An ancestral Wnt–Brachyury feedback loop in axial patterning and recruitment of mesoderm-determining target genes

Received: 26 July 2021
Accepted: 12 September 2022
Published online: 17 November 2022

Transcription factors are crucial drivers of cellular differentiation during animal development and often share ancient evolutionary origins. The T-box transcription factor Brachyury plays a pivotal role as an early mesoderm determinant and neural repressor in vertebrates; yet, the ancestral function and key evolutionary transitions of the role of this transcription factor remain obscure. Here, we present a genome-wide target-gene screen using chromatin immunoprecipitation sequencing in the sea anemone Nematostella vectensis, an early branching non-bilaterian, and the sea urchin Strongylocentrotus purpuratus, a representative of the sister lineage of chordates. Our analysis reveals an ancestral gene regulatory feedback loop connecting Brachyury, FoxA and canonical Wnt signalling involved in axial patterning that predates the cnidarian–bilaterian split about 700 million years ago. Surprisingly, we also found that part of the gene regulatory network controlling the fate of neuromesodermal progenitors in vertebrates was already present in the common ancestor of cnidarians and bilaterians. However, while several endodermal and neuronal Brachyury target genes are ancestrally shared, hardly any of the key mesodermal downstream targets in vertebrates are found in the sea anemone or the sea urchin. Our study suggests that a limited number of target genes involved in mesoderm formation were newly acquired in the vertebrate lineage, leading to a dramatic shift in the function of this ancestral developmental regulator.

T-box genes are a class of transcription factors that evolved in the common ancestor of protists and animals and have been found to play important roles in the development of all animals. The founding member, Brachyury, also called T in mice (and more recently Tbx1), is a key determinant of mesoderm in vertebrates. It is expressed in the presumptive mesoderm and continues to be expressed after gastrulation in the notochord, a defining structure of all chordates. Heterozygous mouse mutants of the T gene locus show a shortened tail (hence the Greek name Brachy, meaning short, and ury, meaning tail), while homozygous mutants die in utero, lacking most of the body posterior to the forelimbs. Knockdown of brachyury in Xenopus also impairs the posterior development of the embryo, while overexpression of brachyury messenger RNA in explanted animal caps leads to mesoderm differentiation. Similar roles have been found in other
and demonstrating the conserved role for Brachyury as a pivotal mesoderm regulator across vertebrates.

Brachyury exhibits a dynamic expression pattern in vertebrate embryos. At the gastrula stage it is expressed around the blastopore and during anterior–posterior axis elongation, at the tailbud stage, brachyury expression is mainly restricted to the most caudal part of the embryo and the notochord. This caudal part of the developing embryo contains a population of bipotent cells called neuromesodermal progenitor cells (NMPs) which exhibit temporal differences in gene expression over developmental time and contribute to the spinal cord, paraxial mesoderm and notochord development. Some studies define these NMPs by the expression of brachyury and sox2, a member of the soxB1 gene family, while others argue that brachyury expression along with Wnt/Fibroblast Growth Factor (FGF) signalling is sufficient for the cells to become NMPs even in the absence of Sox2 protein. Regardless of the definition of NMPs, the mesoderm–defining role of Brachyury has been established within these cell populations. The notochord–specific function of Brachyury also appears to be conserved in non-vertebrate chordates. In the ascidian Ciona intestinalis, representing the sister group to the vertebrates, brachyury expression and function is restricted to the notochord and brachyury knockdown abolishes notochord development. In amphioxus, there are two brachyury paralogues. While to date no functional data are available for amphioxus, the expression pattern of bra1 is very similar to that in vertebrates, suggesting that it probably plays a similar role in this cephalochordate, while bra1 appears to be more restricted to the notochord. Thus, Brachyury is key for mesoderm formation and in particular for notochord development in vertebrates and, possibly, in chordates. However, whether Brachyury has a mesodermal function outside chordates is less clear.

In numerous investigated invertebrate species, brachyury is expressed around the blastopore, similar to the early brachyury expression in vertebrates. For instance, in Echinodermata (for example, sea urchins and sea stars), members of the non-chordate deuterostome group Ambulacaria, brachyury is expressed at the vegetal pole marking the margin of the blastopore. Later, brachyury expression is confined to the most vegetal part of the hindgut surrounding the anus (proctodeum) and the stomodeum. Sea urchin Brachyury has been functionally implicated in the regulation of gastrulation movements. While to date no functional data are available for amphioxus, the expression pattern of bra1 is very similar to that in vertebrates, suggesting that it probably plays a similar role in this cephalochordate, while bra1 appears to be more restricted to the notochord. Thus, Brachyury is key for mesoderm formation and in particular for notochord development in vertebrates and, possibly, in chordates. However, whether Brachyury has a mesodermal function outside chordates is less clear.

To reconstruct the evolution of Brachyury function, we analysed its genomic targets in representatives of phylogenetically informative groups (Fig. 1a), the sea anemone Nematostella and the sea urchin Strongylocentrotus. We generated antibodies against Nematostella and Strongylocentrotus Brachyury proteins and confirmed their specificity (Fig. 1b,c; Methods; Extended Data Fig. 1a–d; Supplementary Information). We next used these antibodies to perform ChiP–seq in early gastrula-stage embryos. For Nematostella, we identified 2,389 putative binding sites in two highly reproducible replicates, which mapped to 1,543 putative target genes (Fig. 2a,b; Methods; Extended Data Fig. 1e). For Strongylocentrotus, we identified 490 binding sites, which corresponded to 391 target genes (Fig. 2b and Extended Data Fig. 1f). The lower number of detected binding sites in the sea urchin might be due to less effective ChiP enrichments or reflect a biological difference. Notably, a large fraction of the Brachyury peaks in Nematostella (850 out of 2,389) overlapped with the previously 2,559 identified enhancers at the gastrula stage as defined by the combination of p300, H3K27ac and H3K4me1, suggesting that about one-third of the identified enhancers are bound by Brachyury at the gastrula stage (Extended Data Fig. 2a,b). Similarly, a large fraction (338 out of 490) of the Brachyury binding sites detected by ChiP–seq in Strongylocentrotus are found in open chromatin as detected by ATAC-seq (Extended Data Fig. 2c). However, we also found a significant fraction in both species that are not associated with active enhancers or open chromatin (Discussion). To compare our findings with other species, we re-analysed previously published ChiP–seq data using our peak calling thresholds and found for mouse and Xenopus 4,000 and 2,497 peaks, respectively. Corroborating the published results, our called mouse peaks corresponded to 3,060 putative target genes, while in Xenopus 1,376 target genes were detected (Fig. 2b).

When analysing the distribution of the Brachyury binding sites with respect to target genes, we found that the distribution reflects the genome sizes of the species. In species with larger genome size the peaks are distributed over a larger range, sometimes beyond 100 kilobases (kb), while in species with smaller genome size the peaks are much more in the vicinity of the transcription start site (TSS) (Fig. 2h–i). This is particularly obvious in the case of Nematostella, where -50% of
Fig. 1 | Lineages sampled in this study and Brachyury protein and mRNA expression. **a**, Simplified phylogenetic tree depicting the major animal lineages. Phyla analysed in this study are highlighted in red. Genome-scale Brachyury target-gene screens from Capsaspora owczarzaki, Nematostella vectensis, Strongylocentrotus purpuratus, Ciona intestinalis, Xenopus tropicalis and Mus musculus, representing filasterea, cnidaria, echinodermata, tunicata and vertebrata (amphibians and mammals), respectively, were analysed.

**b**, Anti-Brachyury stainings and WMISH of Nematostella embryos at the early and late gastrula stage. Left column, merged image of Brachyury protein expression (red), acetylated tubulin staining (green) and nuclei (blue). Middle column, Brachyury protein expression only. Right column, WMISH of brachyury (dark blue) of corresponding stages. Asterisk indicates blastopore. **b(i)** and **b(vii)** show lateral views; **b(iv)** and **b(x)** show oral views.

**c**, Immunofluorescence of Strongylocentrotus anti-Brachyury in 24 hpf blastula (**c(i)** and **c(iv)) and 48 hpf gastrula (**c(vii)**–**(x)**) stages. Left column, merged image of Brachyury protein expression (red), acetylated tubulin staining (green) and nuclei (blue) in the 24 hpf blastula and 48 hpf gastrula stage. Middle column, Brachyury protein only. Right column, WMISH of brachyury at corresponding stages. Orientation, animal is on top. **c(i)** and **c(vii)** are lateral views (oral to the right); **c(iv)** and **c(x)** are oral views. Animal icons are either self-generated (Nematostella, Aplysina) or from Phylopic (http://phylopic.org/); Filasterea (Public Domain Dedication 1.0) was created by T. Michael Keesey; Ctenophora (Creative Commons Attribution 3.0 Unported) was created by M. Kodi; Echinodermata (Creative Commons Attribution 3.0 Unported) was created by H. N. Eyster; Tunicata (Public Domain Dedication 1.0) by C. Schomburg; Amphibia (Public Domain Dedication 1.0) by uncredited person; mammals (Creative Commons Attribution 3.0 Unported) by S. Miranda-Rottmann. Scale bars, 20 µm.
all Brachyury binding sites are located within 1 kb upstream or downstream of the TSS, compared to all three deuterostome species (5–12%) (Fig. 2d–f). The smaller genome size also makes the gene density in Nematostella higher (about one gene per 10 kb) than that of other species concerned. Different transcription factors tend to have strikingly different positional specificity within the enhancer regions. From the binding profile of Brachyury (Fig. 2d–g), Brachyury seems to be strongly concentrated around the TSS in all the species under consideration.

Brachyury binding motifs are highly conserved

The Brachyury DNA-binding motif was originally found to be a palindromic sequence by an in vitro Selex approach, which was supported by ChIP-seq in Xenopus embryos. We detected 700 and 236 palindromic Brachyury motifs in the sea anemone and sea urchin Brachyury ChIP-seq peaks, respectively. This indicates that the capacity for dimerization is an ancestral feature of Brachyury (Fig. 2b,c). We detected significantly higher Brachyury peaks with palindromes compared to peaks with half-palindromes (Extended Data Fig. 4a), suggesting a higher binding affinity of the dimer compared to the half-palindromes. However, we found no significant correlation between the distribution of half-palindromes or palindromes with the expression level or gene ontology (GO) category of the target genes. Thus, the biological meaning of the distribution of half-palindromes versus palindromes remains obscure at this point. However, the detailed comparison of the binding motifs of Brachyury in Nematostella, Strongylocentrotus, Xenopus and mouse showed that the canonical binding motifs are highly conserved throughout metazoans (Fig. 2c). Similar motifs have also been found among the ATAC-seq peaks in the protist Capsaspora. This suggests that the DNA-binding domain and the corresponding binding motif have been conserved from protists to humans, indicating a strong positive selection. This might explain why the overexpression of the cnidarian or even Capsaspora brachyury mRNA in frogs induces the formation of mesoderm.

To detect potential cofactors or competitors of Brachyury in the four species, we searched the peaks for an enrichment of other common motifs, also considering their position with respect to the peak summit. As expected, Brachyury motifs locate centrally, while other transcription factor binding motifs are enriched within ~50 base pairs (bp) from the peak centre, suggesting possible co-binding (Extended Data Fig. 4b). When scanning the peak sequences with the software tool Find Individual Motif Occurrences (FIMO) for known motifs matching to the enriched motifs identified in any species, we found a similar distribution of putative binding sites for homeodomain, bHLH, Pax, HMG (Sox) and Fox proteins in all species (Extended Data Fig. 4b). Especially in Nematostella, a substantial number of putative Sox, Fox and Homeobox motif sites are found together with Brachyury (Extended Data Fig. 4b). This is of particular interest as the expression of foxA and foxB, as well as sox81 and sox82, is partially overlapping with brachyury in Nematostella (see below).
**a** Target selection strategy

| Species       | Peaks | Target genes | Palindrome | Half-palindrome |
|---------------|-------|--------------|------------|-----------------|
| *M. musculus* | 4,000 | 3,063        | 2,555      | 809             |
| *X. tropicalis* | 2,497 | 1,376        | 868        | 1,233           |
| *C. intestinalis* | NA    | 2,285        | NA         | NA              |
| *S. purpuratus* | 490   | 391          | 238        | 119             |
| *N. vectensis* | 2,389 | 1,543        | 700        | 1,071           |
| *C. owczarzaki* | NA    | 1,547        | NA         | NA              |

*b* Brachyury binding sites summary

Targets assigned as per above protocol

**d** Brachyury binding sites distribution

**c** Selection based on the values of Number A and Number B

Details in Supplementary Fig. 2

**h** Distribution of Brachyury binding loci relative to TSS

Feature:
- >100 kb
- 10–100 kb
- 5–10 kb
- 3–5 kb
- 1–3 kb
- 0–1 kb

**b** Graphs showing peak count frequency for different species (e.g., *M. musculus*, *X. tropicalis*, *S. purpuratus*, *N. vectensis*).

