A gene regulatory network directed by zebrafish No tail accounts for its roles in mesoderm formation

Rosalind H. Morley[a,1], Kim Lachani[b,1], Damian Keefe[c], Michael J. Gilchrist[b], Paul Flicek[c], James C. Smith[b,d], and Fiona C. Wardle[a,2]

[a]Department of Physiology, Development and Neuroscience, Cambridge University, Anatomy Building, Downing Street, Cambridge, CB2 3DY, United Kingdom; [b]Wellcome Trust/Cancer Research UK Gurdon Institute; and [d]Department of Zoology, Tennis Court Road, Cambridge University, Cambridge, CB2 1QN, United Kingdom; and [c]European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom

Edited by Michael S. Levine, University of California, Berkeley, CA, and approved January 8, 2009 (received for review August 24, 2008)

Using chromatin immunoprecipitation combined with genomic microarrays we have identified targets of No tail (Ntl), a zebrafish Brachyury ortholog that plays a central role in mesoderm formation. We show that Ntl regulates a downstream network of other transcription factors and identify an in vivo Ntl binding site that resembles the consensus T-box binding site (TBS) previously identified by in vitro studies. We show that the notochord-expressed gene *floating head* (*flh*) is a direct transcriptional target of Ntl and that a combination of TBSs in the *flh* upstream region are required for Ntl-directed expression. Using our genome-scale data we have assembled a preliminary gene regulatory network that begins to describe mesoderm formation and patterning in the early zebrafish embryo.

brachyury | chromatin immunopreciptitation | microarray

Embryonic development proceeds through a series of progressively more restricted cell states in which sets of state-specific genes are expressed in a finely controlled temporal and spatial order. This coordination of gene expression is brought about by the integration of signaling inputs and binding of transcription factors at *cis*-regulatory modules (CRMs) associated with target genes, and can be described by a gene regulatory network (GRN). Such GRNs have been useful in understanding how development proceeds in the early lineages of sea urchins and ascidians (1, 2). However, although GRNs using data from single gene studies have been used to describe aspects of *Xenopus* mesendoderm and lamprey neural crest development (3, 4) a systematic approach to building a GRN by genome-scale analysis has not yet been used to describe early cell fate commitment in developing vertebrate embryos. Understanding how transcriptional regulation drives cell fate commitment in vertebrates is essential not only in understanding their development, but also for informing future efforts to recapitulate cell restriction to different tissue lineages for stem cell-based replacement therapies (5). We have thus set out to assemble a GRN that can describe vertebrate mesoderm development using zebrafish as a model system to identify targets of a key transcriptional regulator, No tail (Ntl).

Ntl is a zebrafish ortholog of *Brachyury*, a T-domain transcription factor that is expressed as an early response to mesoderm induction and plays a central role in mesoderm development in all vertebrates. For instance, studies in mice, *Xenopus*, and zebrafish reveal that *Brachyury* orthologs influence many aspects of mesoderm specification and patterning, being required for formation of the notochord and posterior somites, for normal cell movements during gastrulation and tail outgrowth, and for establishment of left-right asymmetry (reviewed in refs. 6, 7).

Here, we use chromatin immunoprecipitation combined with zebrafish genomic microarrays (ChIP-chip) to survey binding of Ntl at promoter regions. We show that Ntl binds the promoters of transcription factors implicated in posterior identity, muscle specification, cell movements, and notochord development, confirming a role for Ntl as a key transcriptional regulator of mesodermal cell fate and behavior. Using computational methods, we discover an in vivo binding site for Ntl, which resembles the conserved T-box binding site: TCACACCT. We present evidence that Ntl directly regulates expression of *flh*, a transcription factor required for notochord development, through these binding sites. By integrating these data from our genome-scale screen with other studies we have assembled a GRN that begins to describe mesoderm development in zebrafish on a global scale.

Results and Discussion

Ntl Binds Mesodermally Expressed Genes Involved in Transcriptional Regulation of Embryonic Development. To identify direct targets of Ntl responsible for mediating its activity in mesoderm during gastrulation, we performed ChIP-chip on mid-gastrula stage zebrafish embryos [75−85% epiboly; 8–8.5 hours postfertilization (hpf)]. We used a polyclonal antibody that specifically recognizes Ntl, but not its coortholog Bra or other T-domain proteins, and genomic microarrays representing 12 kb around the transcription start site (TSS) of ≈11,000 genes (refs. 8, 9; *SI Text* and Fig. S2*A*). This approach identified bound genomic regions surrounding 218 protein-coding genes, represented by 399 enriched probes (Table S1 and Table S2). A gene was annotated as a target if an enriched probe fell within −9 kb and +3 kb (12 kb) of the TSS. This approach successfully identified all known direct targets of Ntl (*dld*, *tbx6*, and *wnt8*) except *wnt3a*, which was not represented on our microarrays (8, 52). Further information on validation of our approach can be found in *SI Text*.

To assess the roles of these targets, we first classified them by gene ontology terms and identified those terms that were significantly overrepresented in our data set. As might be expected for a developmentally expressed factor, many of the most significantly enriched categories for biological process were those associated with embryonic development (see *SI*). When considering molecular function, the most significantly enriched term was transcription factor activity (GO:0003700; $P < 7.3 \times 10^{-14}$; 4.8-fold enriched; Fig. 1*A*). At least 58 (27%) of the genes in our set encode proteins with transcription factor activity, including well-known developmental genes such as hox, forkhead, and T-box genes (Table S3), suggesting a large aspect of

Author contributions: J.S. and F.C.W. designed research; R.H.M., K.L., and F.C.W. performed research; D.K., M.J.G., P.F., and F.C.W. analyzed data; and D.K., P.F., J.S., and F.C.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE12331).

