CHAPTER ONE

Evolution and Classification of the T-Box Transcription Factor Family

A. Sebé-Pedrós*,[1], I. Ruiz-Trillo[†,‡,§]

*Weizmann Institute of Science, Rehovot, Israel
†Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Spain
‡Universitat de Barcelona, Barcelona, Spain
§Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain
[1]Corresponding author: e-mail address: arnau.sebe-pedros@weizmann.ac.il

Contents

Abstract

T-box proteins are key developmental transcription factors in Metazoa. Until recently they were thought to be animal specific and many T-box classes were considered bilaterian specific. Recent genome data from both early-branching animals and their closest unicellular relatives have radically changed this scenario. Thus, we now know that T-box genes originated in premetazoans, being present in the genomes of some extant early-branching fungi and unicellular holozoans. Here, we update the evolutionary classification of T-box families and review the evolution of T-box function in early-branching animals (sponges, ctenophores, placozoans, and cnidarians) and nonmodel bilaterians. We show that concomitant with the origin of Metazoa, the T-box family radiated into the major known T-box classes. On the other hand, while functional studies are

Current Topics in Developmental Biology, Volume 122
ISSN 0070-2153
http://dx.doi.org/10.1016/bs.ctdb.2016.06.004

still missing for many T-box classes, the emerging picture is that T-box genes have key roles in multiple aspects of development and in adult terminal cell-type differentiation in different animal lineages. A paradigmatic example is that of Brachyury, the founding member of the T-box family, for which several studies indicate a widely conserved role in regulating cell motility in different animal lineages and probably even before the advent of animal multicellularity. Overall, we here review the evolutionary history of T-box genes from holozoans to animals and discuss both their functional diversity and conservation.

1. AN UPDATED PHYLOGENETIC CLASSIFICATION OF THE T-BOX GENE FAMILY

The T-box genes are essential developmental transcription factors (TFs) in Metazoa (Papaioannou, 2001, 2014; Showell, Binder, & Conlon, 2004; Smith, 1999; Wilson & Conlon, 2002). This family is characterized by an evolutionary conserved DNA-binding domain of 180–200 amino acids, known as the T-box domain. The discovery of the first T-box genes (T/Brachyury) in mouse (Herrmann, Labeit, Poustka, King, & Lehrach, 1990), shown to be involved in mesoderm formation, was soon followed by the cloning of T/Brachyury in *Xenopus* (Smith, Price, Green, Weigel, & Herrmann, 1991) and zebrafish (Schulte-Merker, Ho, Herrmann, & Nusslein-Volhard, 1992) and the identification and characterization of members of other T-box classes (Bollag et al., 1994). A T/Brachyury ortholog was soon identified in nonvertebrate species, in particular in insects (Kispert, Herrmann, Leptin, & Reuter, 1994), where it was shown to have conserved expression in developing hindgut. It was also soon found that T/Brachyury encoded a DNA-binding protein involved in transcriptional regulation (Stott, Kispert, & Herrmann, 1993). The first T-box gene described outside Bilateria came in 1999, in which a Brachyury ortholog was identified in the cnidarian *Hydra magnipapillata* (Technau & Bode, 1999). In the early 2000s, several T-box genes were identified in diverse taxa among the earliest-branching metazoan lineages: Brachyury and Tbx2/3 in sponges (Adell, Grebenjuk, Wiens, & Müller, 2003; Adell & Müller, 2005; Larroux et al., 2006, 2008; Manuel, Le Parco, & Borchiellini, 2004), placozoans (Martinelli & Spring, 2003), and ctenophores (Martinelli & Spring, 2005; Yamada, Pang, Martindale, & Tochinai, 2007). Finally, in 2011 T-box genes were identified for the first time outside of animals, in both the filasterean *Capsaspora owczarzaki* (a close unicellular relative of animals) and in the chytrid fungi *Spizellomyces punctatus*

(an early-branching fungal species) (Sebé-Pedrós, de Mendoza, Lang, Degnan, & Ruiz-Trillo, 2011).

Here, we expand our previous survey and evolutionary classification of T-box TF (Sebé-Pedrós et al., 2013). To this end, we analyzed recently sequenced genomes/transcriptomes of key nonmetazoan and early-metazoan taxa (see Supplementary Table 1 (http://dx.doi.org/10.1016/bs.ctdb.2016.06.004)), including seven new ctenophore transcriptomes (Moroz et al., 2014), seven sponge transcriptomes (Riesgo, Farrar, Windsor, Giribet, & Leys, 2014), and five new early-branching fungal genomes (Chang et al., 2015), as well as an increased taxon sampling of bilaterian animals, like the hemichordate *Saccoglossus kowalevski* (Simakov et al., 2015), the myriapod *Strigamia maritima* (Chipman et al., 2014), or early-branching vertebrates like the coelacanth *Latimeria chalumnae* (Amemiya et al., 2013), the elephant shark *Callorhinchus milii* (Venkatesh et al., 2014), and the lamprey *Petromyzon marinus* (Smith et al., 2013). In total, we surveyed 153 complete genomes/transcriptomes representing all major eukaryotic clades, as well as a dense representation of the major metazoan lineages, with particular attention to early-branching animals. The results are summarized in Fig. 1, detailed in Supplementary Table 1, and discussed in the following sections.

2. PREMETAZOAN T-BOX GENES

Our new analysis shows that T-box genes are found in different zoosporic fungal groups, including the Cryptomycota *Rozella allomycis* (representing the earliest-branching fungal lineage (Chang et al., 2015)), the Neocallimastigomycota *Piromyces* sp., the chytrids *S. punctatus* and *Gonapodya prolifera*, and the Zygomycota *Rhizophagus irregularis* and *Mortierella verticillata*. We did not find any T-box genes outside the Opisthokonta, that is, the clade that comprises animals, fungi, and their unicellular relatives (Cavalier-Smith, 2003; Ruiz-Trillo, Roger, Burger, Gray, & Lang, 2008; Torruella et al., 2012). Thus, the T-box family evolved at the root of the Opisthokonta.

Within fungi, T-box genes were secondarily lost in Dikarya (that includes the two major groups of fungi, Ascomycota and Basidiomycota, and most known fungal species, like yeasts and most mushroom-forming fungi) and also in several early-branching fungal lineages, the exact number depending on the still poorly resolved deep nodes of the fungal tree (e.g., three extra losses according to the tree topology of Chang et al., 2015).

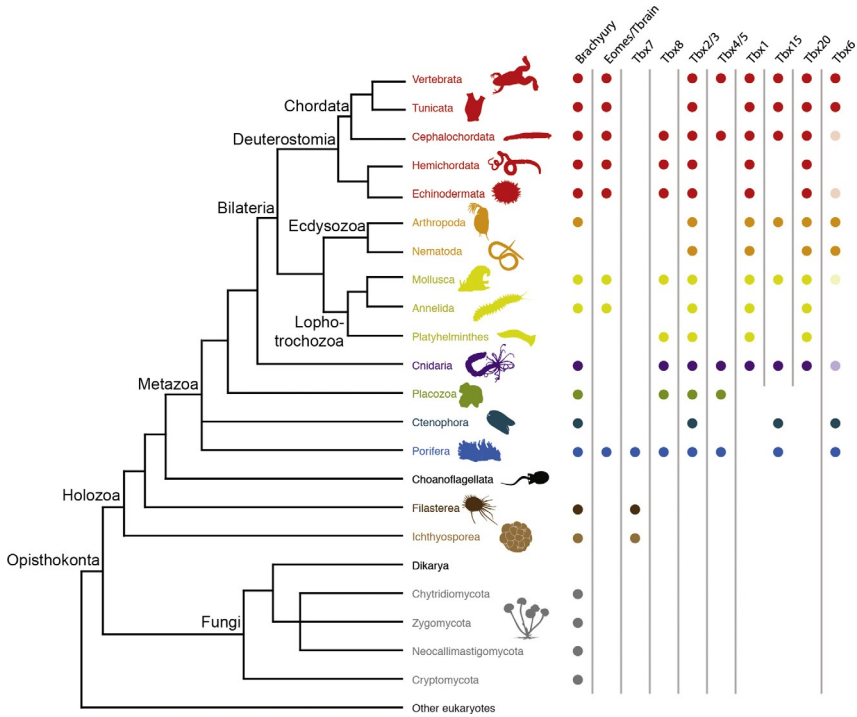

Fig. 1 Phylogenetic distribution of T-box classes. For details, see Supplementary Table 1.

Although highly divergent (Sebé-Pedrós et al., 2013), these fungal T-box genes cluster with the Brachyury family. This indicates that Brachyury is the founding member of the T-box class, being already present before the divergence of animals and fungi.

T-box genes are also present in the genomes of ichthyosporeans and filastereans, two groups of protists closely related to animals. Ichthyosporea have coenocytic development (Mendoza, Taylor, & Ajello, 2002; Suga & Ruiz-Trillo, 2013) and represent the earliest-branching clade inside Holozoa (animals plus three unicellular animal relative lineages) (Torruella et al., 2012). T-box genes are found in the six ichthyosporean taxa for which genomic or transcriptomic data are available, with up to seven copies per genomes (like in *Sphaeroforma arctica*). These ichthyosporean T-box sequences are extremely divergent and lack most of the known functional T-box domain amino acids, but some clearly group inside the Brachyury family (Fig. 1), while others belong to the recently defined Tbx7 family (Sebé-Pedrós et al., 2013). This means that at the root Holozoa (before

the divergence of animals and their immediate unicellular relatives) already two T-box classes existed: Brachyury and Tbx7. Similarly, we identified T-box genes from Brachyury and Tbx7 families in the genomes of the two known filasterean species: *C. owczarzaki* and *Ministeria vibrans*. *M. vibrans* is a marine free-living filopodiated amoeba with a flagellar stalk, while *C. owczarzaki* is a fresh-water filopodiated amoeba with an aggregative multicellular stage (Sebé-Pedrós et al., 2013). *C. owczarzaki* has the most conserved nonmetazoan Brachyury ortholog. It has most of the T-box key DNA-binding and dimerization amino acids, as well as conserved exclusive amino acid motifs of the Brachyury class (Sebé-Pedrós et al., 2013). Moreover, it also seems to be functionally conserved with its animal counterparts, as was shown in functional assays (see later). Finally, T-box genes were secondarily lost in the choanoflagellate lineage (Sebé-Pedrós et al., 2011, 2013).