**c** Genomic region (5’–3’)

- **Intergenic**
- **Upstream** (<1 Kb)
- **UTR**
- **Intron**
- **Exon**
- **Downstream** (<300 bp)

**i** Brachyury binding sites distribution

| Species       | Percentage (%) |
|---------------|----------------|
| *M. musculus* |                |
| *X. tropicalis* |               |
| *S. purpuratus* |               |
| *N. vectensis* |                |
In *Strongylocentrotus*, zygotic *brachyury* expression initiates at the blastopore just before the onset of gastrulation (18 hours post fertilization (hpf)). Shortly before the archenteron breaks through (24 hpf), the future stomodeum (oral ectoderm) of the late gastrula also starts to express *brachyury* (Fig. 1c). Therefore, Brachyury is expected to have different target genes in these two domains.
For instance, different genes are co-expressed with *brachyury* in these two domains (for example, *hox11/13b* is co-expressed with *brachyury* in the blastoporal region, while *gsc* is co-expressed with *brachyury* in the oral ectoderm (Fig. 4c and Extended Data Fig. 5b)). We validated several ChIP target genes that showed a change of expression in the differential gene expression analyses in *brachyury* morphants. Among the *brachyury* co-expressed genes of the presumptive endoderm, *fosa* and *unt* are downregulated, while *hox11/13b* and *otx* are upregulated in *Brachyury* knockdowns. Moreover, the mesodermal markers *ets* and *ese* are ectopically expressed in the presumptive endoderm, suggesting a repressive function of *Brachyury* on these genes in this domain of wild-type embryos (Fig. 4c and Extended Data Fig. 6b). These results show that, as in vertebrates, *Nematostella* and *Strongylocentrotus* *Brachyury* can act both as direct activator and repressor on different target genes. By contrast, some target genes, which are show overlapping expression with *brachyury* in the blastoporal region (for example, *soxC, unt* and *eve*) remain unaffected by the *brachyury* knockdown (Extended Data Fig. 5a), suggesting that other factors than *brachyury* may play a decisive role in their regulation.

Taked together, in both the sea urchin and the sea anemone *Brachyury* activates numerous members of the Wnt–β-catenin and the Wnt–Planar Cell Polarity (PCP) pathway in the blastoporal region of the gastrulating embryo, as well as *foxA*, which in sea urchins is expressed both in the blastopore and the future stomodeum. Since *brachyury* in turn is downstream of canonical Wnt signalling both in *Strongylocentrotus* and in *Nematostella* (as well as other cnidarians), we conclude that in cnidarians, sea urchins and vertebrates, there is a positive feedback loop of *Brachyury* and Wnt signalling at the blastopore and its derivative tissues.

*Nematostella* *Brachyury* represses neuronal genes in the oral domain

One of the surprising recent findings in vertebrates was that *Brachyury* has a dual role in the differentiation of ‘neuromesodermal progenitors’: it activates mesodermal genes but directly also represses neuronal genes, thereby antagonizing the function of Sox2 in the promotion of neural fate. Surprisingly, in *Nematostella*, *Brachyury* also regulates many genes that are expressed at the aboral half, thus not overlapping with *brachyury* expression. During gastrulation, the aboral half is the domain of early neurogenesis. Notably, many of these aborally expressed target genes are involved in neurogenesis, for example the *achaete-scute* homologue *ash-A*, *islet-1*, *tbx2/3*, *masterblind*, *rfx4-like*, *noc* and *ihx-1*, which are expressed in single scattered cells, typical for early neurogenic factors. Interestingly, in *brachyury* morphants, the expression domain of many of these genes expands to the oral domain, suggesting that *Brachyury* might be directly inhibiting early neurogenesis in the oral domain of the blastopore (Fig. 4b and Extended Data Fig. 5). Of note, the putative neuronal markers that show no significant change of expression also do not show a single-cell pattern but rather a global gastrodermal expression pattern, suggesting that the function of *Brachyury* on the expression of neuronal genes is restricted to the ectoderm (compare Fig. 4b and Extended Data Fig. 5).

**Fig. 4 | Validation of Brachyury target genes. a**, Relative spatial expression of *Nematostella* oral markers *brachyury*, *foxA*, *foxB*, *fgf9a*, *unt1*, *unt3* and *unt2* in control morpholino (StdMO) and *Brachyury*-injected embryos (*BraMO*). Note that except for *brachyury*, itself, oral markers are abolished or strongly diminished. *Unt2*, a midbody marker, invades the oral territory of *brachyury* in *BraMO* injected embryos. Blue arrowheads point to domain of expression overlapping with *bra* expression. *B*, Aboral neuronal markers *soxC1, AshA, rfx4-like*, *ihx, ngn-1* and *tbx2/3* expand orally in *Bra* knockdown embryos. The asterisks indicate the oral side of the embryo. Quantitative data are found in *Extended Data Fig. 1d*. All embryos shown in lateral view. Insets show oral (O) or aboral (A) view. Scale bar, 50 µm. c, Schematic representation (left top corner) of *brachyury* expression (green stripes) and regulatory input in the endoderm (yellow), ectoderm (red) and endoderm (blue). Double and single fluorescent WMISH of *Strongylocentrotus* *unt6, foxA, axo, hox11/13b* and *ets1* (magenta) relative to *brachyury* expression (green). Yellow and blue arrows indicate domains of co-expression. Red arrows indicate adjacent/exclusive expression. d, WMISH of *fgf9/16/20*, *nodal*, *bmp2/4*, *fgf5/8*, *six3* and *spec* in *Brachyury* morphants compared to wild-type 24 hpf embryos. Blue arrow indicates *brachyury* expression in oral ectoderm. Pictures are projections of confocal stacks. Nuclei are labelled blue with DAPI. # indicates that the ChIP target is also present in the differential RNA-seq dataset. L/v, lateral view; v/v, vegetal view; o/v, oral view; a/v, aboral view; up, upregulated gene; down, downregulated gene. Scale bar, 20 µm.

**Sea urchin Brachyury is involved in ectodermal patterning**

We also investigated the effect of *Strongylocentrotus* *Brachyury* on genes of the oral ectodermal territory, where *brachyury* is expressed during gastrulation. The oral ectoderm domain is adjacent to the anterior neuroectoderm (ANE) of the embryo that will later give rise to the apical organ of the pluteus larva. The expression of the oral ectoderm (future stomodeum) gene *goosecoid* (ref. 85), which shows an overlapping expression with *brachyury* in this region in wild-type embryos, is severely reduced in *Brachyury* morphants (Extended Data Fig. 6b,d). Similarly, the ANE marker *Nkx2.1* is downregulated in *Brachyury* morphants as revealed by differential RNA sequencing and immunohistochemistry (Extended Data Fig. 6c,d). Notably, several other genes that are also expressed in the ANE (for example, *fgf9/16/20, fgf5/8, six3* and *soxB2*) remain unaffected in *Brachyury* morphants (Fig. 4d and Extended Data Fig. 6b–d), suggesting that *Brachyury* might have an indirect role in patterning the anterior neural ectoderm. Moreover, genes involved in the specification of the oral–aboral (ventral–dorsal) axis such as *nodal* and *bmp2/4* (ref. 86) that partially colocalize with *brachyury* seem to be severely affected, with the expression of *nodal* reduced and the expression of *bmp2/4* abolished (Fig. 4d). *Fgf9/16/20* that is expressed in the oral side is also abolished (Fig. 4d). This suggests that the proper formation of the oral–aboral axis is compromised in *brachyury* morphants. Notably, the expression of the oral ectoderm marker *spec1* (ref. 84) is also reduced in *brachyury* morphants, confirming the general aberrant patterning of the ectoderm. Such a role of a sea urchin *Brachyury* at the oral–aboral axis organizer is in line with what was recently suggested by single-cell transcriptomics data.

To summarize, similar to vertebrates, in *Nematostella* *Brachyury* appears to contribute to the inhibition of neuronal differentiation in the blastoporal domain, thereby restricting it to the aboral part. In *Strongylocentrotus* neuronal markers (for example, *hbn*, *nxa2.1* and *nxa3.2*; Extended Data Fig. 5c,d) appear to be downregulated upon *Bra* knockdown, however, mostly indirectly due to the fact that in the absence of *Brachyury* the ANE is generally mispatterned. This ectodermal mispatterning results not only in loss of distinct ectodermal territories (oral–aboral-ANE) but also in the inability of the ANE to promote neurogenesis and, thus, proper neuronal differentiation.

**Phylogenetic comparisons reveal ancestral Brachyury targets**

To reconstruct the evolutionary changes of *Brachyury* function that led to its mesoderm determination role, conserved among chordates, we aimed to compare the direct *Brachyury* target genes in organisms with available information, ranging from protists to vertebrates. To this end, we first generated a reliable set of orthologues in the species under consideration using OMA (Orthologous Matrix; Methods). We then compared cnidarian and sea urchin *Brachyury* targets identified by ChIP–seq with the published target gene sets from the urochordate *C. intestinalis* and two vertebrates, the mouse *Mus musculus* and the frog *Xenopus tropicalis*. We added to this comparison the putative *Brachyury* targets from the protist *Capsaspora owczarzaki* determined by the search of *Brachyury* binding motifs in ATAC-seq peaks. The OMA allowed us to assign pairwise homologues between these species and thereby reconstruct the ancestral target genes at each of the
phylogenetic branching nodes (Figure 5 and Supplementary Table 1). Since the mouse has the best-annotated genome, we used it as a reference for the functional annotations in the analyses. As in all investigated metazoans regulation of transcription, multicellular organism development and signal transduction are among the most dominant GO categories of the Brachyury targets, we first focused our analysis on
transcription factors. Notably, brachyury is the only target gene that is shared among all organisms, suggesting that self-regulation is an ancestral feature and under strong selection (Extended Data Fig. 7).

There is ample evidence that transcription factor binding sites and hence the corresponding target genes can diverge rapidly. However, important target genes might be under higher selection pressure and maintained over longer evolutionary distances. Thus, by comparison of lineages of various phylogenetic distances, we should be able to identify ancestral target genes of distinct phylogenetic groups. Therefore, to reconstruct how the Brachyury target gene set was conserved or was changed during the course of animal evolution, we sought to determine by pairwise comparisons, which target genes evolved early and have been maintained in several lineages and which genes were recruited only in more recent lineages. A gene was considered an ancestral target for a given ancestor node when a homologue is shared between the earliest branching organism (outgroup) and at least one of the other ingroup species (Methods). To unravel shared ancestral and lineage-specific developmental regulators, we focused the analyses on target genes coding for transcription factors (TFs) (Fig. 5), while the whole set of target genes is found in Supplementary Table 1. The pairwise analysis suggested that six genes were TFs coding targets in the last common ancestor of Capasapora and Metazoa (Fig. 5). Node II is of particular interest, as it reflects the common ancestor of the diploblastic sea anemone Nemastoealla with the deuterostomes. Notably, we found numerous homologous target TFs shared between Nemastoealla and at least one of the deuterostome species (Strongylocentrotus, Ciona, Xenopus and Mus) (Fig. 5).