[1]These authors contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: fcw27@cam.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808382106/DCSupplemental.

DEVELOPMENTAL BIOLOGY

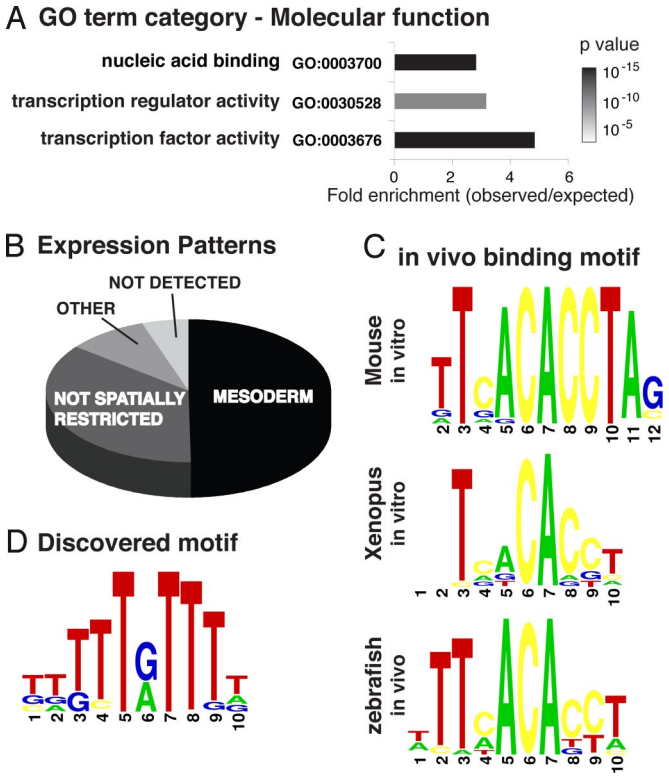

Fig. 1. Ntl regulates developmental transcription factors and binds the conserved T-box binding site. (*A*) GO term analysis shows that Ntl targets are primarily transcription factors. Bars show fold enrichment of the term compared to all annotated genes (observed frequency/expected frequency) and shading indicates the significance of this enrichment based on the calculated *P*-value. Classes with *P*-values $<10^{-4}$ are shown. (*B*) Annotation of target gene expression patterns shows that the majority are expressed in the mesoderm overlapping *ntl* expression. (*C*) The discovered in vivo binding motif for Ntl (from NestedMica) is the same as the in vitro discovered consensus T-binding site for *Xenopus* Brachyury and mouse T (JASPAR database; positions are numbered to correspond with Ntl motif). (*D*) Another enriched motif, resembling a forkhead binding motif, was also discovered in CRMs bound by Ntl.

Ntl's function is to regulate other transcription factors during development.

Brachyury is an activator of transcription (10, 11). We therefore expected genes that are transcriptionally regulated by Ntl to be expressed in the mesoderm, within or overlapping the Ntl expression domain. Our own in situ analyses and a database search for target expression patterns from mid gastrula to early somite stages revealed that, of the 150 genes with known expression patterns, 50% are mesodermally expressed (Fig. 1*B*). At least 11 of these genes have been previously reported to be downregulated in *ntl* mutant embryos (8, 12–21) and we highlight below several other genes whose expression is genetically regulated by Ntl.

However, 5% of genes are not detected (ND) at these stages, 9% have expression in nonmesodermal tissues and 36% of genes are reported as not spatially restricted (NSR). These targets could be false positives either because of the region being called bound when in reality it is not or because enrichment is associated with regulation of a gene elsewhere in the genome rather than the annotated gene. Alternatively, those genes that are not detected or are expressed in other tissues may be negatively regulated by Ntl, although expression of the ectodermal genes *sox19a*, *gata2*, and *foxn4* did not change in *ntl* morphant embryos (data not shown), suggesting that this is not the case. Another explanation for NSR targets is that Ntl regulates these in the mesoderm although other factors are responsible for their regulation in other germ layers. Finally, where a probe falls in a region between 2 head-to-head genes, both will be annotated as targets although only 1 may be regulated by that binding. Indeed we find that 64 of our targets (32 pairs) have enriched probes falling between them and in 70% of cases where the expression of those genes is characterized, 1 gene is mesodermally expressed while the other is NSR, ND, or expressed in another tissue. This raises the possibility that Ntl regulates the mesodermally expressed gene but not the other.

An in Vivo Binding-Site Motif for Ntl. In vitro binding-site selection assays have shown that T-domain proteins recognize the same consensus T-box binding site (TBS; reviewed in ref. 22). To take advantage of having large-scale in vivo binding data for a T-box factor, we used computational methods to discover an enriched motif associated with Ntl binding in the *cis*-regulatory modules (CRMs) of target genes (23, 24). A CRM was defined as a 500-bp sequence flanking an enriched probe (or the most highly enriched probe in a group), resulting in 233 CRMs being identified. These methods discovered the consensus TBS motif (TCACACCT), indicating we have not only identified bona fide targets of Ntl but that the zebrafish protein, like its *Xenopus* and mouse counterparts, binds the previously identified TBS (Fig. 1*C*). We found that 228 CRMs contain a TBS (bit score >0), whilst 119 contain a TBS with more constrained character in which the initial T and central CAC of the motif are invariant (ref. 25; Table S4; Table S5). These 4 bases have been shown to contact amino acids in *Xenopus* Brachyury (26) and because these amino acids are also conserved in Ntl we reasoned they would be important for binding. It should also be noted, however, that putative TBSs can be found in regions where Ntl binding is not detected (see Figs. 2, 3 and Fig. S1).