An interesting finding among premetazoan T-box genes is a T-box with two T-domains present in *C. owczarzaki*. T-box TFs are in general composed of a single central T-domain, with the only known exception of the MGA family, a Tbx6 paralog in vertebrates (see later) with both a T-domain and a basic helix–loop–helix (bHLH) zipper domain (Hurlin, Steingrìmsson, Copeland, Jenkins, & Eisenman, 1999). Interestingly, *Capsaspora* has a large protein (1260 amino acids) with two full T-domains (one central and one C-terminal), which was verified by RT-PCR and RACE-PCR (Sebé-Pedrós et al., 2011). When analyzed separately, these two T-domains cluster inside the Tbx7 family, and one of them groups with a partial sequence of *M. vibrans* (the other filasterean species). This suggests that this unique "double T-box" configuration emerged at the root of the Filasterea. Although this "double T-box" configuration is not found in other species, the presence of two DNA-binding domains is not uncommon in other eukaryotic TF families. It has been hypothesized that multiple DNA-binding domains can increase the length and diversity of DNA motifs recognizable by the limited number of DNA-binding domain families (Charoensawan, Wilson, & Teichmann, 2010; Itzkovitz, Tlusty, & Alon, 2006). Whether this or other explanations account for the presence of this T-box gene in *C. owczarzaki* remains to be elucidated.

In summary, our expanded genomic survey confirms that T-box is an ancient TF that most likely originated early within the Opisthokonts and is present in diverse unicellular opisthokont lineages. T-box genes were subsequently secondarily lost in most fungal lineages and in choanoflagellates. Brachyury is the most ancient T-box class, while the Tbx7 class originated

in the common ancestor of Holozoa. The remaining classes, which are discussed later, appear to be animal specific.

3. METAZOAN T-BOX CLASSES AND FUNCTION

Our updated dataset (Fig. 1, Supplementary Table 1) shows that two major T-box classes emerged before Metazoa (Brachyury and Tbx7), others at the root of Metazoa (Eomes, Tbx1/15/20, Tbx2/3, Tbx4/5, Tbx6, and Tbx8), while some others originated at the Cnidaria + Bilateria clade or within Bilateria. The nonbilaterian sponges, which are an early (if not the earliest) branching animal lineage, have representatives of most T-box genes classes (except those specific to Bilateria or Cnidaria + Bilateria). Depending on the phylogenetic status of sponges this may suggest different scenarios. If sponges are the sister group to the rest of Metazoa, the so-called Porifera-sister hypothesis (Pick et al., 2010; Pisani et al., 2015), then the data will suggest a quick radiation of the T-box family concomitant with the origin of animal multicellularity, followed by secondary loss of some families in ctenophores. Alternatively, if ctenophores are the sister group to the rest of Metazoa (Moroz et al., 2014; Ryan et al., 2013), the so-called Ctenophora sister, then the radiation probably occurred after the divergence of ctenophores from sponges and the rest of animals (even though secondary loss in the lineage leading to extant ctenophores cannot be ruled out). In any case, it seems probable that an expansion of T-box genes occurred somehow early in animal evolution, followed by extensive neo- and sub-functionalization of T-box genes (see later). Below we revise the presence of the different T-box gene families in different animal phyla, plus their inferred function in some nonmodel species.

3.1 Brachyury

The Brachyury T-box class is the founding member of the family. It is present in nonmetazoan taxa and in the genomes of all major metazoan lineages, with the exception of the analyzed taxa representing Nematoda and Platyhelminthes (Fig. 1). In vertebrates, the Brachyury class diversified into the paralog subclasses Tbx19 and Brachyury. The evolution of Brachyury function is discussed later.

3.2 Eomes/Tbrain

The Eomes/Tbrain class is found in sponges, in particular the calcarean sponges *Sycon ciliatum and Leucosolenia complicata*, and in some protostomes

and deuterostomes. This class has been lost in Placozoa and Cnidaria, as well as in Ctenophora (under the sponges-first hypothesis), and in some bilaterian clades, such as Platyhelminthes, Arthropoda, and Nematoda. The Eomes class gave rise to the paralog subclasses Tbx21, Eomes, and Tbrain in vertebrates.

In the calcarean sponge *S. ciliatum*, Eomes is strongly expressed in the oocytes, but not during development. In adult sponges, Eomes is expressed around the opening of the oscular sphincter (Leininger et al., 2014). In the annelid *Hydroides elegans*, Eomes is expressed in animal cap blastomeres and, later, it becomes restricted to the apical tuft cells of the early trochophore larvae (Arenas-Mena, 2008). This pattern of expression in the apical region seems conserved in the tornaria larvae of the hemichordate *Ptychodera flava* (Tagawa, Humphreys, & Satoh, 2000). The authors suggested that this expression pattern might represent an evolutionary link between the apical sensory organ of nonchordate larvae and the vertebrate forebrain. In *Ptychodera*, Eomes is also expressed weakly during development around the blastopore. In the embryo of the sea urchin *Strongylocentrotus purpuratus*, Eomes (ske-T) is expressed only in the skeletogenetic mesenchyme (Croce, Lhomond, Lozano, & Gache, 2001). Finally, in the cephalochordate *Branchiostoma floridae*, Eomes is expressed in the axial and paraxial mesendoderm in early larvae, but no anterior neural domain of expression (similar to that in vertebrate forebrain) is detected in amphioxus (Horton & Gibson-Brown, 2002).

3.3 Tbx7

The Tbx7 class is present in the unicellular filastereans and ichthyosporeans (see earlier), as well as in sponges. Among sponges, Tbx7 is present in the genomes of the calcarean sponges *S. ciliatum* and *L. complicata* and the demosponge *Amphimedon queenslandica*. Therefore, this T-box class was secondarily lost in all other metazoans, under the Porifera-sister hypothesis.

In *Sycon*, Tbx7 is expressed in oocytes and during cleavage and early-stage preinversion embryos. In adult sponges, it is found in cells of the mesohyl (the space between the external pinacoderm and the internal choanoderm, filled with extracellular matrix), particularly at the tips of the uppermost radial chambers (Leininger et al., 2014).

3.4 Tbx8

The Tbx8 class is found in diverse sponges (the demosponges *A. queenslandica* and *Cliona celata*, the calcareans *S. ciliatum* and *L. complicata*, and the homoscleromorph *Oscarella carmela*), placozoans, cnidarians, platyhelminthes,

mollusks, cephalochordates, and ambulacrarians (Echinodermata and Hemichordata). This class was secondarily lost in ctenophores (under the Porifera-sister hypothesis), annelids, tunicates, and vertebrates, and all the ecdysozoans examined (Nematoda and Arthropoda). It is one of the only four families present in Platyhelminthes, together with Tbx2/3, Tbx1, and Tbx20 (these three previously identified in the planarian *Schmidtea polychroa* (Martín-Durán & Romero, 2011)).

3.5 Tbx2/3

This is the most widespread T-box family in animals (Fig. 1), present in all the major lineages examined. The class diverged into Tbx2 and Tbx3 at the root of vertebrates. Unlike most T-box TFs, Tbx2/3 class contains the only examples of T-box acting as transcriptional repressors (He, Wen, Campbell, Wu, & Rao, 1999).

Tbx2/3 expression patterns have been studied for several species. In the demosponge *Suberites domuncula*, Tbx2/3 is expressed in isolated cells of the mesohyl of adults, suggesting a possible role in terminal cell-type differentiation (Adell & Müller, 2005). In the placozoan *Trichoplax adhaerens*, Tbx2/3 is expressed in the periphery of attached animals (Martinelli & Spring, 2003). In the ctenophore *Mnemiopsis leyidi*, Tbx2/3 is expressed in the ctene rows and the apical organ region, suggesting a possible common role in sensory organ formation, as Tbx2/3 is also expressed in the eye of *Drosophila* and chordates (Yamada et al., 2007).

Among bilaterians, the planarian *S. polychroa* (Platyhelminthes) has three Tbx2/3 paralogs that show distinct expression patterns in the embryo (Martín-Durán & Romero, 2011): Tbx2/3a is expressed in the gut, Tbx2/3b in parenchymatic cells (both dorsal and ventral), and Tbx2/3c is found in the embryonic brain. Both in the hemichordate *S. kowalevski* and the annelid *H. elegans*, Tbx2/3 is expressed in the dorsal side of the embryo (Arenas-Mena, 2013; Lowe et al., 2006) and, similarly, Tbx2/3 is expressed in the aboral side of different sea urchin species, including *Paracentrotus lividus* (Croce, Lhomond, & Gache, 2003), *S. purpuratus* (Chen, Luo, & Su, 2011), and *Lytechinus variegatus* (Gross, Peterson, Wu, & McClay, 2003). These findings suggest a conserved role of Tbx2/3 in dorsoventral patterning in bilaterians (Arenas-Mena, 2013).