To understand the functional consequences of shared or species-specific Brachyury target genes, we annotated the target genes encoding TFs by their differential expression in different germ layers and tissues during mouse gastrulation. As a reference, we used a recently published single-cell transcriptome dataset from mouse embryos corresponding to the gastrula stage (stage E8.0 and 8.5; Methods). When we calculated the log, normalized endodermal versus mesodermal as well as neuronal versus mesodermal gene expression levels, we found that ancestral node II genes (Nemastoealla plus deuterostomes) tend to have significantly higher expression levels in endodermal and neuronal cell types compared to mesodermal cell types (Fig. 5). A similar trend of more endodermal targets compared to the earliest branching organisms was detected at node III (deuterostomes) and node IV (chordates), albeit less pronounced than at node II. Only at node V (genes exclusively shared among the two vertebrates), we detected a more mesodermal gene expression, indicating a shift of target genes of Brachyury in the chordate and vertebrate lineage. Node V target genes (32 of which are TFs) are shared between Xenopus and mouse. Since homologues of these genes are not found as targets in the other organisms, they supposedly have evolved newly or have been recruited in vertebrates. Unlike node II genes, we find node V genes to be more highly expressed in mesodermal cell types (Fig. 5 and Extended Data Fig. 9). In addition, several of these genes are known as crucial regulators involved in mesoderm development. Among others, this includes mesp1/2, mesogenin, twist, mef2c and smad6. Thus, these TFs are likely to be involved in conveying the role of Brachyury as a mesoderm determination factor during early gastrulation and a pioneering factor in neuromesodermal progenitor differentiation.

In summary, this phylogenetic analysis of the target genes shows that numerous ancestral target genes that predate the split of cnidarians and bilaterians have a neuronal function at least in vertebrates, whereas a large fraction of key mesodermal target genes was only acquired at the level of the vertebrates.

**Fig. 5.** Apomorph and synapomorphic target genes of Brachyury. To assess the evolutionary origin of vertebrate Brachyury target genes and their function, apomorphic and synapomorphic target genes were assigned to nodes of common ancestors. Transcription factor coding genes targeted by Brachyury that are shared between an outgroup and at least one member of the ingroup are considered ancestral for the respective node. The expression of the target genes in neuronal versus mesodermal cell types was annotated using single-cell RNA-seq data from E8.0 and E8.5 mouse embryos. a, Heatmap of the log2-fold change (logFC) of neuronal (turquoise) versus endodermal (yellow) versus mesodermal (purple) gene expression for each gene using the corresponding mouse genes as they appeared in the single-cell dataset in each node (Methods). b, Boxplot of the log2-fold change (y axis) of endodermal versus mesodermal gene expression for all genes per node (x axis). Note that node II is enriched in endodermally expressed genes while node V is enriched in mesodermally expressed genes (P = 0.008, Wilcoxon rank sum test). c, Boxplot of the log2-fold change (y axis) of neuronal versus mesodermal gene expression for all genes per node (x axis). Note that node II is enriched in neuronal expression while node V is enriched in mesodermal expression (P = 0.008, Wilcoxon rank sum test). The boxes range from the 25th to the 75th percentile and the horizontal lines represent the median. Outliers are shown as dots.
**Node I**
- Eukaryotes
  - Deuterostomes
    - Eumetazoans
      - Chordates
        - Vertebrates
          - X. tropicalis
          - C. owczarzaki
          - S. purpuratus
          - C. intestinalis
          - N. vectensis
          - M. musculus

**Node II**
- log2 endodermal/mesodermal gene expression (cpm)

**Node III**
- log2 neuronal/mesodermal gene expression (cpm)

**Node IV**
- logFC endodermal/mesodermal gene expression

**Node V**
- logFC neuronal/mesodermal gene expression

**Panel a**
- Node I: Eukaryotes → Node II: Deuterostomes → Node III: Chordates → Node IV: Vertebrates → Node V: X. tropicalis, C. owczarzaki, S. purpuratus, C. intestinalis, N. vectensis, M. musculus

**Panel b**
- Graph showing log FC endodermal/mesodermal gene expression with data points for each node.

**Panel c**
- Graph showing log FC neuronal/mesodermal gene expression with data points for each node.
Fig. 6 | Comparison of regulatory functions of Brachyury in the sea anemone, sea urchin and vertebrates. a. The feedback loop of Brachyury and the Wnt signalling pathway is conserved between the sea anemone, sea urchin and vertebrate, although individual Wnt ligands have been changed as targets of Brachyury. Most members of the Wnt pathway are direct targets of Brachyury (Supplementary Table 1). b. Summary of the main functions of Brachyury in axial patterning, neuronal differentiation and germ layer specification in the diploblastic sea anemone, the sea urchin and vertebrates. Silhouette images were self-generated (Nematostella and endoderm) or taken from Phylopic (http://phylopic.org): Echinodermata (Creative Commons Attribution 3.0 Unported) created by H. N. Eyster; mouse (Creative Commons Attribution 3.0 Unported) by S. Miranda-Rottmann. c. Scenario of stepwise evolution of Brachyury function from protists to vertebrates. See text for explanation.
Discussion

In this study, we revealed the genome-wide targets of Brachyury in two invertebrates—the diploblast sea anemone N. vectensis and a basally branching deuterostome, the sea urchin S. purpuratus—and compared them with known sets of target genes from similar ChIP experiments in other organisms.

Limitations of this study

We took particular care to use comparable stages of Nemastella and sea urchin embryos for our ChIP-sequencing experiments at early gastrulation. However, as with any such comparative analyses, we cannot fully rule out that both missing or detected target genes are due to specific differences in the developmental stage, tissue origin and the experimental design. For example, all ChIP-sequencing data were generated using gastrulation stage embryos, except for the mouse dataset, which was generated using in vitro differentiated NMP cells from embryonic stem cells and which might lack notochordal genes. ChIP-sequencing studies focusing on the mouse notochord might therefore add more ancestral genes. Also, each ChIP-sequencing experiment has a specific sensitivity and therefore leads to more or less detected target genes, which is probably the main reason why we detect fewer target genes in the sea urchin. The cross-species analyses of ancestral target genes (Fig. 5) also relies on the annotation of cell clusters of the mouse single-cell RNA-seq data. For instance, others have annotated early notochordal cells as definitive endoderm, which might be considered a false annotation. However, we would like to note that in the protochordate Amphioxus the notochord derives by delamination from the dorsal archenteron midline, independent of the paraxial mesoderm, indicating a close link of notochord and endoderm. In summary, as a cautionary note, we cannot make statements and draw conclusions on individual target genes that might be missing in a given dataset. On the other hand, the mouse and frog datasets share a large number of target genes, many of which have also been validated experimentally at other stages. This suggests that a large fraction of the conserved target genes can be captured.

A conserved blastoporal Brachyury and Wnt feedback loop

Outside metazoans, brachyury and several other T-box genes are found in the protostepan C. owczarzaki, as well as in certain fungi, indicating that it evolved before the emergence of the metazoans (Fig. 6d). The function of brachyury in these organisms is not known but the T-domain of C. owczarzaki is sufficiently conserved to induce mesoderm when expressed in the frog embryo. Our analyses of the binding sites confirm that the binding motif of Brachyury is deeply conserved throughout the animal kingdom and probably even extending to the protostepan C. owczarzaki, hence over >700 million years. As in vertebrates, we do find both palindromic sites indicating binding of dimers, as well as half-palindromes (Extended Data Fig. 3; Supplementary Information). Co-occurring motif analyses suggest either competitive or cooperative binding of other transcription factors with a bias for Sox, Fox and homoeodomain factors (Extended Data Fig. 3; Supplementary Information).

One of the striking features of Brachyury is the highly conserved expression domain around the blastopore in almost all animals studied, with only few exceptions (Supplementary Information). This raises the question of what might stabilize the expression domain over hundreds of millions of years in so many diverging phyla.

In vertebrates, several Wnt pathway members like dkk1, wnt5b, wnt8a, wnt9a and wnt11b (PCP) are targets of Brachyury and brachyury is also regulated by Wnt3 and by EGF in a feedback loop. It was recently shown that Wnt signalling acts upstream of brachyury in Nemastella. In this study, our ChIP-sequencing data in Nemastella have further revealed that numerous wnt ligands, all four frizzled receptors, dishevelled and β-catenin are direct targets of Brachyury, demonstrating the direct regulation of the canonical as well as the putative PCP pathway by Brachyury. Similarly, in Strongylocentrotus, brachyury is activated by the canonical Wnt pathway and, in turn, Brachyury targets several genes of the Wnt pathway, highlighting a deeply conserved feedback loop of Wnt and Brachyury (Fig. 6a). Since brachyury expression and Wnt signalling are colocalized at the blastopore lip in most bilaterian species, we postulate that brachyury forms an ancestral gene regulatory feedback loop at the blastopore. Furthermore, because Wnt signalling is known to play a crucial role in establishing the primary body axis in vertebrates, ambulacrarians (echinoderms and hemichordates) and cnidarians, we propose that one of the ancestral roles of Brachyury in metazoans was in axial patterning (Fig. 6b). In support of this idea, knockdown of brachyury abolishes the axis induction capacity of the blastopore lip in Nemastella. While we do not have evidence for a control of brachyury by FGF signalling at present, we also find FGF ligands (fgf8 in Nemastella and fgf9/16/20 in Strongylocentrotus), as well as foxA (and foxB in Nemastella) as deeply conserved target genes of Brachyury, suggesting that these genes might also belong to the axis patterning kernel (Fig. 6a,b). Animal icons are either self-generated (e.g. Nemastella) or from Phylopic (http://phylopic.org/). Echinodermata (Creative Commons Attribution 3.0 Unported) was created by H. N. Eyster; mammals (Creative Commons Attribution 3.0 Unported) by S. Miranda-Rottmann.

Evolution of the mesoderm inducing role of Bra in chordates

Since Brachyury has a conserved function in mesoderm development in vertebrates and is required for notochord development in urochordates (and possibly in cephalochordates), we looked for conserved downstream mesodermal genes in Nemastella and Strongylocentrotus. Our analysis showed that only few genes with a role in mesoderm formation in vertebrates are also Brachyury target genes in Nemastella (for example, tbx2/3, tbx20, mbnl1 and mnx, gata1/2/3; Fig. 5, Supplementary Table 1 and Extended Data Fig. 6). Instead, a surprisingly large number of mostly negatively regulated TFs that are involved in neurogenesis are shared between Nemastella and vertebrates, suggesting that the negative regulation of neuronal regulators is more ancestral. Likewise, a considerable number of genes that are endodermally expressed in the mouse are targets in Nemastella and hence were supposedly target genes in the common ancestor of chordates and invertebrates (node II). Surprisingly, in Strongylocentrotus, the few mesodermal Brachyury target genes (for example, etsl, gcm and fbxn2/3; Supplementary Table 1 and Fig. 4c,d) are repressed by Brachyury in the presumptive endoderm, in stark contrast to what is observed in chordates. This difference highlights the plasticity of Brachyury in acquiring new functions and targeting new genes.

Some of the key factors driving mesoderm differentiation in chordates, such as tbx6 and myoD, appear to have evolved only in the deuterostome or bilaterian lineage, respectively, since they do not exist in invertebrate genomes. In Strongylocentrotus, two myoD paralogous genes are present (myoD1 and myoD2), both with distinct mesodermal functions in skeletogenesis and myogenesis but there is no evidence that they are target genes of Brachyury. Quite the contrary, our functional studies suggest that, in the sea urchin, brachyury represses mesodermal fate (through repression of the mesoderm specification gene eae) and rather promotes endoderm formation. Tbx6 is a key downstream gene in vertebrates, conveying the mesoderm determination function of Brachyury in feed-forward loops. It is a crucial factor in the development of the paraxial mesoderm and mutation of tbx6 leads to the formation of two supernumerary neural tubes instead of somites. Tbx6 is also a Brachyury target gene in the ascidian C. intestinalis, expressed in the future paraxial muscle, hence mutually exclusive to the brachyury expressing notochord and therefore likely to be negatively regulated by Brachyury. The evolution of chordate tbx6 (arisen through duplication of a member of the tbx2/3/4/5 families) was probably a crucial step in the recruitment of Brachyury to a mesodermal function.
Besides *tbx6*, another crucial evolutionary step was the acquisition of other key mesodermal downstream differentiation determinants in vertebrates, such as *segT*, *mesp1/2*, *twist*, *mesogenin*, *mef2*, *myf5*, *myf6* and *myoD* as direct target genes of Brachyury, linking early mesoderm formation to muscle differentiation. All these genes have been shown to play decisive roles in the specification of the mesoderm in vertebrates. While the myogenic regulatory factors (MRFs) (*myf5*, *myf6*, *mef4* and *myoD*) are not present in the cnidian genomes, *twist* exists but is neither a direct target of Brachyury in *Nematostella* nor in *Strongylocentrotus*, nor does it seem to play a major role in axis formation or germ layer formation.