Using an approach that compares phastCons conservation scores available from the University of California Santa Cruz genome browser, we were able to detect conservation of TBSs between zebrafish and other vertebrate species in a subset of CRMs. The conservation scores are based on a probabilistic model of sequence evolution estimated from multispecies genome alignments (27, 28). We found that for those 37 CRMs with phastCons data available for both the CRM and the motif site, 43% had significantly higher conservation scores over at least 1 of the motif sites in the CRM using a significance threshold of $P < 0.05$ (Table S4). In total, 20 motifs were found to be significantly conserved. In particular, we saw TBS conservation in *cdx4* ($P < 0.0005$); *blf* ($P < 10^{-23}$); *foxd3* ($P < 10^{-20}$); *wnt8a* ($P < 10^{-21}$); and *fgf24* ($P < 10^{-6}$), which we highlight in the GRN below.

T-box factors are known to associate with other transcription factors to regulate gene expression (e.g., ref. 29), so we also searched the CRMs for other enriched motifs that might indicate combinatorial regulation of targets. This identified a motif containing a core consensus sequence of TGTTT (Fig. 1*D*) which resembles the core consensus found in forkhead binding site motifs, suggesting Ntl may interact with a forkhead-related protein in the embryo. There are at least 8 forkhead-related proteins present in the mesoderm of the early zebrafish embryo and investigations are currently underway to identify proteins that may interact with this motif.

Floating Head Is a Direct Target of Ntl. A striking consequence of the *ntl* homozygous mutation is lack of differentiated notochord, deriving from a cell-autonomous requirement for *ntl* in notochord precursor cells (30, 31). To date, however, no direct target of Ntl that mediates this activity has been described. Notochord also fails to develop in *floating head* (*flh*) mutants, which carry a mutation in a homeobox transcription factor related to *Not* (32). Because Ntl acts upstream of *flh*, it may act through *flh* to

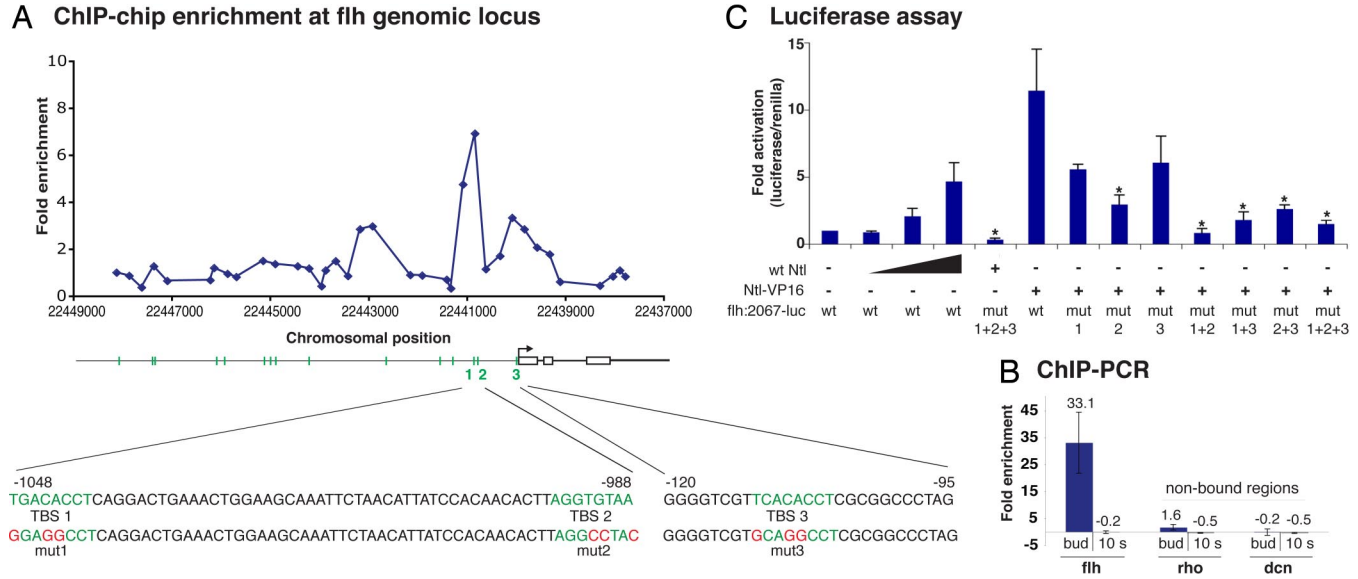

Fig. 2. Direct regulation of *flh* expression by Ntl. (*A*) Ntl binding at the *flh* locus. Plot shows ChIP-enrichment ratio for microarray probes in the *flh* genomic region. Chromosomal position, transcription start site (TSS), intron-exon structure, and putative upstream T-binding sites (TBS) with constrained character are shown below the graph (green lines; see text). Also shown is the sequence surrounding wild-type TBS 1, 2, and 3 (green) and the mutated sequences (red). (*B*) ChIP-PCR on bud (10 hpf) and 10 somite (14 hpf) embryos using primers that amplify TBS 1 + 2 show that *flh* is bound by Ntl at bud stage but not at 10 somite stage. Enrichment of negative genomic regions around *rho* and *dcn* are shown for comparison. (*C*) A 2-kb upstream region of the *flh* locus drives luciferase expression in 75% epiboly-stage embryos in the presence of wild-type Ntl or Ntl-VP16. Increasing doses (150 pg, 375 pg, 750 pg) of Ntl mRNA activate the 2-kb luciferase construct in a dose-dependent manner. Activation in the presence of 150 pg Ntl-VP16 mRNA is decreased when TBS 1, 2, or 3 are mutated individually or in combination. Mutation in TBS 2 alone or combinations of mutations in 2 or 3 TBSs result in a significant reduction of activity (indicated by *, $P < 0.05$). Mutation of all 3 TBSs also results in a significant decrease in luciferase activation when coinjected with 750 pg wild-type Ntl mRNA.