3.6 Tbx4/5

The Tbx 4/5 class, in contrast, is sparsely distributed due to secondary losses in all protostomes analyzed plus in ctenophores and in ambulacrarians

(echinoderms + hemichordates). It is present in sponges, placozoans, cnidarians, cephalochordates, and vertebrates (Fig. 1). The Tbx4/5 class diversified into Tbx4 and Tbx5 subclasses in vertebrates. Nothing is known about the expression patterns of Tbx4/5 in nonbilaterians. In amphioxus, it has been shown to be expressed only in the decentralized cardiac domain of the adults (and not during development), suggesting a common role of Tbx4/5 in cardiogenesis in cephalochordates and vertebrates (Pascual-Anaya et al., 2013).

3.7 Tbx1/15/20

The ancestral Tbx1/15/20 class is present in ctenophores and sponges, and it diversified into three classes (Tbx1, Tbx15, and Tbx20) at the root of Cnidaria + Bilateria. Tbx1 and Tbx20 are present in all major cnidarian and bilaterian lineages examined here, while Tbx15 has been lost in platyhelminthes, annelids, nematodes, and ambulacrarians (and also in some specific taxa in other groups, see Supplementary Table 1). At the root of vertebrates, Tbx1 further diversified into Tbx1 and Tbx10, and Tbx20 diversified into Tbx18, Tbx20, and Tbx22. Interestingly, this Tbx20 diversification occurred after the divergence of the lamprey lineage.

In the ctenophore *M. leyidi*, Tbx1/15/20 shows mesendodermal expression and transient expression along the edge of the blastopore in a biradial pattern (Yamada et al., 2007). In the planarian *S. polychroa*, Tbx1 is localized in discrete dorsal cells, while Tbx20 is in the body margin and in the ventral nerve cords (Martín-Durán & Romero, 2011). Finally, in amphioxus Tbx15 is expressed in the mesendoderm during the gastrula stage and, later, in the forming somites, suggesting a conserved role in chordate segmentation (Beaster-Jones, Horton, Gibson-Brown, Holland, & Holland, 2006). In contrast, amphioxus Tbx20 is expressed, like Tbx4/5, in the precursors of the myocardium, suggesting a conserved role in heart development in chordates (Belgacem, Escande, Escriva, & Bertrand, 2011).

3.8 Tbx6

Finally, the Tbx6 class has been traditionally difficult to identify by phylogeny, with many genes presumptively classified as Tbx6 not grouping together in phylogenies (Holstien et al., 2010; Larroux et al., 2008; Sebé-Pedrós et al., 2013). With the addition of new species, we here recover the monophyly of this class (although with weak nodal support) and, for the first time, assign sponge and ctenophore sequences to this group. Thus, we push the origin of this class to the root of Metazoa. This expanded Tbx6 class includes sponge sequences that were traditionally classified within the

sponge-specific TbxPor class (Holstien et al., 2010; Sebé-Pedrós et al., 2013). Additionally, the Tbx6 class includes sequences from nematodes and arthropods (like the *Drosophila* Dorsocross (Doc) genes) and from tunicates and vertebrates. Some cnidarian and bilaterian genes were previously considered as Tbx6 in other studies (Belgacem et al., 2011; Paps, Holland, & Shimeld, 2012; Pascual-Anaya et al., 2013), but in our analysis they do not clearly cluster within Tbx6.

The Tbx6 class diversified into three subclasses in vertebrates: Tbx6, MGA, and VegT, the latter being secondarily lost in mammals. The vertebrate MGA subclass contains genes with both T-domain and a basic bHLH zipper domain (Hurlin et al., 1999), being, together with the *Capsaspora* Double-T-box (see earlier), the only T-box genes having additional DNA-binding domains.

In the ctenophore *M. leyidi*, Tbx6 (TbxD) functions in ectodermal development of the tentacles (Yamada et al., 2007). In amphioxus, Tbx6 is expressed in the tail epidermis, in some neurons, and in the unsegmented paraxial mesoderm, suggesting a conserved role in posterior mesoderm specification in chordates (Belgacem et al., 2011).

In summary, the T-box TF family has a highly dynamic evolutionary history, with multiple secondary losses along evolution (with the exception of Tbx2/3, present in all metazoan lineages), some fast-evolving members (for example, in sponges and ichthyosporeans), expansions (such as three paralogous eumetazoan classes related to the ancestral Tbx1/15/20), and major structural rearrangements, such as the double T-domain found in *Capsaspora* or the T-domain/bHLH domain fusion in MGA T-box subclass in vertebrates.

4. FUNCTIONAL CONSERVATION OF PREMETAZOAN AND EARLY-METAZOAN BRACHYURY HOMOLOGS

An intriguing question is whether there is some conserved function of any of the T-box classes between premetazoans and metazoans or between bilaterian and nonbilaterian animals. Heterologous experiments, in which a gene is expressed in a different species, have been used to analyze the evolutionary conservation of T-box genes. For example, Satoh et al. showed that ectopic overexpression of different deuterostome (i.e., tunicate, amphioxus, acorn worm, and sea urchin) Brachyury orthologs in the embryos of the tunicate *Ciona intestinalis* had similar effects in inducing the differentiation of notochord cells (Satoh, Harada, & Satoh, 2000). This indicated high

conservation of Brachyury target specificity in between these taxa. In another study, Marcellini et al. used overexpression in *Xenopus* animal caps to demonstrate the specific conserved ability to induce mesoderm of even more evolutionarily distant Brachyury orthologs, including those from the annelid *Platynereis dumerilii* and the cnidarian *H. magnipapillata* (Marcellini, Technau, Smith, & Lemaire, 2003).

An alternative to ectopic overexpression is the use of dominant-negative constructs to analyze the repressive phenotype. For example, *Xenopus* embryos injected with an mRNA encoding a dominant-negative form of Brachyury (XBra_En, consisting of a C-terminal fusion of the Engrailed repressor to Brachyury) showed defective gastrulation and impairment of muscle development (Conlon, Sedgwick, Weston, & Smith, 1996). Using this approach, Yamada et al. studied the functional conservation of the Brachyury ortholog of the ctenophore *M. leyidi* (Yamada, Martindale, Fukui, & Tochinai, 2010). In particular, they injected *Xenopus* embryos with a *M. leyidi*. Bra_En construct and this caused similar defects in gastrulation and mesoderm induction to those observed with the Xbra_En construct and, therefore, suggesting conservation of these distant Brachyury orthologs. Additionally, they analyzed the specific induction of downstream targets of XBra (wnt11, sox17) and also of targets of other T-box genes not activated by Brachyury (chordin, goosecoid) (Conlon, Fairclough, Price, Casey, & Smith, 2001; Xanthos, Kofron, Wylie, & Heasman, 2001). *M. leyidi* Brachyury specifically activated XBra targets, but not the non-XBra targets, revealing high conservation of target specificity between these two distant homologs.

The discovery of T-box genes, and in particular of Brachyury orthologs, outside Metazoa prompted the study of the functional conservation of these nonmetazoan Brachyury genes. To this end, we used coinjection of a dominant-negative XBra_En together with Brachyury mRNA to show, quite surprisingly, that the Brachyury ortholog of the unicellular filasterean *C. owczarzaki* was able to rescue gastrulation and mesoderm induction in *Xenopus* embryos as efficiently as the endogenous Xbra (Fig. 2) (Sebé-Pedrós et al., 2013), thus roughly mimicking the endogenous *Xenopus* Brachyury function (Fig. 2).

Next we evaluated the induction, upon mRNA injection, of downstream targets of XBra (sox17, endodermin, wnt11, and wnt8) and also targets of XVegT (Tbx6) that are known not to be recognized by XBra (such as chordin and pintallavis). Interestingly, the premetazoan *C. owczarzaki* Brachyury ortholog strongly activated all examined T-box targets, not only