Thus, the evolution of several key mesodermal determinants as targets of Brachyury includes both adoption of conserved genes to mesoderm formation as well as the evolution of new genes in the vertebrate lineage.

### Brachyury and the evolution of neuromesodermal progenitors

The tailbud stage of vertebrates has been viewed as secondary axis formation or even as a continuation of the gastrulation process itself. The tailbud harbours a population of NMPs, which has the potential to differentiate into neural tube or into mesoderm, depending on the concentration of Wnt3a. NMPs express both *brachyury* and *sox2*, a marker for pluripotency and early neural fate. Several studies support the view that *sox2* expression drives the cells towards neural fate and a high level of *brachyury* expression promotes mesodermal fate. ChiP–seq as well as functional studies have shown that Brachyury and Sox2 mutually inhibit each other, which eventually leads to a segregation of neuronal and mesodermal cell fates. Brachyury also represses a number of other neuronal target genes, in line with its role in suppressing neuronal fate.

*Sox2* belongs to the subfamily of SoxB1, together with its vertebrate-specific paralogues Sox1 and Sox3. There is no one-to-one orthologue of *sox2* in cnidarians and sea urchin but there is a cnidian and sea urchin *sox2B1* homologue. Interestingly, while there is currently no evidence that *sox2B1* is a target in *Strongylocentrotus* (although both *sox2* and *sox2B1* are differentially expressed in *Bra* knockdowns; Extended Data Fig. 5b,c), at least in *Nematostella*, *sox2B1* is a direct target of Brachyury (Fig. 3d). But unlike in vertebrate NMPs, there is no evidence for mutual inhibition of *SoxB1* and Brachyury in *Nematostella*. Furthermore, although *sox2B1* is expressed in the domain of aboral neurogenesis, functional perturbation of *sox2B1* does not argue for a direct role of *sox2B1* in promoting neural development in *Nematostella* (Extended Data Fig. 4c). However, Brachyury contributes to the repression of neuronal differentiation during embryogenesis. This effect might be in concert with the role of Wnt signalling. Since a significant number of these neuronal genes are also negatively regulated by Brachyury in the NMPs of vertebrates, we conclude that repression of neuronal differentiation is an (unexpected) ancestral role of Brachyury, regardless of its axial position. This is not to say that NMPs have their origin in the common ancestor of cnidarians and bilaterians but that part of the NMP gene regulatory network and the role of Brachyury therein has a deeper origin.

### Concluding remarks

Our comparative approach of genome-wide target genes of a key developmental regulator, Brachyury, has revealed both deeply conserved kernels as well as key acquisitions of new target genes, that were either recruited or evolved newly by gene duplications. Brachyury has a surprisingly conserved expression pattern around the blastopore in most animal phyla. Since this blastoporal expression is part of the ancestral kernel, we postulate that such feedback loops can evolvably stabilize a given expression pattern and its ancestral role. Nevertheless, the function of the gene can still drastically evolve, by the acquisition of only a few, but crucial, new target genes, which in the case of Brachyury, convey the new function in feed-forward loops. At present, the number of comparative studies of genome-wide target-gene screens of conserved transcription factors is very limited but future approaches over a wide range of distantly related organisms may offer insights into the general mechanisms of robustness and evolvability of gene regulatory networks in the evolution of animal body plans.

### Methods

#### Animal culture

*Nematostella* polyps were kept as previously described at 18°C and largely in the dark. Spawning was induced by a combination of light and elevated temperature. Embryos were raised at 21°C. Adult *Strongylocentrotus* individuals were kept in circulating seawater aquaria at Stazione Zoologica Anton Dohrn in Naples. Spawning was induced by vigorous shaking of gravid animals. Embryos were raised at 15°C in filtered Mediterranean seawater diluted 9:1 with deionized water.

### Antibody generation

A 6×His-tagged full-length clone of *Nematostella brachyury* was expressed in *Escherichia coli* strain BL21. After transformation, bacteria were grown overnight in LB supermedium + ampicillin. Then, 10 ml of LB medium/ampicillin were inoculated with 500 µl of overnight culture, grown for 3 h (optical density 600 = 0.5–0.6), isopropyl-β-D-thiogalactoside was added to a final concentration of 1 mM and bacteria culture was incubated for 21 h at 4°C. Brachyury protein was purified using Ni-NTA agarose (Qiagen), followed by polyacrylamide gel electrophoresis. The Brachyury band at ~50 kDa was excised and used to inject two rabbits (PRIMM Biotech). For *Strongylocentrotus* Brachyury, two polyclonal antibodies from two different rabbits were raised against the recombinantly produced C-terminal domain excluding the T-box domain, with the T-box domain. The antibodies were affinity purified by PRIMM Biotech before their use.

For both *Nematostella* and *Strongylocentrotus* the reactivity of polyclonal antibodies from each rabbit were assessed by both Western blot and immunohistochemistry, confirming their specificity (Extended Data Fig. 1). For the ChiP replicates, antibodies from the two different rabbits were used for each.

### Chromatin immunoprecipitation and library preparation

ChiP–seq was carried out essentially as previously published. In brief, embryos at gastrula stage were fixed with 2% formaldehyde for 12 min. Nuclei were collected by Dounce homogenizing the embryos. Chromatin from ~20 nuclei was excised and used to inject two rabbits (PRIMM Biotech). For both *Nematostella* and *Strongylocentrotus* immunoprecipitations were pooled to obtain ~2 ng of DNA, as measured by a Pico Analyzer. This was then used for library preparation according to Illumina ChiP–seq DNA Sample Prep Kit instructions (catalogue no. IP-102-1001). The quantity and quality (size distribution) of the libraries was confirmed using an Agilent Bioanalyzer. Deep sequencing was performed at the Vienna BioCenter Core Facilities (https://www.viennabio-center.org/facilities/) with 50 bp of SE HiSeq 2500 reads. We obtained 45,753,729, 52,253,224, 50,361,646 and 44,089,574 reads for *Strongylocentrotus* *Brachyury* ChiP replicate 1, replicate 2 and input replicate 1, respectively. Of these, 18,744,583, 15,413,040, 21,304,793 and 18,443,726 were uniquely mapped to the genome (see below). We obtained 39,065,464, 33,397,307 and 43,959,250 reads for *Nematostella brachyury* ChiP replicate 1, replicate 2 and input, respectively. Of these, 21,889,877, 17,420,963 and 25,679,308 were uniquely mapped to the genome (see below).

### Orthologue detection

For orthologue detection we used the OMA database. Before we subjected our in silico translated protein sequences to OMA, we subjected our in silico translated protein sequences to OMA.
we excluded peptide sequences smaller than 100 amino acids. To detect the orthologues over the wide evolutionary distance of our investigated organisms, it was necessary to use a less stringent ‘Length tolerance ratio’ = 0.2 parameter of OMA. The other OMA parameters used default values.

**ChIP–seq analysis**

In addition to our ChIP–seq datasets for Brachyury in *Nematostella* and *Strongylocentrotus* we acquired raw reads of published Brachyury ChIP–seq datasets of mouse10 (GEO: GSE93524) and *Xenopus*10 (GEO: GSE48660). Reads from *M. musculus*, *T. tropicalis*, *S. purpuratus* and *N. vectensis* were mapped to genome versions GRCm38 (mm10), v.9.1, v.5.0 and v.1.0, respectively, with BWA using the BWA-MEM algorithm10. ChIP–seq peaks were called using Peakzilla10 using input sequencing as control. Depending on the bimodal distribution of the reads, the Peakzilla algorithm assigns a score to each peak. We examined the relationship between Peakzilla peaks and their scores and the coverage of ChIP–seq reads across the genomes, thereby determining the optimal score cutoff to filter high-quality peaks in each dataset, where the lowest scoring peaks that pass the cutoff are still visually detected as peaks in the coverage tracks. Consequently, we removed peaks with a score <2 for mouse and *Xenopus*, <0.5 for *Strongylocentrotus* and <5 for *Nematostella*. Of note, the number of detected target genes in *Strongylocentrotus* is noticeably lower than in the other species, which could be either due to biological reasons (fewer target genes) or technical reasons (less sensitive antibody) or a combination of both. This may lead to an underestimation of the number of conserved targets in our comparative analyses.

**Target gene selection**

To identify target genes, a commonly used procedure is to select the closest gene to an identified peak. However, in more compact genomes like *Nematostella* or *Strongylocentrotus* (five to ten genes per 100 kb) peaks in the intergenic space are often only marginally less distant to a gene in the other direction. Therefore, we identified the two closest genes to the binding site. Next, using orthologue information obtained earlier with OMA11,12, we determined whether either of the two closest genes has an orthologue that is a target in any other species considered in this study. In about 20% of *Nematostella*, 24% of *Strongylocentrotus*, 3% of *Xenopus* and 28% of mouse peaks, no orthologue information was found; thus, the closest gene to the binding site was assumed to be the target. If both closest genes had orthologous targets, we selected the one with larger number of orthologous target genes (Extended Data Fig. 2d). In case an equal number of the orthologous target genes was found, we kept both putative targets and marked them with ++ (Supplementary Table 1).

**Motif analysis**

To find and compare the binding motifs of Brachyury we used the MEME-ChIP (4.11.2) suite13. For the number of motifs to be predicted in a sequence we selected ‘an’ (any number of repetitions) mode and 1 × 10^10 for the e-value cutoff criterion. The discovered motifs were scanned against the non-redundant sequence profiles obtained from the JASPAR database14. The other parameters were set to default (a detailed list of parameters can be found on https://github.com/technau/brachyury_grn).

**Co-occupancy of motifs**

To investigate what other motifs are present in the vicinity of Brachyury binding motif and their spatial arrangement, we made use of MEME-ChIP with an in-house python script, which finds all the motifs in a peak and their absolute distances from the centre of the peak. It also groups the motifs into motif groups on the basis of motif families in JASPAR15. In case of overlapping motifs, it selects the motif with better False Discovery Rate (FDR)-adjusted P value assigned by FIMO from the MEME-ChIP suite10.

**Brachury knockdowns and RNA-seq**

To knockdown Brachyury function in *Nematostella*, we used two non-overlapping antisense morpholinos (MO) against *brachury*: the translation-blocking morpholino BraATG-MO TCGTCCGAGTGCATGTCGACTATG and the splice-blocking morpholino BraSpliceMO TCCCTGTGGTTGCAACCATACTGCCGCTCC. A total 3–5 pl of both morpholinos were injected at the concentration of 500 µM. A total 50 µg ml^−1 of Dextran-Alexa Fluor 488 MW 10000 (ThermoFisher) was co-injected as tracer to ensure proper delivery of MO16. Standard morpholino (Genetools) StdMo CCTTACTCCTGATTTGAAATA was injected as a control at the same concentration. PolyA-enriched RNA samples were collected and processed for library production using the Lexogen kit and sequencing (50 bp single-end HiSeqV4, -30 million reads per sample (for more details see https://github.com/technau/brachyury_grn).

**Differential gene expression analyses**

For *Nematostella*, we generated six biological replicates of each translational and splice morpholino. Principal component analysis showed that the replicates cluster together (Extended Data Fig. 3a). For *Strongylocentrotus*, we generated three biological replicates (Supplementary Fig. 3b). This was compared to three replicates of morpholino knockdown in *Xenopus*16 (GSE48663) (Extended Data Fig. 3c) and two replicates of mouse knockout mutants17 (GSE93524) (Extended Data Fig. 3d).

**Whole mount in situ hybridization**

WMISH in *Nematostella* was carried out as described before18 with a few minor modifications. Briefly, embryos were fixed in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) for 1 h at room temperature (RT) and then washed 5× with 0.2% Tween in PBS (PTw) at RT. The embryos were then washed once with 50% methanol in PTw and lastly transferred into 100% methanol and stored at −20 °C until further processing. Hybridization was performed at 1 ng riboprobe µl^−1, overnight at 63 °C and detection with alkaline phosphatase conjugated anti-digoxigenin (1:2,000) (anti-digoxigenin-AP, Fab fragments, Roche, REF 11093274910). If no strong signal and mostly background
began to develop in the NBT/BCIP solution, the embryos were washed 5× in PTw and then left to stand in PBS with 2% w/v Triton X100 until the background was at acceptable levels. This was repeated until a strong signal was observed.