specify notochord fate (19, 33). In support of this, we found that Ntl binds at several sites in the upstream region of the *flh* locus (Fig. 2*A*) with strongest binding seen at approximately −1 kb. Inspection of the genomic sequence revealed 2 putative TBS arranged tail-to-tail spaced 45 bp apart in this region (Fig. 2*A*; designated TBS 1 and 2). In addition, a putative TBS is located at −112 bp (TBS 3). Using in situ hybridization on mutant embryos, we confirmed previous reports that initial *flh* expression is independent of Ntl activity but that maintenance of expression is dependent on Ntl during late gastrula and early somite stages (data not shown and Table S6). ChIP-PCR also revealed binding of Ntl over TBS1 + 2 at the end of gastrulation (bud stage, 10 hpf) but not at later stages (10 somite stage, 14 hpf; Fig. 2*B*).

To confirm that *flh* is a direct target of Ntl, we asked whether the upstream region of *flh* is regulated by Ntl activity. A 2-kb region upstream of *flh* was used to drive expression of luciferase in embryos (*flh:2067bp-luc*). Injection of this reporter into embryos resulted in low levels of luciferase activity and we did not detect any significant reduction in this activation when coinjected with Ntl MO or mRNA encoding an Ntl-Engrailed repressor construct, presumably because levels of activation are already low (data not shown). Coinjection of this reporter with wild-type Ntl mRNA showed a dose-dependent increase in luciferase activation at mid-late gastrula stages, as did an activated version of Ntl, Ntl-VP16 (ref. 34; Fig. 2*C*). Because coinjection of the VP16 construct mRNA led to a more robust induction of luciferase activity, we used this in conjunction with reporter constructs to test the effect of mutations in the 3 putative TBSs in this region, either individually or in combination. We observed that coinjection of these mutated constructs with Ntl-VP16 mRNA led to reduced luciferase activation (Fig. 2*C*). Mutation of TBS 2 caused the greatest reduction and mutation of any 2 or more of these TBSs also caused a significant reduction in activity. Similarly, mutation of all 3 sites together resulted in a significant reduction of activity in embryos coinjected with wild-type Ntl mRNA (Fig. 2*C*). These results show that a combination of sites, particularly TBS 2, play an important role in Ntl binding and activation of the *flh* promoter region. Shorter genomic regions (655 bp containing TBS 1 + 2 or 968 bp containing TBS 3 plus the first exon of *flh*) did not drive luciferase expression in embryos, even in the presence of Ntl-VP16 (data not shown), indicating these shorter regions are not sufficient for expression and confirming a combination of sites in the larger 2-kb region are necessary for expression. We also tested the ability of this 2-kb region to drive notochord expression in the whole embryo, but did not observe localized expression (data not shown), suggesting that other regulatory regions in combination with the region we have identified are required to restrict *flh* expression to the notochord.

A Mesodermal GRN Directed by Ntl. The above results suggested that *flh* may mediate the activity of Ntl in notochord formation, so we next asked whether other target genes identified in our screen could mediate the other known activities of Ntl in posterior somite formation, gastrulation movements, and left-right asymmetry. Below, we describe classes of target genes that Ntl binds and genetically regulates and which are involved in these processes (Fig. 3 and Fig. S1). By integrating these data with other published studies we have assembled a preliminary GRN that describes the activities of Ntl in mesoderm formation and how it interacts with several signaling pathways and factors in mesoderm (Fig. 4).

Posterior Fate. Ntl acts in combination with other T-box factors and FGF and Wnt signaling pathways to specify posterior fate, such as trunk and tail somites. Consistent with this known role, our data show that Ntl binds to the upstream regions of *tbx6*, *tbx16*, and many components of both the FGF and Wnt pathways including *wnt8*, a direct target of Ntl (ref. 8; Figs. 3 and 4). Interestingly, we saw binding of Ntl to negative (e.g., *gro2*, *tob1a*,