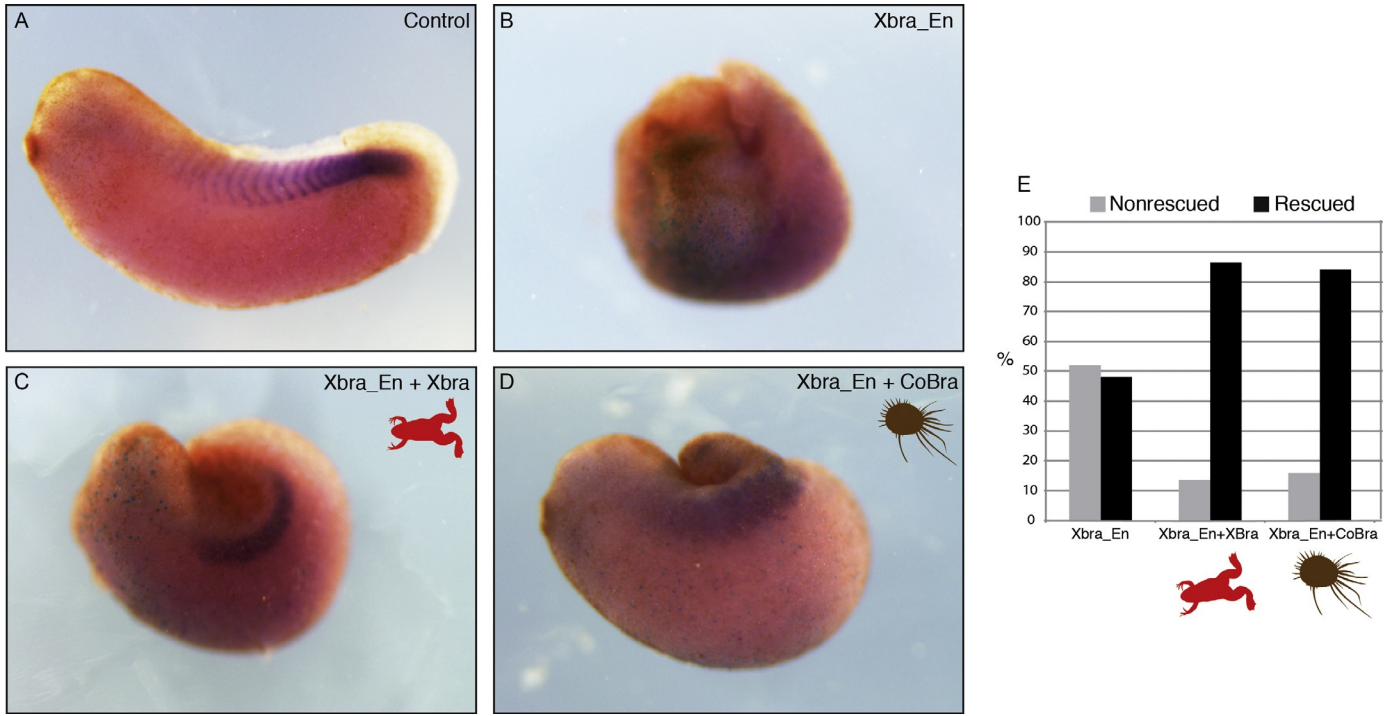

Fig. 2 Rescue assays in *Xenopus laevis* with the *Capsaspora owczarzaki* Brachyury ortholog. (A) Wild-type embryo showing complete trunk formation and full MyoD expression. (B) Embryo injected with Xbra_En (dominant-negative construct), showing no trunk formation and no MyoD expression. (C) Xbra_En-injected embryo rescued by coinjection with endogenous Xbra. (D) Xbra_En-injected embryo rescued by coinjection with *Capsaspora* Brachyury. (E) Barplot summarizing the results of the different control and rescue experiments. *Adapted from Sebé-Pedrós, A., Ariza-Cosano, A., Weirauch, M. T., Leininger, S., Yang, A., Torruella, G., … Ruiz-Trillo, I. (2013). Early evolution of the T-box transcription factor family*. Proceedings of the National Academy of Sciences of the United States of America, 110*(40), 16050–16055. doi:10.1073/pnas.1309748110.*

those XBra-specific. In contrast, Brachyury orthologs of the calcarean sponge *S. ciliatum* and the cnidarian *Nematostella vectensis* fully mimic the specific behavior of endogenous XBra, e.g., not activating chordin (Fig. 3). Protein-binding microarray analysis showed, though, that the DNA-binding motif preference of *C. owczarzaki* Brachyury and those of animals is the same and, in fact, that different T-box families have very similar binding motifs (Fig. 4). This similarity in the binding motifs could explain partial

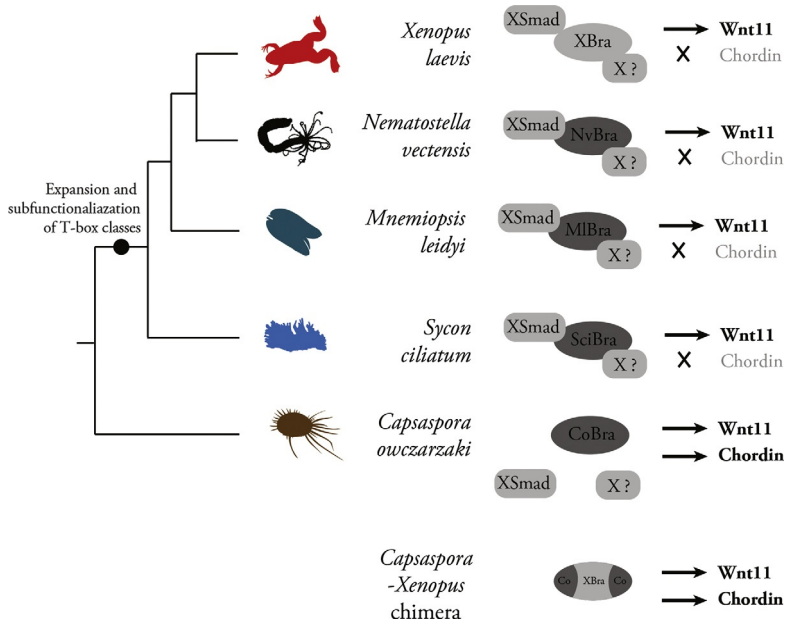

Fig. 3 Functional conservation of Brachyury orthologs in heterologous expression assays. Early-metazoan Brachyury orthologs (from the sponge *Sycon ciliatum* and the cnidarian *Nematostella vectensis*) produce the same molecular phenotype as *Xenopus* Brachyury (activation of Wnt11 and no activation of Chordin). In contrast, nonmetazoan Brachyury orthologs (from the filasteran *Capsaspora owczarzaki*) activate targets of multiple T-box classes, not only of *Xenopus* Brachyury. A chimeric construct with the N- and C-terminal domains (involved in protein–protein interactions) of CoBra and the central T-domain (involved in DNA binding) of XBra shows the same molecular phenotype as wild-type CoBra. These results suggest that the *trans* regulatory interactions between Brachyury and cofactors like Smad (and probably other unknown cofactors, shown as "X?") were established at the onset of Metazoa. *Results from Sebé-Pedrós, A., Ariza-Cosano, A., Weirauch, M. T., Leininger, S., Yang, A., Torruella, G., … Ruiz-Trillo, I. (2013). Early evolution of the T-box transcription factor family.* Proceedings of the National Academy of Sciences of the United States of America, 110*(40), 16050–16055. doi:10.1073/pnas.1309748110.*

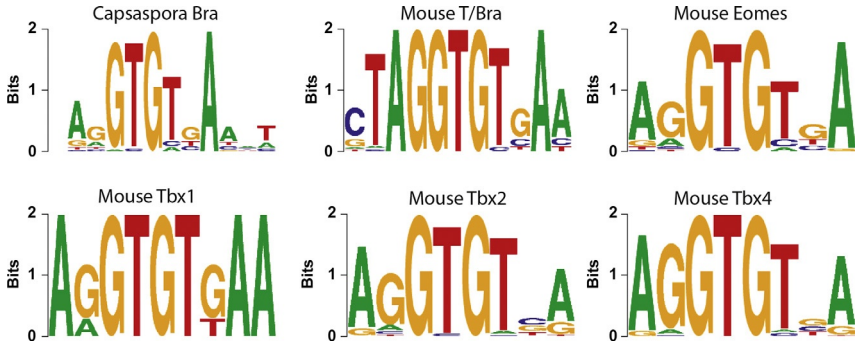

Fig. 4 Highly conserved T-box binding motifs. Protein-binding microarray experiments reveal that the DNA-binding preferences of the mouse and *Capsaspora* Brachyury orthologs are almost identical. Also different mouse T-box classes show similar motifs, indicating conservation across the whole T-box transcription factor family. *Adapted from Sebé-Pedrós, A., Ariza-Cosano, A., Weirauch, M. T., Leininger, S., Yang, A., Torruella, G., … Ruiz-Trillo, I. (2013). Early evolution of the T-box transcription factor family.* Proceedings of the National Academy of Sciences of the United States of America, 110*(40), 16050–16055. doi:10.1073/pnas.1309748110.*

inter-T-box class functional conservation, for example, that observed in rescue experiments (Croce et al., 2003). Croce et al. found that a dominant-negative Tbx2/3 (*coquillette*) in the sea urchin *P. lividus* (Echinodermata) could be rescued not only by coinjecting wild-type Tbx2/3 mRNA but also partially by Brachyury and Eomes mRNA.

Moreover, induction experiments using *C. owczarzaki*–*Xenopus* Brachyury chimeras (i.e., swapping N-terminal, central T- and C-terminal domains) showed that the N- and C-terminal domains, but not the central T-domain, are responsible for the specificity of Bra function (Fig. 3). In line with previous studies (Bielen et al., 2007; Marcellini, 2006; Marcellini et al., 2003), this result indicates that N- and C-terminal regions are essential for Brachyury specificity, by mediating cofactor interactions, e.g., with Smad proteins (Marcellini, 2006).

In summary, the heterologous expression studies of distant Brachyury orthologs suggest that subfunctionalization of Brachyury class was well established at the onset of Metazoa, as evidenced by the ability of sponge and ctenophore Brachyury to mimic endogenous *Xenopus* Brachyury function. The establishment of new cofactor interactions was, probably, an important mechanism in this subfunctionalization process, which occurred concomitantly to the radiation of T-box classes at the root of Metazoa.

5. ANCESTRAL CONSERVED ROLE OF BRACHYURY IN MORPHOGENETIC MOVEMENTS

Heterologous expression experiments tell us about the ability of distant orthologs to mimic endogenous functions. This indicates a certain level of biochemical functional conservation, but tells us nothing about the gene function in its original context. For this, it is necessary to examine expression patterns in developing and adult animals and, when possible, perform perturbation experiments. Since the discovery of Brachyury expression in invaginating hindgut cells in *Drosophila* (Kispert et al., 1994), multiple authors, working on different species, have proposed a conserved role for Brachyury in morphogenetic movements (Gross & McClay, 2001; Scholz & Technau, 2003; Tada & Smith, 2000; Tagawa, Humphreys, & Satoh, 1998; Yamada et al., 2010, 2007). Here, we review the current evidence about Brachyury function in different lineages.