In Strongylocentrotus, WMISH was done as previously described\textsuperscript{132}. Briefly, embryos were fixed in 4% PFA in MOPS buffer overnight at 4 °C, then gradually dehydrated to 70% ethanol and kept at -20 °C until use. Fixed embryos were gradually rehydrated in MOPS buffer, washed several times with the same medium and were prehybridized for 3 h in the hybridization buffer. Hybridization was carried out with 0.1 ng μl\textsuperscript{-1} of probes and incubated for 1 week. The signal for colorimetric in situ hybridization was developed using anti-digoxigenin AP (alkaline phosphatase) conjugated antibody (anti-digoxigenin-AP, Fab fragments, Roche, REF 11093274910) and the AP substrate, while fluorescent in situ hybridization was developed with fluorophore-conjugated tyramide (1:400 reagent diluents, TSA Plus Cyanine 3/Cyanine 5, PerkinElmer, NEL752001KT).

**Immunohistochemistry**

Immunohistochemistry in Nematostella was essentially performed as described in ref.\textsuperscript{134} with the following changes. Briefly, embryos were fixed at 4 °C in 4% PFA in PTwTXD (1× PBS with 0.2% Tween, 0.2% Triton X100 and 0.2% DMSO). All further washing steps were performed with PTwTXD. They were then washed once with PTwTXD and ice-cold acetone and put on ice. Once the embryos had settled down, the acetone was removed and they were washed 10× with the cold PTwTXD. The embryos were then blocked for 2 h at RT in the blocking solution containing 20% v/v of heat-inactivated sheep serum and 80% of 1% w/v BSA in PTwTXD. The blocking solution from the embryos was substituted with the rabbit anti-Bra pre-adsorbed in the blocking solution (1:500) for the time of the blocking and incubated overnight at 4 °C on a table rocker. The next day, the embryos were washed 10× with cold PTwTXD and blocked once again for 2 h at RT in the blocking solution. Blocking solution was then replaced with pre-absorbed goat anti-rabbit (Alexa Fluor 568, 1:2,000, Invitrogen) mixed with DAPI (final concentration 5 μg ml\textsuperscript{-1}) and Alexa Fluor 488 phallidin (final concentration 5 μl l\textsuperscript{-1}, Invitrogen) and incubated overnight at 4 °C. Next day, the embryos were washed 10× in PTwTXD, infiltrated overnight with Vectashield (Vector laboratories) and imaged.

Immunohistochemical detection of Brachury in Strongylocentrotus was performed as described in ref.\textsuperscript{134}. Briefly, the embryos were fixed in 4% PFA in filtered seawater for 15 min at RT, dehydrated in 100% ice-cold methanol for 1 min and washed several times with 1× PBS. PBS was replaced with a blocking solution (4% sheep serum, 1% BSA in 1× PBS) and the samples incubated for 1 h at RT. The blocking solution was replaced with the solution containing the primary antibodies diluted in blocking solution (Brachury 1:100, Nkx2.11:600 (ref.\textsuperscript{134}), Sox2B1:500 (ref.\textsuperscript{134})) and incubated either overnight at 4 °C or 90 min at 37 °C. Embryos were washed several times with PBS, the secondary antibody goat anti-rabbit Alexa Fluor 488 (Invitrogen) diluted in PBS was added for one hour in PBS and then washed again several times.

**Phylogenetic analysis**

Phylogenetic analysis of T-Box, Sox, Zic, TFAP, NR2, RNF and Rfx family members was performed using maximum likelihood methods. Orthologues and paralogues were obtained from the respective proteomes using BLASTP with the protein sequences from Nematostella and mouse as query. The peptide sequence alignments were generated using MAFFT with the parameters –maxiterate 1000 –localpair\textsuperscript{138} and cut to the unambiguously aligned core region. From the alignments maximum likelihood trees were reconstructed using IQ-TREE 1.6.12. Substitution models of LG + I + G4 (T-Box and Sox) and LG + F + G4 (Zic, TFAP, NR2, RNF, Rfx) were chosen after running model selection in IQ-TREE. Support values were generated using UFboot\textsuperscript{139} as implemented in IQ-TREE with 1,000 (T-Box and Sox) and 10,000 (Zic, TFAP, NR2, RNF, Rfx) bootstrap samples, respectively.

**Target gene comparison between species**

We used OMA for its ability to report strict orthologues by verifying pairs of genes\textsuperscript{135,137}. All the translated peptide sequences were filtered against proteins shorter than 100 amino acids. This filtering resulted in 8,381 (C. oeuzarzaki), 25,729 (N. vectensis), 28,659 (S. purpuratus), 25,425 (C. intestinalis), 39,662 (X. tropicalis) and 21,317 (M. musculus) proteins. Our orthology analysis resulted in 1,882 orthologous groups shared by all species (4,791 orthologous groups where at least five species have orthologues, 7,922 orthologous groups where at least four species have orthologues, 13,149 orthologous groups where at least three species have orthologues and 28,918 orthologous groups where at least two species have orthologues).

We then selected the target genes coding for transcription factors using InterProScan\textsuperscript{138} annotation on the basis of the presence of the DNA-binding domain from the Pfam database, then sorted target genes into five nodes on the basis of evolutionary lineages representing Metazoa, Bilateria, Deuterostomia, Chordata and Vertebrata. Genes were annotated according to the mouse gene annotations. Homologous target genes were determined by OMA. Afterwards the OMA annotation was manually improved for several genes by more detailed phylogenetic analyses (see sections above). Homologous target genes that are shared between at least two organisms were placed at the node where the two lineages diverged, regardless of how many other members of the ingroup also shared this target gene. For instance, a target gene detected in Nematostella and at least one of the bilaterian species would be placed at node II (common ancestor of cnidarians and bilaterians). Vertebrates often have multiple paralogues that have only one homologue in the earlier diverging species, which can lead to double assignments in the analysis. For instance, foxA, which is a conserved target in Nematostella, sea urchin and vertebrates, has duplicated into several paralogues in vertebrates\textsuperscript{139}. OMA finds foxA2 as a shared target of Nematostella and mouse and foxA3 as a shared target of sea urchin and mouse. In such cases, we only kept the homologue at the lower node.

To compare the expression of ancestral target genes in different nodes in mouse gastrulation stage cells of neuronal, endodermal and mesodermal origin, we extracted the mouse ENSEMBLE gene IDs for all genes in each node. In the case where we extended the OMA annotation by phylogenetic analysis and one ancestral gene was present in node II (for example, foxA) we used all mouse paralogues (for example, foxA1, foxA2 and foxA3). We then used the Mouse Gastrulation Data R package (v.1.8.0) to extract single-cell RNA-seq data for mouse embryos at stages E8.0 and E8.5 for all node target genes described above. We used the gene counts to calculate log normalized expression values (logNormCounts function of R package scuttle) to categorize genes as neuronal, endodermal or mesodermal (and others). Then we generated a pseudo-bulk expression matrix per mesodermal, endodermal and neuronal cell category for all genes and normalized the resulting counts to counts per million reads (cpm). Node genes were selected from this pseudo-bulk matrix and the log-fold-change of neuronal versus mesodermal expression or endodermal versus mesodermal expression was calculated as log, of neuronal (or endodermal) cpm + 1/ mesodermal cpm + 1.

**Reporting summary**

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The raw files from the ChIP–seq and RNA-seq experiments have been deposited to the NCBI GEO database under the accession
numbers GSE182573 and GSE198320 for *N. vectensis* and *S. purpuratus*, respectively.

**Code availability**

All scripts developed for this study are available at GitHub at https://github.com/technau/brachyury_grn, https://github.com/xxxmichixxx/MouseEmbryoSingleCell and https://github.com/fmi-basel/gbuehler-MiniChip.

**References**

1. Smith, J. C., Price, B. M. J., Green, J. B. A., Weigel, D. & Herrmann, B. G. Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response to mesoderm induction. *Cell* **67**, 79–87 (1991).

2. Cunliffe, V. & Smith, J. C. Ectopic mesoderm formation in Xenopus embryos caused by widespread expression of a Brachyury homologue. *Nature* **358**, 427–430 (1992).

3. Wilson, V., Manson, L., Skarnes, W. C. & Beddington, R. S. The T gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development* **121**, 877–886 (1995).

4. O’Reilly, M. A. J., Smith, J. C. & Cunliffe, V. Pattern of the mesoderm in Xenopus: dose-dependent and synergistic effects of Brachyury and Pintallavis. *Development* **121**, 1351–1359 (1995).

5. Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. & Lehrach, H. Cloning of the T gene required in mesoderm formation in the mouse. *Nature* **343**, 617–622 (1990).

6. Kispert, A., Ortner, H., Cooke, J. & Herrmann, B. G. The chick Brachyury gene: developmental expression pattern and response to axial induction by localized activin. *Dev. Biol.* **168**, 406–415 (1995).

7. Stemple, D. L. Structure and function of the notochord: an essential organ for chordate development. *Development* **132**, 2503–2512 (2005).

8. Wilkinson, D. G., Bhatt, S. & Herrmann, B. G. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657–659 (1990).

9. Beddington, R. S., Rashbass, P. & Wilson, V. Brachyury—a gene affecting mouse gastrulation and early organogenesis. *Dev. Suppl.* **165**, 157–165 (1992).

10. Stott, D., Kispert, A. & Herrmann, B. G. Rescue of the tail defect of Brachyury mice. *Genes Dev.* **7**, 197–203 (1993).

11. Conlon, F. L. & Smith, J. C. Interference with Brachury function inhibits convergent extension, causes apoptosis, and reveals separate requirements in the FGF and activin signalling pathways. *Dev. Biol.* **213**, 85–100 (1999).

12. Schulte-Merker, S., van Eeden, F. J. M., Halpern, M. E., Kimmel, C. B. & Nüsslein-Volhard, C. No tail (ntl) is the zebrafish homologue of Brachyury (T) is an immediate-early response to mesoderm induction. *Cell* **100**, 1009–1015 (1994).

13. Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N. & McMahon, A. P. T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev.* **13**, 3185–3190 (1999).

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

15. Wymeersch, F. J. et al. Position-dependent plasticity of distinct progenitor types in the primitive streak. *eLife* **5**, e10042 (2016).

16. Gouti, M. et al. A gene regulatory network balances neural and mesoderm specification during vertebrate trunk development. *Dev. Cell* **41**, 243–261 (2017).

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

18. Gouti, M. et al. In vitro generation of neureomesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol.* **12**, e1001937 (2014).

19. Henrique, D., Abranches, E., Verrier, L. & Storey, K. G. Neureomesodermal progenitors and the making of the spinal cord. *Development* **142**, 2864–2875 (2015).

20. Selleck, M. A. & Stern, C. D. Fate mapping and cell lineage analysis of Hensen’s node in the chick embryo. *Development* **112**, 615–626 (1991).

21. Koch, F. et al. Antagonistic activities of Sox2 and Brachyury control the fate choice of neuro-mesodermal progenitors. *Dev. Cell* **42**, 514–526 (2017).

22. Mugele, D. et al. Genetic approaches in mice demonstrate that neuro-mesodermal progenitors express T/Brachyury but not Sox2. Preprint at bioRxiv https://doi.org/10.1101/503854 (2018).

23. Takahashi, H. et al. Brachyury downstream notochord differentiation in the ascidian embryo. *Genes Dev.* **13**, 1519–1523 (1999).

24. Holland, P. W. H., Koschorz, B., Holland, L. Z. & Herrmann, B. G. Conservation of Brachyury (T) genes in amphioxus and vertebrates: developmental and evolutionary implications. *Development* **121**, 4283–4291 (1995).

25. Satoh, N. et al. A preliminary single-cell RNA-seq analysis of embryonic cells that express Brachyury in the amphioxus, *Branchiostoma japonicum*. *Front. Cell Dev. Biol.* **9**, 696875 (2021).