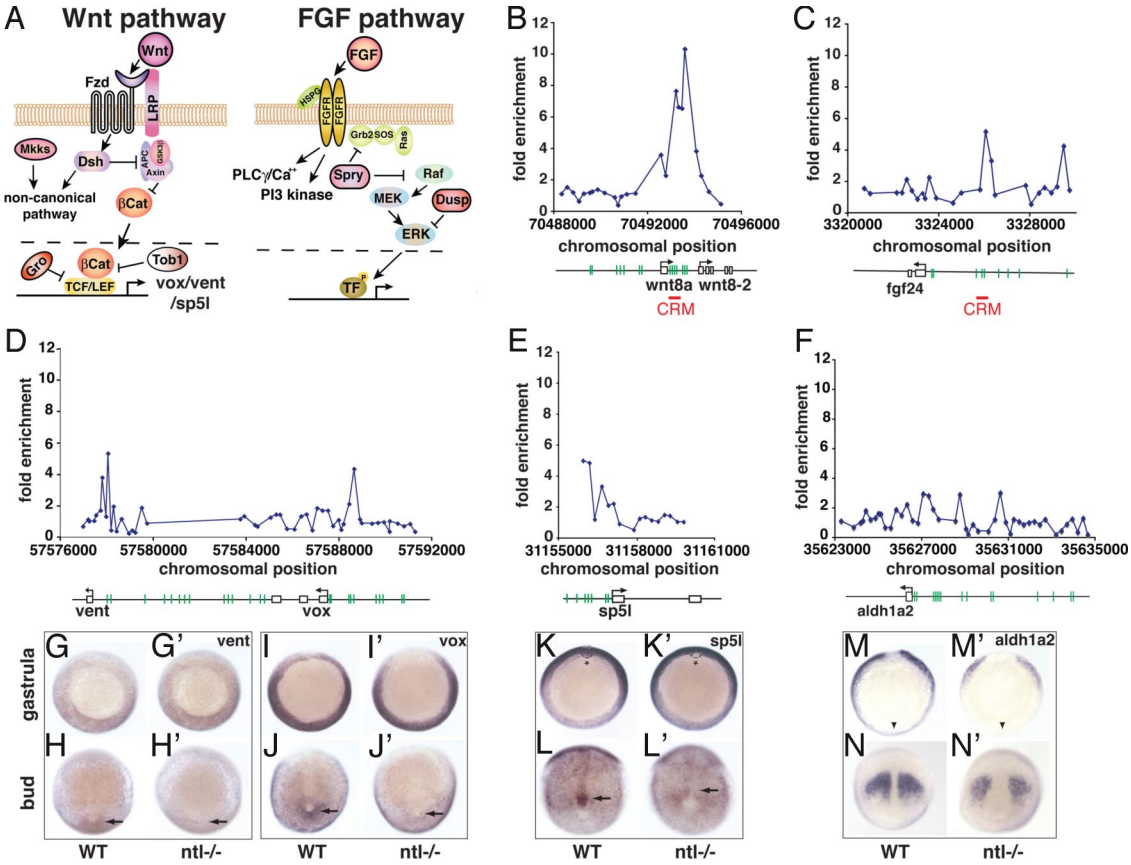

Fig. 3. Ntl binds and regulates the expression of genes involved in posterior identity. (*A*) Diagrams to show fgf and wnt signaling pathways. Components outlined in black were identified as Ntl targets by ChIP-chip. (*B–F*) Ntl binding in genomic regions around *wnt8*, *fgf24*, *vent*, *vox*, *sp5l*, and *aldh1a2*. Plots show ChIP-enrichment ratio for microarray probes in the genomic regions. Chromosomal position, TSS, intron-exon structure, putative upstream TBSs with constrained character (green lines; see text), and conserved CRMs (red bar; see text) are shown below the graphs. (*G–N′*) In situ hybridization of *ntl* mutants compared to wild types for *vent*, *vox*, *sp5l*, *and aldh1a2* (posterior views, ventral down). Expression of *vent*, *vox*, and *sp5l* shows downregulation at bud stages (arrows). In addition expression of *sp5l* is absent from the dorsal forerunner cells in *ntl* mutant embryos (asterisk; DFCs outlined by dashed line). *aldh1a2* expression in the ventral margin is downregulated at gastrula stages.

spry4, *dusp1*, *dusp6*) and positive effectors (e.g., *wnt8*, *fgf24*, *fgfr4*) of these signaling pathways (Table S2; Fig. 3*A*), raising the possibility that Ntl regulates the balance of FGF and Wnt signaling in the posterior of the embryo. This may be especially important because Ntl is involved in a feedback loop with Wnt and FGF signaling (e.g., ref. 8). Our data also revealed that Ntl acts in several feed-forward and feedback loops (Fig. 4). For instance Ntl binds and regulates the expression of *wnt8* target genes *vox*, *vent*, and *sp5l* and *wnt8* itself (Fig. 3). In situ hybridization of *ntl* mutant embryos showed that the posterior expression of *vox*, *vent*, and axial expression of *sp5l* is absent or reduced by bud stage (Fig. 3 *G*−*L′*). *sp5l* is also a target of FGF signaling and itself regulates *ntl* expression (35). We have also included in the GRN other transcription factors that are involved in posterior patterning and are bound and regulated by Ntl and/or Wnt signaling: *cdx1a*, *cdx4*, and *eve1* and *fgf8*, which works in combination with *fgf24* in the posterior of the embryo (13, 36). Tbx16 also regulates the expression of some of these targets, revealing another feed-forward loop, and we have indicated these interactions in Fig. 4 (12, 13, 15, 16, 20, 52).

Retinoic acid (RA) signaling is also implicated in posterior formation of the embryo and we see binding to the upstream region of *aldh1a2* (Fig. 3*F*). In situ hybridization analysis of *aldh1a2* expression shows that at mid-gastrula stages expression is decreased in the ventral margin in mutant embryos (Fig. 3 *M* and *M′*; Table S6).

We were also struck by the number of Ntl targets present in the Notch pathway that are expressed in the margin during gastrulation and tailbud at somite stages. The Notch pathway is essential for tailbud outgrowth, for differentiating between the hypochord and notochord, and for correct segmentation of the paraxial mesoderm into somites (37, 38). *dld* is a known direct target of Ntl (52), and we also see binding to the upstream regions of *dlc*, *notch3*, *her1*, *her11/5*, and *her12* (data not shown). The Notch pathway also regulates itself through negative feedback, and we have indicated these interactions (Fig. 4). However, we did not see any changes in gene expression for these targets during gastrulation or early somite stages, indicating that either they are not direct targets of Ntl or that Ntl is regulating the activity of these genes in a redundant fashion with other factors, such as Tbx16 (12).