In sponges, Brachyury function has been studied only in two species. Adell et al. used immunostainings to evaluate the expression of Brachyury in different culture stages of the demosponge *S. domuncula* (Adell & Müller, 2005). The highest levels of Brachyury protein were detected in adherent aggregates of cells, after sponge dissociation. Hence, the authors proposed a possible role for *S. domuncula*-Bra in morphogenetic movements, through the regulation of cell motility and adhesion.

In the calcarean sponge *S. ciliatum*, the two Brachyury paralogs are expressed in the oocytes and in the micromere cells of postinversion embryos (Leininger et al., 2014). In adults of *S. ciliatum*, Brachyury is expressed in the choanocytes. The authors propose that this expression pattern (together with other markers) indicates a possible homology of the micromers/choanoderm and eumetazoan endomesoderm (Leininger et al., 2014).

In ctenophores, Brachyury is expressed around the blastopore in *M. leyidi* (Yamada et al., 2007), after the blastopore is formed. It is also found in cells of the invaginating tentacular bulbs and in the floor of the apical organ. In another key study, Yamada et al. injected *M. leyidi* embryos with morpholino to knockdown Brachyury, which resulted in a failure to invaginate the ectodermal cells surrounding the blastopore (Yamada et al., 2010). This effect was rescued by coinjection of the morpholino with endogenous Brachyury mRNA. The results of this study strongly suggest a primitive conserved role of Brachyury in morphogenetic movements, such as those involved in gastrulation.

In placozoans, Brachyury is expressed in discrete groups of cells in the periphery of adult animals, suggesting a role in cell differentiation (Martinelli & Spring, 2003).

In cnidarians, the role of Brachyury has been extensively studied in several species. In the anemone *N. vectensis*, Brachyury is expressed around the blastopore in gastrulating embryos and in the developing mesenteries, although not in the adult muscle cells (Scholz & Technau, 2003). In contrast, in *Podocoryne carnea*, Brachyury is expressed in the adult muscle cells (Spring et al., 2002). During the development of the coral *Acropora digitifera*, Brachyury is expressed around the blastopore, in cells moving inward during gastrulation (Hayward, Grasso, Saint, Miller, & Ball, 2015). Finally, *H. magnipapillata* has two Brachyury paralogs that are differentially expressed (Bielen et al., 2007; Technau & Bode, 1999). HyBra1 is expressed in the endoderm of the hypostome, the tissue surrounding the adult mouth, while HyBra2 is expressed in the ectoderm of the hypostome. *H. magnipapillata* lacks classical gastrulation with a well-defined blastopore (instead, embryos of *H. magnipapillata* undergo multipolar ingression); therefore, circumblastoporal expression of Brachyury cannot be evaluated.

In multiple bilaterian lineages, the circumblastoporal expression of Brachyury is conserved, for example, in annelids (Lartillot, Lespinet, Vervoort, & Adoutte, 2002), echinoderms (Croce, Lhomond, & Gache, 2001; Gross & McClay, 2001; Peterson, Harada, Cameron, & Davidson, 1999; Rast, Cameron, Poustka, & Davidson, 2002), hemichordates (Lowe et al., 2006; Tagawa et al., 1998), priapulids (Martín-Durán, Janssen, Wennberg, Budd, & Hejnol, 2012), and cephalochordates (Onai et al., 2009). In the annelid *H. elegans* Brachyury is expressed in the invaginating blastomers that lead to gastrulation (Arenas-Mena, 2013). Later in development, Brachyury expression is retained in the hindgut and/or the foregut in different lineages (reviewed by Hejnol & Martín-Durán, 2015). For example, Brachyury is expressed both in mouth and anus in echinoderms, hemichordates, annelids, and molluscs, but only in the hindgut in arthropods and priapulids (nematodes have lost Brachyury) (Hejnol & Martín-Durán, 2015). Brachyury is also known to be a key regulator of notochord development in tunicates and cephalochordates (Katikala et al., 2013; Onai et al., 2009). Further support for the role of Brachyury in morphogenetic movements comes from key experiments by Gross et al. in the sea urchin *L. variegatus* (Gross & McClay, 2001). Like in cnidarians, ctenophores, and many bilaterians, LvBra is expressed around the blastopore and, in later stages, in the stomodeum and the anal region of the

pluteus larva hindgut. Interestingly, blocking of Brachyury function (by injection of a dominant-negative LvBra_EN construct) completely abolished gastrulation movements, but it did not affect the expression of endodermal and mesodermal marker genes. Similarly, morpholino knockdown of Brachyury also blocks gastrulation in another sea urchin species, *S. purpuratus* (Rast et al., 2002).

What about the function of Brachyury in nonmetazoans? A recent comparative transcriptomic analysis of two life stages of the ichthyosporean *Creolimax fragrantissima* showed that several of the highly divergent Brachyury paralogs in this species were upregulated in the amoeboid dispersal stage (compared with the multinucleated, cell-walled, osmotrophic stage) (De Mendoza, Suga, Permanyer, Irimia, & Ruiz-Trillo, 2015). Additionally, in an analysis of the regulatory genome of the filasterean *C. owczarzaki*, the downstream network of Brachyury was inferred (Fig. 5) (Sebé-Pedrós et al., 2016). Interestingly, multiple gene orthologs are conserved between the *Capsaspora* Brachyury and the mouse Brachyury downstream target networks (Lolas, Valenzuela, Tjian, & Liu, 2014). These conserved orthologs are enriched in functions associated to cell motility, amoeboid movement, and actin cytoskeleton (Fig. 5). This result points to an ancestral role for Brachyury in regulating a core network of genes associated with cell motility, a function that was already present before the advent of animal multicellularity.

Overall, in the past two decades extensive evidence has accumulated in different animal species that support a conserved ancestral involvement of Brachyury in morphogenetic movement, with recent evidence even suggesting a premetazoan role of Brachyury in regulating cell motility. Moreover, we have seen examples of a myriad of additional roles of Brachyury in other developmental and adult contexts.

6. THE EVOLUTION OF T-BOX REGULATION

The evolution of TF function goes beyond the diversification of families/classes and the number of paralog members. Specific TF binding to DNA sites, either in enhancer or promoter elements, depends on multiple layers of regulation, including TF translocation to the nucleus, interaction of other TFs and cofactors, TF-binding affinities to specific sequences, and the chromatin context (nucleosome occupancy and modifications, DNA methylation, and chromatin folding) of these sequences (Spitz & Furlong, 2012). Evolutionary changes affecting TF function can be in *trans* or in *cis*. The ones

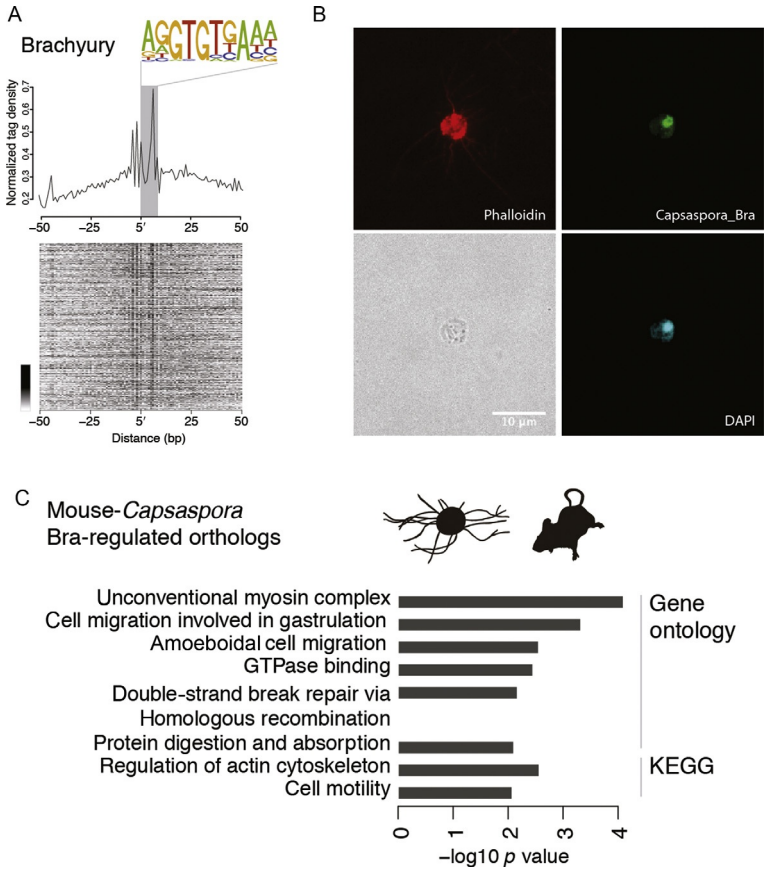

Fig. 5 A unicellular Brachyury regulatory network. (A) Plot of read density centered around *Bra* motifs (see Fig. 4) and heatmap of signal around individual sites in *Capsaspora* ATAC-seq experiments. (B) *Capsaspora* filopodial stage cell stained with phalloidin (*red*, actin cytoskeleton), DAPI (*blue*, nucleus) and *Capsaspora-Brachyury* antibody (*green*). Notice *Bra* localization in the nucleus. (C) Enriched GO terms and KEGG pathways among genes associated with *Bra* regulatory sites with shared orthologs regulated by *Bra* in mouse. *Adapted from Sebé-Pedrós et al. (2016).*

in *trans* affect the primary TF coding sequence, while the ones in *cis* affect regulatory sequences controlling the expression of the TF itself or changes in the TF DNA-binding sites in the genome, which results in new downstream targets controlled by a TF and, therefore, the rewiring of the TF network.