26. Technau, U. Brachyury, the blastopore and the evolution of the mesoderm. *BioEssays* **23**, 788–794 (2001).

27. Croce, J. et al. Wnt6 activates endoderm in the sea urchin gene regulatory network. *Development* **138**, 3297–3306 (2011).

28. Harada, Y., Yasuo, H. & Satoh, N. A sea urchin homologue of the chordate Brachyury (T) gene is expressed in the secondary mesenchyme founder cells. *Development* **121**, 2747–2754 (1995).

29. Shoguchi, E., Satoh, N. & Maruyama, Y. K. Pattern of Brachyury gene expression in starfish embryos resembles that of hemichordate embryos but not of sea urchin embryos. *Mech. Dev.* **82**, 185–189 (1999).

30. Croce, J., Lhomond, G. & Gache, C. Expression pattern of Brachyury in the embryo of the sea urchin *Paracentrotus lividus*. *Dev. Genes Evol.* **211**, 617–619 (2001).

31. Annunziata, R. & Arnone, M. I. A dynamic regulatory network explains ParaHox gene control of gut patterning in the sea urchin. *Development* **141**, 2462–2472 (2014).

32. Gross, J. M. & McClay, D. R. The role of Brachyury (T) during gastrulation movements in the sea urchin *Lytechinus variegatus*. *Dev. Biol.* **239**, 132–147 (2001).

33. Arndt, D., Technau, U. & Wittbrodt, J. Evolution of the bilaterian larval foregut. *Nature* **409**, 81–85 (2001).

34. Boyle, M. J., Yamaguchi, E. & Seaver, E. C. Molecular conservation of metazoan gut formation: evidence from expression of endomesoderm genes in *Capitella teleta* (Annelida). *Evodevo* **5**, 39 (2014).

35. Kin, K., Kakoi, S. & Wada, H. A novel role for dpp in the shaping of bivalve shells revealed in a conserved molluscan developmental program. *Dev. Biol.* **329**, 152–166 (2009).

36. Lartillot, N., Lespinet, O., Vervoort, M. & Adoutte, A. Expression pattern of Brachyury in the mollusc *Patella vulgata* explains ParaHox gene control of gut patterning in the sea urchin. *Development* **121**, 1009–1015 (2001).

37. Technau, U. & Scholz, C. B. Origin and evolution of endoderm and mesoderm. *Int. J. Dev. Biol.* **47**, 531–539 (2003).

38. Technau, U. & Bode, H. R. HyBrA, a Brachyury homologue, acts during head formation in Hydra. *Development* **126**, 999–1010 (1999).
39. Technau, U. et al. Parameters of self-organization in Hydra aggregates. Proc. Natl Acad. Sci. USA 97, 12127–12131 (2000).
40. Scholz, C. B. & Technau, U. The ancestral role of Brachury expression of NemBr1 in the basal cnidarian Nemastella vectensis (Anthozoa). Dev. Genes Evol. 212, 563–570 (2003).
41. Servetnick, M. D. et al. Cas9-mediated excision of Nemastella Brachury disrupts endoderm development, pharynx formation and oral–aboral patterning. Development 144, 2951–2960 (2017).
42. Yasuoka, Y., Shinnato, C. & Satoh, N. The mesoderm-forming gene brachury regulates ectoderm–endoderm demarcation in the coral Acropora digitafer. Curr. Biol. 26, 2885–2892 (2016).
43. Marcellini, S., Technau, U., Smith, J. C. & Lemaire, P. Evolution of Brachury proteins: identification of a novel regulatory domain conserved within Bilateria. Dev. Biol. 260, 352–361 (2003).
44. Bielen, H. et al. Divergent functions of two ancient Hydra Brachury paralogues suggest specific roles for their C-terminal domains in tissue fate induction. Development 134, 4187–4197 (2007).
45. Schweiger, M. et al. Evolutionary conservation of the eumetazoan gene regulatory landscape. Genome Res. 24, 639–650 (2014).
46. Gentsch, G. E. et al. InVivo T-Box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. Cell Rep. 4, 1185–1196 (2013).
47. Putnam, N. H. et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317, 86–94 (2007).
48. Zimmermann, B. et al. Sea anemone genomes reveal ancestral metazoan chromosomal macrosynteny. Preprint at bioRxiv https://doi.org/10.1101/2020.10.30.359448 (2021).
49. Grossman, S. R. et al. Positional specificity of different transcription factor classes within enhancers. Proc. Natl Acad. Sci. USA 115, E7222–E7230 (2018).
50. Müller, C. W. & Herrmann, B. C. Genetic and morphological structure of the T-domain-DNA complex of the Brachury transcription factor. Nature 389, 884–888 (1997).
51. Kispert, A. & Herrmann, B. G. The Brachury gene encodes a novel DNA binding protein. EMBO J. 12, 3211–3220 (1993).
52. Sebé-Pedrós, A. et al. The dynamic regulatory genome of capsaspora and the origin of animal multicellularity. Cell 165, 1224–1237 (2016).
53. Sebé-Pedrós, A. et al. Early evolution of the T-box transcription factor family. Proc. Natl Acad. Sci. USA 110, 16050–16055 (2013).
54. Garrick, R. J. et al. Lineage tracing of neuromesodermal progenitors reveals novel wnt-dependent roles in trunk progenitor cell maintenance and differentiation. Development 142, 1628–1638 (2015).
55. Dermitzakis, E. T. & Clark, A. G. Evolution of transcription factor binding sites in mammalian gene regulatory regions: conservation and turnover. Mol. Biol. Evol. 19, 1114–1121 (2002).
56. Pijuan-Sala, B. et al. A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 566, 490–495 (2019).
57. Grosswendt, S. et al. Epigenetic regulator function through mouse gastrulation. Nature 584, 102–108 (2020).
58. Tosic, J. et al. Eomes and Brachury control pluripotency exit and germ-layer segregation by changing the chromatin state. Nat. Cell Biol. 21, 1518–1531 (2019).
59. Goto, H., Kimmey, S. C., Row, R. H., Matus, D. Q. & Martin, B. L. FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition. Development 144, 1412–1421 (2017).
60. Martin, B. L. & Kimelman, D. Brachury establishes the embryonic neuromesodermal progenitor niche. Genes Dev. 24, 2778–2783 (2010).
61. Davidson, E. H. et al. A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev. Biol. 246, 162–190 (2002).
62. Kasupower, A. et al. Unexpected complexity of the Wnt gene family in a sea anemone. Nature 433, 156–160 (2005).
63. Logan, C. Y., Miller, J. R., Ferkowicz, M. J. & McClay, D. R. Nuclear beta-catenin is required to specify vegetal cell fates in the sea urchin embryo. Development 126, 345–357 (1999).
64. Lebedeva, T. et al. Cnidarian–bilaterian comparison reveals the ancestral regulatory logic of the β-catenin dependent axial patterning. Nat. Commun. 12, 4032 (2021).
65. Röttinger, E., Dahlin, P. & Martindale, M. Q. A framework for the establishment of a cnidarian gene regulatory network for ‘endomesoderm’ specification: the inputs of β-catenin/TCF signaling. PLoS Genet. 8, e1003164 (2012).
66. Kirillova, A. et al. Germ-layer commitment and axis formation in sea anemone embryonic cell aggregates. Proc. Natl Acad. Sci. USA 115, 1813–1818 (2018).
67. Tu, Q., Brown, C. T., Davidson, E. H. & Oliveri, P. Sea urchin Forkhead gene family: phylogeny and embryonic expression. Dev. Biol. 300, 49–62 (2006).
68. Cui, M., Siro, L., Li, E., Davidson, E. H. & Peter, I. S. Specific functions of the Wnt signaling system in gene regulatory networks throughout the early sea urchin embryo. Proc. Natl Acad. Sci. USA 111, E5029–E5038 (2014).
69. Arenas-Mena, C., Cameron, R. A. & Davidson, E. H. Hindgut specification and cell-adhesion functions of Sphox1/13b in the endoderm of the sea urchin embryo. Dev. Growth Differ. 48, 463–472 (2006).
70. Peter, I. S. & Davidson, E. H. The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage. Dev. Biol. 340, 188–199 (2010).
71. Rizzo, F., Fernandez-Serra, M., Squarzon, P., Archimandritis, A. & Arnone, M. I. Identification and developmental expression of the ets gene family in the sea urchin (Strongylocentrotus purpuratus). Dev. Biol. 300, 35–48 (2006).
72. McClay, D. R., Miranda, E. & Feinberg, S. L. Neurogenesis in the sea urchin embryo is initiated uniquely in three domains. Development 145, dev16742 (2018).
73. Wickramayake, A. H. et al. Nuclear β-catenin-dependent Wnt8 signaling in vegetal cells of the early sea urchin embryo regulates gastrulation and differentiation of endoderm and mesodermal cell lineages. Genesis 39, 194–205 (2004).
74. Nakashishi, N., Renfer, E., Technau, U. & Rentzsch, F. Nervous systems of the sea anemone Nemastella vectensis are generated by ectoderm and endoderm and shaped by distinct mechanisms. Development 139, 347–357 (2012).
75. Layden, M. J. & Martindale, M. Q. Non-canonical Notch signaling represents an ancestral mechanism to regulate neural differentiation. Evodevo 5, 30 (2014).
76. Layden, M. J., Boekhout, M. & Martindale, M. Q. Nemastella vectensis achaete-scute homolog NvashA regulates embryonic ectodermal neurogenesis and represents an ancient component of the metazoan neural specification pathway. Development 139, 1013–1022 (2012).
77. Rentzsch, F., Layden, M. & Manuel, M. The cellular and molecular basis of cnidarian neurogenesis. WIREs Dev. Biol. 6, e257 https://doi.org/10.1002/wde2.257 (2017).
78. Angerer, L. M. et al. Sea urchin goosecoid function links fate specification along the animal–vegetal and oral–aboral embryonic axes. Development 126, 4393–4404 (2001).
79. Wei, Z., Yaguchi, J., Yaguchi, S., Angerer, R. C. & Angerer, L. M. The sea urchin animal pole domain is a Six3-dependent neurogenic patterning center. Development 136, 1179–1189 (2009).
80. Range, R. C. Canonical and non-canonical Wnt signaling pathways define the expression domains of Frazzled 5/8 and Frazzled 1/2/7 along the early anterior–posterior axis in sea urchin embryos. *Dev. Biol.* **444**, 83–92 (2018).

81. Anishchenko, E., Arnone, M. I. & D’Aniello, S. SoxB2 in sea urchin development: implications in neurogenesis, ciliogenesis and skeletal patterning. *EvoDevo* **9**, 5 (2018).

82. Duboc, V. et al. Nodal and BMP2/4 pattern the mesoderm and endoderm during development of the sea urchin embryo. *Development* **137**, 223–235 (2010).

83. Röttinger, E. et al. FGF steps guide migration of mesenchymal cells, control skeletal morphogenesis and regulate gastrulation during sea urchin development. *Development* **135**, 353–365 (2008).

84. Otim, O., Amore, G., Minokawa, T., McClay, D. R. & Davidson, E. H. A single-cell RNA-seq analysis of embryos with mutations in T-box gene Tbx6. *Nucleic Acids Res.* **46**, D260–D266 (2018).

85. Satoh, N. et al. Functional FGF binding sites establish a temporal read-out of gene expression in the Ciona notochord. *PLoS Biol.* **11**, e1001697 (2013).

86. Kubo, A. et al. Genomic cis-regulatory networks in the early Ciona intestinalis embryo. *Development* **137**, 1613–1623 (2010).

87. Khan, A. et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46**, 3375–3387 (2017).

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

89. Griffin, K. J., Amacher, S. L., Kimmel, C. B. & Kimelman, D. Molecular identification of spadetail: regulation of zebraloph trunk and tail mesoderm formation by T-box genes. *Development* **125**, 3379–3388 (1998).

90. Chapman, D. L., Cooper-Morgan, A., Harrelson, Z. & Papaioannou, V. E. Critical role for Tbx6 in mesoderm specification in the mouse embryo. *Mechanisms Dev.* **120**, 837–847 (2003).

91. Chapman, D. L., Agulnik, I., Hancock, S., Silver, L. M. & Papaioannou, V. E. Tbx6 is a mouse T-box gene implicated in paraxial mesoderm formation at gastrulation. *Dev. Biol.* **180**, 534–542 (1996).