Muscle Cell Fate. We also saw binding to the upstream regions of genes that specify muscle cell fate including *myod*, *msgn1*, and *foxd3* (Fig. 4 and Fig. S1 D–F). *foxd3* in turn directly regulates *myf5* expression and is itself regulated by *pax3* (39). In situ hybridization analysis showed expression of *foxd3* in the axis and margin of the embryo at gastrula stages was weak in *ntl* mutant embryos compared to wild type (Fig. S1 *B* and *B′*). By bud stage, *foxd3* expression was also seen in the neural crest region in both mutant and wild-type embryos, although the posterior expression was absent in mutant embryos (Fig. S1 *C* and *C′*; Table S6). *msgn1* and *myod* expression fail to initiate in *ntl* mutant embryos, although expression levels recover during gastrulation or somi-

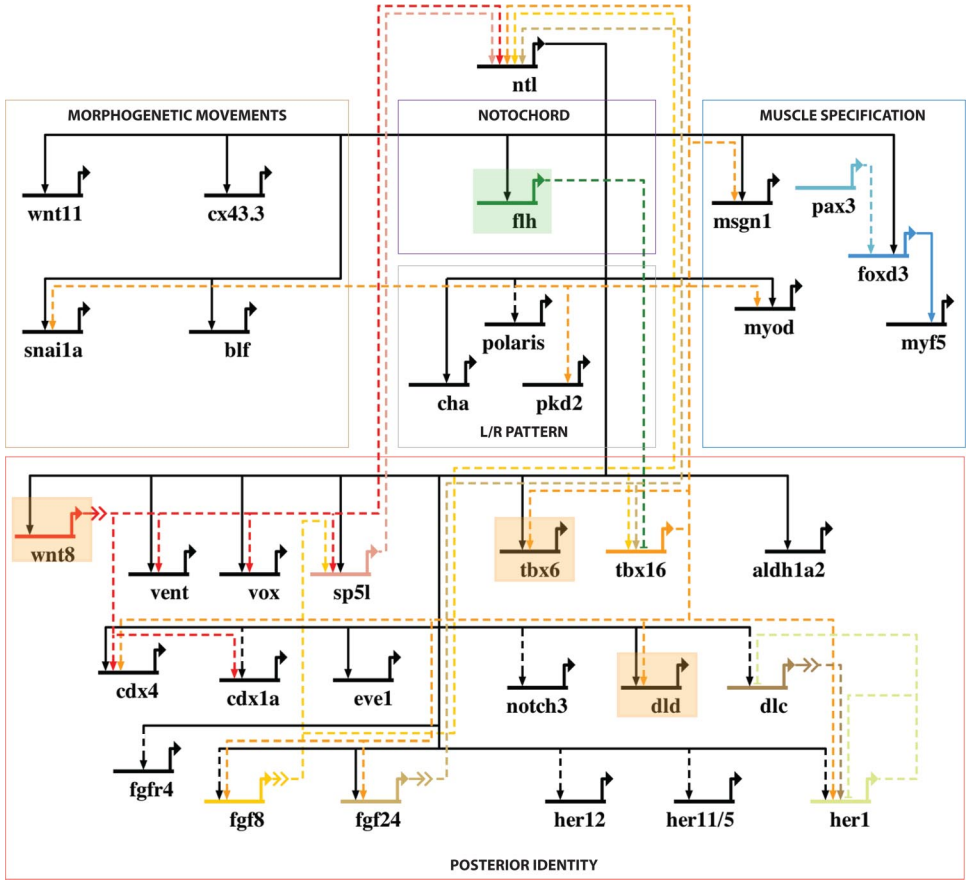

Fig. 4. Gene regulatory network for mesoderm specification and patterning in zebrafish. Ntl directs a network of transcription factors, signals, and differentiation genes. Genes are grouped together to represent different aspects of Ntl activity in the mesoderm: morphogenetic movements, notochord specification, muscle specification, posterior identity, and left/right patterning. Double arrowheads indicate an input through intracellular signaling. Solid lines indicate binding of target promoter *and* genetic regulation, while dashed lines indicate genetic regulation of target *or* binding of target promoter has been shown. Boxed genes indicate that additional assays have shown direct regulation of the target by Ntl.

togenesis, respectively (15, 21). Regulation of these myogenic factors provides a mechanism by which Ntl can influence muscle cell fate in the posterior of the embryo in addition to a general posterior identity.

Gastrulation Movements. Another important activity of Ntl, and other Brachyury orthologs, is to regulate cell movements during gastrulation and tailbud outgrowth (40, 41). Components of the planar cell polarity Wnt pathway, including *wnt11*, have been implicated in this activity (42, 43). Accordingly, we saw binding of Ntl to the upstream region of *wnt11* and several other genes implicated in cell movements including *snai1a*, *blf*, *cx43.3*, and *tbx16* (Fig. 4 and Fig. S1 *H–M*). In situ hybridization analysis revealed that *blf* is genetically regulated by Ntl, because its expression in the lateral margin and paraxial mesoderm was considerably reduced in mutant embryos compared to wild type (Fig. S1 *I–J′*; Table S6). Similarly, *snai1a*, *cx43.3*, and *tbx16* are also genetically regulated by Ntl (14, 16, 18, 20).

Left-Right Patterning. Ntl is required in the dorsal forerunner cells (DFCs) for formation of Kupffer's vesicle, the organ of asymmetry in zebrafish. Tbx16 is also required in these cells for normal left-right patterning and genetic regulation of *pkd2* expression (44). However, other than *tbx16*, our screen did not identify any other left-right patterning genes, including *cha* and *polaris*, which are genetic targets of Ntl (refs. 44, 45; Fig. 4). This may be because of the small number of cells in which this interaction is present and which could not be detected by our approach, or because these are not in fact direct targets of Ntl. However, we identified at least 6 target genes that are expressed in the DFCs in addition to other mesodermal tissues, including *sp5l* whose expression in DFCs is absent in *ntl* mutant embryos (Fig. 3 *K* and *K′*).

Previous screens have identified targets of Brachyury in sea urchins and ascidians (46, 47) and it is notable that we see almost no target overlap with these screens. This may be because both screens were based on subtractive hybridizations and may have identified targets that are several steps removed from early Ntl regulation. For instance, sea urchin *bra* is involved in cell movements during gastrulation, and we might expect to see some overlap in targets. However, the sea urchin screen predominantly identified structural proteins and enzymes rather than regulatory genes such as transcription factors, suggesting that a later stage of regulation was being assayed. Similarly in ascidians, where *Ci-bra* is involved in notochord formation, the screen identified many structural proteins such as collagens, and very few transcription factors.