An example of *trans* changes in T-box function is the case of Brachyury cofactor interaction (Marcellini, 2006; Marcellini et al., 2003). It is known

that many T-box TFs act together with other TFs and cofactors. For example, in mammals, Tbx5 interacts through its T-domain with the homeobox TF Nkx2.5 and also Gata4 during cardiomyocyte differentiation, Tbx2 with Rb1 protein through its C-terminal region, and Tbx18 interacts with the TF Pax3 in the regulation of AP somite polarity (reviewed by Papaioannou, 2014). In the case of Brachyury, heterologous overexpression experiments (described earlier) suggested that Brachyury target specificity in metazoans arose through changes in cofactor interactions (such as Smad, and probably others), affecting both the N- and C-terminal domains of the protein. These interactions were probably established at the onset of Metazoa. Instead, a premetazoan Brachyury ortholog (that of *C. owczarzaki*) behaves as a "panT-box" gene, strongly activating downstream targets of different animal T-box classes (in this case, Brachyury and Tbx6). This *trans* regulatory change occurred concomitantly with the explosive diversification of T-box classes at root of Metazoa, resulting in the subfunctionalization of rapidly duplicated T-box TFs.

Trans changes can also affect the DNA-binding specificities of a TF, for example, restricting its ability to bind to particular sites (Hudson et al., 2015). An example in the T-box family is found in Eomes/Tbrain of echinoderms. Jarvela et al. reported that the Eomes orthologs of sea urchin (*S. purpuratus*) and sea star (*Patiria miniata*) have differences in their secondary binding motifs, which also differ from the secondary motifs in vertebrate Eomes (Jarvela et al., 2014). These differences likely derive from evolutionary changes in the DNA-contacting amino acids and result in important differences in the role of Eomes in the development of these two echinoderms: in the sea urchin Eomes functions in skeletogenesis (see earlier), while in the sea star Eomes has roles in the endomesoderm and also in the ectoderm (Jarvela et al., 2014).

An example of *cis* changes affecting the domain of expression of T-box TFs is found in Tbx4/5 of chordates (Minguillón, Gibson-Brown, & Logan, 2009). Tbx4/5 is expressed in the cardiac region of the cephalochordate amphioxus, suggesting a conserved role in myogenesis in chordates; while in vertebrates Tbx4 and Tbx5 paralogs have well-studied roles in limb development. Heterologous rescue experiments in mouse by Minguillón et al. showed that amphioxus Tbx4/5 is able to induce limb growth, suggesting that no major changes in Tbx4/5 proteins occurred during the vertebrate transition. Instead, the authors propose that the newly evolved *cis*-regulatory elements (in this case the LPM enhancer) changed the Tbx4/5 expression domain, providing the basis for the acquisition of paired appendages during

vertebrate evolution. A similar example of change in enhancer function is the one reported by Infante et al. in snake Tbx4 regulation. In snakes, this gene lost Tbx4 hindlimb expression through changes in the HLEB enhancer that is known to drive hindlimb and genitalia Tbx4 expression in mouse (Infante et al., 2015).

Finally, changes in the *cis*-regulatory target sequences throughout the genome can cause rapid evolution in the downstream network of a TF (Sorrells & Johnson, 2015). In a recent study, Lolas et al. analyzed the Brachyury downstream network in mouse (Lolas et al., 2014), and compared it to those known for zebrafish (Morley et al., 2009) and *Xenopus* (Gentsch et al., 2016). By comparing orthologous targets genes between species, they showed relatively little conservation (∼10–15% of the genes) in the Brachyury downstream network of these vertebrate species. This shows the rapid divergence to the Brachyury network, likely mostly through changes in the Bra *cis*-target sequences in the genome.

ACKNOWLEDGMENTS

We thank José M. Martín-Durán for critical discussion about T-box function in different bilaterian lineages and Meritxell Antó for help with the figures. This work was supported by an Institució Catalana de Recerca i Estudis Avançats contract, a European Research Council Consolidator grant (ERC-2012-Co-616960), and grant BFU2014-57779-P from Ministerio de Economía y Competitividad (MINECO) to I.R.-T., funded partly by Fondo Europeo de Desarrollo regional (FEDER). We also acknowledge financial support from Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya (Project 2014 SGR 619).

REFERENCES

Adell, T., Grebenjuk, V. A., Wiens, M., & Müller, W. E. G. (2003). Isolation and characterization of two T-box genes from sponges, the phylogenetically oldest metazoan taxon. *Development Genes and Evolution*, *213*(9), 421–434. http://dx.doi.org/10.1007/s00427-003-0345-5.

Adell, T., & Müller, W. E. G. (2005). Expression pattern of the Brachyury and Tbx2 homologues from the sponge Suberites domuncula. *Biology of the Cell/under the Auspices of the European Cell Biology Organization*, *97*(8), 641–650. http://dx.doi.org/10.1042/BC20040135.

Amemiya, C. T., Alföldi, J., Lee, A. P., Fan, S., Philippe, H., Maccallum, I., … Lindblad-Toh, K. (2013). The African coelacanth genome provides insights into tetrapod evolution. *Nature*, *496*(7445), 311–316. http://dx.doi.org/10.1038/nature12027.

Arenas-Mena, C. (2008). The transcription factors HeBlimp and HeT-brain of an indirectly developing polychaete suggest ancestral endodermal, gastrulation, and sensory cell-type specification roles. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, *310*, 567–576. http://dx.doi.org/10.1002/jez.b.21225.

Arenas-Mena, C. (2013). Brachyury, Tbx2/3 and sall expression during embryogenesis of the indirectly developing polychaete Hydroides elegans. *International Journal of Developmental Biology*, *57*(1), 73–83. http://dx.doi.org/10.1387/ijdb.120056ca.

Beaster-Jones, L., Horton, A. C., Gibson-Brown, J. J., Holland, N. D., & Holland, L. Z. (2006). The amphioxus T-box gene, AmphiTbx15/18/22, illuminates the origins of chordate segmentation. *Evolution & Development*, *8*(2), 119–129. http://dx.doi.org/10.1111/j.1525-142X.2006.00083.x.

Belgacem, M. R., Escande, M. L., Escriva, H., & Bertrand, S. (2011). Amphioxus Tbx6/16 and Tbx20 embryonic expression patterns reveal ancestral functions in chordates. *Gene Expression Patterns*, *11*(3–4), 239–243. http://dx.doi.org/10.1016/j.gep.2010.12.006.

Bielen, H., Oberleitner, S., Marcellini, S., Gee, L., Lemaire, P., Bode, H. R., … Technau, U. (2007). Divergent functions of two ancient Hydra Brachyury paralogues suggest specific roles for their C-terminal domains in tissue fate induction. *Development*, *134*(23), 4187. http://dx.doi.org/10.1242/dev.010173.

Bollag, R. J., Siegfried, Z., Cebra-Thomas, J. A., Garvey, N., Davison, E. M., & Silver, L. M. (1994). An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. *Nature Genetics*, *7*(3), 383–389.

Cavalier-Smith, T. (2003). Phylogeny of Choanozoa, Apusozoa, and other Protozoa and early eukaryote megaevolution. *Journal of Molecular Evolution*, *56*(5), 540–563. http://dx.doi.org/10.1007/s00239-002-2424-z.

Chang, Y., Wang, S., Sekimoto, S., Aerts, A., Choi, C., Clum, A., … Berbee, M. (2015). Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. *Genome Biology and Evolution*, *7*(6), 1590–1601. http://dx.doi.org/10.1093/gbe/evv090.

Charoensawan, V., Wilson, D., & Teichmann, S. a. (2010). Genomic repertoires of DNA-binding transcription factors across the tree of life. *Nucleic Acids Research*, *38*(21), 7364–7377. http://dx.doi.org/10.1093/nar/gkq617.

Chen, J.-H., Luo, Y.-J., & Su, Y.-H. (2011). The dynamic gene expression patterns of transcription factors constituting the sea urchin aboral ectoderm gene regulatory network. *Developmental Dynamics*, *240*(1), 250–260. http://dx.doi.org/10.1002/dvdy.22514.

Chipman, A. D., Ferrier, D. E. K., Brena, C., Qu, J., Hughes, D. S. T., Schröder, R., … Richards, S. (2014). The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede Strigamia maritima. *PLoS Biology*, *12*(11), e1002005. http://dx.doi.org/10.1371/journal.pbio.1002005.

Conlon, F. L., Fairclough, L., Price, B. M., Casey, E. S., & Smith, J. C. (2001). Determinants of T box protein specificity. *Development*, *128*(19), 3749–3758.

Conlon, F. L., Sedgwick, S. G., Weston, K. M., & Smith, J. C. (1996). Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. *Development*, *122*(8), 2427–2435.

Croce, J., Lhomond, G., & Gache, C. (2001a). Expression pattern of Brachyury in the embryo of the sea urchin Paracentrotus lividus. *Development Genes and Evolution*, *211*(12), 617–619. http://dx.doi.org/10.1007/s00427-001-0200-5.

Croce, J., Lhomond, G., & Gache, C. (2003). Coquillette, a sea urchin T-box gene of the Tbx2 subfamily, is expressed asymmetrically along the oral-aboral axis of the embryo and is involved in skeletogenesis. *Mechanisms of Development*, *120*(5), 561–572. S0925477303000224 [pii].

Croce, J., Lhomond, G., Lozano, J. C., & Gache, C. (2001b). ske-T, a T-box gene expressed in the skeletogenic mesenchyme lineage of the sea urchin embryo. *Mechanisms of Development*, *107*(1–2), 159–162. S0925477301004701 [pii].

De Mendoza, A., Suga, H., Permanyer, J., Irimia, M., & Ruiz-Trillo, I. (2015). Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *eLife*, *4*, e08904.