92. Takatori, N. et al. T-box genes in the ascidian Ciona intestinalis: characterization of cDNAs and spatial expression. *Dev. Dynam.* **230**, 743–753 (2004).

93. Sebé-Pedrós, A. & Ruiz-Trillo, I. Evolution and classification of the T-box transcription factor family. *Curr. Top. Dev. Biol.* **122**, 1–26 (2017).

94. Sebe-Pedros, A., de Mendoza, A., Lang, B. F., Degnan, B. M. & Ruiz-Trillo, I. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan Capsaspora owczarzaki. *Mol. Biol. Evol.* **28**, 1241–1254 (2011).

95. Arnold, S. J. et al. Brachury is a target gene of the Wnt/β-catenin signaling pathway. *Mechanisms Dev.* **91**, 249–258 (2000).

96. Sethi, A. J., Wikramanyake, R. M., Angerer, R. C., Range, R. C. & Angerer, L. M. Sequential signaling crosstalk regulates endomesoderm segregation in sea urchin embryos. *Science* **335**, 590–593 (2012).

97. Niehrs, C. On growth and form: a Cartesian coordinate system of Wnt and BMP signaling specifies bilaterian body axes. *Development* **137**, 845–857 (2010).

98. Andrikou, C., Jovene, E., Rizzo, F., Oliveri, P. & Arnone, M. I. Myogenesis in the sea urchin embryo: the molecular fingerprint of the myoblast precursors. *EvoDevo* **4**, 33 (2013).

99. Sethi, A. J., Wikramanyake, R. M., Angerer, R. C., Range, R. C. & Angerer, L. M. Sequential signaling crosstalk regulates endomesoderm segregation in sea urchin embryos. *Science* **335**, 590–593 (2012).
118. Varet, H., Brillet-Guéguen, L., Coppée, J.-Y. & Dillies, M.-A. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 15, 550 (2014).

119. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 40, 4288–4297 (2012).

120. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419 (2017).

121. Genikovich, G. & Technau, U. In situ hybridization of starlet sea anemone (Nematostella vectensis) embryos, larvae, and polyps. Cold Spring Harb. Protoc. 4, pdb.prot5282 (2009).

122. Perillo, M., Paganos, P., Spurrell, M., Arnone, M. I. & Wessel, G. M. in Developmental Biology of the Sea Urchin and Other Marine Invertebrates: Methods and Protocols: 2219, 195–216 (Humana Press, 2021).

123. Takacs, C. M. et al. Expression of an NK2 homeodomain gene in the apical ectoderm defines a new territory in the early sea urchin embryo. Dev. Biol. 269, 152–164 (2004).

124. Garner, S. et al. Neurogenesis in sea urchin embryos and the diversity of deuterostome neurogenic mechanisms. Development 143, 286–297 (2016).

125. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780 (2013).

126. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195 (2013).

127. Roth, A. C. J., Gonnet, G. H. & Dessizmoz, C. Algorithm of OMA for large-scale orthology inference. BMC Bioinf. 9, 518 (2008).

128. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 15, 550 (2014).

129. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014).

130. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 41, e108 (2013).

131. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419 (2017).

132. Perillo, M., Paganos, P., Spurrell, M., Arnone, M. I. & Wessel, G. M. in Developmental Biology of the Sea Urchin and Other Marine Invertebrates: Methods and Protocols: 2219, 195–216 (Humana Press, 2021).

133. Takacs, C. M. et al. Expression of an NK2 homeodomain gene in the apical ectoderm defines a new territory in the early sea urchin embryo. Dev. Biol. 269, 152–164 (2004).

134. Garner, S. et al. Neurogenesis in sea urchin embryos and the diversity of deuterostome neurogenic mechanisms. Development 143, 286–297 (2016).

135. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780 (2013).

136. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195 (2013).

137. Roth, A. C. J., Gonnet, G. H. & Dessizmoz, C. Algorithm of OMA for large-scale orthology inference. BMC Bioinf. 9, 518 (2008).

138. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014).

139. Carlsson, P. & Mahlapuu, M. Forkhead transcription factors: key players in development and metabolism. Dev. Biol. 250, 1–23 (2002).

Acknowledgements

We are grateful to L. Blitz for critically reading the manuscript. We would like to thank the members of the Technau and the Arnone laboratory for discussions. This work was supported by the EU-ITN EVONET to U.T. and M.I.A. and by grants of the Austrian Science Fund FWF to U.T. (P34404, P31018) and by EU-ITN EVOCELL (grant no. 766053) to M.I.A.

Author contributions

U.T. and M.I.A. conceived the study. M.S., C.A., P.F.M., P.P. and T.L. performed experiments. M.S., R.D., D.V., B.Z., H.A.S., G.G., G.B., M.I.A. and U.T. analysed the data. M.S., C.A., R.D., M.I.A. and U.T. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-022-01905-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-022-01905-w.

Correspondence and requests for materials should be addressed to Maria Ina Arnone or Ulrich Technau.

Peer review information Nature Ecology & Evolution thanks Noriyuki Satoh, Jesse Veenvliet and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | Specificity of the Brachyury antibodies and target gene detection strategy. Specificity of the Brachyury antibodies and target gene detection strategy. (A) Ectopic expression of Brachyury at aboral pole of Nematostella confirms the specificity of the antibody. Zygotes were injected with a plasmid of EF1a::mCherry-p2A-Brachyury. Embryos with mosaic expression were stained for mCherry and Brachyury antibodies. (B) Western Blot of Brachyury antibody in control and Morpholino-mediated knockdown of brachyury in Nematostella (C) Immunocytochemistry of gastrula stage embryos of Nematostella (oral views) in controls and after morpholino-mediated knockdown of Brachyury. (D) Quantitative summary of Brachyury knockdown with morpholino oligonucleotides. (E) Western blot of anti-Brachyury in Strongylocentrotus. The estimated protein size is approx. 50kD. Developmental stages tested: 6h, 12h, 48h, 72h. Recombinant protein (RP) size: 9 kD. (F, G) Heatmap of Brachyury binding sites in Bra_AB1 and Bra_AB2 of Nematostella and Strongylocentrotus related to the TSS of all target genes. (H) Summary of the sources of the datasets used in this study.
Extended Data Fig. 2 | See next page for caption.
Extended Data Fig. 2 | Brachyury binding sites and relationship to chromatin modifications. Brachyury binding sites and relationship to chromatin modifications. (A) Heatmap of Nematostella Brachyury binding sites from this study with chromatin modification sites earlier identified in Schwaiger et al. 2014. (B) Venn diagram showing the overlap of Brachyury binding sites in Nematostella identified in this study overlaps with the enhancer/promoter sites previously identified in51. (C) Venn diagram of Brachyury binding sites identified in this study in Strongylocentrotus showing the overlap with the previously identified open chromatin sites identified with ATAC-seq data. (D) Brachyury target selection strategy. Two closest genes on either side of the binding site were considered and their respective orthologs in the species under study were identified. A gene was prioritized as a target, if it was also a target gene in one or more species.
Extended Data Fig. 3 | PCA analyses and summary of DEG analyses of RNA-seq experiments. (A–D) Principal Component Analysis (PCA) of RNA-seq datasets in Nematostella, Strongylocentrotus, Xenopus and Mouse respectively. The blue dots represent the control/WT samples while as red dots indicate KD/KO samples. In case of Nematostella both pre-mRNA splicing (spl) and translation-blocking (tra) morpholinos were used. (E) Summary of differentially expressed genes after Bra KD/KO. Differential expression analysis was done using DEseq2 R package with 0.05 alpha value. Data for Nematostella, Strongylocentrotus and Xenopus is from morpholino induced knockdown of Bra transcripts while mouse data is a result of Bra knockout. (F) Overlap of direct (ChIP–seq detected) and indirect targets (DEGs) across different species. Each species was used as a query species (query dataset) and the genes determined that are differentially expressed, and also ChIP targets (‘column ‘direct targets in query’) or not ChIP targets (column ‘indirect targets in query’). The numbers refer to the number of orthologs in each one of the other species. The numbers are low, since the overlap between ChIP–seq targets and DEGs from RNA-seq experiments is only 1–10% within a given species.
Extended Data Fig. 4 | See next page for caption.
Extended Data Fig. 4 | Motif analyses of Brachyury ChIP-peaks and Gene Ontology analyses of target genes. Motif analyses of Brachyury ChIP-peaks and Gene Ontology analyses of target genes. (A) Brachyury ChIP metaplots around Brachyury peaks containing palindrome or half-palindrome (single) or no Brachyury binding motifs.

The average read count (normalized to a million reads) was calculated for Brachyury ChIP–seq reads for regions around peak summits spanning 2kb in 20bp windows. The shaded area around the lines represents the 95% confidence interval across peaks in a category. Note that in all four species, peaks with palindromes show significantly higher ChIP–seq read counts, which may serve as a proxy for the strength of Brachyury binding. (B) Intersection of motifs with the peak region. The Brachyury peak region (as identified using the Peakzilla algorithm) was scanned for presence of other transcription factor binding sites using Fimo (MEME-ChIP suite) with default settings. The resulting binding motifs were grouped by the transcription factor families Paired box (Pax), basic helix–loop–helix (bHLH), Forkhead box (Fox), homeobox (homeo), High mobility group (hmg), T-box (tbox). Presence of these motifs together with the Brachyury motif is highlighted in orange in the upset plots (iii, vi, ix, xii). Distance of these motifs from the peak centre was also tracked and is shown in adjoining plots (i, ii, iv, v, vii, viii, x, xi). (C) ChIP/DE gene set overlap and GO analysis. Overlap (dark grey) between ChIP targets (black) and differentially expressed (light grey) genes in M. musculus (i), X. tropicalis (iv), S. purpuratus (vi), N. vectensis (x). In C. owczarzaki (xiii) only the number of ATAC-seq peaks with Bra motifs is shown. Gene ontology analysis of Brachyury ChIP targets for M. musculus (ii), X. tropicalis (v), S. purpuratus (vii), N. vectensis (xi) and C. owczarzaki (xiv) and of differentially expressed genes after Bra KO/KD (iii, vi, ix, xii). Only gene ontology terms for biological process and molecular function were reported. The colour of the dot represents the score (−log(p-value)) assigned by topGO while the size of the dot represents the number of genes associated with the term. (D) GO analyses of target genes of peaks with or without a Bra consensus motif. Note that no significant difference can be detected.
Extended Data Fig. 5 | Expression analysis of Brachyury ChIP targets by WMISH after knockdown in Nematostella. Expression analysis of Brachyury ChIP targets by WMISH after knockdown in Nematostella. (A) Morpholino-mediated knockdown of target genes with complex expression pattern show partial down or upregulation in the ectodermal layer. (B) Neuronal target genes that are not affected by Brachyury knockdown. Note that all unaffected genes are expressed in the inner layer of the embryo. (C) Knockdown of SoxB1 (a homologue of vertebrate Sox2) shows no effect on neuronal target genes regulated by Brachyury.
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Expression analyses of Brachyury target genes in Strongylocentrotus upon knockdown. Expression analyses of Brachyury target genes in Strongylocentrotus upon knockdown. (A) ChIP target genes that are not affected by Brachyury knockdown. (B) Expression analysis of Brachyury RNA-seq targets by WISH after morpholino induced knockdown. (C) Expression analysis of Brachyury RNA-seq targets by immunohistochemistry after morpholino induced knockdown. Arrows show the embryonic domain in which we see an effect (red arrow: mesodermally derived; blue ectodermally derived).