We also compared our GRN to one created for *Xenopus* mesendoderm formation (3). During gastrulation, Xbra directly regulates expression of its targets, *Xwnt11* and *fgf4*, which are involved in gastrulation and posterior patterning in a similar way to zebrafish *wnt11* and *fgf24*. However, 2 other Xbra targets, *Xegr1* and the *bix* genes, were not found in our screen. This may be because these genes act differently in zebrafish compared to *Xenopus*, because neither has direct sequence or functional orthologs. Notwithstanding this observation, much of the core

regulation of mesoderm formation is likely to be similar because many signaling pathways and transcription factors, especially the activity of Brachyury, are conserved across vertebrates. Further studies into the targets of other transcription factors and signaling pathways are needed to assess fully the degree of overlap between regulation of mesoderm formation in *Xenopus*, zebrafish, and other vertebrates.

To summarize, we have shown how Ntl can orchestrate aspects of mesoderm development through regulating a network of transcription factors and signaling pathways. The GRN we describe here provides the basis for building a larger network that can be used to further understand mesoderm specification and patterning. For instance, Ntl regulates mesoderm development in combination with other factors, including T-box factors such as Spt, Tbx6, and Bra, and a future goal is to identify direct targets of these factors and place them in the mesodermal GRN.

Materials and Methods

Detailed materials and methods can be found in *SI*. ChIP-chip assays and analysis were performed as described (in ref. 48). GO term analysis was performed using GOToolBox (49). Motif enrichment analysis was performed using NestedMICA and Trawler (23, 24). For luciferase assays, 40 pg luciferase construct + 0.75 pg pCS2 + Renilla were injected in 1-cell embryos with 150, 450, or 750 pg Ntl mRNA or 150 pg Ntl-VP16 mRNA where indicated. Embryos were processed using Dual Luciferase Assay kit (Promega) according to manufacturer's instructions. In situ hybridization was carried out by standard methods (50). Primer sequences are given in Table S7 GRN was created using Biotapestry software (51).

ACKNOWLEDGMENTS. We are grateful to Stefan Schulte-Merker for his generous gift of anti-Ntl antibody. We thank David Kimelman for *Bra* and Ashley Bruce for *Ntl-VP16* and *Ntl-EnR*. We acknowledge funding from a Medical Research Council career development award and Lister Institute research prize (F.C.W.), a Wellcome Trust program grant and the EU program European Transcriptome, Regulome and Cellular Commitment Consortium (J.C.S.).

