopus laevis (New York: Garland Publishing).

Olivera-Martinez, I., Harada, H., Halley, P.A., and Storey, K.G. (2012). Loss of FGF-dependent mesoderm identity and rise of endogenous retinoid signalling determine cessation of body axis elongation. PLoS Biol. *10*, e1001415.

Schulte-Merker, S., and Smith, J.C. (1995). Mesoderm formation in response to Brachyury requires FGF signalling. Curr. Biol. *5*, 62–67.

Sharov, A.A., and Ko, M.S. (2009). Exhaustive search for over-represented DNA sequence motifs with CisFinder. DNA Res. *16*, 261–273.

Showell, C., Binder, O., and Conlon, F.L. (2004). T-box genes in early embryogenesis. Dev. Dyn. *229*, 201–218.

Sparrow, D.B., Jen, W.C., Kotecha, S., Towers, N., Kintner, C., and Mohun, T.J. (1998). Thylacine 1 is expressed segmentally within the paraxial mesoderm of the Xenopus embryo and interacts with the Notch pathway. Development *125*, 2041–2051.

Tada, M., and Smith, J.C. (2000). Xwnt11 is a target of Xenopus Brachyury: regulation of gastrulation movements via Dishevelled, but not through the canonical Wnt pathway. Development *127*, 2227–2238.

Tada, M., O'Reilly, M.A., and Smith, J.C. (1997). Analysis of competence and of Brachyury autoinduction by use of hormone-inducible Xbra. Development *124*, 2225–2234.

Teo, A.K., Arnold, S.J., Trotter, M.W., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., and Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. Genes Dev. *25*, 238–250.

Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., and Nicolas, J.F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev. Cell *17*, 365–376.

White, R.J., Sun, B.I., Sive, H.L., and Smith, J.C. (2002). Direct and indirect regulation of derrière, a Xenopus mesoderm-inducing factor, by VegT. Development *129*, 4867–4876.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Wilm, B., James, R.G., Schultheiss, T.M., and Hogan, B.L. (2004). The forkhead genes, Foxc1 and Foxc2, regulate paraxial versus intermediate mesoderm cell fate. Dev. Biol. *271*, 176–189.

Wittler, L., Shin, E.H., Grote, P., Kispert, A., Beckers, A., Gossler, A., Werber, M., and Herrmann, B.G. (2007). Expression of Msgn1 in the presomitic mesoderm is controlled by synergism of WNT signalling and Tbx6. EMBO Rep. *8*, 784–789.

Yoon, J.K., and Wold, B. (2000). The bHLH regulator pMesogenin1 is required for maturation and segmentation of paraxial mesoderm. Genes Dev. *14*, 3204–3214.

Yoon, S.J., Wills, A.E., Chuong, E., Gupta, R., and Baker, J.C. (2011). HEB and E2A function as SMAD/FOXH1 cofactors. Genes Dev. *25*, 1654–1661.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.