(D) Differentially expressed genes after Morpholino induced knockdown that are also ChIP targets at 24h. Key genes playing a crucial role in endoderm development are highlighted in yellow, key genes playing a role in mesoderm development are highlighted in red while key genes playing a role in ectoderm development are highlighted in blue. Asterisks indicate genes that are also ChIP targets. l/v: lateral view; v/v vegetal view; o/v oral view; a/v aboral view. up: upregulated gene; down: downregulated gene. The scale bar is 20 µm.
Extended Data Fig. 7 | Shared Brachyury targets between lineages of Metazoa, Bilateria, Chordata and Vertebrata. Shared Brachyury targets between lineages of Metazoa, Bilateria, Chordata and Vertebrata. Upset plot of shared orthologous genes as detected by OMA between different lineage combinations. Note that brachyury is the only target gene found in all investigated organisms. The large number of shared target genes between mouse and Xenopus indicates that this screen is robust against slight differences in developmental staging, source of cells, experimental design and sensitivity.
Extended Data Fig. 8 | See next page for caption.
Extended Data Fig. 8 | Protein phylogenies of selected target genes. (A) T-box family phylogenetic tree. T-box family tree constructed with T-box genes from Apis mellifera (Ame), Branchiostoma floridae (Bfl), Capsaspora owczarzaki (CAOG), Ciona intestinalis (Ciona), Lottia gigantean (Lgi), Mus musculus (Mmus), Nematostella vectensis (NVE), Strongylocentrotus purpuratus (Spu), Xenopus tropicalis (Xtro). This protein maximum likelihood tree was constructed with T-box genes from Apis mellifera (Ame), Branchiostoma floridae (Bfl), Capsaspora owczarzaki (CAOG), Ciona intestinalis (Ciona), Lottia gigantean (Lgi), Mus musculus (Mmus), Nematostella vectensis (NVE), Strongylocentrotus purpuratus (Spu), Xenopus tropicalis (Xtro). The tree was rooted on a T-box gene from the fungus Paramicrosporidium sacamoeae. Brachyury target genes in Nematostella and Strongylocentrotus are indicated by arrows. (B) Sox family phylogenetic tree. Sox family tree constructed with Sox genes from Apis mellifera (Ame), Acropora millepora (Ami), Amphimedon queenslandica (Aq), Ciona intestinalis (ci), Mus musculus (mmu), Nematostella vectensis (NVE), Strongylocentrotus purpuratus (Spu), Xenopus tropicalis (Xtro). This tree was constructed with a maximum likelihood method with 1000 UFboot samples, the values at the nodes represent the support values. UFboot values below 50% are not shown and the nodes are marked with a red circle. The values at nodes with 100% support are also not shown. The tree was rooted with a sponge A. queenslandica Sox protein. Brachyury target genes in Nematostella and Strongylocentrotus are indicated by arrows. (C) Zic family phylogeny. ZIC family tree constructed from ZIC genes from Amphimedon queenslandica (Aque), Branchiostoma floridana (Bflfo), Capitella teleta (Ctel), Capsaspora owczarzaki (Cowe), Ciona intestinalis (ci), Drosophila melanogaster (Dme), Homo sapiens (Hsap), Mus musculus (Mmus), Nematostella vectensis (Nve), Strongylocentrotus purpuratus (Spu), Tribolium castaneum (Tcas), Xenopus tropicalis (Xtro). This protein maximum likelihood tree was constructed using IQ-Tree with 10000 UFboot samples. The values at the nodes represent the UFboot support, where values below 50% are not shown. The tree is shown midpoint-rooted with Figtree. Sequences from Nematostella are marked bold in green, those from Strongylocentrotus in bold and blue. Brachyury target genes in Nematostella and Strongylocentrotus are indicated by arrows. (D) TFAP2 family phylogeny. TFAP2 family tree constructed from TFAP2 genes from Amphimedon queenslandica (Aque), Branchiostoma floridana (Bflfo), Capitella teleta (Ctel), Ciona intestinalis (ci), Drosophila melanogaster (Dme), Homo sapiens (Hsap), Mus musculus (Mmus), Nematostella vectensis (Nve), Strongylocentrotus purpuratus (Spu), Tribolium castaneum (Tcas), Xenopus tropicalis (Xtro). This protein maximum likelihood tree was constructed using IQ-Tree with 10000 UFboot samples. The values at the nodes represent the UFboot support, where values below 50% are not shown. The tree is shown midpoint-rooted with Figtree. Sequences from Nematostella are marked bold in green, those from Strongylocentrotus in bold and blue. Brachyury target genes in Nematostella and Strongylocentrotus are indicated by arrows.
Extended Data Fig. 9 | Expression of apomorphic and synapomorphic target genes of Brachyury in mouse E8.5 neuronal, endodermal, and mesodermal cell types. Expression of apomorphic and synapomorphic target genes of Brachyury in mouse E8.5 neuronal, endodermal, and mesodermal cell types. This is the same analysis as shown in Figure 6, except that definitive endoderm cells were removed from the analysis (A–C) or the single-cell gene expression dataset from Grosswendt et al.121 (E, F) was used. A–C: The expression of the target genes in neuronal versus mesodermal cell types was annotated using single-cell RNA-seq data from E8.0 and E8.5 mouse embryos, as in Figure 6. However, since definitive endoderm cells in this dataset express notochord marker genes, we decided to remove these cells from the analysis before continuing as in Figure 6. (A) Heatmap of the log2 fold change (logFC) of neuronal (turquoise) or endodermal (yellow) vs mesodermal (purple) gene expression for each gene using the corresponding mouse gene symbols as they appeared in the single-cell dataset in each node (see Methods for details). (B) Boxplot of the log2 fold change (y-axis) of endodermal vs mesodermal gene expression for all genes per node (x-axis). Note that node II is enriched in endodermal expression while node V is enriched in mesodermal expression (p-value = 0.003, Wilcoxon Rank Sum Test). (C) Boxplot of the log2 fold change (y-axis) of neuronal vs mesodermal gene expression for all genes per node (x-axis). Note that node II is enriched in neuronal expression while node V is enriched in mesodermal expression (p-value = 0.008, Wilcoxon Rank Sum Test). D–F: The expression of the target genes in neuronal versus mesodermal cell types was annotated using single-cell RNA-seq data from E8.5 mouse embryos (read counts per gene/cell from GEO accession GSE122187). To annotate cells as neuronal, endodermal or mesodermal, we used the information from Supplementary Table 2 of Grosswendt et al. For endodermal, we used Lineage = Eendo, for mesodermal we used Lineage = Emeso, and for neuronal we used the following cell states: 1,11,24, and 39, which, according to Supplementary Fig. II of Grosswendt et al corresponds to: neural ectoderm anterior, neural ectoderm posterior, fore/midbrain, and future spinal cord. (D) Heatmap of the log2 fold change (logFC) of neuronal (turquoise) or endodermal (yellow) vs mesodermal (purple) gene expression for each gene using the corresponding mouse gene symbols as they appeared in the single-cell dataset in each node (see Methods for details). (E) Boxplot of the log2 fold change (y-axis) of neuronal vs mesodermal gene expression for all genes per node (x-axis). Note that node II is enriched in neuronal expression while node V is enriched in mesodermal expression (p-value = 0.037, Wilcoxon Rank Sum Test). (F) Boxplot of the log2 fold change (y-axis) of endodermal vs mesodermal gene expression for all genes per node (x-axis). Note that when, as is the case in this dataset, ‘gut’ is the only annotated endodermal cell type, node II is not more enriched in endodermally expressed genes compared to node V (p-value = 0.23, Wilcoxon Rank Sum Test). The boxes range from the 25th to the 75th percentile and the horizontal lines represent the median. Outliers are shown as dots.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|-----|-----------|
| ☑   | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☑   | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☑   | The statistical test(s) used AND whether they are one- or two-sided |
| ☑   | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
| ☑   | A description of all covariates tested |
| ☑   | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☑   | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☑   | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted |
| ☑   | Give P values as exact values whenever suitable. |
| ☑   | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☑   | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☑   | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

| Data collection |
|-----------------|
| https://github.com/dnyansagar/gene_regulatory_network |
| https://github.com/xxxxmichixxx/MouseEmbryoSingleCell |

| Data analysis |
|---------------|
| https://github.com/dnyansagar/gene_regulatory_network |
| https://github.com/xxxxmichixxx/MouseEmbryoSingleCell |
| https://github.com/fmi-basel/gbuehler-MiniChip |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.
Data

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The raw files from the ChIP-seq and RNA-seq experiments have been deposited to the NCBI GEO database under the accession number GSE182573 and GSE198320 for N. vectensis and S. purpuratus respectively.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender
Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.

Population characteristics
Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment
Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight
Identify the organization(s) that approved the study protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size
Sample sizes of 2 replicates for ChIP seq are common and showed good correlation of peaks. For RNAseq we used 6 replicates for Nematostella and 3 for sea urchin which was sufficient when analyzing the data using PCA.

Data exclusions
No exclusions.

Replication
All replications were successful according to correlations/PCA. One ChIP replicate of the sea urchin Brachyury chip was less strong compared to the other, but still agreed very well where strong enough.

Randomization
Embryos were randomly allocated to control or Brachyury morpholino groups.

Blinding
Blinding was not necessary for deep sequencing data analysis since all samples were processed in parallel automatically.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.
Antibodies

A 6xHis-tagged full-length clone of Nematostella Brachyury was expressed in E. coli strain BL21. After transformation, bacteria were grown overnight in LB supermedium + Ampicillin. Then, 10 ml LB medium / Ampicillin were inoculated with 500 μl of overnight culture, grown for 3 h (O.D. 600 = 0.5-0.6), IPTG was added to a final concentration of 1 mM, and bacteria culture was incubated for 21 h at 4°C. Brachyury protein was purified using Ni-NTA agarose (Qiagen), followed by polyacrylamide gel electrophoresis. The Brachyury band at ~50 kDa was excised and used to inject two rabbits (PRIMM Biotech). For Strongylocentrotus Brachyury, two polyclonal antibodies from 2 different rabbits were raised against the recombinantly produced C-terminal domain excluding the T-box domain. The antibodies were affinity purified by PRIMM Biotech prior to their use.

Validation

For both Nematostella and Strongylocentrotus the reactivity of polyclonal antibodies from each rabbit were assessed by both Western blot and immunohistochemistry, confirming their specificity (Supplemental Fig. 1).

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals Nematostella vectensis and Strongylocentrotus purpuratus. Gastrulation stage embryos.

Wild animals Study did not involve wild animals.

Reporting on sex No sex based analysis could be done in embryos.

Field-collected samples No field collections.

Ethics oversight Ethics oversight not required for Nematostella vectensis and Strongylocentrotus purpuratus.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

Confirm that both raw and final processed data have been deposited in a public database such as GEO.

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE198320
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE182573

Files in database submission

10435_gastrula.fastq
10436_gastrula.fastq
10437_Input_gastrula.fastq

Genome browser session (e.g. UCSC)

Not applicable.

Methodology

Replicates 2 biological replicates with polyclonal antibodies from different rabbits. Very good agreement of replicates based on called peaks and number of reads over peak regions.

Sequencing depth We obtained 39065464, 33397307 and 43959250 50bp single end reads for Nematostella Brachyury ChIP replicate 1, replicate 2, and Input, respectively. Of these, 21889877, 17420963, and 25679308 were uniquely mapped to the genome. We obtained 45753729, 52253224, 50361646 and 44089574 reads for Strongylocentrotus Brachyury ChIP replicate 1, replicate 2, Input replicate 1 and Input replicate 2, respectively. Of these, 18744583, 15413040, 21304793, and 18443726 were uniquely mapped to the genome.
| Antibodies | Brachyury antibodies as described above. For the ChIP replicates, antibodies from the two different rabbits was used for each. |
|---|---|
| Peak calling parameters | ChIP-seq peaks were called using Peakzilla [Bardet et al., 2013] for each replicate using the Input sequencing as control. Depending on the bimodal distribution of the reads, the Peakzilla algorithm assigns a score to each peak. We examined the relationship between Peakzilla peaks and their scores, and the coverage of ChIP-seq reads across the genomes, thereby determining the optimal score cutoff to filter high-quality peaks in each dataset, where the lowest scoring peaks that pass the cutoff are still visually detected as peaks in the coverage tracks. Consequently, we removed peaks with a score less than 2 for mouse and Xenopus, less than 0.5 for Strongylocentrotus and less than 5 for Nematostella. |
| Data quality | Heatmaps of Brachyury ChIP signal were generated to ensure that there is a correct distribution and enrichment of reads over called peaks in both replicates. Despite being of different overall signal strength, the replicates agreed very well with each other overall. |
| Software | https://github.com/dnyansagar/gene_regulatory_network  
https://github.com/fmi-basel/gbuehler-MiniChip |