1. Imai KS, Levine M, Satoh N, Satou Y (2006) Regulatory blueprint for a chordate embryo. *Science* 312:1183–1187.
2. Oliveri P, Tu Q, Davidson EH (2008) Global regulatory logic for specification of an embryonic cell lineage. *Proc Natl Acad Sci USA* 105:5955–5962.
3. Loose M, Patient R (2004) A genetic regulatory network for Xenopus mesendoderm formation. *Dev Biol* 271:467–478.
4. Nikitina NV, Bronner-Fraser M (2008) Gene regulatory networks that control the specification of neural-crest cells in the lamprey. *Biochim Biophys Acta*, 10.1016/j.bbagrm.2008.03.006.
5. Smith JC, Wardle FC, Loose M, Stanley E, Patient R (2007) in *Current Protocols in Stem Cell Biology*, eds Bhatia M, Eggan K, Elefanty A, Fisher S, Patient R, Pera M (John Wiley & Sons, New York).
6. Wardle FC, Papaioannou VE (2008) Teasing out T-box targets in early mesoderm. *Curr Opin Genet Dev* 18:418–425.
7. Showell C, Binder O, Conlon FL (2004) T-box genes in early embryogenesis. *Dev Dyn* 229:201–218.
8. Martin BL, Kimelman D (2008) Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. *Dev Cell* 15:121–133.
9. Schulte-Merker S, Ho RK, Herrmann BG, Nusslein-Volhard C (1992) The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development* 116:1021–1032.
10. Conlon FL, Sedgwick SG, Weston KM, Smith JC (1996) Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. *Development* 122:2427–2435.
11. Kispert A, Koschorz B, Herrmann BG (1995) The T protein encoded by Brachyury is a tissue-specific transcription factor. *EMBO J* 14:4763–4772.
12. Amacher SL, Draper BW, Summers BR, Kimmel CB (2002) The zebrafish T-box genes no tail and spadetail are required for development of trunk and tail mesoderm and medial floor plate. *Development* 129:3311–3323.
13. Draper BW, Stock DW, Kimmel CB (2003) Zebrafish fgf24 functions with fgf8 to promote posterior mesodermal development. *Development* 130:4639–4654.
14. Essner JJ, Laing JG, Beyer EC, Johnson RG, Hackett PB, Jr (1996) Expression of zebrafish connexin43.4 in the notochord and tail bud of wild-type and mutant no tail embryos. *Dev Biol* 177:449–462.
15. Goering LM, *et al.* (2003) An interacting network of T-box genes directs gene expression and fate in the zebrafish mesoderm. *Proc Natl Acad Sci USA* 100:9410–9415.
16. Griffin KJ, Amacher SL, Kimmel CB, Kimelman D (1998) Molecular identification of spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes. *Development* 125:3379–3388.
17. Joly JS, Joly C, Schulte-Merker S, Boulekbache H, Condamine H (1993) The ventral and posterior expression of the zebrafish homeobox gene eve1 is perturbed in dorsalized and mutant embryos. *Development* 119:1261–1275.
18. Makita R, Mizuno T, Koshida S, Kuroiwa A, Takeda H (1998) Zebrafish wnt11: pattern and regulation of the expression by the yolk cell and No tail activity. *Mech Dev* 71:165–176.
19. Melby AE, Kimelman D, Kimmel CB (1997) Spatial regulation of floating head expression in the developing notochord. *Dev Dyn* 209:156–165.
20. Thisse C, Thisse B, Schilling TF, Postlethwait JH (1993) Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development* 119:1203–1215.
21. Weinberg ES, *et al.* (1996) Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos. *Development* 122:271–280.
22. Tada M, Smith JC (2001) T-targets: clues to understanding the functions of T-box proteins. *Dev Growth Differ* 43:1–11.
23. Down TA, Hubbard TJ (2005) NestedMICA: sensitive inference of over-represented motifs in nucleic acid sequence. *Nucleic Acids Res* 33:1445–1453.
24. Ettwiller L, Paten B, Ramialison M, Birney E, Wittbrodt J (2007) Trawler: de novo regulatory motif discovery pipeline for chromatin immunoprecipitation. *Nat Methods* 4:563–565.
25. Durbin R, Eddy SR, Krogh A, Mitchison G (1998) in *Biological Sequence Analysis.* (Cambridge Univ Press, Cambridge, UK).
26. Muller CW, Herrmann BG (1997) Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. *Nature* 389:884–888.
27. Miller W, *et al.* (2007) 28-way vertebrate alignment and conservation track in the UCSC Genome Browser. *Genome Res* 17:1797–1808.
28. Siepel A, *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 15:1034–1050.
29. Messenger NJ, *et al.* (2005) Functional specificity of the Xenopus T-domain protein brachyury is conferred by its ability to interact with smad1. *Dev Cell* 8:599–610.
30. Halpern ME, Ho RK, Walker C, Kimmel CB (1993) Induction of muscle pioneers and floor plate is distinguished by the zebrafish no tail mutation. *Cell* 75:99–111.
31. Schulte-Merker S, van Eeden FJ, Halpern ME, Kimmel CB, Nusslein-Volhard C (1994) No tail (ntl) is the zebrafish homologue of the mouse T (Brachyury) gene. *Development* 120:1009–1015.
32. Talbot WS, *et al.* (1995) A homeobox gene essential for zebrafish notochord development. *Nature* 378:150–157.
33. Halpern ME, *et al.* (1997) Genetic interactions in zebrafish midline development. *Dev Biol* 187:154–170.
34. Bruce AE, *et al.* (2003) The maternally expressed zebrafish T-box gene eomesodermin regulates organizer formation. *Development* 130:5503–5517.
35. Zhao J, Cao Y, Zhao C, Postlethwait J, Meng A (2003) An SP1-like transcription factor Spr2 acts downstream of Fgf signaling to mediate mesoderm induction. *EMBO J* 22:6078–6088.
36. Shimizu T, Bae YK, Muraoka O, Hibi M (2005) Interaction of Wnt and caudal-related genes in zebrafish posterior body formation. *Dev Biol* 279:125–141.
37. Holley SA (2007) The genetics and embryology of zebrafish metamerism. *Dev Dyn* 236:1422–1449.
38. Latimer AJ, Appel B (2006) Notch signaling regulates midline cell specification and proliferation in zebrafish. *Dev Biol* 298:392–402.
39. Lee HC, Huang HY, Lin CY, Chen YH, Tsai HJ (2006) Foxd3 mediates zebrafish myf5 expression during early somitogenesis. *Dev Biol* 290:359–372.
40. Glickman NS, Kimmel CB, Jones MA, Adams RJ (2003) Shaping the zebrafish notochord. *Development* 130:873–887.
41. Marlow F, Gonzalez EM, Yin C, Rojo C, Solnica-Krezel L (2004) No tail co-operates with non-canonical Wnt signaling to regulate posterior body morphogenesis in zebrafish. *Development* 131:203–216.
42. Heisenberg CP, *et al.* (2000) Silberblick/Wnt11 mediates convergent extension movements during zebrafish gastrulation. *Nature* 405:76–81.
43. Tada M, Smith JC (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.
44. Bisgrove BW, Snarr BS, Emrazian A, Yost HJ (2005) Polaris and Polycystin-2 in dorsal forerunner cells and Kupffer's vesicle are required for specification of the zebrafish left-right axis. *Dev Biol* 287:274–288.
45. Gourronc F, Ahmad N, Nedza N, Eggleston T, Rebagliati M (2007) Nodal activity around Kupffer's vesicle depends on the T-box transcription factors Notail and Spadetail and on Notch signaling. *Dev Dyn* 236:2131–2146.
46. Rast JP, Cameron RA, Poustka AJ, Davidson EH (2002) Brachyury target genes in the early sea urchin embryo isolated by differential macroarray screening. *Dev Biol* 246:191–208.
47. Takahashi H, *et al.* (1999) Brachyury downstream notochord differentiation in the ascidian embryo. *Genes Dev* 13:1519–1523.
48. Wardle FC, *et al.* (2006) Zebrafish promoter microarrays identify actively transcribed embryonic genes. *Genome Biol* 7:R71.
49. Martin D, *et al.* (2004) GOToolBox: functional analysis of gene datasets based on Gene Ontology. *Genome Biol* 5:R101.
50. Nusslien-Volhard C, Dahm R (2002) *Zebrafish* (Oxford Univ Press, Oxford, UK).
51. Longabaugh WJ, Davidson EH, Bolouri H (2005) Computational representation of developmental genetic regulatory networks. *Dev Biol* 283:1–16.
52. Garnett AT, et al. (2009) Identification of direct T-box target genes in the developing zebrafish mesoderm. *Development* 136:749–760.