Gentsch, G. E., Owens, N. D. L., Martin, S. R., Piccinelli, P., Faial, T., Trotter, M. W. B., … Smith, J. C. (2016). In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell Reports*, *4*(6), 1185–1196. http://dx.doi.org/10.1016/j.celrep.2013.08.012.

Gross, J. M., & McClay, D. R. (2001). The role of Brachyury (T) during gastrulation movements in the sea urchin Lytechinus variegatus. *Developmental Biology*, *239*, 132–147. http://dx.doi.org/10.1006/dbio.2001.0426.

Gross, J. M., Peterson, R. E., Wu, S.-Y., & McClay, D. R. (2003). LvTbx2/3: A T-box family transcription factor involved in formation of the oral/aboral axis of the sea urchin embryo. *Development*, *130*(9), 1989–1999.

Hayward, D. C., Grasso, L. C., Saint, R., Miller, D. J., & Ball, E. E. (2015). The organizer in evolution–gastrulation and organizer gene expression highlight the importance of Brachyury during development of the coral, Acropora millepora. *Developmental Biology*, *399*(2), 337–347. http://dx.doi.org/10.1016/j.ydbio.2015.01.006.

He, M., Wen, L., Campbell, C. E., Wu, J. Y., & Rao, Y. (1999). Transcription repression by Xenopus ET and its human ortholog TBX3, a gene involved in ulnar-mammary syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(18), 10212–10217. http://dx.doi.org/10.1073/pnas.96.18.10212.

Hejnol, A., & Martín-Durán, J. M. (2015). Getting to the bottom of anal evolution. *Zoologischer Anzeiger—A Journal of Comparative Zoology*, *256*, 61–74. http://dx.doi.org/10.1016/j.jcz.2015.02.006.

Herrmann, B. G., Labeit, S., Poustka, A., King, T. R., & Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. *Nature*, *343*(6259), 617–622.

Holstien, K., Rivera, A., Windsor, P., Ding, S., Leys, S. P., Hill, M., & Hill, A. (2010). Expansion, diversification, and expression of T-box family genes in Porifera. *Development Genes and Evolution*, *220*(9–10), 251–262. http://dx.doi.org/10.1007/s00427-010-0344-2.

Horton, A. C., & Gibson-Brown, J. J. (2002). Evolution of developmental functions by the eomesodermin, T-brain-1, Tbx21 subfamily of T-box genes: Insights from amphioxus. *The Journal of Experimental Zoology*, *294*(2), 112–121. http://dx.doi.org/10.1002/jez.10151.

Hudson, W. H., Kossmann, B. R., Mitchelle, I., De Vera, S., Chuo, S., Weikum, E. R., … Ortlund, E. A. (2015). Distal substitutions drive divergent DNA specificity among paralogous transcription factors through subdivision of conformational space. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(2), 1–6. http://dx.doi.org/10.1073/pnas.1518960113.

Hurlin, P. J., Steingrìmsson, E., Copeland, N. G., Jenkins, N. A., & Eisenman, R. N. (1999). Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif. *The EMBO Journal*, *18*(24), 7019–7028. http://dx.doi.org/10.1093/emboj/18.24.7019.

Infante, C. R., Mihala, A. G., Park, S., Wang, J. S., Johnson, K. K., Lauderdale, J. D., & Menke, D. B. (2015). Shared enhancer activity in the limbs and phallus and functional divergence of a limb-genital cis-regulatory element in snakes. *Developmental Cell*, *35*(1), 107–119. http://dx.doi.org/10.1016/j.devcel.2015.09.003.

Itzkovitz, S., Tlusty, T., & Alon, U. (2006). Coding limits on the number of transcription factors. *BMC Genomics*, 7, 239. http://dx.doi.org/10.1186/1471-2164-7-239.

Jarvela, A. M., Brubaker, L., Vedenko, A., Gupta, A., Armitage, B. A., Bulyk, M. L., & Hinman, V. F. (2014). Modular evolution of DNA-binding preference of a Tbrain transcription factor provides a mechanism for modifying gene regulatory networks. *Molecular Biology and Evolution*, *31*(10), 2672–2688. http://dx.doi.org/10.1093/molbev/msu213.

Katikala, L., Aihara, H., Passamaneck, Y. J., Gazdoiu, S., José-Edwards, D. S., Kugler, J. E., … Di Gregorio, A. (2013). Functional Brachyury binding sites establish a temporal read-out of gene expression in the Ciona notochord. *PLoS Biology*, *11*(10), e1001697. http://dx.doi.org/10.1371/journal.pbio.1001697.

Kispert, A., Herrmann, B. G., Leptin, M., & Reuter, R. (1994). Homologs of the mouse Brachyury gene are involved in the specification of posterior terminal structures in Drosophila, Tribolium, and Locusta. *Genes and Development*, *8*, 2137–2150. http://dx.doi.org/10.1101/gad.8.18.2137.

Larroux, C., Fahey, B., Liubicich, D., Hinman, V. F., Gauthier, M., Gongora, M., … Degnan, B. M. (2006). Developmental expression of transcription factor genes in a demosponge: Insights into the origin of metazoan multicellularity. *Evolution & Development*, *8*(2), 150–173. http://dx.doi.org/10.1111/j.1525-142X.2006.00086.x.

Larroux, C., Luke, G. N., Koopman, P., Rokhsar, D. S., Shimeld, S. M., Degnan, B. M., & Shimeld, M. (2008). Genesis and expansion of metazoan transcription factor gene classes. *Molecular Biology and Evolution*, *25*(5), 980–996. http://dx.doi.org/10.1093/molbev/msn047.

Lartillot, N., Lespinet, O., Vervoort, M., & Adoutte, A. (2002). Expression pattern of Brachyury in the mollusc Patella vulgata suggests a conserved role in the establishment of the AP axis in Bilateria. *Development*, *129*(6), 1411–1421.

Leininger, S., Adamski, M., Bergum, B., Guder, C., Liu, J., Laplante, M., … Adamska, M. (2014). Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans. *Nature Communications*, *5*, 3905. http://dx.doi.org/10.1038/ncomms4905.

Lolas, M., Valenzuela, P. D. T., Tjian, R., & Liu, Z. (2014). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 4478–4483. http://dx.doi.org/10.1073/pnas.1402612111.

Lowe, C. J., Terasaki, M., Wu, M., Freeman, R. M., Runft, L., Kwan, K., … Gerhart, J. (2006). Dorsoventral patterning in hemichordates: Insights into early chordate evolution. *PLoS Biology*, *4*(9), e291. http://dx.doi.org/10.1371/journal.pbio.0040291.

Manuel, M., Le Parco, Y., & Borchiellini, C. (2004). Comparative analysis of Brachyury T-domains, with the characterization of two new sponge sequences, from a hexactinellid and a calcisponge. *Gene*, *340*(2), 291–301. http://dx.doi.org/10.1016/j.gene.2004.07.021.

Marcellini, S. (2006). When Brachyury meets Smad1: The evolution of bilateral symmetry during gastrulation. *Bioessays*, *28*(4), 413–420. http://dx.doi.org/10.1002/bies.20387.

Marcellini, S., Technau, U., Smith, J., & Lemaire, P. (2003). Evolution of Brachyury proteins: Identification of a novel regulatory domain conserved within Bilateria. *Developmental Biology*, *260*(2), 352–361.

Martín-Durán, J. M., Janssen, R., Wennberg, S., Budd, G. E., & Hejnol, A. (2012). Deuterostomic development in the protostome Priapulus caudatus. *Current Biology*, *22*(22), 2161–2166. http://dx.doi.org/10.1016/j.cub.2012.09.037.

Martín-Durán, J. M., & Romero, R. (2011). Evolutionary implications of morphogenesis and molecular patterning of the blind gut in the planarian Schmidtea polychroa. *Developmental Biology*, *352*(1), 164–176. http://dx.doi.org/10.1016/j.ydbio.2011.01.032.

Martinelli, C., & Spring, J. (2003). Distinct expression patterns of the two T-box homologues Brachyury and Tbx2/3 in the placozoan Trichoplax adhaerens. *Development Genes and Evolution*, *213*(10), 492–499. http://dx.doi.org/10.1007/s00427-003-0353-5.

Martinelli, C., & Spring, J. (2005). T-box and homeobox genes from the ctenophore Pleurobrachia pileus: Comparison of Brachyury, Tbx2/3 and Tlx in basal metazoans and bilaterians. *FEBS Letters*, *579*(22), 5024–5028. http://dx.doi.org/10.1016/j.febslet.2005.08.008.

Mendoza, L., Taylor, J. W., & Ajello, L. (2002). The class mesomycetozoea: A heterogeneous group of microorganisms at the animal-fungal boundary. *Annual Review of Microbiology*, *56*, 315–344. http://dx.doi.org/10.1146/annurev.micro.56.012302.160950.

Minguillón, C., Gibson-Brown, J. J., & Logan, M. P. (2009). Tbx4/5 gene duplication and the origin of vertebrate paired appendages. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(51), 21726–21730. http://dx.doi.org/10.1073/pnas.0910153106.

Morley, R. H., Lachani, K., Keefe, D., Gilchrist, M. J., Flicek, P., Smith, J. C., & Wardle, F. C. (2009). A gene regulatory network directed by zebrafish No tail accounts for its roles in mesoderm formation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(10), 3829–3834. http://dx.doi.org/10.1073/pnas.0808382106.

Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S., … Kohn, A. B. (2014). The ctenophore genome and the evolutionary origins of neural systems. *Nature*, *510*(7503), 109–114. http://dx.doi.org/10.1038/nature13400.

Onai, T., Lin, H. C., Schubert, M., Koop, D., Osborne, P. W., Alvarez, S., … Holland, L. Z. (2009). Retinoic acid and Wnt/beta-catenin have complementary roles in anterior/posterior patterning embryos of the basal chordate amphioxus. *Developmental Biology*, *332*(2), 223–233. http://dx.doi.org/10.1016/j.ydbio.2009.05.571. S0012-1606(09)00901-4 [pii]\r.

Papaioannou, V. E. (2001). T-box genes in development: From hydra to humans. *International Review of Cytology*, *207*, 1–70.

Papaioannou, V. E. (2014). The T-box gene family: Emerging roles in development, stem cells and cancer. *Development*, *141*, 3819–3833. http://dx.doi.org/10.1242/dev.104471.

Paps, J., Holland, P. W. H., & Shimeld, S. M. (2012). A genome-wide view of transcription factor gene diversity in chordate evolution: Less gene loss in amphioxus? *Briefings in Functional Genomics*, *11*(2), 177–186. http://dx.doi.org/10.1093/bfgp/els012.

Pascual-Anaya, J., Albuixech-Crespo, B., Somorjai, I. M. L., Carmona, R., Oisi, Y., Alvarez, S., … Garcia-Fernàndez, J. (2013). The evolutionary origins of chordate hematopoiesis and vertebrate endothelia. *Developmental Biology*, *375*(2), 182–192. http://dx.doi.org/10.1016/j.ydbio.2012.11.015.

Peterson, K. J., Harada, Y., Cameron, R. A., & Davidson, E. H. (1999). Expression pattern of Brachyury and Not in the sea urchin: Comparative implications for the origins of mesoderm in the basal deuterostomes. *Developmental Biology*, *207*(2), 419–431. http://dx.doi.org/10.1006/dbio.1998.9177.

Pick, K. S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D. J., Wrede, P., … Wörheide, G. (2010). Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Molecular Biology and Evolution*, *27*(9), 1983–1987. http://dx.doi.org/10.1093/molbev/msq089.

Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H., … Wörheide, G. (2015). Genomic data do not support comb jellies as the sister group to all other animals. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(50), 15402–15407. http://dx.doi.org/10.1073/pnas.1518127112.

Rast, J. P., Cameron, R. A., Poustka, A. J., & Davidson, E. H. (2002). Brachyury target genes in the early sea urchin embryo isolated by differential macroarray screening. *Developmental Biology*, *246*(1), 191–208. http://dx.doi.org/10.1006/dbio.2002.0654.

Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., & Leys, S. P. (2014). The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Molecular Biology and Evolution*, *31*(5), 1102–1120. http://dx.doi.org/10.1093/molbev/msu057.

Ruiz-Trillo, I., Roger, A. J., Burger, G., Gray, M. W., & Lang, B. F. (2008). A phylogenomic investigation into the origin of metazoa. *Molecular Biology and Evolution*, *25*(4), 664–672. http://dx.doi.org/10.1093/molbev/msn006.

Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A.-D., Moreland, R. T., Simmons, D. K., … Baxevanis, A. D. (2013). The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. *Science*, *342*(6164), 1242592. http://dx.doi.org/10.1126/science.1242592.

Satoh, G., Harada, Y., & Satoh, N. (2000). The expression of nonchordate deuterostome Brachyury genes in the ascidian Ciona embryo can promote the differentiation of extra notochord cells. *Mechanisms of Development*, *96*(2), 155–163.

Scholz, C. B., & Technau, U. (2003). The ancestral role of Brachyury: Expression of NemBra1 in the basal cnidarian Nematostella vectensis (Anthozoa). *Development Genes and Evolution*, *212*(12), 563–570. http://dx.doi.org/10.1007/s00427-002-0272-x.

Schulte-Merker, S., Ho, R. K., Herrmann, B. G., & Nusslein-Volhard, C. (1992). The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development (Cambridge, England)*, *116*(4), 1021–1032.

Sebé-Pedrós, A., Ariza-Cosano, A., Weirauch, M. T., Leininger, S., Yang, A., Torruella, G., … Ruiz-Trillo, I. (2013). Early evolution of the T-box transcription factor family. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(40), 16050–16055. http://dx.doi.org/10.1073/pnas.1309748110.

Sebé-Pedrós, A., Ballaré, C., Parra-Acero, H., Chiva, C., Tena, J. J., Sabidó, E., … Ruiz-Trillo, I. (2016). The dynamic regulatory genome of *Capsaspora* and the origin of animal multicellularity. *Cell*, *165*(5), 1224–1237. http://dx.doi.org/10.1016/j.cell.2016.03.034.

Sebé-Pedrós, A., de Mendoza, A., Lang, B. F., Degnan, B. M., & Ruiz-Trillo, I. (2011). Unexpected repertoire of metazoan transcription factors in the unicellular holozoan Capsaspora owczarzaki. *Molecular Biology and Evolution*, *28*(3), 1241–1254. http://dx.doi.org/10.1093/molbev/msq309.

Showell, C., Binder, O., & Conlon, F. L. (2004). T-box genes in early embryogenesis. *Developmental Dynamics*, *229*(1), 201–218. http://dx.doi.org/10.1002/dvdy.10480.

Simakov, O., Kawashima, T., Marlétaz, F., Jenkins, J., Koyanagi, R., Mitros, T., … Gerhart, J. (2015). Hemichordate genomes and deuterostome origins. *Nature*, *527*(7579), 459–465. http://dx.doi.org/10.1038/nature16150.

Smith, J. (1999). T-box genes: What they do and how they do it. *Trends in Genetics*, *15*(4), 154–158. http://dx.doi.org/10.1016/S0168-9525(99)01693-5.

Smith, J. J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M. S., … Li, W. (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nature Genetics*, *45*(4), 415–421. http://dx.doi.org/10.1038/ng.2568. 421e1–2.

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

Sorrells, T. R., & Johnson, A. D. (2015). Making sense of transcription networks. *Cell*, *161*, 714–723.

Spitz, F., & Furlong, E. E. M. (2012). Transcription factors: From enhancer binding to developmental control. *Nature Reviews. Genetics*, *13*(9), 613–626. http://dx.doi.org/10.1038/nrg3207.

Spring, J., Yanze, N., Jösch, C., Middel, A. M., Winninger, B., & Schmid, V. (2002). Conservation of Brachyury, Mef2, and Snail in the myogenic lineage of jellyfish: A connection to the mesoderm of bilateria. *Developmental Biology*, *244*(2), 372–384. http://dx.doi.org/10.1006/dbio.2002.0616.

Stott, D., Kispert, A., & Herrmann, B. G. (1993). Rescue of the tail defect of Brachyury mice. *Genes & Development*, *7*(2), 197–203.

Suga, H., & Ruiz-Trillo, I. (2013). Development of ichthyosporeans sheds light on the origin of metazoan multicellularity. *Developmental Biology*, *377*, 284–292. http://dx.doi.org/10.1016/j.ydbio.2013.01.009.

Tada, M., & Smith, J. C. (2000). Xwnt11 is a target of Xenopus Brachyury: Regulation of gastrulation movements via Dishevelled, but not through the canonical Wnt pathway. *Development*, *127*(10), 2227–2238.

Tagawa, K., Humphreys, T., & Satoh, N. (1998). Novel pattern of Brachyury gene expression in hemichordate embryos. *Mechanisms of Development*, *75*(1–2), 139–143. http://dx.doi.org/10.1016/S0925-4773(98)00078-1.

Tagawa, K., Humphreys, T., & Satoh, N. (2000). T-brain expression in the apical organ of hemichordate tornaria larvae suggests its evolutionary link to the vertebrate forebrain. *Journal of Experimental Zoology*, *288*(1), 23–31. http://dx.doi.org/10.1002/(SICI)1097-010X(20000415)288:1<23::AID-JEZ3>3.0.CO;2-H.

Technau, U., & Bode, H. (1999). HyBra1, a Brachyury homologue, acts during head formation in Hydra. *Development*, *126*, 999–1010.

Torruella, G., Derelle, R., Paps, J., Lang, B. F., Roger, A. J., Shalchian-Tabrizi, K., & Ruiz-Trillo, I. (2012). Phylogenetic relationships within the opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. *Molecular Biology and Evolution*, *29*(2), 531–544.

Venkatesh, B., Lee, A. P., Ravi, V., Maurya, A. K., Lian, M. M., Swann, J. B., … Warren, W. C. (2014). Elephant shark genome provides unique insights into gnathostome evolution. *Nature*, *505*(7482), 174–179. http://dx.doi.org/10.1038/nature12826.

Wilson, V., & Conlon, F. L. (2002). The T-box family. *Genome Biology*, *3*(6). reviews 3008.1–3008.7.

Xanthos, J. B., Kofron, M., Wylie, C., & Heasman, J. (2001). Maternal VegT is the initiator of a molecular network specifying endoderm in Xenopus laevis. *Development*, *128*(2), 167–180.

Yamada, A., Martindale, M. Q., Fukui, A., & Tochinai, S. (2010). Highly conserved functions of the Brachyury gene on morphogenetic movements: Insight from the early-diverging phylum Ctenophora. *Developmental Biology*, *339*(1), 212–222. http://dx.doi.org/10.1016/j.ydbio.2009.12.019.

Yamada, A., Pang, K., Martindale, M. Q., & Tochinai, S. (2007). Surprisingly complex T-box gene complement in diploblastic metazoans. *Evolution and Development*, *9*(3), 220–230. http://dx.doi.org/10.1111/j.1525-142X.2007.00154.x.